

NÚMERO DE ALUMNOS POR CURSO: UN ESTADO DEL ARTE

por

MONIKA EUGENIA DOCKENDORFF AGUILERA

Proyecto de Magíster presentado a la Facultad de Educación de la
Pontificia Universidad Católica de Chile para optar
al grado de Magíster en Educación, con Mención en
Dirección y Liderazgo Educacional

Profesor Guía: Gabriel Gutiérrez

Enero 2015

Santiago, Chile

AGRADECIMIENTOS

La presente Tesis es un trabajo del cual han participado muchas personas de forma directa o indirecta, leyendo, opinando, corrigiendo, explicando y sobre todo acompañándome y motivándome durante su desarrollo.

Agradezco especialmente a mi pareja Fernando Pérez por su constante e inquebrantable apoyo, que significó escuchar y opinar sobre prácticamente todas las secciones de este trabajo. Además estuvo siempre a mi lado cuando la motivación aflojaba y me brindó el apoyo emocional que más de alguna vez necesité.

A Gabriel Gutiérrez, mi profesor guía, por haber confiado en mi persona, por la paciencia y por la dirección de este trabajo. Además es el responsable de mi interés por este tema que surgió en el curso de *Políticas Educativas y Sistemas Educativos* el primer semestre de 2012.

A la fundación *Elige Educar* por haber financiado este Magíster además del Programa de Formación Pedagógica, lo que me permitió cambiar el rumbo y dedicarme a la educación, donde he encontrado mucha satisfacción y sentido. A Hernán Hochschild, su director ejecutivo, por interesarse en el tema, debatir sobre ésta y otras medidas, y abrir canales de comunicación e información en esta materia.

Al profesor Ernesto San Martín del Departamento de Estadística, Facultad de Matemáticas, quien fue entre otros mi profesor en Medición I y II, y que me brindó apoyo para comprender ciertos aspectos metodológicos asociados a inferencias múltiples.

A mis padres y hermana que me acompañaron en esta aventura que significó la Tesis y que, de forma incondicional, entendieron mis ausencias y mis malos momentos. A mi madre un reconocimiento de honor por intentar leerse el documento completo.

Gracias a todos.

Tabla de Contenidos

I.	INTRODUCCIÓN	6
1.1.	Relevancia: OCDE y el caso Chileno	7
II.	METODOLOGÍA.....	9
III.	ANTECEDENTES	11
3.1.	Número de alumnos por curso versus proporción de alumnos por docente	13
3.2.	Proyecto STAR (Student/Teacher Achievement Ratio)	15
3.2.1.	Programas de reducción del número de alumnos por curso en EEUU	18
IV.	PROBLEMAS METODOLÓGICOS	20
4.1.	Análisis Crítico del Proyecto STAR.....	21
4.1.1.	Efecto Hawthorne	22
4.2.	Diseños de Investigación Alternativos	23
4.3.	Fuentes de Sesgos	27
4.3.1.	Omisión de Variables.....	27
4.3.2.	Problema de Clasificación o <i>Sorting</i>	28
4.4.	Efectos en Grupos Minoritarios	35
4.4.1.	Composición del Curso	38
4.5.	Externalidades y Efectividad	40
4.6.	Validez y Capacidad Predictiva.....	43
V.	RESULTADOS INTERNACIONALES.....	45
5.1.	Estudios Comparativos a Gran Escala	46
5.2.	Estudios por Región.....	48
5.2.1.	Asia	48
5.2.2.	Latinoamérica.....	51
VI.	PROCESOS DE AULA	54
6.1.	El Docente en el Aula	57
6.1.1.	Calidad Docente y Eficacia	61
6.1.2.	Condiciones Laborales.....	67
6.2.	Conducta en el Aula	69
6.3.	Habilidades No Cognitivas y Efectos de Largo Plazo	76
6.3.1.	La Percepción de los Estudiantes	83

VII.	CONCLUSIONES	86
7.1.	Desafíos para la investigación en Chile	89
7.2.	Recomendaciones para la implementación de reducciones de alumnos por curso.....	91
VIII.	APÉNDICE: Dirección y Liderazgo Escolar.....	93
	Bibliografía	95

RESUMEN

Partiendo de la gran relevancia que ha tenido el impacto del número de alumnos por curso en el debate de política educativa internacional y sus futuras implicancias a nivel nacional, este trabajo describe y explica las principales dificultades y hallazgos en este ámbito. Basado en una amplia revisión de la literatura disponible, organiza y analiza la información proporcionando un ordenamiento que permite una comprensión global del tema. Primero intenta aclarar desde una perspectiva metodológica, las fuentes de discrepancia y confusión que han alimentado la gran cantidad de investigaciones con resultados inconclusos y mixtos. Se describen los principales hallazgos provenientes de las dos líneas de investigación que se han delineado en los últimos 30 años: la más antigua y principal establece relaciones causales entre la reducción del tamaño del curso y los resultados académicos medidos como puntajes de test. Un enfoque alternativo y más reciente examina el impacto de la medida en los procesos de aula, distinguiendo las variables que explican los beneficios de la intervención. Entendiendo que los cursos reducidos ofrecen mejores oportunidades de aprendizaje pero que no causan una mejor enseñanza, se identifican las circunstancias bajo las cuales esta intervención surte el mayor efecto. Finalmente se plantean algunas conclusiones y desafíos para el desarrollo de la investigación en Chile.

Palabras clave: número de alumnos por curso – resultados académicos – procesos de aula – reducción – oportunidades de aprendizaje.

ABSTRACT

Acknowledging the great relevance that class size has had in the international educational debate, and its forthcoming national implications, this paper describes and explains the main difficulties and findings in this field. Based on an extensive literature review, this work organizes and analyzes the information providing a structure that allows a global understanding of this topic. It first attempts to clarify from a methodological perspective, the sources of discrepancy and confusion that have led to a great deal of research with inconclusive and mixed results. It discusses the core outcomes delivered by the two lines of research that have arisen in the last 30 years: the primary approach has examined the relationship between class size and academic achievement, measured as test scores. An alternative and newer perspective analyzes the impact of class size on classroom processes, identifying the variables responsible for the benefits of the intervention. Understanding that smaller classes provide better learning opportunities, but they do not cause better teaching, this study identifies the circumstances under which this reform is most effective and draws conclusions and research challenges for the Chilean context.

Keywords: class size reduction – academic achievement – classroom processes – review – small classes – learning opportunities

I. INTRODUCCIÓN

A lo largo de las últimas décadas la expansión de los sistemas educativos y la universalización de la educación pública, se tradujo en la incorporación de niños y jóvenes provenientes de sectores históricamente excluidos a la educación inicial, primaria y secundaria, lo que ha generado nuevos desafíos en términos pedagógicos e institucionales. La efectividad del sistema en su totalidad se ha visto cuestionada dada la escasa movilidad social que ha generado la escuela pública, disputándose su calidad y su capacidad de generar aprendizajes significativos.

En 1966, el Informe sobre igualdad de oportunidades en educación de Coleman concluyó que no existían evidencias significativas del efecto que las variables de la escuela ejercían sobre el rendimiento de los estudiantes, es decir, las escuelas estarían reproduciendo las condiciones desiguales de origen de sus alumnos. En consecuencia, se instala la idea de que las escuelas públicas eran ineficientes y poco efectivas, y que allegar recursos adicionales para estos fines sería un desperdicio de fondos estatales. Sin embargo, el gasto en educación continuó incrementándose en las siguientes décadas.

Es en este contexto que se enmarcan numerosas investigaciones que examinan el fenómeno educativo como una función de producción y que buscan establecer el impacto de mayores recursos en el desempeño escolar¹. Políticos e investigadores han intentado identificar intervenciones específicas que permitan mejorar los resultados académicos de los estudiantes, como aumentar los sueldos de los docentes, mejorar su formación y reducir el tamaño de los cursos. En particular, la reducción de la cantidad de alumnos por curso y su impacto en resultados de aprendizaje fue objeto de múltiples estudios y análisis, en un intento por validar la presumida eficacia de la medida como mecanismo para incrementar los resultados de aprendizaje. Es así como se convierte en una de las áreas más investigadas en educación (Finn, 2002). Sin embargo, la inconsistencia de los resultados obtenidos, la escasa evidencia confiable de los beneficios que se le atribuyen junto a los altos costos que implica su implementación han generado posturas extremas en esta discusión, transformándose en uno de los tópicos más controvertidos y altamente politizados en el debate educativo internacional.

Por un lado la intuición y el sentido común sugieren que a menor cantidad de estudiantes por curso mejor será la educación que reciba cada uno de ellos. Los testimonios de profesores y profesionales de la educación son favorables a reducciones de alumnos por curso, y en su mayoría defienden esta medida como un mecanismo eficaz para mejorar los resultados de aprendizaje. Los cursos reducidos son también un aspecto muy valorado por padres y apoderados, convirtiéndolos en un factor relevante a la hora de elegir un establecimiento educativo para sus hijos. Esto se asocia a la noción de una enseñanza más individualizada y de mejor calidad, y es justamente una de las variables que utilizan algunos establecimientos educacionales de elite para agregar valor y diferenciarse. Por lo mismo, los programas de reducción de alumnos por curso en la educación

¹ Revisar el meta análisis de Greenwald, Hedges & Laine (1996) sobre el impacto de una variedad de recursos educativos en logros de aprendizaje.

pública generan una mayor confianza en el sistema². Además de lo anterior, se les atribuye la capacidad de generar mayores beneficios en los estudiantes más vulnerables, aspecto que los posiciona como una intervención doblemente favorable: por un lado, promueve mayores logros educativos en todos los estudiantes; por otro, beneficia más a los alumnos desaventajados, disminuyendo la brecha entre los de mejor y peor rendimiento, contribuyendo así a una mayor equidad educativa.

Es, sin embargo, una de las intervenciones educativas más onerosas, ya que requiere de mayor capital humano e infraestructura para ser implementada. El salario de los docentes y la cantidad de alumnos por curso son los principales factores contribuyentes del costo por estudiante y por consiguiente del sistema. Los programas de reducción de alumnos por curso van de la mano con los ciclos económicos, favoreciéndose su instauración en épocas de bonanza y sufriendo recortes durante períodos de recesión. Sus detractores arguyen que las millonarias inversiones que precisan dichas intervenciones no se condicen con los supuestos beneficios que han resultado difíciles de comprobar y cuya evidencia no es categórica. Además sus detractores perciben en general a la educación pública como ineficiente y poco efectiva, y consideran gastos adicionales de esta naturaleza como un despilfarro de fondos públicos. Aquellos contrarios a reducir alumnos por curso buscan promover alternativas de intervención más costo-efectivas que produzcan efectos similares³, sin embargo no ha surgido hasta ahora una propuesta alternativa lo suficientemente convincente y efectiva, que reemplace la tan popular y controvertida reducción de alumnos por curso.

1.1. Relevancia: OCDE y el caso Chileno

Analizar la situación actual de los países de la OCDE y las tendencias a nivel mundial que se han observado en el tamaño de los cursos, es relevante para Chile dado que desde 2009 forma parte de este exclusivo grupo que se ha convertido en su marco de referencia. Chile ha logrado excelentes índices de cobertura educativa y los problemas de infraestructura, de acceso a materiales y recursos educativos se han ido resolviendo en las últimas décadas, sin embargo se ubica dentro de los peor evaluados en términos de resultados. Hoy lo que está en juego es mejorar la calidad de la educación y el número de alumnos por curso ha sido un elemento importante de esa discusión en los países mas avanzados.

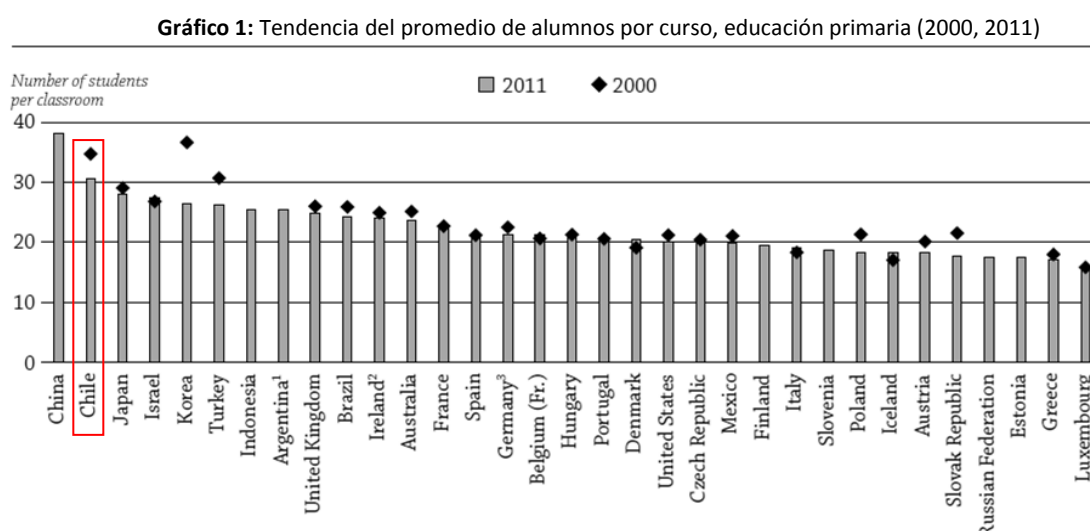
De acuerdo a los últimos datos entregados por la OCDE, el número de alumnos por curso promedio de los países miembros en 2011 era apenas superior a 21 estudiantes en Educación Básica. Este indicador aumenta levemente al incorporar a los países miembros del G20 como Argentina, Brasil, China e India. Considerando a todos los países con datos disponibles, en un extremo destacan China y Chile por tener más de 30 alumnos por curso en promedio, en

² A diferencia de reformas que implican exclusivamente cambios en el proceso de enseñanza-aprendizaje, las reducciones del número de alumnos por curso, son relativamente fáciles de evaluar en cuanto a su implementación.

³ Mejorar la formación y calidad docente, discrecionalidad del gasto adicional, mejoras de gobernanza y gestión y el aumento de la frecuencia de las evaluaciones son propuestas de mecanismos alternativos para mejorar los logros de aprendizaje.

contraposición a países como Luxemburgo que cuenta con menos de 16 alumnos por curso. Casi la mitad de los países tienen menos de 20 estudiantes promedio por curso. Los países que más disminuyeron este indicador en los últimos años son Corea y Turquía, que exhibían cursos muy numerosos y se alejaban considerablemente de los índices promedio de la OCDE.

De acuerdo a los indicadores entregados (OCDE, 2013) entre 2005 y 2011 el número de alumnos por curso en los países miembros con datos para ambos años disminuyó en un 18% en primaria y en un 6% en el nivel secundario inicial. El tamaño del curso promedio estimado para los países miembros es de 16,1 en primaria, 17,3 en el nivel secundario inicial y de 19,7 para el nivel secundario superior⁴. Tal como se observa en el gráfico 1, la tendencia muestra una clara disminución en la cantidad de alumnos por curso, que se puede explicar parcialmente por reformas educativas orientadas a reducir el tamaño de los cursos en este período.



En el caso particular de Chile, se observa una disminución en el tamaño de los cursos desde 1990 para todo el sistema educativo (Paredes & Pinto, 2009). Ésta ha sido especialmente notoria desde el 2000, producto de la baja sostenida de matrícula en el sector municipal, sin embargo, es posible advertir que la diferencia entre el sector subvencionado y el particular pagado es de al menos 10 alumnos por curso en promedio.

Tabla 1: Número de alumnos por curso, por nivel y dependencia, 2012, Chile. Elaboración propia.

Dependencia	alumnos por curso		
	Municipal	Particular Subvencionada	Particular Pagada
Educación Básica	33	38	24
Educación Media	37	35	26

⁴ Equivalencias para el caso chileno. Primaria: de 1° a 6° básico; secundaria inicial: de 7° Básico a I° Medio; secundaria superior: de II° a IV° Medio.

La tabla 1 fue construida a partir de datos entregados por el MINEDUC en su serie evidencias (2012) y caracteriza las diferencias por nivel educativo y dependencia a nivel nacional⁵. Si comparamos la situación chilena con la del resto de los países de la OCDE, sólo el sector particular pagado estaría acercándose a índices internacionalmente aceptables en relación a la cantidad de alumnos por curso. Esta significativa brecha es especialmente importante en el contexto nacional actual que comienza a implementar una gran reforma educativa, y que inicia un debate que busca establecer mecanismos para mejorar la calidad y la equidad del sistema educativo chileno. Como es de público conocimiento, el programa de gobierno 2014-2018 de la presidenta Michelle Bachelet señala la implementación de una *reducción gradual del número máximo de estudiantes por curso en la educación subvencionada* (Chile de Todos, 2014, p. 20). Es importante señalar que actualmente el número máximo de alumnos por curso, que se rige por el decreto supremo de educación Nº 8144 de 1980, alcanza los 45 estudiantes por curso. Además de lo anterior, institucionalmente el sistema de financiamiento -que entrega una subvención por alumno- incentiva a tener cursos lo más numerosos posible, primando el criterio económico por sobre el educativo⁶.

Conocer los hallazgos de la investigación en este ámbito, sus beneficios y limitaciones, así como incorporar el aprendizaje que ha dejado la implementación y ejecución de varios programas de reducción del número de alumnos por curso en el mundo, es fundamental para que esta medida sea evaluada en su mérito y pueda contribuir de forma efectiva a alcanzar los objetivos que la reforma se ha planteado.

II. METODOLOGÍA

La metodología de trabajo que se describe a continuación responde principalmente al objetivo de abordar el tema estudiado en toda su complejidad y desde las múltiples aristas que ella implica. Tradicionalmente el efecto del número de alumnos por curso se ha estimado desde una perspectiva de resultados académicos estandarizados y de corto plazo. El presente estudio intenta superar la mirada estrecha del impacto de esta intervención incorporando la mayor cantidad de evidencia posible que permita evaluar su potencial alcance en todas sus dimensiones. En consecuencia la literatura seleccionada fue amplia e incluyó estudios de diversa naturaleza, representativos de distintos contextos y realidades que pudieran aportar elementos respecto de los beneficios y limitaciones de la medida analizada.

Palabras clave: La selección bibliográfica que dio origen a este trabajo, se llevó a cabo inicialmente como una búsqueda sistemática e iterativa de artículos en la base de datos *Web of Science* (WoS) entre Julio de 2013 y Junio de 2014, utilizando las palabras clave *Class Size*, *Class Size Effect*, *Class Size Reduction*, *CSR*, *Achievement*, *Academic Outcome*, *Educational Attainment* y *Small Classes*.

⁵ Se utilizan las medianas y no el promedio que se ve muy distorsionado por valores extremos.

⁶ Más detalle en la sección 4.2. Diseños de Investigación Alternativos.

La revisión de esta selección inicial de la literatura permitió incorporar trabajos adicionales de autores que frecuentemente aparecían citados, los que constituyen el núcleo de la investigación en esta primera línea, incluyéndose también capítulos de libros, tesis y *papers* presentados en conferencias.

A partir de esta lectura preliminar, emergieron dos aproximaciones distintas a la investigación de los efectos del tamaño del curso: una más antigua que examina directamente la relación entre el tamaño del curso y los resultados de aprendizaje medidos como puntajes de test y la otra más reciente que vincula su impacto a los procesos de aula. En esta última línea los procesos pedagógicos y sus actores -profesores y alumnos- se tornan clave en el análisis. En consecuencia, la búsqueda de bibliografía posterior respondió a cubrir estos dos enfoques de manera equilibrada, incorporando palabras clave adicionales como: *Classroom Processes*, *Classroom Quality*, *Teacher Pupil Interaction*, *Teacher Efficacy*, *School Engagement*, y *Non Cognitive*. De esta forma aproximadamente un 50% de las publicaciones se concentran en procesos de aula o al menos delinean hipótesis respecto de aquellos mecanismos que se desencadenan en cursos reducidos para explicar los efectos de la medida estudiada. Asimismo se incluyeron estudios que relacionan la intervención al desarrollo de habilidades no cognitivas y a resultados de largo plazo, vinculados a esta segunda línea de investigación.

Filtros: Como filtro de selección se estableció el año 1999 por ser un año clave en la publicación de varios estudios. Fueron considerados criterios de inclusión como la especificación del tamaño de los cursos, la estimación del tamaño del efecto, el análisis desagregado por características del alumno, la rectificación de sesgos de medición y el rol de los distintos actores en la intervención. También fueron contemplados trabajos que examinan la experiencia recogida de la implementación de este tipo de intervención. Como criterio general de inclusión se consideraron las referencias en otras publicaciones. Se excluyeron artículos que se referían al tamaño del establecimiento y estudios que se centraban en la educación universitaria, así como trabajos exclusivamente enfocados en el costo de la intervención. Estudios de temáticas o resultados similares a los seleccionados también fueron descartados.

Enfoque: De acuerdo a su enfoque metodológico un 56% de las investigaciones utiliza una aproximación exclusivamente cuantitativa, un 32% emplea un enfoque mixto y el 12% restante es cualitativo. Los enfoques van directamente asociados a las distintas líneas de investigación, observándose una vinculación directa del enfoque cuantitativo a la estimación correlacional del tamaño del curso y los resultados de aprendizaje; por su parte, los enfoques mixto y cualitativo se asocian mayormente a trabajos sobre procesos de aula.

Escala: Respecto del impacto y la profundidad de los trabajos, seis de estas publicaciones corresponden a revisiones de la literatura en torno al tema de los efectos del número de alumnos por curso. Cinco investigaciones utilizan datos de pruebas estandarizadas internacionales como PISA y TIMSS para estimar los efectos de la medida en forma comparativa. Sin embargo, la gran mayoría de los estudios corresponden a estimaciones del impacto de esta intervención, en contextos locales y nacionales, a mediana y gran escala.

Ámbito: De los estudios publicados en revistas especializadas, 25 de ellos aparecen en publicaciones vinculadas al ámbito de la economía y 34 en revistas sobre educación, lo que revela la alta proporción de economistas en la investigación de este tema. Su presencia en el debate ha relevado y enfatizado temas fundamentalmente económicos como rentabilidad, incentivos y costos. Seis investigaciones examinan los beneficios de largo plazo que tiene la medida. Cabe mencionar que sólo 4 de los estudios han sido publicados en revistas sobre psicología, lo que por un lado confirma la innegable dimensión socio comunitaria del tema, pero por el otro evidencia la falta de desarrollo de modelos conceptuales que sirvan de evidencia para nutrir políticas educativas.

Origen: Dada la alta proporción de trabajos desarrollados en EEUU, se favoreció la incorporación de investigaciones de países y contextos culturales diversos. De las 78 publicaciones revisadas en profundidad en este trabajo, un 55% se originó en EEUU y sólo 35 provienen de otras regiones y países: principalmente de Europa (19), algunas del contexto latinoamericano (7), otras de Asia (6), Canadá (2) e Israel (1).

Adicionalmente a la concentración geográfica, se observa que el emblemático *experimento STAR*, centra la atención de la investigación en este ámbito constituyendo una referencia para prácticamente la totalidad de las investigaciones posteriores. Así, 14 estudios utilizan sus datos, ya sea para reportar sus resultados, para estimar efectos específicos o reestimar ciertos coeficientes con técnicas estadísticas más avanzadas.

III. ANTECEDENTES

Hasta mediados de los años '80 se habían llevado a cabo múltiples estudios sobre el efecto de la cantidad de alumnos por curso, sin embargo se caracterizaban por ser diseños no experimentales de investigación, por utilizar muestras más bien pequeñas y ser de corta duración, lo que ponía en duda sus hallazgos. Estos estudios buscaban validar la noción -más bien intuitiva- del beneficio que reporta estudiar en cursos reducidos en resultados académicos, sobre todo para los estudiantes con mayores dificultades y desventajas. Dos interesantes estudios resumen bien la situación que se vivía en regiones muy diversas del planeta: Canadá y Sudáfrica.

Por un lado el estudio experimental conducido en Toronto, Canadá destaca por su diseño, alcance y resultados (Shapson, Wright, Eason, & Fitzgerald, 1980). Fue llevado a cabo en la década de los '70 y se prolongó por dos años. Comparaba el rendimiento de los alumnos en cuatro tamaños de curso diferentes: de 16, 23, 30 y 37 estudiantes, en los niveles de 5° y 6° de Educación Básica. A pesar de que el desempeño en matemática fue débilmente mejor en cursos reducidos, sus resultados no fueron sistemáticos en las otras asignaturas evaluadas⁷. Los testimonios de profesores revelaron actitudes positivas hacia cursos reducidos. El experimento concluyó que los

⁷ Lectura, matemática, arte, inglés y auto percepción.

mayores logros de aprendizaje se explicaban por una *maximización del tiempo de instrucción* más que por cambios de prácticas pedagógicas, hipótesis clave del debate posterior.

Contrastante con la experiencia descrita para Canadá, es la situación extrema que vivía Sudáfrica en los últimos años del apartheid (1991-93), donde estudiantes de raza negra asistían a escuelas centralmente administradas por blancos en sectores residenciales limitados, con cursos numerosos de entre 20 y 80 alumnos por profesor. Las condiciones y los recursos con que contaban los estudiantes negros y blancos eran incomparables, y la inamovilidad residencial que enfrentaban las familias negras impedía la búsqueda de mejores opciones de escuela para esos hogares. Los resultados del estudio conducido por Case & Deaton (1999), muestran grandes diferencias en logros educativos entre estos dos grupos, que sin embargo resultan difíciles de atribuir exclusivamente a la cantidad de alumnos por curso⁸.

Pese a la escasa consistencia de los resultados de los estudios disponibles y los problemas metodológicos que presentaban, Glass & Smith (1978) publican un meta-análisis, estableciendo convergencia en cuatro proposiciones generales⁹:

- a) Es esperable que la disminución de alumnos por curso produzca aumentos en logros académicos.
- b) Los mayores beneficios se logran al disminuir la cantidad de alumnos por debajo de 20 estudiantes.
- c) Cursos pequeños reportan un mayor beneficio en matemática y lectura en niveles iniciales.
- d) Alumnos vulnerables y de grupos minoritarios tienen un mejor desempeño académico en clases reducidas.

Estas proposiciones fueron cuestionadas por detractores de programas de reducción de alumnos por curso al considerarlas inconsistentes con otras investigaciones que establecían una correlación inexistente o no significativa entre el número de alumnos por curso y el desempeño académico de éstos.

En los estudios disponibles de la época, la conceptualización de esta relación usualmente correspondía a una función de producción educativa, en que se estima la correlación entre los *inputs* (tamaño del curso) y los *outputs* (resultados académicos medidos como puntaje de test). Eric Hanushek, connotado investigador educativo y economista americano, contrario a intervenciones que requieran de un mayor gasto y promotor de soluciones estructurales y de cambios en la gobernanza de los establecimientos, en sucesivas revisiones literarias y monografías incluye una serie de estudios de resultados no concluyentes, como evidencia para desestimar este tipo de medida¹⁰. Por su parte Alan Krueger, economista, político e investigador ha estudiado y

⁸ Se controla por variables del hogar, pero se vislumbran problemas de clasificación que se discuten con posterioridad en la sección 4.3. Fuentes de Sesgos.

⁹ Extraído de Finn & Achilles (1999).

¹⁰ De 277 investigaciones revisadas señala que el 14% presenta impactos positivos y significativos, 14% muestra impactos negativos y significativos y el 72% restante sugiere efectos no significativos (Hanushek,

defendido las reducciones de alumnos por curso en varias oportunidades. En el documento, *The Class Size Debate*, Krueger y Hanushek confrontan argumentos a favor y en contra de la medida. Krueger alega desproporcionada ponderación de estudios con múltiples estimaciones y a su vez, Hanushek, discute que las investigaciones que defiende Krueger son de baja calidad (Mishel & Rothstein, 2002). Hanushek concluye que no existe un efecto sistemático entre la reducción del tamaño del curso y los resultados de aprendizaje, ya sea en países desarrollados o en vías de desarrollo¹¹.

La principal crítica que se le ha hecho a las revisiones bibliográficas de Hanushek, y que constituye la fuente inicial de confusión en este ámbito, es que utilizan la *proporción de alumnos por docente* como constructo evaluado de manera intercambiable con el concepto de *número de alumnos por curso*.

3.1. Número de alumnos por curso versus proporción de alumnos por docente

Una de las complicaciones más tempranas que aparece en la estimación del efecto del tamaño del curso sobre los logros de aprendizaje es la tendencia del uso de diferentes constructos como variable explicativa por grupos ideológicamente opuestos. La equiparación de estos dos conceptos ha distorsionado los resultados de las estimaciones, ya que estas dos medidas se calculan de manera diferente y tienen implicancias educativas distintas.

El *número de alumnos por curso* es la cantidad de estudiantes asignados a un curso y de los cuales el profesor es responsable durante un año. Está directamente relacionado con el tiempo que el docente destina a la instrucción y a fomentar la dedicación de sus alumnos al aprendizaje. La cantidad de alumnos en un curso limita las interacciones profesor-alumno que pueden ocurrir, el tiempo que se dedica a cada estudiante y la individualización de la enseñanza (Finn & Achilles, 1999, p. 104). Afecta también las interacciones entre estudiantes, el involucramiento a nivel social, el tipo de actividades que puede promover el docente y la asignación de tiempo a éstas. Determina asimismo las posibilidades de generar discusiones, debates abiertos y las posibilidades de retroalimentación. El número de estudiantes en un curso está vinculado directamente a la frecuencia de conductas disruptivas, las que inhiben el aprendizaje del grupo curso. Estas características establecen la incidencia de la acción educativa relativa al concepto de zona de desarrollo próximo, introducida por Levy Vygotski, que caracterizan el desarrollo potencial que un estudiante puede alcanzar con la ayuda de un adulto o persona más capaz. Esta cantidad impacta la organización de la instrucción y afecta directamente a profesores, apoderados y estudiantes.

A diferencia del *número de alumnos por curso*, la *proporción de alumnos por docente* es el ratio entre el número de estudiantes de una unidad educativa y el número de profesionales de la educación – equivalentes de tiempo completo – asignados a esa unidad educativa. Sin embargo, raramente esta unidad educativa es la sala de clase. Comúnmente lo son unidades más extensas

1999b). Posteriormente el autor reconoce el beneficio de cursos reducidos en 4° y 5° básico en lenguaje y matemática, pero cuestiona su costo-efectividad (Rivkin, Hanushek, & Kain, 2005).

¹¹ Más tarde matiza estas afirmaciones (Hanushek & Woessmann, 2010). Ver sección 5.1. Estudios Comparativos a Gran Escala.

como escuelas o distritos escolares, e incluso estados o naciones. Agrega mayor distorsión al constructo la inclusión de otros profesionales que no son profesores de aulas regulares como paradocentes, asistentes de la educación, educadores diferenciales y otros profesionales como especialistas en lectura, música o arte, bibliotecarias, profesores con responsabilidad compartida en ciertos subsectores, profesores de reemplazo, personal para actividades remediales, etc¹². En la medida que aumentan las demandas a las escuelas y se crean programas especiales de integración escolar subvencionados por el estado, mayor es el número de educadores diferenciales y psicopedagogos que son contratados, disminuyendo la *proporción de alumnos por docente*, lo que sin embargo no cambia en absoluto la realidad de los alumnos de aulas regulares. Lo central respecto de este concepto es que no describe el contexto inmediato del proceso de enseñanza aprendizaje de la mayoría de los alumnos en el aula regular y no refleja las diferencias del tamaño de los cursos que existen al interior de la unidad académica analizada, introduciendo un error de medición en la estimación de los efectos de la variable de interés (Ehrenberg, Brewer, Gamoran, & Willms, 2001). Esta proporción es un indicador global de los recursos humanos dispuestos de forma directa o indirecta para el aprendizaje. Permite asegurar una distribución equitativa de estos recursos y en esa medida es relevante para administradores, legisladores y políticos.

Se estima que en las escuelas primarias de EEUU la diferencia entre la *proporción de alumnos por docente* y el *número de alumnos por curso* es de entre 9 y 10 estudiantes (Sharp, 2002). En consecuencia, el uso erróneo de estos conceptos de forma intercambiable, ha introducido distorsiones en la estimación del impacto de la medida, lo que a su vez ha generado diferentes interpretaciones de los posibles beneficios y diversas recomendaciones de política pública. A pesar de que los constructos actualmente están claramente diferenciados, a medida que se suceden los experimentos, investigaciones y estudios -donde se cuestiona la eficacia de esta intervención por sus elevados costos de implementación- se vuelven a evaluar medidas que combinan ambos conceptos¹³.

De hecho, el primer experimento masivo llevado a cabo en el estado de Indiana, EEUU entre 1981 y 1983 -Proyecto *Prime Time*-, apuntó a una reducción en la *proporción de alumnos por docente* en las escuelas. Si bien sus resultados fueron mixtos -con algunos beneficios en lectura y matemática- éste estableció la viabilidad de la realización de experimentos a gran escala, donde la organización del curso es modificado en pos de mejores resultados educativos. Se convierte así en el precedente del hito que marca la historia de la investigación en torno a los efectos del número de alumnos por curso: el Proyecto STAR *Student/Teacher Achievement Ratio*, experimento de campo, aleatorio, a gran escala¹⁴, con seguimiento tanto de corto como largo plazo que se llevó a cabo entre los años 1985 y 1989 en el estado de Tennessee, EEUU, y cuyos positivos resultados fueron ampliamente divulgados, local e internacionalmente. Los evaluadores de este proyecto estaban especialmente interesados en obtener respuestas definitivas y establecer si la

¹² En 2009 ¼ de la fuerza docente en UK la representaban los paradocentes (Blatchford, 2011). En EEUU alcanzaba el 43% (NEA, 2013), y en Nueva Zelandia a ½ de la fuerza de trabajo (Educational Leaders, 2011).

¹³ Para una contrastación entre estos constructos y sus implicancias ver el apéndice A de Achilles, C. (2003).

¹⁴ El diseño aleatorio del experimento permite establecer causalidad al introducir exogeneidad en las variaciones del tamaño del curso. Ver sección IV. Problemas Metodológicos.

intervención analizada –una reducción del número de alumnos por curso– tendría mayores efectos sobre grupos minoritarios de estudiantes.

3.2. Proyecto STAR (Student/Teacher Achievement Ratio)

Resulta necesario caracterizar este experimento ya que la mayoría de las investigaciones posteriores a esta intervención aluden a su metodología y resultados, y son de alguna u otra forma intentos por confirmar lo allí hallado o por rectificar aquello.

El Proyecto STAR consistió en la asignación aleatoria de un total de 12.000 alumnos entrantes a Kindergarten en tres tipos de cursos: Pequeños (13-17 alumnos); Regulares (22-26 alumnos) y Regulares con apoyo de un docente adicional¹⁵. Se pasó de una mediana de 24 alumnos por curso a 15 estudiantes, en los cursos reducidos. Estos alumnos permanecieron por 4 años consecutivos -hasta finalizar 3° de Educación Básica- en cada uno de los tipos de curso asignados en un comienzo. Los docentes fueron situados aleatoriamente en los tres tipos de curso cada año, de manera de controlar el efecto del profesor en los resultados de aprendizaje que evidenciaran los estudiantes. Éstos se midieron mediante pruebas estandarizadas SAT *Stanford Achievement Tests* para lectura, comprensión oral de textos, reconocimiento de palabras, y matemática. Adicionalmente, y comenzando al inicio del 1° de Educación Básica se aplicó a cada alumno un set de pruebas asociadas al currículo desarrolladas por el estado de Tennessee llamadas *Basic Skills First* (BSF). Con el objeto de medir aspectos asociados a motivación y auto concepto académico se administró el instrumento *The Self-Concept and Motivation Inventory* (SCAMIN). Su grado de confiabilidad sin embargo resulta bajo -coeficiente *alpha* entre 0.56 y 0.69- para niveles iniciales como los aquí evaluados. Por su parte profesores y ayudantes respondieron cuestionarios y completaron bitácoras para registrar sus percepciones y experiencias. El diseño del estudio permitió el análisis de los efectos en distintos grupos de estudiantes: por raza, género y nivel socioeconómico (Finn & Achilles, 1990).

El diseño estadístico consideraba a las escuelas como una dimensión aleatoria, anidada en las cuatro zonas geográficas establecidas: zona marginal, urbana, suburbana y rural. Otras diferencias inter escuela controladas por la aleatoriedad del experimento corresponden a la diversidad de la población atendida, gasto por alumno, recursos educativos disponibles y composición de personal de los distintos establecimientos, entre otros. Si bien se encuentran amplias diferencias entre zonas geográficas, éstas se atribuyen a la composición racial del alumnado en cada una de esas comunidades.

Con el propósito de comprobar si los beneficios de haber pertenecido a cursos reducidos en los niveles educativos iniciales permanecían en el tiempo, es que se condujo el estudio *The Lasting Benefits Study*, seguimiento del desempeño académico y comportamiento de los alumnos participantes del experimento original que volvieron a cursos regulares a partir de 4° de Educación Básica. En estos niveles intermedios -5° a 8° de Educación Básica- se administró el CTBS *Comprehensive Test of Basic Skills*, pudiéndose demostrar que las ventajas de las clases reducidas

¹⁵ Con el objeto de diferenciar los posibles efectos de la *proporción de alumnos por docente* respecto del *número de alumnos por curso*.

perduraban en los niveles siguientes, con una leve disminución del tamaño de los efectos¹⁶. De acuerdo a lo que reportan sus autores Finn & Achilles (1999), los principales resultados de este experimento son:

- a) Se encontraron diferencias significativas entre los tres tipos de cursos en todas las medidas de desempeño y en todos los subsectores evaluados para cada año del experimento. En promedio, los alumnos de los cursos reducidos evidenciaron un desempeño académico superior que los estudiantes de los otros tipos de curso.
- b) Los efectos son atribuibles a la diferencia entre el desempeño promedio de los estudiantes de cursos reducidos respecto de los otros tipos de curso. No se encontraron diferencias significativas entre los cursos regulares y aquellos con un docente adicional de apoyo.
- c) No hay correlación con género, es decir los beneficios se aprecian tanto en niños como en niñas.
- d) En cada nivel se evidencia cierta correlación con raza/etnia o con la zona geográfica de la escuela. Los beneficios resultan sustancialmente mayores para estudiantes de grupos minoritarios o de barrios marginales.
- e) La ventaja de los cursos reducidos resultó significativa para todos los subsectores en cada año consecutivo (4° de Educación Básica y superiores).
- f) Estudiantes de cursos reducidos evidenciaron comportamientos superiores de compromiso en 4° de Educación Básica, como por ejemplo mayor esfuerzo en actividades de aprendizaje, mayor iniciativa, y menor disrupción y comportamiento desatento.

Respecto del tamaño de los efectos, éstos fueron calculados a partir de la diferencia entre el desempeño promedio en test psicométricos referidos a normas de los cursos reducidos respecto del promedio del desempeño de los otros dos tipos de curso, dividido por la desviación estándar de cursos regulares. El tamaño del efecto para mediciones edumétricas referidas a criterio se realizó como diferencias en porcentajes de alumnos que aprueban el test.

De la tabla 2 tomada de Finn & Achilles (1999) se desprende que para el nivel de Kindergarten la ventaja de cursos reducidos fluctúa entre 0.15σ y 0.18σ , lo que se traduce en un progreso de casi un mes de clase respecto de sus compañeros en cursos regulares. Al finalizar 1° de Educación Básica la ventaja fluctúa entre 0.22σ y 0.27σ , lo que se traduce en un adelanto de dos meses respecto de los otros alumnos, y ya para finales del 3° de Educación Básica, el efecto es de entre 0.23σ y 0.26σ . A finales de 5° de Educación Básica, los alumnos que fueron parte del programa de clases reducidas entre Kindergarten y 3° de Educación Básica, muestran una ventaja de casi medio año (5 meses) por sobre los alumnos de cursos no reducidos. En promedio el tamaño del efecto entre 4° y 7° de Educación Básica fluctuó entre 0.13σ y 0.22σ . Concluyen los

¹⁶ El debate sobre la perdurabilidad de los efectos de cursos reducidos en los niveles iniciales, es análogo a la importancia de la educación pre escolar. Ver sección 6.3. Habilidades No Cognitivas y Efectos de Largo Plazo.

autores que se requieren entre 3 y 4 años de intervención para que los mayores logros de aprendizaje perduren en el largo plazo (Finn, Gerber, Achilles & Boyd-Zaharias, 2001).

Tabla 2: tamaño del efecto de cursos reducidos, niveles K-3. Fuente: Finn & Achilles (1999)

Scale	Group	Grade level			
		K (N = 5,738)	1 (N = 6,572)	2 (N = 5,148) ^a	3 (N = 4,744) ^a
Word Study Skills	White	0.15	0.16	0.11	
	Minority	0.17	0.32	0.34	
	All	0.15	0.22	0.20	
Reading	White	0.15	0.16	0.11	0.16 ^b
	Minority	0.15	0.35	0.26	0.35 ^b
	All	0.18	0.22	0.19	0.25^b
Total Reading	White		0.17	0.13	0.17
	Minority		0.37	0.33	0.40
	All	0.18	0.24	0.23	0.26
Basic Skills First (BSF)– Reading	White		4.8%	1.6%	4.0%
	Minority		17.3%	12.7%	9.3%
	All		9.6%	6.9%	7.2%
Total Mathematics	White	0.17	0.22	0.12	0.16
	Minority	0.08	0.31	0.35	0.30
	All	0.15	0.27	0.20	0.23
Basic Skills First (BSF)– Mathematics	White		3.1%	1.2%	4.4%
	Minority		7.0%	9.9%	8.3%
	All		5.9%	4.7%	6.7%

La ventaja para los grupos minoritarios, principalmente alumnos afroamericanos, fue de dos a tres veces mayor que para estudiantes blancos. Se traduce en una disminución de la brecha en la tasa de aprobación del test de lectura, el que por ejemplo, para 1° de Educación Básica corresponde a un 14,3% en cursos regulares versus un 4,1% en cursos reducidos¹⁷.

La hipótesis general consistía en que la reducción de alumnos por curso afectaba indirectamente los logros educativos y que estaba mediada por eventos en el aula. En un intento por determinar cuáles eran estos mecanismos que explicaban los beneficios de cursos reducidos y conciliarlos con ideas presentes en el debate, se centró el argumento en que las clases reducidas son superiores, no porque los profesores modifiquen sus prácticas pedagógicas –hallazgo sustentado por el experimento llevado a cabo en Toronto, Canadá en los años '70-, sino porque los docentes se dedican mayormente a aquellas estrategias que han usado comúnmente, es decir, se *maximiza el tiempo de instrucción*. Sin embargo afirman que los principales cambios suceden en el comportamiento de los alumnos, ya que el contexto de cursos reducidos promueve la participación de los alumnos en el proceso de aprendizaje, incluyendo a aquellos que no estarían dispuestos a participar de estar en cursos más numerosos. Es decir, la clave de los beneficios radica en un *mayor compromiso hacia el aprendizaje por parte de los alumnos*, ya que contextos de mayor visibilidad les restringen la posibilidad de desvincularse y difundir la responsabilidad de interrupciones en subgrupos dentro del curso.

¹⁷ Este resultado, ha sido objeto de constantes discusiones y debates. Se desarrolla en la sección 4.4. Efectos en Grupos Minoritarios.

3.2.1. Programas de reducción del número de alumnos por curso en EEUU

Como consecuencia directa de los resultados obtenidos del proyecto STAR, se comienzan a implementar en varios estados de EEUU programas de reducción de alumnos por curso (*CSR Class Size Reduction*). Hacia el año 2000 alrededor de 35 estados habían implementado en este tipo de programas. Éstos, si bien fueron materializados como intervenciones educativas cuya eficacia había sido comprobada previamente, fueron documentados y evaluados extensamente.

En el condado de *Burke*, perteneciente al estado de Carolina del Norte, se implementó entre en los años 1991-92 un programa piloto, que redujo el número de alumnos por curso entre 1° y 3° de Educación Básica -de 25 o más- a 15. Los buenos resultados obtenidos llevaron a su extensión en 1995-96 a todas las escuelas básicas del condado y para el año 2000 ya se había implementado una reducción a 17 alumnos por curso en los niveles iniciales. Todos los resultados de STAR se confirmaron, incluyendo la perdurabilidad de los efectos y mayores beneficios para grupos minoritarios. El tamaño del efecto resultó ser de 0.45σ en matemática y 0.56σ en lectura al finalizar 2° de Educación Básica (Averett & McLennan, 2004). Lo interesante de esta iniciativa fue que administradores implementaron estas intervenciones sin aumentar el gasto por alumno – principal crítica a la medida– sino que reasignando recursos ya disponibles, principalmente designando cursos pequeños a profesores que previamente no enseñaban a tiempo completo (Finn, 2002).

SAGE: *Wisconsin's Student Achievement Guarantee in Education Program*. A partir del año 1996-97 este programa se implementa en el estado de Wisconsin como un diseño cuasi experimental¹⁸, dirigido a comunidades con altos índices de pobreza. Participaron aproximadamente 3.000 alumnos de Kindergarten y 1° de Educación Básica. Las intervenciones consistieron en diversas configuraciones: cursos reducidos de 15 alumnos por profesor, ya fuera en salas individuales o en espacios compartidos (con división temporal), y cursos con 2 profesores para 30 alumnos, para todas las asignaturas o sólo para las materias principales (*floating teacher*). Las escuelas de control tenían entre 21 y 25 alumnos por curso. Los resultados obtenidos confirmaron lo observado en STAR. El tamaño del efecto fluctuó entre 0.1σ y 0.2σ en puntajes de test, en comparación con escuelas similares. Sin embargo no se encuentra diferencia entre aquellas configuraciones de 15:1 y las de 30:2, cuestionando la superioridad de la reducción del número de alumnos por curso por sobre disminuciones en la *proporción de alumnos por docente*. Se confirma un mayor beneficio para estudiantes afroamericanos, reduciéndose la brecha entre este grupo y estudiantes blancos. En este caso, adicionalmente a los instrumentos de evaluación de desempeño estudiantil (CTBS), se utilizaron cuestionarios y entrevistas a profesores además de la observación de clases, de manera de conceptualizar aquellos mecanismos que median los procesos de aula en cursos reducidos, y poder evaluar las hipótesis que dejó STAR. Del análisis de esta información cualitativa se concluyó que el gran factor interviniente es la *individualización de la enseñanza*, entendida como ayuda directa a los alumnos en la adquisición de ciertas habilidades o contenidos (Molnar, Smith, Zahorik, Palmer, Halbach & Ehrle, 1999).

¹⁸ No fue posible asignar de manera aleatoria alumnos y profesores a los distintos tipos de cursos, ni mantener los cohortes intactos durante los años que duró el proyecto.

Este análisis además permitió caracterizar todos aquellos procesos de aula¹⁹ mediadores que intervienen en cursos reducidos, ya sea en el número de alumnos por curso o en la proporción de alumnos por docente, que surgen de los distintos instrumentos aplicados:

Proyecto SAGE: Resultados de cuestionarios y entrevistas
Conocimiento más profundo de los estudiantes: personalidad y habilidades. Más tiempo de reunión con apoderados. Alumnos se conocen mejor y mejora disposición de compartir sus opiniones y problemas con los demás.
Disciplina: ocurren menos problemas de disciplina por la menor cantidad de estudiantes, pero también la atmósfera más cercana y familiar contribuye a disminuir el comportamiento inapropiado.
Utilización eficaz del tiempo en instrucción: versus control disciplinar y tareas administrativas (registros)
Individualización: se abordan dudas particulares, se expresa en voz alta el entendimiento individual, alumnos reciben correcciones y retroalimentación frecuente y directa. Docentes conocen fortalezas y debilidades de alumnos y saben la etapa del ciclo de aprendizaje donde se encuentra cada uno. Se organiza la clase en grupos pequeños basados en sus necesidades, lo que beneficia a los más desaventajados.
Más actividades centradas en el alumno: mayor independencia y autonomía. Clase reducida permite control aunque se desarrollen actividades más libres.
Aprendizaje: avances más rápidos en contenidos, más cobertura del texto, mayor profundidad y extensión en temas, nuevos tópicos de interés de estudiantes.
Menos estrés y trabajo más placentero: el trabajo docente se concentra en la enseñanza en vez de problemas de comportamiento, papeleo y otros.

Hacia el año escolar 1999-2000, el Programa Federal de Reducción de Alumnos por Curso otorgó fondos por 1.3 billones de dólares para implementar este tipo de medida en las escuelas de más bajo rendimiento y con mayores necesidades. Se destaca en su primer reporte el propósito del proyecto, que describe el mecanismo de la siguiente forma: “fondos para reducir el tamaño de los cursos han ayudado a crear cursos más manejables, de manera que los profesores puedan centrarse en los procesos de enseñanza aprendizaje y destinar menos tiempo a la disciplina” (Cohen, Miller, Stonehill, & Geddes, 2000, p. 1). El 87% de los fondos se utilizaron en nuevas contrataciones y las configuraciones de clase implementadas fueron variadas, como las descritas para el proyecto SAGE anteriormente. Es decir, a diferencia de los hallazgos de STAR, que no encontró beneficios al disminuir la *proporción de alumnos por docente*, este programa federal financió diversos tipos de configuraciones de cursos alternativos. Sin embargo, y antes de contar con los análisis más profundos, hacia el 2001 el entrante presidente Bush eliminó estos fondos centrando su gestión en el controvertido plan *No Child Left Behind*. No obstante, hay un vínculo entre estos dos programas de reforma, ya que la implementada por Bush exige que las escuelas progresen en cuanto a reducir la brecha de desempeño entre los alumnos más avanzados y aquellos de menor rendimiento, lo que se ha traducido en la posterior implementación de programas de reducción de alumnos por curso como mecanismo para lograr dicho requerimiento.

¹⁹ Ver sección VI. Procesos de Aula.

IV. PROBLEMAS METODOLÓGICOS

Como se mencionó anteriormente, el debate sobre la reducción del tamaño de los cursos previo al proyecto STAR era sumamente confuso. Se caracterizaba por múltiples estudios de dudosa calidad, con bajos niveles de confiabilidad y validez, con resultados distorsionados por sesgos de medición, de escaso alcance y desprovistos de estrategias de identificación adecuadas al problema, los que introdujeron por mucho tiempo imprecisiones en la discusión de la polémica medida. Producto de los resultados de STAR y los programas que le siguieron, las teorías se hicieron más robustas y aumentó la cantidad de oportunidades para estimar los efectos de la intervención, lo que hizo más evidente los problemas que acarreaban varias de las estimaciones previas.

Los múltiples errores asociados a la estimación del efecto de la reducción de alumnos por curso, se resolvieron sólo en parte con la clarificación del concepto que se utiliza como variable de interés: *número de alumnos por curso*. La principal dificultad metodológica radicaba en una correcta aislación de la variable independiente de otros efectos, con el propósito de establecer relaciones causales entre el tamaño del curso y los resultados de aprendizaje medidos como puntajes de test. Esta dificultad fue superada en gran medida por los contundentes resultados del experimento STAR, que validado por su diseño experimental aleatorio, garantizaba al menos la ausencia de las principales fuentes de sesgos, al introducir **exogeneidad de las variaciones** en la cantidad de alumnos por curso²⁰.

La contundencia y solidez de los resultados de STAR, que daban respuestas claras a las interrogantes sobre el impacto de la medida analizada, generaron una ola de críticas a su implementación, metodología y diseño (1). Estos cuestionamientos abrieron los espacios a la utilización de modelos y diseños de investigación alternativos (2). A su vez, la focalización en los métodos y técnicas para la estimación de la medida, permitió identificar las principales fuentes de sesgo, comúnmente presentes en los estudios e introducir correcciones a éstos (3). Por otra parte, el constante cuestionamiento de un mayor beneficio para alumnos de grupos minoritarios - hallazgo más sensible y controvertido de STAR-, motivó su reestimación con nuevas aproximaciones en múltiples ocasiones y en diversos contextos (4). Paralelamente, la implementación y ejecución de varios programas de reducción de alumnos por curso dejaban entrever ciertas consecuencias no intencionadas advirtiendo posibles externalidades. Así surgen estudios que evalúan comparativamente la eficacia de ésta y otras medidas alternativas (5).

A continuación se desarrollan estos cinco puntos vinculados a problemas metodológicos. Se caracterizan los problemas enunciados y adicionalmente se explican y describen hallazgos asociados a cada uno de ellos. Éstos se enmarcan en la primera línea de investigación enunciada anteriormente, que establece relaciones causales entre la reducción del tamaño del curso y los

²⁰ La exogeneidad (cambio generado por causas externas o aleatorias) de las variaciones en la cantidad de alumnos por curso permite vincular causalmente el efecto del tamaño del curso con los resultados de aprendizaje.

resultados académicos medidos como puntajes de test. Al final de cada sección se agrega un cuadro resumen de los estudios descritos en ella.

4.1. Análisis Crítico del Proyecto STAR

Los resultados del proyecto STAR fueron tan rotundos e influyentes en EEUU y a nivel internacional que han sido sujeto de múltiples análisis posteriores, con nuevas y sofisticadas técnicas estadísticas, tanto en un intento por rechazar su validez como para reafirmar sus hallazgos. Las principales críticas al proyecto y a la validez de sus hallazgos han sido metodológicas. A pesar que hasta ahora los experimentos aleatorios son considerados los más confiables en este ámbito, las condiciones experimentales del proyecto también se transformaron en una fuente de cuestionamientos.

Las condiciones de implementación del proyecto han sido debatidas, dado que el nivel de Kindergarten no era obligatorio al momento de inicio del experimento, por lo que muchos niños ingresaron al estudio en 1° de Educación Básica, transgrediendo los 4 años de intervención establecidos en el proyecto. Además se observó cierta movilidad de estudiantes entre escuelas que pertenecían al experimento y otras no participantes. También hubo cambios de alumnos entre cursos al interior de las escuelas durante el experimento, lo que causó una superposición en el tamaño real de los cursos respecto del tipo de curso definido. Mueller (2013) identifica y caracteriza estas fallas en la implementación del proyecto, y aborda estos problemas estableciendo ciertos supuestos y redefiniendo el modelo de regresión. De esta forma valida el diseño de STAR y obtiene resultados reestimados bajo el nuevo modelo, que vinculan los positivos efectos de cursos reducidos y años de ejercicio docente, sólo a profesores experimentados²¹.

Los cuestionamientos más categóricos alegan una sobreestimación de los resultados del experimento²², al provenir esta multiplicidad de hallazgos del mismo tratamiento de reducción de alumnos. Al reportar *múltiples resultados* de forma independiente –puntajes de test en varias asignaturas²³– a partir de un tratamiento individual, es probable que se encuentren impactos causales significativos donde no existen, ya que aumenta la probabilidad de falsos positivos. Es decir, podría haber más casos donde se consideran significativos ciertos efectos que no lo son²⁴.

Al abordar estos problemas los americanos Shin & Raudenbush (2011) que advierten una sobreestimación de los resultados del experimento con la metodología utilizada originalmente, reestiman los resultados, abordando además el problema de datos faltantes. La estrategia utilizada fue el uso de un modelo estructural de ecuaciones simultáneas multivariable de tres niveles donde el tamaño del curso actuaba como regresor externo y el tipo de curso como uno endógeno. Los resultados de los cuatro test fueron analizados por ausencia de datos, con el propósito de estimar de manera más precisa los efectos causales de la variable de interés. Los

²¹ Revisar sección 6.1.1. Calidad Docente y Eficacia.

²² Reducciones de alumnos por curso producen mayores logros de aprendizaje en todas las asignaturas evaluadas, resultando estos beneficios más significativos para los grupos de estudiantes desaventajados.

²³ Lectura, matemática, comprensión oral de textos y reconocimiento de palabras.

²⁴ Se considera la posibilidad de que los múltiples resultados correlacionen entre sí, corrigiendo el valor p de las estimaciones en base a la cantidad de resultados reportados.

autores señalan que al superar estos problemas metodológicos han proporcionado la evidencia más efectiva que respalda que una reducción de alumnos por curso causa mayores logros de aprendizaje y que, al mismo tiempo, han maximizado la interpretabilidad de los resultados, al haber aislado las estimaciones de los resultados por asignatura. Por ejemplo, la reducción de un alumno en 1° de Educación Básica genera aumentos en lectura, matemática, comprensión oral de textos y reconocimiento de palabras en 1.09, 1.09, 0.70, y 1.29 puntos en promedio, respectivamente. Las estimaciones que entregan los autores serían incluso mayores que las proporcionadas por Krueger anteriormente²⁵.

Por su parte, el análisis empírico de los canadienses Ding & Lehrer (2011) que igualmente alega una posible sobreestimación de los efectos, provee tres nuevos hallazgos: encuentran que las estimaciones del impacto general del tamaño del curso son robustas, aún corregidas por problemas de inferencia múltiple. Sin embargo estas correcciones anulan los efectos diferenciales para grupos minoritarios en Kindergarten y revelan la heterogeneidad del impacto de la medida, favoreciendo más a alumnos de mejor desempeño.

Resumen de los estudios descritos en la sección 4.1. Análisis Crítico del Proyecto STAR						
Título	Autor/Año/ País	Publicación	Alcance	Especificidad	Enfoque	Hallazgo Clave
<i>Teacher experience and the class size effect — Experimental evidence.</i>	Mueller, S. 2013, Alemania	<i>Journal of Public Economics</i>	Local: datos del proyecto STAR, Tennessee, EEUU.	Específico: examina el efecto de CSR ²⁶ mediado por calidad docente definida como experiencia.	Cuantitativa: caracterización de docente según años de experiencia, dummy.	Solo docentes experimentados logran un efecto producto de CSR. Beneficia a todos los deciles, pero menos a deciles bajos.
<i>The Causal Effect of Class Size on Academic Achievement: Multivariate Instrumental Variable Estimators With Data Missing at Random.</i>	Shin, Y., & Raudenbush, S. 2011, EEUU	<i>Journal of Educational and Behavioral Statistics</i>	Local: datos del proyecto STAR, Tennessee, EEUU.	Específico: examina la posible sobreestimación de los resultados de STAR, al provenir la multiplicidad de hallazgos del mismo tratamiento.	Mixto: utiliza un modelo de ecuaciones simultáneas multivariadas de tres niveles con estimaciones por máxima verosimilitud.	Mayor precisión de las estimaciones del impacto de CSR por asignatura, las que confirman los resultados originales, ampliando la interpretabilidad de éstas.
<i>Experimental estimates of the impacts of class size on test scores robustness and heterogeneity.</i>	Ding, W., & Lehrer, S. 2011, Canadá	<i>Education Economics</i>	Local: datos del proyecto STAR, Tennessee, EEUU.	Específico: evalúa la validez de multiplicidad de resultados de STAR y mayor beneficio para grupos minoritarios.	Cuantitativa: corrección de estimaciones por multiplicidad de resultados producto de CSR.	Validan los resultados generales de STAR pero encuentran mayores beneficios para alumnos aventajados.

4.1.1. Efecto Hawthorne

Otros cuestionamientos al proyecto STAR abarcan temas más generales, como el supuesto de que el número de alumnos por curso pueda modelarse a partir de una función de producción educativa. Aquello supone que los establecimientos educacionales operan bajo estrictos

²⁵ Reporta -0.71 (0.15), -0.49 (0.23), -0.24 (0.21), y -0.66 (0.21) para lectura, matemática, comprensión oral de textos y reconocimiento de palabras respectivamente.

²⁶ CSR: *Class Size Reduction*.

propósitos de maximización de logros académicos²⁷, en circunstancias que las reducciones de alumnos por curso sólo generarían *oportunidades* de mejoras en resultados educativos. El que estas oportunidades se aprovechen al máximo dependerá, señalan, de los *incentivos* que perciban los diferentes actores afectados por la intervención, ya sean padres o apoderados²⁸, profesores, sindicatos, administradores o sostenedores. Estos grupos de interés, favorecen la intervención por múltiples razones, las que no necesariamente se condicen con mejorar los logros educativos; por ejemplo alivianar el ejercicio de la profesión, sustituir *inputs* educativos del hogar a la escuela, aumentar la demanda de profesores, incrementar el poder y dominio en el sistema educativo, ampliar los ingresos por subvención, entregar enseñanza individualizada, etc. Estas distorsiones ligadas a *intereses e incentivos*, asociadas al diseño experimental del proyecto STAR conocidas en la literatura como efecto *Hawthorne*, son las que ponen en duda la magnitud de sus resultados y develan las desventajas de su implementación en relación a otras estrategias alternativas (Hoxby, 2000).

Principalmente, el hecho de ser parte de un experimento que busca encontrar evidencia que sustente y justifique la implementación de reducciones de alumnos por curso, hace que los involucrados estén más dispuestos a modificar su comportamiento de manera de lograr mejores resultados en pos de que dicha intervención sea promulgada. Esta posible alteración²⁹, produciría una *sobreestimación* de los resultados de intervenciones que alteran los incentivos en condiciones experimentales. El efecto *Hawthorne* supone que una vez que el programa de reducción de alumnos por curso -producto de los buenos resultados del experimento- se implemente a nivel general, los incentivos por mejorar los resultados disminuyen y se relaja el comportamiento.

4.2. Diseños de Investigación Alternativos

Considerando que esta posible alteración sobreestima los resultados de la intervención estudiada en *experimentos aleatorios* -en particular los del proyecto STAR- y contemplando además el alto costo que significa este tipo de experimentos, surgen diseños de investigación alternativos que se basan en **variaciones naturales de la población**³⁰, los que aseguran la exogeneidad en los cambios del tamaño de los cursos analizados. Otros modelos alternativos utilizan un **diseño de regresión discontinua**³¹, estrategia de identificación que aprovecha reglas de

²⁷ La postura de los economistas sostiene que los incentivos perversos causan el derroche de recursos en la educación pública. Incentivos como un mayor salario asociado a un mejor desempeño y mayor competencia serían más efectivos (Rivkin, Hanushek, & Kain, 2005).

²⁸ La popularidad del gobernador de California se duplicó luego de anunciar una reducción del número de alumnos por curso (Blatchford, 2011).

²⁹ Efecto Hawthorne: docentes mejorarían su productividad en pos del éxito del experimento. El efecto opuesto, Henry, señala que los individuos de los grupos de control mejorarían su desempeño en competencia con quienes reciben el tratamiento.

³⁰ Compara cohortes en el tiempo. Ej.: Cohorte del año escolar 1990-91 pequeña (curso reducido), versus cohorte del año escolar 1991-92 numerosa en comparación. Nota: si las zonas geográficas o distritos comparados tienen diferentes umbrales máximos, se invalida la conclusión.

³¹ Aprovecha la relación discontinua entre la matrícula y el tamaño de los cursos que resultan de los umbrales máximos establecidos de estudiantes por curso. Un pequeño cambio en una variable continua como matrícula, genera un gran cambio en una variable discreta como número de cursos, y a su vez en la cantidad de alumnos en cada curso. Compara desempeño entre cursos de distinto tamaño.

umbrales máximos y mínimos de alumnos por curso, los que determinan cambios en la cantidad de cursos de acuerdo a aumentos o disminuciones de la matrícula, y que por consiguiente establece variaciones en la cantidad de estudiantes por curso de forma exógena. Investigaciones basadas en estas estrategias de identificación se han implementado en varios países del mundo con resultados muy disímiles.

La estrategia de identificación que utiliza el *modelo de regresión discontinua* fue la elegida para el estudio de los efectos de la cantidad de alumnos por curso en los logros de aprendizaje de los estudiantes en **Israel** entre los años 1991-92. La regla utilizada en este país se remonta al Talmud Babilónico del siglo VI, que establece 25 alumnos por profesor. Maimónides, filósofo rabínico del siglo XII, interpreta el antiguo texto determinando que cursos de entre 25 y 40 alumnos precisan de dos docentes por aula, y que al sobrepasar este límite superior debe crearse un curso adicional, asignando así a 20,5 estudiantes promedio en cada uno. Esta regla ha estado vigente en Israel para determinar el tamaño de los cursos desde 1969.

Como se observa en el gráfico 2, el tamaño del curso que predice la regla de Maimónides de acuerdo a la matrícula, modela bastante bien el tamaño real de los cursos en Israel. Aprovechando que las fuentes que determinan variaciones en el número de estudiantes por curso eran exógenas, se pudo estudiar el efecto de esta variable sobre los resultados académicos de los alumnos.

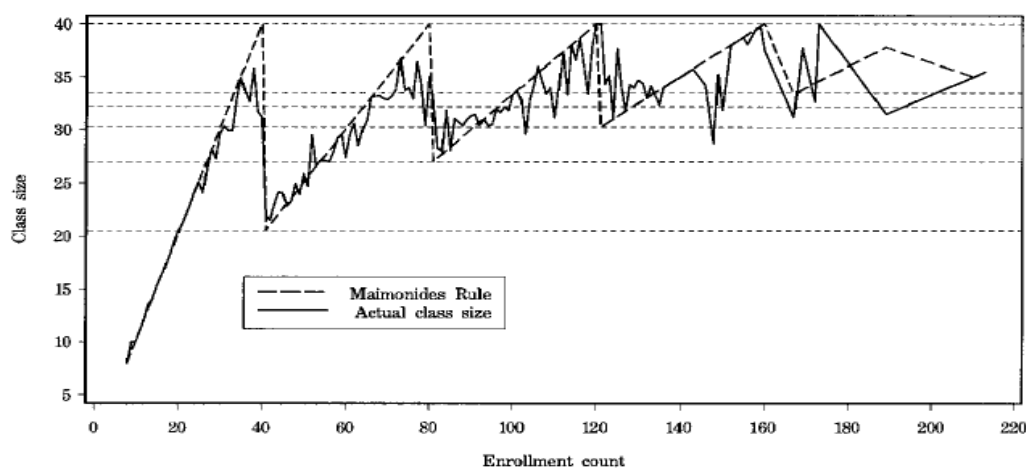


Gráfico 2: Tamaño del curso en 1991 por matrícula inicial, tamaño del curso promedio real y cantidad de estudiantes en base a la predicción según la regla de Maimónides, Israel.

Los resultados del estudio sugieren que una reducción en la cantidad de alumnos por cursos induce a significativos y sustanciales aumentos en los puntajes de los test nacionales, a partir de la muestra de 2.000 alumnos de los niveles 4° y 5° de Educación Básica que fue analizada. El tamaño del efecto reportado, producto de una reducción de 10.5 alumnos por curso, fluctuó entre 0.1σ y 0.2σ de la distribución de puntajes (Angrist & Lavy, 1999).

Recientemente se llevó a cabo un estudio que utiliza la metodología que aprovecha la aleatoriedad de los nacimientos en las áreas geográficas del estado de **Minnesota** basándose en *variaciones naturales de la población*. Esta investigación evalúa los resultados en matemática y

lectura en alumnos de 3° y 5° de Educación Básica, encontrando que el tamaño del curso tiene efectos positivos sobre el desempeño, pero que dichos efectos son muy pequeños -entre 0.04σ y 0.05σ de la distribución de puntajes, producto de una reducción de 10 alumnos por curso-. Los autores descartan que este tipo de intervención pudiera generar significativas mejoras en logros de aprendizajes, dado el pequeño tamaño del efecto estimado (Cho, Glewwe, & Whitler, 2012).

Utilizando las dos estrategias de identificación alternativas mencionadas, el *modelo de regresión discontinua y variaciones naturales de la población* el estudio que presenta Hoxby utiliza datos de panel correspondientes al estado de **Connecticut** entre los años 1992-93 y 1997-98 donde analiza los efectos sobre el desempeño académico en cursos de entre 10 y 30 alumnos, que es el rango relevante que permite la ley en EEUU. La investigadora no encuentra efectos en logros académicos asociados a reducciones en el número de alumnos por curso, independiente de la estrategia de identificación utilizada. Los resultados son consistentes entre ambas metodologías. (Hoxby, 2000)

Asimismo los noruegos Leuven, Oosterbeek & Rønning (2008) utilizaron ambos métodos - el *diseño de regresión discontinua* y el que aprovecha la *variación demográfica*- para estimar el efecto de la cantidad de alumnos por curso en **Noruega** al finalizar la educación secundaria inicial³². Al término de esta etapa cada alumno rinde un examen que determinará sus posibilidades futuras de educación. Los cursos tienen 20 alumnos en promedio y rige un umbral máximo de 30 alumnos por curso. El único efecto, aunque débil, es que aquellos alumnos de madres y docentes menos educados tienden a beneficiarse más de reducciones en el tamaño de los cursos. Sin embargo ambas estrategias de identificación arrojan efectos que no pueden diferenciarse de cero, es decir, no habría efecto del tamaño del curso sobre el desempeño académico en Noruega³³. Las posibles explicaciones que entregan los autores para un tamaño del efecto cero, son la sustitución de *inputs* parentales, interpretada como una menor dedicación o esfuerzo del hogar, al ubicarse al alumno en un curso reducido. Por otro lado, un estilo de enseñanza uniforme, reflejaría la incapacidad de los docentes de transformar el mayor tiempo disponible en mejores resultados. Los académicos intentan dar una explicación de sus resultados, proponiendo la inexistencia de un *efecto universal* del tamaño del curso (validez externa)³⁴.

Estos diseños de investigación alternativos, sin embargo precisan de ciertas condiciones para ser aplicados. Urquiola & Verhoogen -de la Universidad de Columbia, NY- advierten que la utilización del *diseño de regresión discontinua* tiene limitaciones, y no sería recomendable en sistemas educativos donde los supuestos del modelo no se cumplen³⁵. El caso de **Chile** sería el más representativo ya que prácticamente la mitad de los establecimientos son privados, compiten en un *mercado educativo* y muchos de ellos persiguen fines de lucro. Los establecimientos tienen la

³² 9° nivel, equivalente a 1° de Educación Media.

³³ Anterior a este estudio, se habían obtenido estimaciones del efecto de la cantidad de alumnos por curso fluctuantes entre 0.13σ y 0.26σ para una reducción de 10 estudiantes por curso en Noruega, pero las especificaciones del modelo eran más laxas (Bonesrønning, 2003).

³⁴ El cuestionamiento sobre la generalizabilidad de los resultados resulta fundamental en este tema.

³⁵ Los modelos de RD asumen exogeneidad en las variaciones del número de alumnos por curso.

capacidad de establecer precios e influir en su matrícula e impera la *libre elección* para escoger establecimiento, al menos de forma declarativa. Considerando estas características los autores construyen un modelo de selección de diversos tamaños de curso, en donde los hogares y los establecimientos son heterogéneos. El tamaño de los cursos se considera un componente de la calidad de la escuela, ofrecido como un aspecto diferenciador que agrega valor. El ingreso y la disposición a pagar por educación de *calidad* de las familias que buscan escuela, corresponde a la variable independiente que explica el número de alumnos por curso.

El modelo arroja dos predicciones empíricas que tienen sustento en los datos. Primero, como se observa en el gráfico 3, hay una relación de U invertida entre los ingresos y el tamaño de los cursos. Los hogares más ricos seleccionan escuelas de alta productividad y calidad como es esperable, y la relación de U invertida surge de la interacción de dos efectos: la alta productividad permite copar la capacidad y crear nuevos cursos más pequeños satisfaciendo los requerimientos de los hogares que atiende. A bajos niveles de productividad sucede lo contrario, es decir, no se logran copar las aulas.

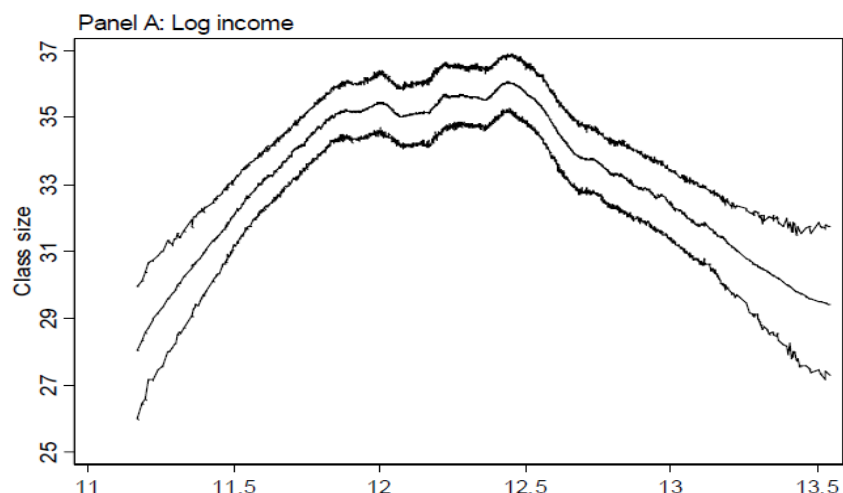


Gráfico 3: Tamaño del curso e ingreso, para escuelas privadas con *voucher* de sectores urbanos en Chile, 2002.

Cabe destacar que esta es la primera relación *no-lineal* propuesta entre cantidad de alumnos por curso e ingreso del hogar. Segundo, dado que hay un tope máximo de 45 alumnos por curso, los establecimientos tienden a estabilizar su matrícula en múltiplos de 45, para evitar crear un nuevo curso con menos alumnos, maximizando así sus ingresos por matrícula producto de la subvención³⁶. Esta predicción es la que imposibilita la utilización del modelo de regresión discontinua en sistemas educativos de libre mercado como el chileno. Con respecto a la variable de interés y sus efectos, en sistemas educativos como el chileno estas regresiones tenderían a subestimar el efecto de reducciones al tamaño del curso en sectores vulnerables y a sobreestimarlos en sectores más acomodados, dado que el aprendizaje tiende a estar altamente

³⁶ Para el caso de Chile existen al menos 5 veces más datos justo por debajo del umbral de 45 alumnos que por sobre este valor, lo que evidencia la manipulación de la variable (Fredriksson, Öckert, & Oosterbeek, 2012).

correlacionado con el nivel socioeconómico (Urquiola & Verhoogen, 2009). De acuerdo a los autores estos resultados enfatizan el valor de *experimentos aleatorios* en este tipo de contextos.

Resumen de los estudios descritos en la sección 4.2. Diseños de Investigación Alternativos						
Título	Autor/Año /País	Publicación	Alcance	Especificidad	Enfoque	Hallazgo Clave
<i>Using Maimonides' Rule To Estimate the Effect of Class Size on Scholastic Achievement.</i>	Angrist, J., & Lavy, V. 1999, Israel	<i>The Quarterly Journal of Economics</i>	Nacional: datos de CSR en Israel entre 1991-92, en niveles 4° y 5°.	General: impacto en resultados de aprendizaje producto de CSR.	Cuantitativo: estrategia de identificación modelo de regresión discontinua.	CSR induce a mejoras en logros de aprendizaje cuantificados entre 0.1σ y 0.2σ
<i>Do reductions in class size raise students' test scores? Evidence from population variation in Minnesota's elementary schools.</i>	Cho, H., Glewwe, P., & Whitley, M. 2012, EEUU	<i>Economics of Education Review</i>	Local: 17 años de datos de CS en Minnesota, EEUU, niveles 3° y 5°.	Específico: uso de variaciones naturales de la población para estimar CSR en Minnesota.	Cuantitativa: estrategia de variación natural de la población, CS ³⁷ y resultados de test.	Encuentra impacto positivo de CSR sobre logros de aprendizaje pero el tamaño del efecto es muy pequeño.
<i>The effects of class size on student achievement: New evidence from population variation.</i>	Hoxby, C. 2000, EEUU	<i>Quarterly Journal of Economics</i>	Local: datos de panel de Connecticut para 4°, 6° y 8°, entre los años 1992-93 y 1997-98, EEUU	Específico: impacto de CSR en desempeño en Connecticut	Cuantitativo: estrategia de variación natural de la población y modelo de regresión discontinua.	Las estimaciones indican que CSR no tiene un impacto significativo en logros de aprendizaje.
<i>Quasi-experimental estimates of the effect of class size on achievement in Norway.</i>	Leuven, E., Oosterbeek, H., & Rønning, M. 2008, Noruega	<i>Scandinavian Journal of Economics</i>	Nacional: dos cohortes, año 2001-02 y 2002-03, 1° Medio secundaria de Noruega.	Específico: estima el impacto de CSR sobre logros de aprendizaje en Noruega.	Cuantitativo: utiliza modelo de regresión discontinua y de variación natural de la población.	El efecto de CSR en Noruega resulta igual a cero. Hipótesis: no existe un efecto universal. Cuestiona validez externa.
<i>Class-Size Caps, Sorting, and the Regression-Discontinuity Design.</i>	Urquiola, M., & Verhoogen, E. 2009, EEUU	<i>American Economic Review</i>	Nacional: datos Mineduc Chile, y puntajes Simce (Chile). 4° básico, 2002.	Específico: efecto de características del sistema educativo en estimaciones de CSR con modelos de RD ³⁸ .	Cuantitativa: modelo de regresión discontinua. Correlación entre CS e ingreso del hogar.	Modelo arroja tamaños de curso asociados a ingreso del hogar y efectividad escolar. Revalorización de experimentos aleatorios.

4.3. Fuentes de Sesgos

Las principales fuentes de sesgo en la estimación del impacto del tamaño del curso presentes en los estudios disponibles son la omisión de variables y la clasificación. La omisión de variables genera un sesgo en la estimación al no poder explicar adecuadamente el impacto de otros factores asociados a la variable dependiente, esto es, los logros de aprendizaje. La clasificación por su parte sesga las estimaciones al no considerar que estudiantes de distinto nivel de habilidad estén siendo asignados a cursos de tamaño diferente.

4.3.1. Omisión de Variables

Para que los resultados de las investigaciones sean válidos, insesgados y permitan una inferencia causal (validez interna), es necesario incorporar la mayor cantidad de variables predictoras de logros de aprendizaje de los estudiantes, de manera de aislar el efecto de la

³⁷ CS: *Class Size*.

³⁸ RD: Regresión Discontinua.

variable de interés. De esta forma, todo aquello que no sea causado por las variables explicativas comunes como características individuales de los alumnos (género y desempeño previo); variables de la familia (nivel socioeconómico, etnia, escolaridad de los padres, ingreso, barrio, etc.), variables docentes (calidad, comportamiento, experiencia, certificación, prácticas, etc.) y variables escolares (gasto, composición, recursos, infraestructura, antecedentes distritales, etc.) pueda ser atribuido al tamaño del curso.

En este sentido la calidad y disponibilidad de datos no fue la más adecuada hasta mediados de los '90. Común ha sido la utilización de mediciones agregadas (a nivel de escuela, distrito o estado), que implica el uso de promedios agregados en vez de datos que reflejen la realidad del aula, o estimaciones aproximadas que distorsionan los resultados, sobre o subestimando el tamaño del efecto. Además del problema de la disponibilidad de los datos a nivel individual, muchas de estas medidas no son observables o accesibles para el investigador, como por ejemplo la habilidad y motivación innata de estudiantes y profesores. Otras veces los instrumentos utilizados para medir el desempeño, por ejemplo, no son los más indicados³⁹, lo que también pone en duda la validez de los resultados.

Estrategias para corregir y mitigar el efecto del problema de la omisión de variables son las estimaciones de modelos estadísticos de **efecto fijo**⁴⁰ y **efecto aleatorio**⁴¹ (Averett & McLennan, 2004). La decisión entre uno y otro tipo de efecto depende del supuesto sobre la naturaleza de la heterogeneidad entre las unidades de observación.

4.3.2. Problema de Clasificación o *Sorting*

Hasta ahora hemos considerado que si las variaciones en la cantidad de alumnos por curso son generadas de forma externa o aleatoria, ya sea por *diseños experimentales* de investigación como el proyecto STAR, por variaciones en la matrícula que generan discontinuidades en el tamaño de los cursos producto de *reglas de umbrales máximos* o por *variaciones naturales de la población*, la exogeneidad estaría garantizada. Sin embargo un potencial sesgo -que impide establecer causalidad- y que posiblemente está presente en diseños de investigación tanto experimentales como cuasi-experimentales es el **problema de la clasificación** o *sorting problem*⁴².

Este problema consiste en asignar a los estudiantes de forma endógena –en vez de exógena o aleatoria– a cursos de diferentes tamaños **al interior de las escuelas**, de acuerdo a aspectos observables y no observables, como por ejemplo rendimiento, conducta o alguna otra característica o necesidad. Así un alumno puede ser asignado a un curso reducido porque su

³⁹ Desempeño absoluto, versus aumentos en resultados de test. Resultados de test con o sin consecuencias para los alumnos.

⁴⁰ El efecto fijo, asume que la especificidad de cada escuela está correlacionada con la variable explicativa. Este efecto mide las diferencias no observables entre escuelas que no han sido incorporadas al modelo, mejorando la bondad de ajuste de las estimaciones.

⁴¹ El efecto aleatorio supone heterogeneidad entre escuelas o docentes que no está correlacionada con las variables involucradas como contexto familiar, características individuales, de los profesores y escuelas, entre otras.

⁴² También aparece en la literatura como problema de selección.

rendimiento es bajo o es muy disruptivo, imposibilitando la inferencia de variaciones en el desempeño causadas por diferentes tamaños de cursos. Si las escuelas usan los cursos reducidos en forma compensatoria para remediar dificultades de aprendizaje pasadas de alumnos individuales, una correlación positiva entre el tamaño del curso y el desempeño puede simplemente reflejar la asignación de estudiantes y no el real efecto causal de la variable. Cursos avanzados tienden a ser más pequeños, lo que arrojaría también correlaciones espurias entre desempeño y tamaño del curso.

Adicionalmente, este problema de clasificación se puede dar **entre escuelas**, y en ese caso se lo denomina *Tiebout sorting*. Éste pone de manifiesto el hecho de que los alumnos no están distribuidos aleatoriamente en las distintas escuelas del país. Los padres que valoran la educación podrían elegir una *buena* escuela que ofrezca cursos reducidos, o adquirir una propiedad en un barrio donde los cursos de las escuelas son más pequeños⁴³. Si estos padres hacen inversiones educativas no observables como contratar tutores o leer con sus hijos, es muy probable que los resultados académicos reflejen un mayor capital cultural y social, y un mayor compromiso de estos apoderados con la educación de sus hijos, en vez de -o en combinación con- los efectos de una menor cantidad de alumnos por curso. Asimismo el impacto en los logros de aprendizaje al concentrar alumnos de hogares de mayor capital cultural, podría atribuírsele erróneamente a una menor cantidad de alumnos por curso, cuando ese resultado está influido por el efecto par⁴⁴.

Para corregir los sesgos de la clasificación lo más efectivo es el diseño de *experimentos aleatorios*, en que la asignación de los individuos al grupo experimental o al de control es realizada al azar, lo que garantiza que las variables mencionadas anteriormente no estén correlacionadas con los logros de aprendizaje que reporten estudiantes luego del tratamiento, permitiendo aislar la variable estudiada –número de alumnos por curso– y sus efectos. Sin embargo, algunos investigadores como Hanushek y Hoxby señalan que los experimentos nunca se implementan de manera completamente aleatoria. Cuestionan la posibilidad de generalizar los resultados allí obtenidos a otras poblaciones; alegan que las reducciones implementadas en STAR exceden las aplicadas de forma masiva. Señalan además que los impactos positivos de la medida estarían circunscritos exclusivamente al nivel de Kindergarten y posiblemente a 1° de Educación Básica. Además, como ya se describió anteriormente, los efectos Hawthorne o Henry podrían influir en los resultados de estos experimentos (Hanushek, 1999a).

Existen algunas técnicas que permiten corregir estimaciones sesgadas e inconsistentes producto de la endogeneidad de la variable de interés. Una técnica de medición que permite mitigar en parte este efecto es el **análisis de valor agregado** en vez de la utilización del rendimiento absoluto en la medición de los resultados educativos. En este caso lo que se mide es la diferencia entre la posición académica inicial del alumno y la evaluación subsecuente, posterior

⁴³ En algunos países como EEUU, Canadá, Grecia y Polonia se prioriza la asignación a la escuela de acuerdo a la proximidad residencial.

⁴⁴ *Tiebout*: la clasificación residencial de hogares, implica una asignación no aleatoria de habilidades en las diferentes escuelas, donde aquellos establecimientos de altos ingresos se benefician en mayor medida de factores no financieros como el efecto par (Brewer & McEwan, 2010).

al tratamiento. Es decir, se observa el efecto acumulativo en los logros de aprendizaje más que el nivel absoluto de éstos, reduciendo el efecto de la clasificación. Sin embargo, si la posición inicial está afectada al problema de clasificación, entonces también lo estará el análisis de valor agregado posterior⁴⁵ (Krueger, 1999).

La técnica que utiliza **variables instrumentales (IV)** para corregir la endogeneidad del tamaño del curso, es otra alternativa disponible. Las variables instrumentales elegidas deben estar altamente correlacionadas con la variable explicativa de interés – tamaño del curso –, pero no estar relacionada con la variable dependiente – logros de aprendizaje–. Con la introducción de variables instrumentales, cuyos residuos no correlacionan con la variable estimada, se corrige este sesgo⁴⁶; sin embargo aquí la dificultad reside en identificar instrumentos que no se vinculen a la clasificación. Por ejemplo, el gasto promedio por alumno como variable instrumental, es un buen predictor del tamaño de los cursos, sin embargo es altamente probable que los hogares con mayores recursos hayan elegido escuelas con un mayor gasto por alumno, lo que vincula dicha variable instrumental con la variable dependiente, desempeño (Averett & McLennan, 2004).

Por su parte los modelos de **efecto fijo de la escuela**, son los que controlan la posibilidad de que *mejores* escuelas tengan además cursos de menor tamaño, lo que corresponde al problema de clasificación entre escuelas -generada por elección residencial o modalidad educativa-⁴⁷; la que se aborda agregando una variable *dummy* a cada escuela de la muestra. Así, de existir alguna característica sistemática pero no observable en estas escuelas, este estimador la corregirá, ya que excluye cualquier variación persistente de desempeño o del tamaño de los cursos.

Para ilustrar el impacto que tienen estos controles sobre problemas de clasificación, tanto al interior de las escuelas como entre escuelas en la estimación del efecto del tamaño del curso, se presenta la siguiente figura del estudio comparativo internacional de Wössmann & West (2006a), que contrasta los coeficientes estimados obtenidos a partir de las tres estrategias de identificación utilizadas.

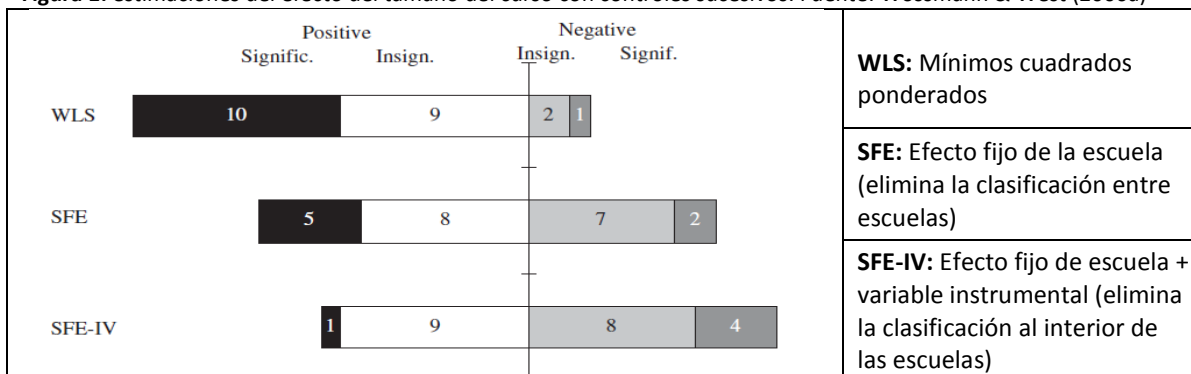
Al considerar sólo los resultados obtenidos mediante la técnica de mínimos cuadrados ponderados WLS, se concluye que en prácticamente la mitad de los países analizados, los cursos más numerosos generan mayores resultados de aprendizaje.

⁴⁵ Si a un alumno, que pertenece a un curso reducido, se le mide su avance en desempeño entre el segundo y tercer año, el diferencial reportado subestimaré el efecto del curso reducido en el primer año, que es donde probablemente ocurre el mayor beneficio de ser parte de un curso con menos estudiantes.

⁴⁶ La asignación no aleatoria de estudiantes, tanto a nivel de escuela como de curso, es probable que genere correlaciones entre el tamaño del curso y el error de estimación ε de la función de producción, lo que distorsiona la estimación del parámetro asociado a la variable de interés. De este modo se debe identificar un *proxie* de la variable, que sea un buen predictor de ésta, pero que no correlacione con los residuos.

⁴⁷ Científico-humanista; técnico profesional, artístico, vocacional, etc.

Figura 1: estimaciones del efecto del tamaño del curso con controles sucesivos. Fuente: Wössmann & West (2006a)



Al incorporar el efecto fijo de la escuela SFE, los estimadores positivos caen a la mitad y los negativos aumentan débilmente, sin embargo al incluir además la variable instrumental IV, sólo persiste un país con un estimador positivo (contra intuitivo) y emergen cuatro con coeficientes negativos significativos, sugiriendo que el tamaño del curso tiene un impacto en los aprendizajes de los alumnos⁴⁸. Nótese que la incorporación de las correcciones a los sesgos de medición hace que la **conclusión sea manifiestamente opuesta a la inicial**, lo que explica en parte la inconsistencia de los estudios e investigaciones en este ámbito.

Como se mencionó anteriormente, los problemas metodológicos han sido una dificultad constante en la estimación del efecto del tamaño del curso. Algunas investigaciones enfocadas exclusivamente en advertir estas dificultades presentes en estudios precedentes y motivados por superarlas, han aportado con hallazgos clave en este ámbito.

Los autores Borland, Howsen, & Trawick (2005) afirmaron que la gran cantidad de intentos fallidos por identificar empíricamente la relación entre el número de alumnos por curso y los logros de aprendizaje estaban siendo entorpecidos por una de cuatro causas: primero la utilización indistinta de la *cantidad de alumnos por curso* como de la *proporción de estudiantes por docente* como variable independiente; segundo, modelos con múltiples variables omitidas (sobre todo habilidad); tercero, la imposibilidad de considerar la endogeneidad del tamaño del curso respecto del desempeño, y cuarto, el uso de una función incorrecta al especificar la relación buscada (monotónica lineal o no lineal).

En un intento por corregir estos errores y obtener estimaciones insesgadas, los autores abordan los diferentes problemas señalados tomando el valor observado del tamaño del curso en vez de los datos agregados, incluyen una medida de la habilidad como variable explicativa y consideran el tamaño del curso como endógeno. Sin embargo, respecto del problema de la relación monotónica introducen una innovación. Proponen que al aumentar el número de alumnos por curso, y considerando que los estudiantes aprenden tanto de sus pares como del docente, habría dos efectos esperados combinados. Por un lado, el profesor ofrece menores oportunidades de aprendizaje individuales al tener más alumnos a su cargo. Por el otro, aumentan

⁴⁸ Sin embargo, sólo para 2 de los 11 países se encontró un efecto negativo importante y significativo del tamaño del curso en los aprendizajes de los alumnos.

las oportunidades de aprendizaje entre pares, al haber mayor posibilidad de interacción entre ellos, lo que se ve fortalecido por una mayor competencia al interior del curso. Considerando que estos dos efectos son opuestos, los autores proponen un modelo no monotónico que relaciona el desempeño con la cantidad de estudiantes por curso. Los resultados sugieren un *tamaño óptimo* del curso de entre 21.3 y 23.24 dependiendo de las distintas asignaturas, confirmándose una *relación no lineal ni monotónica*. Es decir, el signo del efecto de reducciones de alumnos por curso va a depender del tamaño inicial del curso relativo al tamaño óptimo.

Es posible advertir que las dificultades de medición y estimación presentes en este ámbito han incentivado a los académicos e investigadores a utilizar estrategias y aproximaciones innovadoras que desafían y renuevan los supuestos tradicionales, y que muchas veces han resultado en descubrimientos y significativos aportes a la discusión.

4.3.2.1. Educación Compensatoria

El problema de clasificación está directamente vinculado a las prácticas compensatorias en educación, y la evidencia de su uso tiene implicancias directas en el estudio del efecto del tamaño del curso.

La educación compensatoria se refiere a aquellas iniciativas orientadas a atender y satisfacer las carencias de alumnos con necesidades educativas especiales, como por ejemplo cursos compensatorios, actividades remediales y educación bilingüe en países donde la proporción de estudiantes inmigrantes extranjeros es alta. A medida que aumentan los programas de integración escolar, cuantiosos recursos son asignados a estas iniciativas y mayor es la cantidad de cursos que atienden a alumnos en programas compensatorios, los que tienden a ser pequeños. Alumnos vulnerables, de familias más pobres, de grupos étnicos y raciales minoritarios presentan un probabilidad desproporcionadamente alta de ser asignados a estos cursos, lo que evidencia un *problema de clasificación*.

La evidencia empírica muestra que a nivel internacional, se implementan fuertes prácticas compensatorias en la asignación o clasificación de estudiantes a cursos reducidos en 18 sistemas educativos de distintos países que participan en TIMSS⁴⁹ (Wössmann & West, 2006b). Se identifican este tipo de patrones principalmente **entre escuelas**, más que al interior de las escuelas. Es decir, se observa con mayor frecuencia la distribución de estudiantes desaventajados a escuelas con cursos de un menor tamaño promedio, que la selección de este tipo de estudiantes a cursos de tamaños reducidos al interior de los establecimientos educacionales.

Sin embargo, estas prácticas de clasificación compensatoria pueden resultar en la asignación de recursos en forma *compensatoria* o *regresiva* de acuerdo a cómo y quién las establezca. La hipótesis sugiere que en la *clasificación entre escuelas*, la capacidad de un sistema educativo de identificar las habilidades innatas de los estudiantes resulta en una asignación de recursos *compensatoria*, y que al contrario, la libertad de residencia combinada con la

⁴⁹ Australia, Bélgica (Flandes y Francesa), Canadá, República Checa, Francia, Grecia, Hong Kong, Islandia, Japón, Corea, Portugal, Rumania, Escocia, Singapur, Eslovenia, España y Estados Unidos.

descentralización financiera del sistema, profundiza la asignación de recursos de forma *regresiva*⁵⁰. Al *interior de las escuelas* los autores proponen que la clasificación será *regresiva* cuando dependa más bien de los profesores jefes de departamento – que podrían asignarse los cursos pequeños y aventajados– y en cambio será *compensatoria*, si está en manos de la dirección del establecimiento.

La evidencia sobre prácticas de clasificación compensatoria tiene importantes implicancias en los análisis causales que se realicen respecto de los efectos de mayores recursos en logros de aprendizaje, por lo que se hace necesario abordar el problema de asignación de recursos endógena en investigaciones relativas al número de alumnos por curso y sus efectos. Así lo hacen los estudios que a continuación se describen:

Los antecedentes del estudio NELS, *National Education Longitudinal Study of 1988*, igualmente evidencian el uso de políticas educativas compensatorias en **EEUU**. Los datos revelan que los estudiantes talentosos y de desempeño regular asisten a cursos de 22 alumnos promedio por curso; más que el doble del tamaño de los cursos compensatorios, que no supera los 10 estudiantes⁵¹. Se encontró también que en general, afroamericanos e hispanos asisten a escuelas con cursos levemente más numerosos que los blancos. Para medir las diferencias del tamaño de los cursos por raza y etnia -netas del efecto de selección- y estimar el efecto de la cantidad de alumnos por curso en logros de aprendizaje, los autores abordan el problema de clasificación endógena⁵² al interior de las escuelas mediante la inclusión de variables instrumentales⁵³. Las estimaciones de la variable de interés obtenidas, corregidas por los efectos de las políticas redistributivas, proyectan que las diferencias del tamaño del curso por raza y etnia entre 8° de Educación Básica y 2° de Educación Media, explicarían entre un 18% y un 47% de las diferencias en mejoras de desempeño en evaluaciones sucesivas entre blancos y negros e hispanos. En términos del tamaño del efecto, éste se estima en 0.07σ por cada estudiante menos por curso (Boozer & Rouse, 2001).

En esta misma línea, investigadores italianos estudiaron los efectos de reducciones de alumnos por curso en escuelas públicas, producto de una intervención que estipula un máximo de 25 alumnos por curso en Educación Básica. El movimiento de reforma educativa que surge en 2008 nace a propósito del mal desempeño relativo de **Italia** en pruebas como PISA y TIMMS, y de diferencias en la proporción de alumnos por docente entre diversas regiones del país. Su implementación generó discontinuidades en el tamaño de los cursos que permitieron la utilización del *modelo de regresión discontinua* en su versión *fuzzy*⁵⁴ para estudiar los impactos de la variable

⁵⁰ Sólo Estados Unidos, el país con mayor descentralización financiera y movilidad residencial, presenta un patrón fuerte de prácticas *regresivas* de clasificación entre escuelas. Cabe señalar que el sistema Chileno, ausente en el estudio, es aún más liberal, pro mercado y de libre elección que el de EEUU.

⁵¹ Probablemente otros inputs educativos, como la calidad docente haya sido asignada en forma remedial.

⁵² Previo a la corrección por clasificación endógena, los resultados arrojan efectos positivos entre cantidad de alumnos por curso y desempeño, lo que confirma la tesis de que los alumnos talentosos son ubicados en cursos más numerosos.

⁵³ Se utilizó la proporción de estudiantes por profesor como variable instrumental.

⁵⁴ No requiere que la probabilidad de recibir el tratamiento pase de 0 a 1; caso de cumplimiento imperfecto.

de interés en 5° de Educación Básica. Los supuestos del modelo se ven débilmente transgredidos, para la muestra reducida de escuelas que respetan la norma. Esto porque se observan prácticas de *retención* de la matrícula en múltiplos del umbral (25, 50, 75, etc.) y prácticas de clasificación al sobrepasar estos umbrales de matrícula, incluyendo más alumnos de entornos socioeconómicos desfavorables, como estrategia para obtener más recursos del estado, principalmente docentes. Señalan que si bien las estimaciones del efecto del número de alumnos por curso sobre los logros de aprendizaje no son significativas, observan que el tamaño del curso se está usando en primaria como un tipo de política educativa compensatoria (Minzyuk & Russo, 2012).

Resumen de los estudios descritos en la sección 4.3. Fuentes de Sesgos						
Título	Autor/Año /País	Publicación	Alcance	Especificidad	Enfoque	Hallazgo Clave
<i>Exploring the Effect of Class Size on Student Achievement: What Have We Learned Over the Past Two Decades?</i>	Averett, S., & McLennan, M. 2004, EEUU	<i>Edward Elgar Publishers</i>	Internacional: síntesis de investigación de CSR a nivel internacional, sus hallazgos y problemas metodológicos.	General: revisión investigaciones de los últimos 20 años. Evidencia empírica agrupada por tipo de estudio.	Cualitativa: focaliza la revisión en estudios que examinan la relación entre tamaño del curso y puntajes en test.	Al controlar por endogeneidad surge impacto CSR. ¿Medida es costo-efectiva? Externalidades en implementación masiva.
<i>Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects.</i>	Hanushek, E. 1999, EEUU	<i>Educational Evaluation and Policy Analysis</i>	Internacional: contrasta evidencia internacional como nacional, no experimental y econométrica con los resultados de STAR.	General: analiza evidencia disponible sobre CSR y examina críticamente el diseño del experimento STAR.	Mixto: detalla falencias y supuestos de STAR que no se cumplen en otras condiciones lo que limita su validez externa.	La evidencia no experimental como experimental no ofrece razones que sustenten efectos sistemáticos de CSR.
<i>Experimental estimates of education production function.</i>	Krueger, A. B. 1999, EEUU	<i>Quarterly Journal of Economics</i>	Local: resultados de STAR, Tennessee, EEUU, entre 1985-89, nivel K-3. Impacto internacional.	General: impacto en resultados de aprendizaje producto de CSR.	Cuantitativo: análisis econométrico de los resultados. Corrección por abandono y cambios de curso.	Confirma los resultados de STAR; efecto Hawthorne poco probable. Medida de valor agregado se pone en duda.
<i>Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS.</i>	Wössmann, L., & West, M. 2006, Alemania	<i>European Economic Review</i>	Internacional: datos de TIMSS (matemática y ciencias) para 11 sistemas educativos 1994-95 en nivel 7° y 8°.	Específico: impacto de CSR en resultados prueba TIMSS a los 13 años en 11 países.	Cuantitativo: usa variación de tamaño de cursos en 7° y 8° para ambas asignaturas. Utiliza efecto fijo de escuela y IV.	2 de los 11 países muestran efectos negativos significativos de CSR; países con cuerpo docente mal remunerado, menos efectivo.
<i>An investigation of the effect of class size on student academic achievement</i>	Borland, M., Howsen, R., & Trawick, M. 2005, EEUU	<i>Education Economics</i>	Local: base de datos de Kentucky, EEUU, nivel 3°, 1990.	Específico: Revisión de errores estadísticos previos, reestimación insesgada.	Cuantitativa: Estimación libre de problemas de constructo, endogeneidad, omisión de variables y función.	Reestimación del efecto libre de errores comunes. Encuentra tamaño óptimo de 20 aporx. Relación no monotónica ni lineal.
<i>Which school systems sort weaker students into smaller classes? International evidence.</i>	Wössmann, L., & West, M. 2006, Alemania	<i>European Journal of Political Economy</i>	Internacional: evidencia sobre prácticas de clasificación en 18 sistemas educativos, 7° y 8° en 1994/95.	Específico: identifica prácticas de clasificación entre y al interior de escuelas.	Cuantitativo: utiliza puntajes y datos de la prueba TIMSS 1994/95.	Identifica efectos compensatorios o regresivos asociados a prácticas de selección.

<i>Intraschool Variation in Class Size: Patterns and Implications.</i>	Boozer, M., & Rouse, C. 2001, EEUU	<i>Journal of Urban Economics</i>	Nacional: panel de datos nacional longitudinal NELS: 88, nivel 8° y II Medio.	Específico: efecto de CSR medido en puntajes de test y tamaño de curso, por raza y	Cuantitativa: análisis de puntajes de test por raza y etnia asociado a CS.	Cursos reducidos aumentan puntajes de test. Brecha de negros hispanos puede explicarse por CS
<i>The Causal Effect of Class Size on Pupils' Performance: Evidence from Italian Primary Schools.</i>	Minzyuk, L., & Russo, F. 2012, Italia	<i>Società italiana di economia</i>	Nacional: datos de desempeño en 5° en escuelas Italianas, en el año 2008/09	Específico: estima impacto de CSR en logros de aprendizaje.	Cuantitativo: utiliza el modelo de RD en su versión fuzzy. Máximo 25 alumnos por curso.	No encuentra evidencia que sustente CSR, pero se advierte que CS se utiliza como política compensatoria.

4.4. Efectos en Grupos Minoritarios

La validez de los postulados de STAR, sobre todo de aquel que establece mayores beneficios para alumnos vulnerables, ha sido ampliamente cuestionada por investigadores preocupados de aportar con evidencia rigurosa respecto de una efectiva asignación de los recursos en educación. Resulta central verificar la capacidad de esta intervención de generar un mayor progreso en los estudiantes de peor desempeño, ya que al beneficiar más a los alumnos desaventajados, disminuye la brecha entre los de mejor y peor rendimiento, contribuyendo así a una mayor equidad educativa. Varios estudios en la literatura se centran en este aspecto de la investigación, sin embargo los resultados son mixtos, encontrándose evidencia a favor y en contra de este resultado⁵⁵.

Inicialmente los autores Nye, Hedges, & Konstantopoulos (2002) declaran inválidos los efectos diferenciales reportados para los subgrupos -blancos versus hispanos y afroamericanos- analizados en STAR. Señalan que éstos sólo habrían sido significativos para logros de aprendizaje en lectura. Adicionalmente, el estudio que hizo el seguimiento a STAR, *The Lasting Benefits Study*, declara mayores beneficios para grupos minoritarios de bajo desempeño, sin embargo los autores alegan que los niveles de significación de estos hallazgos no fueron reportados. Los investigadores vuelven a analizar estos efectos diferenciales utilizando dos criterios para categorizar a alumnos de bajo rendimiento; a) alumnos por debajo de la media y b) alumnos por debajo del 1er cuartil. Los resultados obtenidos apoyan la conclusión general que postula STAR, esto es, que cursos pequeños en niveles iniciales conducen a un mayor desempeño académico, sin embargo señalan que la evidencia de beneficios diferenciales para alumnos de bajo rendimiento es contradictoria y débil. Como hipótesis los autores postulan que los beneficios observados en el largo plazo podrían explicarse por una mayor probabilidad de que en cursos reducidos los profesores identifiquen a aquellos alumnos en riesgo de presentar bajos niveles de logro. Al identificar estos casos de forma temprana el docente está en condiciones de ofrecer estrategias de instrucción diferenciada como

⁵⁵ Una posible explicación es la poca claridad de la categoría de alumno evaluada, con múltiples posibilidades: afroamericanos, hispanos, alumnos prioritarios, vulnerables de familias de escasos recursos, con dificultades para el aprendizaje, de peor desempeño, desvinculados del quehacer escolar, etc. Es probable que los estudiantes en estas categorías compartan ciertas características, pero es también posible que sus diferencias, la metodología utilizada y los diversos contextos donde se ha estudiado este efecto diferencial sea la causa de las inconsistencias de los resultados.

educación compensatoria, lo que podría haber prevenido el seguimiento de una trayectoria escolar de bajo rendimiento o fracaso.

Dos años más tarde, el mismo equipo de investigadores publica un estudio que da cuenta de la perdurabilidad de los beneficios de STAR. En particular buscan comprobar que los beneficios generados por los cuatro años de tratamiento permanecieron al menos los cinco años siguientes. Llama la atención que a diferencia del estudio anterior, documentan la existencia de mayores beneficios en estudiantes de grupos minoritarios y además acreditan que este beneficio diferencial se mantiene en el tiempo en vez de disminuir. Los autores señalan que los resultados obtenidos entregan evidencia significativa de que los beneficios diferenciales para grupos minoritarios, por raza y etnia, son mayores en lectura. La evidencia para matemática, sostiene que la brecha de desempeño de género (hombres por sobre mujeres) se reduce. Afirman que estos nuevos análisis respaldan la idea de que la reducción de alumnos por cursos es una medida que mejora los resultados de todos los estudiantes y que a la vez reduce la inequidad (Nye, Hedges, & Konstantopoulos, 2004).

Importante evidencia de efectos diferenciales fuera de los EEUU arroja el estudio a gran escala realizado en el **Reino Unido**, CSPAR *Class Size and Pupil Adult Ratio* que utiliza un diseño longitudinal naturalista⁵⁶. Éste analizó el desempeño de más de 10.000 estudiantes comenzando en año escolar 1996-97 (primera cohorte de 4/5 años de edad) haciendo un seguimiento por 3 años en una primera etapa⁵⁷, completándose la Educación Básica en una segunda etapa posterior. Adicionalmente se observaron distintas prácticas de aula involucradas en el proceso de enseñanza-aprendizaje que se veían impactadas por estas variaciones⁵⁸. Se comparó el desempeño de cursos reducidos de menos de 25 estudiantes con el de cursos numerosos⁵⁹. El tamaño del efecto en los resultados académicos posterior al primer año de tratamiento, fue claro y similar a lo reportado en STAR en lenguaje y matemática. Una diferencia de 10 alumnos por curso, pasando de 30 a 20 estudiantes, resultó en un mejor desempeño de 0.35σ para el grupo de bajo rendimiento inicial, 0.2σ para aquellos alumnos de rendimiento previo regular y de 0.15σ para el grupo aventajado (Blatchford, 2011). Es decir, se corrobora lo hallado en STAR, encontrándose un efecto mayor para los estudiantes más desaventajados. Sin embargo al tercer año del proyecto los efectos ya no eran claros, lo que -a diferencia de STAR- sugiere que éstos se desvanecen con el tiempo. Si se evidencia una mayor *perdurabilidad* del efecto cuando el estudiante continúa en un curso reducido, lo que sugiere que éste es *acumulativo*. La segunda etapa del proyecto que evaluó el desempeño entre los 7 y 11 años, no arrojó resultados

⁵⁶ La ventaja respecto de estudios experimentales es que el efecto del tamaño del curso se examina a través de la distribución real del número de alumnos por curso.

⁵⁷ Se controló por habilidad previa aplicando un pre test.

⁵⁸ Revisar sección VI. Procesos de Aula.

⁵⁹ Estos valores difieren de lo que en STAR se considera reducido y numeroso. En 1997 se fija en 30 la cota máxima de alumnos por curso en niveles iniciales en UK. Contradictoriamente a lo esperado los cursos de Educación Básica presentaban más alumnos promedio que los de Educación Media en 1996. (27.5 vs 21.9).

significativos respecto del efecto de variaciones en la cantidad de alumnos por curso, confirmando que el nivel más susceptible de conseguir beneficios del tratamiento es el inicial⁶⁰.

Estos hallazgos se corroboran en el estudio de Bosworth (2011) que toma los datos de alumnos de 5° de Educación Básica de **Carolina del Norte**. El autor encuentra evidencia de que los estudiantes son ubicados en cursos de diferentes tamaños de acuerdo a características observables. Mediante un *efecto fijo* del docente y controles de la composición del curso, el investigador corrige los sesgos de selección observados. Sus conclusiones señalan que el efecto del tamaño del curso es heterogéneo al beneficiar en mayor medida a los alumnos más desaventajados que a aquellos en la parte superior de la distribución de puntajes. Observa que la diferencia de puntajes es menor en cursos reducidos que en aquellos más numerosos y que las reducciones de alumnos por curso acortarían estas brechas al interior del curso, más que mejorar el nivel promedio de desempeño.

A diferencia de los estudios anteriores, los canadienses Ding & Lehrer (2011) señalan que los modelos estadísticos empleados en STAR limitan el tratamiento de la heterogeneidad del impacto de cursos reducidos en la distribución de puntajes, al reportar sólo resultados promedio, ya sea ignorando la heterogeneidad al interior de las escuelas o tratándola como un *efecto aleatorio*. Los autores emplean el método más flexible para evaluar la existencia de heterogeneidad a través de distintos subgrupos, estimando un modelo completamente saturado que contiene todas las posibles interacciones entre estudiantes, tipo de curso y covariables del profesor. Los resultados muestran gran discrepancia con los postulados de STAR, ya que no habría evidencia de beneficios adicionales para alumnos desaventajados o de grupos minoritarios; más aún, se encuentra que los estudiantes con mayores puntajes en test son aquellos que más se benefician de la intervención, lo que sugiere una gran influencia del entorno familiar. Destaca la substancial heterogeneidad en los impactos de reducciones de alumnos por curso en la contribución a la producción de resultados educativos y sus efectos para políticas públicas.

Buscando identificar al tipo de estudiante que resulta mayormente beneficiado de una reducción de alumnos por curso, de acuerdo al lugar que ocupa en la distribución de puntajes, el autor holandés Levin (2001) utiliza una técnica de regresión por cuantiles para estimar el efecto del número de alumnos por curso sobre el desempeño académico. El investigador utiliza un panel de datos muy completo PRIMA, que permite controlar por una gran cantidad de características observables de la escuela y de los alumnos de 4°, 6° y 8° de Educación Básica en el año escolar 1994-95 en **Holanda**. Utiliza además una variable instrumental para controlar por sesgo de clasificación. El autor encuentra sólo un efecto significativo del efecto del número de alumnos por curso en 8° de Educación Básica para lenguaje y matemática, pero que únicamente afecta a alumnos del *percentil 50*, por lo que concluye, dada la heterogeneidad del efecto, podría ser una medida potencialmente regresiva.

⁶⁰ Cursos reducidos pueden prevenir la aparición de problemas en niveles iniciales, pero su capacidad de aminorarlos en etapas posteriores no es clara.

Del mismo modo Mueller (2013) utiliza el método de regresión por cuantiles (incondicionales) permitiendo efectos diferenciados por desempeño, para estimar el efecto del tamaño del curso basándose en los datos de STAR. Encuentra que el efecto de cursos reducidos, favorecido por docentes experimentados, beneficia a todos los deciles de la distribución de logros, pero es mayor para los cuantiles medios y altos, lo que de nuevo corrobora el impacto heterogéneo pero regresivo de la medida.

4.4.1. Composición del Curso

Vinculado a los problemas de selección y prácticas compensatorias que agrupan a los estudiantes por habilidad y de la potencial capacidad de esta medida de favorecer mayormente a alumnos vulnerables y de grupos minoritarios, es que surge la pregunta por los efectos de la composición del curso. Las múltiples investigaciones que arrojan resultados inconclusos o efectos heterogéneos no deseables, sugieren que bajo ciertas condiciones, el efecto de reducciones de alumnos por curso no tendrían un impacto significativo y que en cambio modificaciones en la composición del curso que impactan los resultados vía el efecto par, podrían arrojar resultados más contundentes. La noción general detrás del efecto par es que la composición de los alumnos que comparten el aula afecta los logros de aprendizaje obtenidos por dichos estudiantes y que en consecuencia, diferentes configuraciones de estudiantes alcanzarán logros educativos distintos.

La teoría del aprendizaje cognitivo social señala que estudiantes aprenden tanto del profesor como de sus pares, por lo que contar con una cantidad de alumnos suficiente y de ciertas habilidades, es necesario para que este doble intercambio se dé de manera efectiva.

Cuantificar la influencia de la composición del curso en los logros de aprendizaje, o *efecto par*, fue lo que motivó el estudio de Sojourner (2013). Aprovechando la aleatoriedad en la composición de los cursos del proyecto STAR, se logra estimar el efecto del desempeño de los pares en Kindergarten, en los resultados de 1° de Educación Básica. El estudio no busca vincular la reducción de alumnos por curso con el *efecto par*, sino más bien aprovechar estos datos para obtener estimaciones más precisas de la influencia de los compañeros de curso sobre el desempeño. Es decir, se analiza el efecto del desempeño del grupo en promedio, más que la contribución que pueda hacer cada estudiante, con sus características únicas, en forma individual. Como se sabe, el experimento STAR fue diseñado aleatoriamente desde su inicio, sin embargo los alumnos que fueron asignados a curso *regulares* o *regulares con apoyo*, fueron reasignados aleatoriamente, incluyendo a nuevos alumnos entrantes, al comenzar 1° de Educación Básica. Por este motivo no se cuenta con datos del desempeño previo de los pares para el total de la muestra, alcanzando al 40% los datos faltantes. Para estimar el *efecto par*, los autores abordan el problema de los datos faltantes descomponiendo las variables causales del efecto en porciones observables y no observables, asumiendo separabilidad aditiva del desempeño promedio de los pares y linealidad⁶¹. Por su parte el problema de la composición de grupos se supera por diseño. Las

⁶¹ Los cursos se conforman aleatoriamente con algunos estudiantes para cuyos pares se conocen los datos de desempeño y otros desconocidos. La varianza de los datos observados a través de los cursos permite medir el efecto par de la forma usual, sin embargo la aleatoriedad de la asignación garantiza que el valor esperado promedio de los datos faltantes sea la misma a través de los grupos.

estimaciones sugieren que en promedio, los logros de aprendizaje en 1° de Educación Básica aumentan en 0,35 percentiles por cada aumento de percentil del desempeño promedio de los pares.

Algunos estudios, que buscando estimar el tamaño del efecto del número de alumnos por curso no obtienen resultados significativos, optan por incluir variables que dan cuenta de la influencia de la composición del curso, logrando explicar de esta forma los mejores resultados. Es así como los autores Dobbelsteen, Levin, & Oosterbeek (2002) se dispusieron a comprobar mediante una metodología similar al diseño de regresión discontinua de Angrist & Lavy, el efecto del tamaño de los cursos en **Holanda**, utilizando las reglas que vinculan la matrícula con el sueldo correspondiente a la cantidad de profesores necesarios para atender dicha cantidad de alumnos. Estas reglas generan discontinuidades en el tamaño de los cursos, menores que en Israel, pero significativas. Los resultados no sugieren efectos positivos en los logros de aprendizaje para alumnos de cursos reducidos, al contrario, en ocasiones éstos son negativos. Los autores interpretan estos resultados, tal como propone Borland et al., (2005) evaluando en conjunto dos efectos. Por un lado, al disminuir la cantidad de estudiantes, el profesor dispone de más tiempo para cada uno de ellos y eventualmente podría ofrecer mayores y mejores oportunidades de aprendizaje. Sin embargo también es cierto que cada individuo tiene menos compañeros con los que interactuar. Basados en la teoría del aprendizaje cognitivo social, que postula que los alumnos se benefician de un mayor número de compañeros de similar nivel de habilidades cognitivas, los autores incorporan una variable adicional que cuantifica la *cantidad de compañeros de similar CI*⁶². Esta variable adicional correlaciona positivamente con desempeño, esto es, a mayor número de compañeros de similar coeficiente intelectual mejor son los resultados⁶³. La conclusión general es que reducciones de alumnos por curso como actualmente están siendo implementadas en Holanda con un umbral de 28 alumnos por curso, no parecieran ser las medidas más efectivas para mejorar el desempeño de los estudiantes. Este indicador se vería más afectado por la composición del curso en términos de las competencias cognitivas de sus alumnos y su cantidad.

Resumen de los estudios descritos en la sección 4.4. Efectos en Grupos Minoritarios						
Título	Autor/Año /País	Publicación	Alcance	Especificidad	Enfoque	Hallazgo Clave
<i>Do Low-Achieving Students Benefit More from Small Classes? Evidence from the Tennessee Class Size Experiment</i>	Nye, B., Hedges, L., & Konstantopoulos, S. 2002, EEUU	<i>Educational Evaluation and Policy Analysis</i>	Local: reestimación del efecto de CSR con datos de STAR para grupos minoritarios.	Específico: busca evidencia sobre beneficio diferencial en grupos minoritarios y desaventajados.	Cuantitativo: reclasifica alumno de bajo rendimiento y reestima el progreso de este grupo.	Evidencia general CSR, pero resultados contradictorios y débiles de mayores beneficios en alumnos de bajo desempeño.
<i>Do Minorities Experience Larger Lasting Benefits From Small Classes? Evidence from a five-year follow-up of the Tennessee class size experiment.</i>	Nye, B., Hedges, L. V., & Konstantopoulos, S. 2004, EEUU	<i>Journal of Educational Research</i>	Local: seguimiento de 5 años de experimento STAR, Tennessee.	Específico: persistencia de mayores beneficios para grupos minoritarios (raza, etnia, género) en el tiempo.	Cuantitativo: regresión de puntajes de test por raza, etnia y género, asociado a CSR.	Los beneficios de CSR en K-3 perduran al menos los 5 años posteriores. Se reduce la brecha racial y étnica en lectura y de género en matemática.

⁶² En la misma línea argumentativa de los investigadores Borland, Howsen, & Trawick (2005), que proponen una relación no monotónica.

⁶³ Betts & Shkolnik (1999) encuentran una relación entre habilidades y logros de aprendizaje. Sección 6.1.

<i>Three generations of research on class size effects.</i>	Blatchford, P. 2011, UK	<i>APA Educational psychology handbook</i>	Internacional: revisión de estudios sobre CSR a nivel internacional.	General: impacto de CSR en puntajes de test, procesos de aula y estrategias en cursos de tamaño distinto.	Mixto: revisión de la literatura; tipos y líneas de investigación.	Distingue las falencias de dos enfoques y propone caminos de cooperación para futuros estudios en el contexto de la 3era generación.
<i>Class size, class composition, and the distribution of student achievement</i>	Bosworth, R.2011, EEUU	<i>Education Economics</i>	Local: Datos de escuelas públicas de Carolina del Norte, nivel 4° y 5°.	Específico: efecto de CSR medido en puntajes de test y tamaño de curso por características de alumnos.	Cuantitativo: control de efecto del profesor y controles de composición del aula.	Efectos heterogéneos de CSR, mayor para alumnos de bajo desempeño. Medida acorta la brecha más que subir el promedio.
<i>Experimental estimates of the impacts of class size on test scores robustness and heterogeneity.</i>	Ding, W., & Lehrer, S. 2011, Canadá	<i>Education Economics</i>	Local: reestimación de resultados de STAR, Tennessee, EEUU.	Específico: examina validez de multiplicidad de resultados de STAR y beneficios para grupos minoritario.	Cuantitativa: corrección de estimaciones por multiplicidad de resultados producto de CSR.	Validan los resultados generales de STAR pero encuentran mayores beneficios para alumnos aventajados.
<i>For whom the reductions count: A quantile regression analysis of class size and peer effects on scholastic achievement.</i>	Levin, J. 2001, Holanda	<i>Empirical Economics</i>	Nacional: panel de datos PRIMA, para 4°, 6° y 8°, año 1994-95, Holanda.	Específico: estima el impacto de CSR por cuantiles de logro en Holanda.	Cuantitativo: técnica de regresión por cuantiles.	Sólo encuentra un efecto significativo en 8°, pero afecta a estudiantes del percentil 50, por lo que concluye es una medida regresiva.
<i>Teacher experience and the class size effect — Experimental evidence.</i>	Mueller, S. 2013, Alemania	<i>Journal of Public Economics</i>	Local: datos de STAR en función de calidad docente.	Específico: efecto de CSR mediado por calidad (años de experiencia)	Cuantitativa: caracterización de docente según años de experiencia, dummy.	Solo docentes experimentados logran un efecto producto de CSR. Beneficia a todos los deciles, pero menos a deciles bajos.
<i>Identification of Peer Effects with Missing Peer Data: Evidence from Project STAR.</i>	Sojourner, A. 2013, EEUU	<i>The Economic Journal</i>	Local: estimación del impacto de efecto par en nivel 1°, usa datos de STAR.	Específico: estima el efecto par en logros de aprendizaje individual en el contexto de STAR.	Cuantitativo: utiliza pre test del 60% de la muestra. Supera problema de datos faltantes.	Desempeño en 1° aumenta en 0,35 percentiles por cada aumento de percentil del desempeño promedio de los pares.
<i>The causal effect of class size on scholastic achievement: distinguishing the pure class size effect from the effect of changes in class composition.</i>	Dobbelsteen, S., Levin, J., & Oosterbeek, H. 2002, Holanda	<i>Oxford Bulletin of Economics and Statistics</i>	Local: datos de la encuesta PRIMA de Holanda para nivel, 4°, 6° y 8°, año 1994/95.	General: estima el efecto de CSR en logros de aprendizaje en Holanda. Analiza la composición del curso y su impacto.	Cuantitativa: estima efecto de CSR con modelo de RD sin éxito. Variable composición del curso.	CSR no impacta logros de aprendizaje en Holanda. Encuentra efecto par motivado por la teoría del aprendizaje social (N° de compañeros con similar CI)

4.5. Externalidades y Efectividad

La implementación y ejecución de programas de reducción del número de alumnos por curso, que han sido extensamente documentados y evaluados, entregan evidencia sobre ciertas consecuencias no intencionadas o externalidades asociadas a la medida, que pueden contrarrestar los efectos esperados.

El masivo programa de reducción de alumnos por curso iniciado en 1996 en el estado de California, que disminuyó de 30 a 20 estudiantes por sala, dejó al descubierto los efectos indirectos asociados del cambio en la composición docente, dado el importante aumento de profesores sin experiencia y/o con certificación incompleta que fue necesario contratar para poder implementar la medida. En los dos primeros años se crearon 25.000 nuevas plazas para profesores, lo que produjo un *debilitamiento en la calidad del profesorado*. La mayor parte de este nuevo grupo de docentes fue asignado a escuelas que atienden una población mayoritariamente

vulnerable y con gran presencia de grupos minoritarios, lo que significa que los beneficios del programa fueron asignados en forma regresiva. En parte, estas vacantes fueron generadas por la partida de docentes experimentados, motivados por nuevas plazas creadas en comunidades más prósperas y acomodadas. La evaluación de este programa reportó que producto de disminuciones de 10 alumnos por curso en promedio, los resultados académicos en matemática y lectura aumentaron del orden de 0.1σ y 0.06σ respectivamente. Sin embargo, el efecto indirecto de la peor calidad relativa de los docentes adicionales, contrarrestó en parte estos logros. Sus autores señalan que el efecto de un profesor sin experiencia respecto de uno con al menos dos años de práctica se traduce en menores logros académicos correspondientes a 0.1σ y 0.07σ para matemática y lectura respectivamente, prácticamente idénticos a los beneficios del programa en cuestión (Jepsen & Rivkin, 2009).

Otra externalidad que se observó en la implementación del programa de reducción de alumnos por curso en California fue un *aumento del número de estudiantes por curso en niveles superiores* de Educación Básica con el propósito de cumplir las reglas del programa establecido para los niveles iniciales K-3. Esta compensación significó aumentos de 2 alumnos promedio por curso en los niveles superiores de 4° y 5° de Educación Básica, lo que se tradujo en menores logros académicos del orden de 0.07σ y 0.03σ en matemática para 5° y 4° de Educación Básica respectivamente (Sims, 2009). El autor señala que la disminución en logros de aprendizaje en cursos superiores, producto del aumento de alumnos por curso, proporciona mayor evidencia a favor de la medida favorece mayores logros de aprendizaje.

Asimismo, la experiencia que dejó la implementación del programa de reducción de alumnos por curso en Ontario, Canadá, dice relación con las consecuencias no intencionadas en la distribución de los recursos. En particular se observó un *menor acceso a profesionales especialistas y docentes de apoyo en cursos superiores* de Educación Básica, al estar estos recursos concentrados en los niveles iniciales producto de la intervención (Bascia, 2010).

La implementación del programa SAGE en el estado de Wisconsin, también contribuye a una mejor comprensión de las implicancias de este tipo de intervenciones. Su análisis caracteriza la adaptación de escuelas y profesores involucrados a las nuevas configuraciones de aula⁶⁴ producidas por la intervención, la que –como alegan sus evaluadores– se lleva a cabo *sin un desarrollo profesional* específico. El financiamiento del programa cubrió la contratación de docentes adicionales, pero se vio limitado por la infraestructura, lo que significó en ciertos casos reasignar personal y espacio físico. En algunas escuelas, por falta de espacio, tuvieron que constituir duplas de trabajo altamente capacitadas a ejercer en modalidad de equipo, pero donde uno de los dos profesores lidera la enseñanza y el otro realiza trabajo administrativo. Esta estrategia resultó en aumentos reales del número de alumnos por profesor, para los estudiantes involucrados y en dificultades en la implementación de la co-enseñanza. Se advierte así la existencia de amplias diferencias en las capacidades y habilidades de escuelas y docentes para incorporar nuevas estrategias y adaptarse a los nuevos escenarios. Los autores señalan que es

⁶⁴ Tanto 15:1 como 30:2.

poco probable que dos profesores trabajen a la par y logren mayores aprendizajes en todos sus alumnos sin un entrenamiento específico. Indican que, ya sea en cursos reducidos o en situaciones de co-enseñanza, los profesores pueden mejorar su desempeño si se les apoya de forma específica, desarrollando estrategias para enseñar en contextos diversos, y evitar que apliquen las mismas herramientas a estos nuevos escenarios. Afirman que el no asumir esta intervención como una medida tanto programática como instruccional, arriesga el éxito de ésta, al no contemplar un desarrollo profesional específico (Graue, Hatch, Rao, & Oen, 2007).

A medida que surgen las externalidades descritas, aparecen estudios que comparan la efectividad de esta medida con el impacto de recursos equivalentes u otras intervenciones alternativas como ejemplifican los siguientes estudios.

El estado de Florida en 2002, dicta una ley que modifica su constitución e implementa obligatoriamente un umbral máximo de 18 alumnos por curso entre Kindergarten y 3° de Educación Básica, 22 alumnos entre 4° y 8° de Educación Básica, y 25 para Educación Media, a partir del año 2004. Producto de la implementación de la medida se observa un débil pero positivo efecto sobre la tasa de ausentismo y violencia escolar; aspectos no cognitivos deseables que se asocian a reducciones de alumnos por curso. Con el propósito de estimar la relación costo-efectividad académica de la intervención⁶⁵, se toman como grupo de control a escuelas que recibieron fondos adicionales equivalentes, pero sin enfrentar la obligatoriedad de utilizarlos en reducir alumnos por curso, sino que con la libertad de asignarlos discrecionalmente (Chingos, 2012). Los resultados sugieren que luego de tres años de tratamiento, el desempeño en las escuelas que redujeron la cantidad de alumnos por curso fue prácticamente idéntico al que obtuvieron las escuelas que recibieron los recursos adicionales para libre disposición. Así, el autor concluye que la efectividad de esta intervención sería equivalente a otras medidas que puedan ser financiadas con recursos adicionales que reciban las escuelas, desestimando su superioridad en términos de efectividad⁶⁶.

Más drástica es la conclusión de Yeh (2009) al sostener que las evaluaciones formativas intensivas -como medida alternativa a reducciones de alumnos por curso- como intervención son 124 veces más costo-efectivas que disminuir el tamaño de los cursos. De acuerdo al autor, para alcanzar resultados similares es necesario evaluar el progreso de los estudiantes en lectura y matemática entre tres a cinco veces por semana. Este sistema de evaluaciones focaliza la actividad educativa en torno a la evaluación para el aprendizaje, en vez a la medición y la responsabilización.

Resumen de los estudios descritos en la sección 4.5. Externalidades y Efectividad						
Título	Autor /Año/País	Publicación	Alcance	Especificidad	Enfoque	Hallazgo Clave
<i>Class Size Reduction and Student Achievement: The Potential Tradeoff between Teacher Quality and Class Size.</i>	Jepsen, C., & Rivkin, S. 2009, EEUU	<i>Journal of Human Resources</i>	Local: datos del programa de CSR en California, 1997-2001, nivel K-3.	Específico: impacto de CSR en desempeño contrarrestado por menor calidad docentes entrantes.	Cuantitativo: parametrización de características docentes, puntajes de test, tamaño del curso.	La peor calidad de los 25000 nuevos docentes necesarios para implementar CSR, contrarrestó los logros producto de CSR.

⁶⁵ Se concentra en los niveles de 4° a 8° de Educación Básica, con una reducción promedio de 2 a 3 alumnos por curso. Las escuelas de control ya cumplían con las cuotas máximas establecidas.

⁶⁶ Una limitación de este estudio es la modesta reducción en la cantidad de alumnos evaluada.

<i>Crowding Peter to educate Paul: Lessons from a class size reduction externality.</i>	Sims, D. R. 2009, EEUU	<i>Economics of Education Review</i>	Local: datos del programa de CSR en California, 1997-2001, nivel K-3.	Específico: cambios en tamaño de cursos superiores producto de CSR en California.	Cuantitativa: puntajes de test en cursos de niveles reducidos y aumentados (4° y 5°)	Al reducir cursos en K-3 aumentó la cantidad de alumnos en niveles superiores con impacto negativo en logros de aprendizaje.
<i>Ontario's Primary Class Size Initiative: Report on Early Implementation.</i>	Bascia, N. 2010, Canadá	<i>Canadian Education Association</i>	Nacional: datos del programa de CSR a 20 alumnos en Ontario, Canadá, 2008-2009.	General: evaluación general del programa (beneficios y consecuencias)	Mixto: análisis de datos y entrevistas.	Exitoso programa, pero debe mejorar en desarrollo docente, instrucción en cursos combinados, agrupar estudiantes y asignar docentes; equilibrio entre ámbito local y regional.
<i>The Wisdom of Class-Size Reduction</i>	Graue, E., Hatch, K., Rao, K., & Oen, D. 2007, EEUU	<i>American Educational Research Journal</i>	Local: datos programa de CSR en Wisconsin SAGE, 96-97 por 5 años, nivel K-3.	General: implementación y ejecución del programa y sus consecuencias.	Cualitativa: estudio de caso.	Para lograr efectividad en CSR se requiere de desarrollo profesional específico al nuevo contexto.
<i>The impact of a universal class size reduction policy: Evidence from Florida's statewide mandate.</i>	Chingos, M. M. 2012, EEUU	<i>Economics of Education Review</i>	Local: resultados CSR y recursos equivalentes en Florida, EEUU, segundo ciclo, 2006-9.	General: evaluación de la efectividad de CSR relativa a otras intervenciones.	Cuantitativa: contrastación de desviaciones de tendencias en desempeño producto de CSR y recursos equivalentes.	El efecto adicional de CSR es casi nulo en 2do ciclo. Si hay evidencia de efectos en resultados no cognitivos.
<i>Class size reduction or rapid formative assessment? A comparison of cost-effectiveness.</i>	Yeh, S. S. 2009, EEUU	<i>Educational Research Review</i>	Nacional: datos de meta análisis EEUU y STAR, SAGE, California. Estimaciones costos RAND	Específico: comparación de costo-efectividad de CSR y RFA.	Mixto: calculo de costos y beneficios, entrevistas, comparación entre diferentes medidas.	CSR es 124 veces menos costo efectivo que la evaluación formativa intensiva (entre 3 y 5 veces por semana).

4.6. Validez y Capacidad Predictiva

La validez interna y externa de los diseños de investigación hasta aquí analizados caracteriza su sensibilidad para detectar relaciones causales entre las variables y la posibilidad de generalizar sus efectos y causas a otros grupos o contextos.

Los primeros *estudios correlacionales* que examinan vínculos entre el número de alumnos por curso y resultados de aprendizaje medidos como puntajes de test, encuentran asociaciones positivas entre la cantidad de alumnos y su desempeño: esto es, a mayor cantidad de alumnos mejor desempeño. Su principal falla fue su inhabilidad para aislar el tamaño del curso de otros factores que impactan el desempeño de los alumnos. Si bien controlaron por ciertas variables que explican en parte los resultados -como la ocupación de los padres o el tamaño de la escuela- no advirtieron que podrían deberse a que comúnmente los estudiantes con mayores dificultades son asignados a cursos pequeños y los más avanzados a cursos más numerosos, con profesores más experimentados y eficaces haciéndose cargo de estos últimos. Así, la concentración de alumnos más aventajados en cursos grandes junto a la agrupación de aquellos de peor rendimiento en cursos pequeños, explicaría la relación positiva entre cantidad de alumnos por curso y desempeño. En consecuencia, un diseño de investigación correlacional que no considera la endogeneidad en la asignación de los alumnos a cursos de diferentes tamaños, no presenta validez interna en cuanto es incapaz de establecer relaciones causales válidas entre las variables.

Así como resulta vital especificar la endogeneidad en la asignación de alumnos a cursos de tamaños distintos en el modelo, también es imprescindible controlar los otros factores que explican la variabilidad en los logros de aprendizaje que evidencian los estudiantes, para así aislar la variable de interés y poder estimar su real impacto. Sin embargo la omisión de variables explicativas ha sido otro problema que ha introducido distorsiones y confusión en la comprensión del fenómeno analizado.

La validez del constructo fue otra fuente de problemas tempranos. Inicialmente se usó de forma intercambiable el *número de alumnos por curso* y la *proporción de estudiantes por docente*, en circunstancias que ambos constructos representan situaciones educativas muy diferentes, cuyas repercusiones en el desempeño no tienen por qué ser idénticas. La primera refleja la realidad del aula y determina el contexto en que se dan los procesos de enseñanza-aprendizaje entre alumnos y profesores. La segunda en cambio expresa la cantidad de recursos humanos asignados a educación, pero a escalas comúnmente mayores que el aula, como son la escuela, el distrito, la región e incluso el país, contabilizando a todos los profesionales de la educación de forma proporcional. Por consiguiente, las estimaciones que utilizan la proporción de estudiantes por profesor como *proxy* del número de alumnos por curso también ven amenazada su validez interna en cuanto establecen relaciones espurias entre las variables.

Los diseños de investigación de *experimentos aleatorios* caracterizados en este trabajo, abordan las amenazas a la validez interna descritas anteriormente. El asignar de forma aleatoria a los sujetos a los distintos grupos (tratamiento y control) hace exógena la distribución de los alumnos a los cursos de tamaños diferentes y por lo tanto controla otras posibles causas que expliquen diferencias en desempeño. Por otra parte al definir el tamaño de los cursos como parte del diseño del experimento, el constructo queda determinado internamente. Sin embargo hay algunas amenazas a la validez interna y externa asociadas a experimentos aleatorios que se han descrito en la literatura. Es más probable que un diseño de investigación con inferencias múltiples, que reporta varios resultados o hallazgos a partir de un mismo tratamiento, encuentre impactos causales significativos donde no existen, ya que aumenta la probabilidad de falsos positivos. En consecuencia podría sobreestimar los efectos amenazando su validez interna. Por otra parte, una amenaza de factor situacional a la validez externa es el efecto reactivo del ambiente o situación de investigación, llamado efecto Hawthorne. Este postula que los sujetos tienden a responder al ambiente y reaccionar de modo diferente cuando están siendo observados, lo que podría sobreestimar el impacto de la intervención en condiciones experimentales.

Estrategias de identificación cuasi experimentales como el *diseño de regresión discontinua* o las *variaciones naturales de la población*, determinan causas exógenas en las variaciones del tamaño de los cursos. El *diseño de regresión discontinua* lo hace mediante las reglas de umbrales máximos y mínimos de alumnos por curso y la estrategia que utiliza *variaciones naturales de la población* se basa en las diferencias entre cohortes a lo largo de los años analizados. Si bien estos diseños externalizan la variación del tamaño del curso, continúan afectos a los problemas de clasificación que deben rectificarse, sobre todo en sistemas educativos mercantilizados donde la endogeneidad se da tanto entre escuelas como al interior de ellas. En consecuencia estas

aproximaciones requieren de métodos que corrijan estos sesgos, siendo el uso de variables instrumentales y/o técnicas de efecto fijo los más comunes. Sin embargo la literatura advierte lo problemático que es hallar variables instrumentales adecuadas y la escasez de bases de datos longitudinales apropiadas. Los resultados de este tipo de estudios son en muchos casos contradictorios, revelando la dificultad que implica especificar correctamente el modelo con datos no aleatorios.

Otra aproximación utilizada para examinar el número de alumnos por curso son los estudios naturalísticos longitudinales a gran escala que exploran las relaciones entre el tamaño del curso y los resultados en logros de aprendizaje, mediante observaciones directas de los procesos tal como suceden en aulas reales y cubriendo todo el rango de tamaños de cursos existentes. Las posibles variables de interferencia como desempeño previo, nivel de pobreza y características del profesor se ajustan posteriormente. Esta estrategia de identificación evita los problemas de validez externa de situaciones experimentales y tiene la ventaja que permite observar los procesos de aula intervinientes.

Como se describe en esta sección, las dificultades metodológicas que impactan la validez y en consecuencia la capacidad predictiva de los estudios en torno al número de alumnos por curso son de diversa índole y han alimentado nociones más bien imprecisas sobre el real impacto de la intervención.

No cabe duda que los estudios más confiables son aquellos cuyos diseños de investigación corresponden a experimentos aleatorios y estudios naturalísticos, en particular STAR y CSPAR. Principalmente porque abordan la endogeneidad, ya sea mediante la aleatorización de los sujetos a cursos pequeños y regulares, o porque al observar la realidad tal como se da en aulas de distintos tamaños, particularizan los impactos por tipo de alumno de acuerdo a sus características, aportando en autenticidad. Además estos estudios incorporan la observación de los procesos de aula⁶⁷ intervinientes y en ese sentido enriquecen el análisis y las estimaciones, proporcionando un modelo sobre los mecanismos que explican los potenciales beneficios.

V. RESULTADOS INTERNACIONALES

Investigadores y académicos han pasado varias décadas intentando estimar el tamaño del efecto de reducciones en la cantidad de alumnos por curso en los resultados educativos. En esta primera etapa la investigación se centró en establecer relaciones directas entre el tamaño del curso y el desempeño de los estudiantes medido como puntajes en test, lo que buscaba reunir evidencia y dar respuesta a la pregunta ¿La reducción de alumnos por curso genera mayores logros de aprendizaje? Varios hallazgos han sido ya descritos en la sección anterior, asociados a investigaciones que se focalizan en algún aspecto metodológico o temático específico analizado.

⁶⁷ Desarrollado en detalle en la sección VI. Procesos de Aula.

Como hemos visto hasta aquí, las fuentes de variación del tamaño de los cursos que han permitido estudiar su impacto se producen por razones de diversa naturaleza: realización de experimentos aleatorios como STAR, por variaciones naturales de la población (Hoxby, 2000), por discontinuidades producto de reglas de umbrales máximos y mínimos de alumnos por curso (Angrist & Lavy, 1999), por la implementación de programas de reducción del tamaño de los cursos como SAGE, por aumentos de alumnos en niveles no intervenidos (Sims, 2009) y por comparación de tamaño de cursos de niveles contiguos que rinden TIMSS (Wössmann & West, 2006a). En las secciones siguientes aparecerán otros diseños de investigación que explotan la diferencia del tamaño de los cursos de distintas asignaturas analizadas (Dee & West, 2011), la transición a cursos regulares posterior a la intervención (Babcock & Betts, 2009; Krueger & Whitmore, 2001), la demanda del curso analizado según hora del día (Cortes, Moussa, & Weinstein, 2013) y el aumento de cobertura educativa con aumentos en el tamaño de los cursos (Medina & Bríñez, 2007). Sin embargo, y a pesar de que surjan nuevas y mejores oportunidades para estimar el impacto del tamaño del curso, con métodos y técnicas más sofisticados que permitan controlar los posibles sesgos, la idea de que este efecto sea absoluto y universal parece dudosa (Leuven et al., 2008). Como se discute a continuación, las grandes diferencias entre las realidades educativas analizadas, los disímiles puntos de partida y el contexto cultural en que están inmersos parecieran ser -en conjunto las dificultades metodológicas descritas- la causa de los resultados mixtos e inconclusos de un sinnúmero de investigaciones conducidas en esta línea.

5.1. Estudios Comparativos a Gran Escala

Actualmente, y gracias a las pruebas internacionales como PISA y TIMSS aplicadas en varios países del mundo, se cuenta con datos que permiten comparar el efecto de esta intervención a través de distintas regiones. La intuición sugiere que una reducción de alumnos por curso debiera tener un mayor impacto en países más atrasados y con cursos más numerosos, esto es, en aquellos sistemas que ofrezcan amplias oportunidades de avanzar y mejorar en logros de aprendizaje.

La tercera versión de la prueba TIMSS 94-95 de matemática y ciencias tuvo una característica distintiva: el test se aplicó a los dos niveles contiguos con mayor presencia de estudiantes de 13 años (7° y 8° de Educación Básica en la mayoría de los países)⁶⁸, lo que hizo posible comparar el desempeño relativo de estudiantes de 7° y 8° de Educación Básica al interior de escuelas, con cursos de entre 20 y 33 estudiantes en promedio. Los autores Wössmann & West (2006a) analizan el efecto del tamaño del curso para 11 sistemas educativos distintos⁶⁹. Mediante una estimación de mínimos cuadrados en dos etapas eliminan las fuentes endógenas de variación tanto al interior de las escuelas como entre escuelas⁷⁰. Al considerar la estrategia de identificación que elimina ambas fuentes de endogeneidad, sólo perdura un caso contra intuitivo con coeficientes positivos y significativos: Portugal en matemática. A su vez aparecen 4 casos con

⁶⁸ La prueba TIMSS posteriormente sólo midió los resultados de alumnos de 8° de Educación Básica.

⁶⁹ Bélgica, Grecia, Portugal, España, Francia, Islandia, Rumania, República Checa, Singapur, Eslovenia y Canadá. Sólo estos 11 de los 38 países de la muestra inicial cumplen los requisitos del modelo.

⁷⁰ Variable instrumental: tamaño del curso promedio; efecto fijo de la escuela.

coeficientes negativos significativos, revelando el beneficio de cursos reducidos en particular para Grecia e Islandia. Sin embargo prevalecen estimaciones de coeficientes estadísticamente no significativos en la muestra, lo que sugiere la ausencia de un impacto causal del tamaño del curso en logros de aprendizaje en la mayoría de los países de la muestra. Al comparar la magnitud del tamaño del efecto encontrado en este estudio para los 4 casos negativos, eso es, un aumento del desempeño en 0.024σ al disminuir un alumno por curso, con los resultados reportados por Krueger para STAR⁷¹ se verifica que son bastante similares.

Otra importante conclusión es que la medición del tamaño del efecto de reducciones de alumnos por curso sobre aprendizajes, debe ser abordada por separado para los distintos sistemas educativos, lo que significa que hallazgos en un país no necesariamente pueden ser generalizados a los de otro sistema educativo (validez externa).

Se observa que los países más favorecidos por cursos reducidos presentan un desempeño menor que el resto, incluso teniendo cursos de menor o igual tamaño que la media de la muestra. Por otro lado, para Singapur que presenta los mejores resultados y los cursos más numerosos de la muestra, un efecto beneficioso de la intervención queda totalmente descartado, lo que sugiere que la medida es más productiva para sistemas educativos menos efectivos. La efectividad de un sistema educativo va de la mano con el gasto en educación per cápita en cada país, lo que a su vez se refleja en las remuneraciones de los profesores. Un nivel de salarios bajo en promedio indica que los profesores provienen de la parte baja de la distribución de capacidades de los profesionales del país. Países como Grecia e Islandia tienen en términos relativos mayor cantidad de profesores pero mal remunerados, que aquellos sistemas insensibles a reducciones de alumnos por cursos, que tienen relativamente una menor cantidad de profesores, pero bien pagados. En términos de las calificaciones de los docentes, en sistemas educativos menos efectivos los profesores tienen a lo más un BA⁷², a diferencia del resto de los países de la muestra donde al menos un 50% cuenta con niveles educativos mayores que un BA. Es decir, profesores más aptos son capaces de promover el aprendizaje independientemente del tamaño del curso, sin embargo docentes menos hábiles, si bien pueden enseñar razonablemente bien en cursos reducidos no lo logran en cursos numerosos.

Aprovechando los resultados de TIMSS 2003 para 8° de Educación Básica en matemática y ciencias, los autores Altinok & Kingdon (2012) analizan el efecto del tamaño del curso en los logros de aprendizaje en 33 países. Abordan el problema de endogeneidad en la medición, relacionando la diferencia en el desempeño de un estudiante en las distintas áreas evaluadas, con el número de alumnos por curso en los correspondientes subsectores, en vez de a través del tiempo. Dado que las clases se dictan en salas diferentes, a cada asignatura evaluada le corresponde un curso de tamaño distinto⁷³. Para efectos comparativos (validez externa) los países participantes se agruparon en tres categorías: 1. países desarrollados, 2. países de Europa del Este y de Asia

⁷¹ 0.22σ al disminuir 7 u 8 alumnos por curso ó 0.03σ por cada alumno menos.

⁷² *Bachelor of Arts*, aproximadamente equivalente a una licenciatura.

⁷³ Se eliminaron de la muestra aquellos países que no cumplían con los requisitos de presentar diferencias en el tamaño de los cursos en estas asignaturas.

Central, y 3. países en vías al desarrollo. Para el primer grupo, 11 de los 14 países presentan una estimación positiva y significativa del tamaño del curso sugiriendo contra intuitivamente que un aumento en la cantidad de estudiantes mejora los logros de aprendizaje; para el segundo grupo, de los 13 países 5 presentan coeficientes positivos y 4 negativos, débilmente significativos. El último grupo, los países en vías de desarrollo, presentan mayoritariamente una relación negativa entre el tamaño del curso y el desempeño, lo que sugiere potenciales beneficios de una disminución de alumnos por cursos. Chile, parte de este último grupo, es el país que presenta el mayor efecto de la medida en este estudio, donde una disminución de 7.3 alumnos en promedio, reportaría mayores logros de aprendizaje cuantificados en $0,16\sigma$. Comentarios de Eric Hanushek sobre el trabajo de Altinok & Kingdon sugieren que los resultados serían negativos y significativos sólo para países en vías de desarrollo y donde la calidad de los docentes es relativamente baja (Hanushek & Woessmann, 2010).

Resumen de los estudios descritos en la sección 5.1. Estudios Comparativos a Gran Escala						
Título	Autor/Año /País	Publicación	Alcance	Especificidad	Enfoque	Hallazgo Clave
<i>Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS.</i>	Wössmann, L., & West, M. 2006, Alemania	<i>European Economic Review</i>	Internacional: datos de TIMSS ciencias y matemática, para 11 sistemas educativos 1994-95, nivel 7° y 8°.	Específico: impacto de CSR en resultados prueba TIMSS a los 13 años en 11 países.	Cuantitativo: usa variación de tamaño de cursos en 7° y 8° para ambas asignaturas. Utiliza efecto fijo de escuela y IV.	2 de los 11 países muestran efectos negativos significativos de CSR; países con cuerpo docente mal remunerado, menos efectivo.
<i>New evidence on class size effects: A pupil fixed effects approach</i>	Altinok, N., and, & Kingdon, G. 2012, UK	<i>Oxford Bulletin of Economics and Statistics</i>	Internacional: puntajes TIMSS (ciencias y matemática) 2003 para 8° en 33 países.	Específico: efecto de CSR medido en puntajes de matemática y ciencias y tamaño de curso.	Cuantitativa: efecto fijo del alumno (compara puntajes entre matemática y ciencias) en vez de a través del tiempo.	Sólo en 1/3 de los países el efecto de CSR es negativo y pequeño. Países en vías de desarrollo, con docentes menos efectivos son más sensibles a CSR.
<i>The Economics of International differences in Educational Achievement.</i>	Hanushek, E., & Woessmann, L. 2010, EEUU	<i>National Bureau of Economic Research.</i>	Internacional: estudio comparativo internacional.	General: revisa la literatura económica sobre diferencias en logros educativos internacionales. * Uno de los temas es el efecto de CSR.	Cuantitativo: análisis de resultados de pruebas estandarizadas internacionales.	* La presencia de un efecto de CSR a nivel internacional está sistemáticamente vinculado a países cuyo cuerpo docente es peor remunerado y de menor nivel educativo.

5.2. Estudios por Región

Las investigaciones, estudios y experimentos relativos al tamaño del curso y sus efectos se han concentrado principalmente en EEUU y Europa, sin embargo es interesante revisar las iniciativas llevadas adelante en países asiáticos dadas sus diferencias culturales, tradición de cursos numerosos y buenos resultados relativos. Por su parte Latinoamérica, como región emergente, también ha avanzado en esta materia y cuenta con estudios que contribuyen al debate actual del tema.

5.2.1. Asia

El contexto cultural asiático sugiere importantes diferencias en el impacto de un tratamiento de la naturaleza de la reducción de alumnos por curso: por ejemplo, la proporción de tiempo dedicado en tareas de aprendizaje es en promedio mayor al 90%, a diferencia de los países

occidentales donde este índice fluctúa entre un 60% y 70%. No obstante, algunos países de Asia oriental -como China, Hong Kong, Singapur, Corea y Japón- y otros del sur de Asia como India han tendido a favorecer políticas de reducción de alumnos por curso últimamente (Blatchford, Bassett, & Brown, 2011).

En **China** a raíz de la baja de la natalidad y de un debate político muy intenso, se inicia la implementación de un programa de reducción de estudiantes por curso en escuelas básicas en el año escolar en 2009-2010. Previo a esta reforma, el gobierno encargó un estudio a gran escala llamado *SCT Small Class Teaching* en que basarse, el más reciente y único de estas características en el contexto asiático, conducido entre 2004 y 2008. Si bien sus resultados entre cursos experimentales y de control no se distinguieron claramente, la principal desventaja de su diseño de investigación es que junto con la reducción del tamaño de los cursos, se implementaron una serie de iniciativas de desarrollo profesional, por lo que no es posible saber a qué se atribuyen exactamente los resultados. Sin embargo sus hallazgos incluyen la constatación de que profesores en cursos numerosos enseñan al grupo curso con más frecuencia que en cursos reducidos; que en contextos de cursos reducidos se formulan preguntas más desafiantes a los alumnos; se utiliza más el trabajo en parejas y grupo; y que los docentes interactúan sostenidamente con estudiantes en forma individual más intensamente, entre otros (Galton & Pell, 2009).

En 2008 **Japón** contaba con 28 alumnos por curso promedio, con un umbral máximo de 40. El año 2011 se decreta que el límite superior no puede superar los 35 estudiantes. Los autores Akabayashi & Nakamura (2012) aprovechan esta disminución para estudiar el efecto en el desempeño en lenguaje y matemática de los alumnos de 6° de Educación Básica y 1° de Educación Media de escuelas públicas. Utilizando un modelo de regresión discontinua, análogo al de Angrist & Lavy, y mediante una metodología de valor agregado -que permite controlar por características no observables y calidad docente, entre otros- encuentran que la reducción de alumnos por curso tiene un gran impacto positivo en los resultados de los puntajes de evaluaciones del idioma japonés sólo en 6° de Educación Básica. La disminución de un alumno genera una mejora de los puntajes en $0.0112\sigma^{74}$. Sin embargo, al correlacionar estos aumentos de desempeño con el valor del terreno en donde se ubica la escuela, se revela que éstos son significativos sólo en áreas relativamente acomodadas. Este hallazgo implica a juicio del autor, que la medida no es un buen mecanismo para disminuir las brechas de desempeño entre escuelas en Japón.

A diferencia de las experiencias en China y Japón, el programa implementado en **India** durante 2006 y 2007, que aleatoriamente asignó a 100 escuelas la posibilidad de contratar un docente adicional (paradocente⁷⁵), fue exitoso. La iniciativa permitió reducir entre 5 y 10 alumnos por curso⁷⁶ en los niveles de 1° a 5° de Educación Básica. Los resultados del tratamiento son concluyentes: al finalizar el segundo año del programa los estudiantes de escuelas participantes mostraron un rendimiento superior a alumnos de las escuelas de control en 0.15σ y 0.13σ en

⁷⁴ $\frac{1}{3}$ y $\frac{1}{2}$ del tamaño del efecto encontrado por Angrist & Lavy y Krueger respectivamente.

⁷⁵ Recibe un sueldo correspondiente a la quinta parte del sueldo de docentes regulares.

⁷⁶ 5 (de 34 a 28) y 10 (36 a 26) alumnos por curso.

matemática y lenguaje respectivamente. Lo anterior revela que, al menos en países en vías al desarrollo, en contextos rurales y con cursos numerosos, la implementación de reducciones de alumnos por curso es suficientemente efectiva, incluso con profesores sin certificación. Posteriormente en 2010 se promulga la ley “Derecho a Educación” en India, la que estipula una reducción en la proporción profesor/alumno de 40:1 a 30:1, sin embargo establece la prohibición de contratar profesores sin certificación o paradocentes. Los autores sugieren un mayor debate respecto de la prohibición, en circunstancias que la implementación de este tipo de medida significa un gran costo para el país. Argumentan que la implementación y ejecución de la nueva ley a nivel nacional puede lograrse contratando paradocentes, cuya efectividad fue comprobada en el programa descrito (Muralidharan & Sundararama, 2010).

Por su parte G. J. Harfitt (2012a) en un estudio exploratorio a pequeña escala, busca dar cuenta del impacto que puede tener una reducción de los cursos de educación secundaria en **Hong Kong** en el aprendizaje de un segundo idioma. En culturas de herencia confuciana, la ansiedad del aprendizaje del inglés hace que los alumnos adopten un rol pasivo, lo que a su vez afecta su desempeño y su capacidad de comunicación. Se realizaron múltiples estudios de casos, analizando las experiencias de clases de inglés a estudiantes chinos de cursos reducidos (21-25 alumnos) y cursos numerosos (38-41 alumnos), donde cada profesor enseñaba un curso de cada tipo, de manera de controlar por calidad docente. De acuerdo a lo que reportaron los estudiantes, los cursos reducidos promueven una mayor sensación de seguridad al interior del aula, lo que disminuye el miedo a recibir una mala evaluación de sus profesores y pares, a quienes identifican como su grupo de estudio. Este estudio releva la voz del estudiante como actor fundamental del proceso educativo y el potencial de esta intervención para romper las barreras culturales que entorpecen el aprendizaje de una segunda lengua en países de Asia Oriental.

Resumen de los estudios descritos en la sección 5.2.1. Asia						
Título	Autor/Año /País	Publicación	Alcance	Especificidad	Enfoque	Hallazgo Clave
<i>Study on Small Class Teaching in Primary Schools in Hong Kong: Final Report.</i>	Galton, M., & Pell, T. 2009, Hong Kong	<i>Hong Kong: Education Bureau and Cambridge University.</i>	Local: datos de CSR de estudio a gran escala SCT conducido entre 2004-09 en Hong Kong, niveles 1-3.	General: desempeño y apreciaciones de CSR y desarrollo profesional combinado.	Mixto: análisis de datos, observación de clases, y entrevistas a alumnos y profesores.	Se constatan ciertas prácticas pedagógicas que se ven favorecidas en SC: individualización, desafío, trabajo en grupo y parejas, más interacción docente-alumno.
<i>Can Small Class Policy close the gap? An empirical analysis of class size effects in Japan</i>	Akabayashi, H., & Nakamura, R. 2012, Japón	<i>Tokyo Center for Economic Research</i>	Nacional: CSR en Japón y Yokohama, para 6° y I Medio, 2008-09.	Específico: impacto de CSR asociado al valor del terreno de las escuelas.	Cuantitativo: puntajes de test, CS, valor del terreno donde se ubica la escuela.	CSR tiene gran impacto en lenguaje en 6° nivel, en áreas más acomodadas. Medida regresiva.
<i>Contract Teachers: Experimental Evidence from India.</i>	Muralidharan, K., & Sundararama, V. 2010, India	<i>NBER Working Paper No. 19440.</i>	Local: datos de contratación de paradocente adicional en sector rural de India, años 2006-07. Nivel 1°-5°.	Específico: impacto del paradocente adicional (menor certificación), implementado como CSR.	Cuantitativo: puntajes de test, estima efectividad relativa de paradocentes y docentes regulares.	Producto de CSR, luego de 2 años, estudiantes mejoran en matemática y lenguaje en 0.15σ y 0.13σ respectivamente. Paradocente es igualmente efectivo que docente regular.
<i>Class size and language learning in Hong Kong: the</i>	Harfitt, G. J. 2012, Hong Kong	<i>Educational Research</i>	Local: datos sobre 4 docentes de inglés de	Específico: examina el efecto de CSR sobre la	Cualitativo: observación de clase, transcripciones,	CSR promueve una sensación de seguridad y favorece la participación; debilita el temor de ser

<i>students' perspective.</i>			secundaria en Hong Kong, cursos reducidos y numerosos.	ansiedad en el aprendizaje de una segunda lengua en Asia.	conducción de entrevistas; docentes y alumnos.	mal evaluado por pares y docentes. CSR es efectivo en disminuir la ansiedad de aprender otro idioma.
-------------------------------	--	--	--	---	--	--

5.2.2. Latinoamérica

El principal actor que ha contribuido a estudiar este tema en la región ha sido la UNESCO, que consistente con su misión por mejorar la educación en todo el mundo, con la convicción de que es la clave del desarrollo económico y social, influye en la generación de mejoras y promueve cambios en la Educación. Por su parte Bolivia y Colombia han sido escenario de estudios a nivel país en la región. Chile cuenta con dos tesis de magíster en este tema.

Un reciente estudio de la Unesco sobre el impacto del tamaño del curso en logros de aprendizaje, analiza la situación actual de la región y los desafíos que enfrenta. En base a los resultados de los Estudios Regionales Comparativos y Explicativos aplicados⁷⁷ que evalúan la calidad de la educación, concluye que esta medida puede beneficiar a los países del continente. Señala que a nivel latinoamericano se observa una transición demográfica caracterizada por una disminución de la tasa de natalidad en la región y una disminución relativa de la cantidad de alumnos en edad escolar. Se presenta a su juicio una oportunidad educativa, en cuanto disminuye la razón de dependencia educativa⁷⁸, lo que hace aumentar los recursos disponibles para los actuales escolares (Bilagher & Holz, 2012). A nivel general en la región, se sugiere inicialmente mejorar los indicadores de cobertura educativa y luego aumentar los recursos por estudiante, donde una disminución del número de alumnos por curso adquiere un rol protagónico como política pública en pos de mejoras de logros de aprendizaje. Los datos de PERCE sugerían que una disminución de la cantidad de estudiantes por curso mejoraba débilmente los puntajes de las pruebas, presentando un efecto no lineal; el mayor beneficio se obtendría alrededor de 20 estudiantes por curso. Hallazgos preliminares utilizando los datos de SERCE indican, según los autores, que mayores logros educativos se relacionan negativamente con el número de alumnos por curso, esto es, una menor cantidad de estudiantes por curso resulta en mayores puntajes, en promedio. Se concluye que si bien los efectos en resultados de aprendizaje son débiles, esta intervención podría ser la más recomendable para América Latina si se implementa de manera correcta, ya que además de los logros de aprendizaje existen otros factores que contribuyen al mejor desempeño de los estudiantes que se ven impactados por este tipo de medida. Sería esperable que los datos de TERCE pudieran complementar estos hallazgos, señalan.

Por su parte Urquiola, preocupado de utilizar estrategias de identificación adecuadas para la realidad de la región⁷⁹ y en particular a la situación del sistema educativo de **Bolivia**, aprovecha las condiciones de ruralidad de este país para estudiar el efecto del tamaño de los cursos en el desempeño de los estudiantes de 3° de Educación Básica. Considerando la significativa correlación entre cantidad de alumnos por curso y matrícula, y a su vez al nivel socioeconómico en Bolivia,

⁷⁷ PERCE (1997), SERCE (2006) y TERCE (2013): Primer, Segundo y Tercer Estudio Regional Comparativo y Explicativo, UNESCO.

⁷⁸ Población en edad escolar dividido por población en edad laboral.

⁷⁹ Advierte que el modelo de regresión discontinua no es aplicable a sistemas educativos como el chileno.

descarta el análisis en sectores urbanos, con el propósito de evitar el sesgo de clasificación por capital cultural y social respecto de logros de aprendizaje. Por otra parte, la proporción profesor/alumno que se establece mediante reglas institucionales en sectores rurales de Bolivia, permite la implementación de dos diseños de investigación que aíslan los efectos de la variable de interés. El primero utiliza un patrón de asignación de profesores como variable instrumental, que permite a los directores solicitar un profesor adicional cuando sus cursos superan los 30 alumnos⁸⁰, lo que genera una discontinuidad en la relación de la matrícula respecto del tamaño de los cursos. El segundo se basa en comparaciones entre escuelas remotas con menos de 30 alumnos por curso, estos es, con un curso por nivel. En este contexto, no es posible asignar estudiantes más débiles a cursos más pequeños, lo que elimina el problema de clasificación. En estos casos de pequeños pueblos alejados, los hogares enfrentan una única opción de escuela, igualándose la matrícula por nivel a la cantidad de alumnos por curso. Ambos diseños arrojan una relación negativa entre cantidad de alumnos por curso y logros de aprendizaje; una reducción del tamaño del curso de una σ (8 alumnos aprox.) mejora los resultados en aproximadamente 0.3σ , lo que es equivalente a mover un estudiante del percentil 25 a la media de la distribución de los puntajes en lenguaje (Urquiola, 2006).

Colombia ha llevado a cabo una serie de reformas educativas, cuya implementación ha generado las condiciones para estudiar el impacto del tamaño del curso sobre los logros de aprendizaje. La reforma educativa implementada en el año 2003, buscaba promover la cobertura educativa del sector público, aumentando la matrícula y a su vez el tamaño de los cursos. Si bien el estudio revela un deterioro de la calidad de la educación pública respecto de la privada –cuyos cursos se mantuvieron estables–, sus autores señalan que no existiría evidencia de que la variable explicativa fuera el aumento del número de alumnos por curso (Medina & Bríñez, 2007). Sin embargo recientemente otro estudio con datos para Colombia, metodológicamente más robusto, que utiliza los resultados de TIMSS 2007, concluye que el número de alumnos por curso es una variable altamente explicativa de los resultados de aprendizaje (Breton, 2012). En promedio los alumnos colombianos de 4° de Educación Básica obtuvieron 355 puntos, donde la media de la distribución de puntajes corresponde a 500 y la desviación estándar a 100. Aumentos de 19 a 60 alumnos por curso⁸¹ significaron una disminución de 80 puntos en el resultado del test. El estudio controla por características de los alumnos, escuelas y endogeneidad.

En el caso de **Chile** las indagaciones respecto de la medida se han dado solo a nivel de tesis de magíster con dos de ellas publicadas en la revista Docencia. En su ejemplar de Agosto del año 2000 aparecen los resultados de la Tesis de Magister en Educación de Nelson Gutiérrez de la Universidad de Concepción titulada *“El número de alumnos por curso y su relación con la calidad de los procesos pedagógicos”* donde concluye que el peso de este factor es de un 15% de los resultados de Simce de 4° de Educación Básica⁸², y mediante un análisis de experiencias de aula identifica los fenómenos que se ven impactados por la variable estudiada (Gutiérrez, Castro, &

⁸⁰ En la práctica, el profesor adicional es asignado al alcanzar en promedio 38 alumnos por curso.

⁸¹ Media de 36.7 estudiantes por curso.

⁸² Muestra $n=4$, establecimientos de Concepción.

Segure, 2000). Cuatro años más tarde, la misma revista publica los resultados de la Tesis de Magíster en Educación de Jefferson Myers de la Universidad de Santiago de Chile, titulada “*La relación entre el número de alumnos por curso y el rendimiento en la Prueba de Matemáticas del SIMCE 2000*” donde concluye que existe un significativo impacto de reducciones de alumnos por curso cuantificadas en aproximadamente 10 puntos SIMCE en 4° de Educación Básica para el subsector de matemática, producto de diferencias de 11 o más alumnos (Myers, 2004).

Resumen de los estudios descritos en la sección 5.2.2. Latinoamérica						
Título	Autor/Año /País	Publicación	Alcance	Especificidad	Enfoque	Hallazgo Clave
<i>Class Size: Can Demographic Transition Help Quality Of Education In Latin America And The Caribbean? Evidence From Serce.</i>	Bilagher, M., & Holz, M. 2012, Unesco	<i>UNESCO, Latin America and the Caribbean</i>	Regional: datos de Serce, Estudio Regional Comparativo y Explicativo	General: evaluación de la situación actual y desafíos que enfrenta la región en educación.	Mixto: análisis de datos y recomendaciones de política pública.	CSR sería una buena estrategia para mejorar logros de aprendizaje en países que ya han logrado buenos indicadores de cobertura educativa.
<i>Identifying Class Size Effects in Developing Countries Evidence from Rural Bolivia.</i>	Urquiola, M. 2006, Bolivia	<i>Review of Economics and Statistics</i>	Local: datos de escuelas rurales de un curso por nivel en Bolivia, nivel 3°, año 1996.	Específico: estimar efecto de CSR en logros de aprendizaje libre de sesgo de selección y nivel socioeconómico.	Cuantitativo: utiliza modelo de RD y ruralidad donde matrícula por nivel es igual al tamaño del curso.	CSR de una σ (8 alumnos aprox.) mejora los resultados en 0.3 σ , equivalente a pasar del percentil 25 al 50 de la distribución en lenguaje.
<i>The Effects of Class Size on Quality of Education in Colombia.</i>	Medina, C., & Bríñez, O. 2007, Colombia	<i>The Quality of Education in Latin America and the Caribbean.</i>	Nacional: datos de tamaño del curso y resultados de escuelas públicas, 2003.	Específico: estima impacto de aumentar el tamaño de los cursos en logros de aprendizaje.	Cuantitativo: análisis del tamaño de cursos, puntajes de test, antes y después de la fusión.	Se observa un deterioro de los resultados en escuelas públicas que aumentaron el tamaño de sus cursos; los autores descartan CS como causa.
<i>Evidence that Class Size Matters in 4th Grade Mathematics: An Analysis of TIMSS 2007 Data for Colombia</i>	Breton, T. R. 2012, Colombia	<i>International Journal of Educational Development</i>	Nacional: datos de TIMSS para Colombia, nivel 4°, 2007.	Específico: efecto de CSR en Colombia en resultados de habilidades matemáticas en 4°.	Cuantitativa: correlación entre tamaño del curso y puntaje TIMSS.	Cursos numerosos tienen efectos adversos en los resultados de aprendizaje en Colombia.
<i>El número de alumnos por curso y su relación con la calidad de los procesos pedagógicos.</i>	Gutiérrez, N., Castro, A., & Segure, T. 2000, Chile	<i>Docencia</i>	Local: utiliza puntajes Simce 4°, 1996 de 4 escuelas de Concepción.	General: impacto de CSR en desempeño SIMCE de 4° en Chile y su vinculación con procesos de aula.	Mixto: verificación de procesos de aula, análisis de puntajes Simce.	CSR (menos de 26 alumnos) tiene un impacto del 15% en los resultados de Simce. Mayor frecuencia de conductas que favorecen aprendizaje en aula reducida.
<i>La relación entre el número de alumnos por curso y el rendimiento en la prueba de Matemáticas del SIMCE 2000.</i>	Myers, J. 2004, Chile	<i>Docencia</i>	Nacional: datos del Simce de matemática del año 2000, nivel 4°. 46 escuelas críticas urbanas.	Específico: impacto de CSR en resultados de prueba estandarizada, subsector matemática.	Cuantitativo: regresión de puntajes de test por grupos de tamaño de curso (20-29, 30-39, 40-45)	Diferencias de 10 puntos SIMCE en 4° de Educación Básica para el subsector de matemática, producto de diferencias de 11 o más alumnos

VI. PROCESOS DE AULA

Hasta aquí, los estudios e investigaciones analizados abordan el tema desde una perspectiva principalmente cuantitativa, buscando establecer un vínculo causal directo entre la cantidad de alumnos por curso y los logros de aprendizaje medidos como puntajes de test. Las dificultades han sido esencialmente metodológicas, donde los métodos y técnicas estadísticas junto a las estrategias de identificación utilizadas han tomando un rol central.

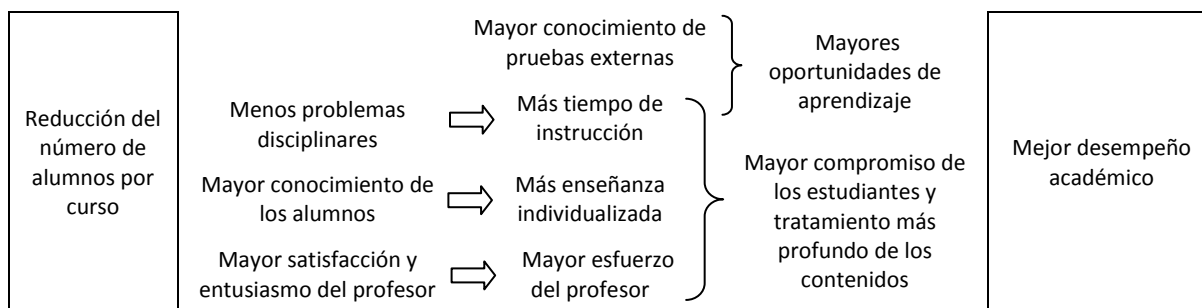
Más recientemente se ha desarrollado una línea de investigación alternativa que ha intentado indagar en los mecanismos que se desencadenan al interior de aulas reducidas enfocándose en el estudio de los vínculos entre el tamaño del curso y los efectos tanto en alumnos y docentes como factores determinantes de resultados de aprendizaje. Se busca evaluar, cómo el tamaño del curso interactúa con otros elementos del contexto local para facilitar la enseñanza y el aprendizaje, y cómo afecta resultados de largo plazo explicados por habilidades cognitivas y no cognitivas facilitadas por experiencias educativas en contextos de aulas reducidas.

Bajo esta perspectiva, se asume que la relación entre las variables es indirecta y que está mediada por procesos de aula intervinientes que brindan oportunidades para mejorar los aprendizajes de los alumnos. En consecuencia, el propósito de las investigaciones más recientes y de las evaluaciones de implementación de los programas de reducción de alumnos por curso, es identificar aquellos mecanismos que operan en cursos reducidos para distinguir las variables que explican los beneficios de la intervención y potenciar aquellas circunstancias bajo las cuales la reducción de alumnos por curso surte el mayor efecto. En esta línea encontramos estudios que buscan comprobar que *el docente* -en este nuevo contexto- es capaz de generar un mayor impacto sobre los aprendizajes, mientras otros argumentan que es el *cambio en la conducta y actitud de los alumnos*, lo que explicaría los mejores desempeños, donde el compromiso hacia el aprendizaje de los estudiantes se torna esencial para el éxito escolar. Aproximaciones alternativas buscan dar cuenta del impacto de una multiplicidad de factores intervinientes en forma conjunta e interrelacionada, mientras otras han intentado caracterizar estrategias efectivas para cursos de distintos tamaños.

Lorin Anderson fue el primer investigador que conceptualizó un modelo de impacto indirecto de la variable de interés mediado por procesos de aula protagonizados por profesores y alumnos (Anderson, 2000). La hipótesis sugería que menos problemas de disciplina, mayor conocimiento de los alumnos por parte del docente, y mayor satisfacción y entusiasmo del profesor fomentan una enseñanza mas individualizada lo que a su vez se traduce en mayor cobertura de contenidos, generando un mejor desempeño académico. La evidencia empírica a la fecha validaba una menor cantidad de interrupciones, sustentaba también el supuesto de un mayor conocimiento de habilidades y dificultades de los alumnos, sin embargo existía poca evidencia de la individualización de la instrucción y respecto de una mayor cobertura, los resultados eran mixtos. Incorporando esta evidencia reestructuró el modelo agregando variables, redefiniendo

algunas ya existentes y cambiando las relaciones e impactos de los procesos de aula que generan posibles beneficios. El diagrama 1 resume el modelo reformulado:

Diagrama 1: Modelo reformulado de Anderson. Fuente: (Anderson, 2000), elaboración propia.



De acuerdo al autor, luego de reexaminar la evidencia en el contexto del modelo reformulado⁸³, surgen tres temas centrales para el análisis:

1. De la observación de clases ha sido bien documentada la ausencia de diferencias significativas en las prácticas pedagógicas docentes en cursos numerosos y reducidos. Para aprovechar el potencial de cursos reducidos, los profesores deben cambiar la sustancia más que la forma de su quehacer docente, es decir deben aprovechar las oportunidades para enseñar mejor, no necesariamente de forma distinta.
2. El mayor compromiso de los estudiantes tampoco es relevante en una dimensión superficial, sino que debe abarcar elementos cualitativos críticos, en términos de actividad e involucramiento emocional.
3. Por otra parte, independiente del tamaño de la clase, los profesores recurren a distintas estrategias de enseñanza: curso completo, individualizada y en pequeños grupos. La oportunidad que brindan los cursos reducidos es que el balance entre cobertura y tratamiento profundo de contenidos se puede lograr de mejor manera al combinar eficazmente las estrategias antes mencionadas.

Por último, el autor sugiere que el incorporar la variable *mayor esfuerzo del profesor* al modelo, centraliza la búsqueda en entender cómo pueden los profesores enseñar mejor, sin necesariamente hacerlo de forma distinta. Concluye que los cursos reducidos ofrecen *oportunidades* a los profesores para enseñar mejor, pero que hacer clases a un curso reducido *no causa* una mejor enseñanza.

Los autores Ehrenberg et al., (2001) al igual que Anderson abordan los temas relativos al número de alumnos por curso en forma global y plantean una serie de interrogantes relativas al desempeño, a la existencia de umbrales, a su impacto en distintos niveles y asignaturas, contextos y estudiantes, a su potencial efecto acumulativo, su perdurabilidad en el largo plazo, a los procesos mediadores intervinientes, a su relación costo beneficio y a su importancia respecto de

⁸³ Señala que la discusión conceptual está mucho más atrasada que la investigación educativa en este tema.

otros factores⁸⁴. El tamaño del curso importaría en cuanto se pueden lograr mayores aprendizajes por una *mayor efectividad de las prácticas pedagógicas* usualmente utilizadas, o por cambios en las estrategias de enseñanza que utilizan los docentes. Señalan que para mantener el orden los profesores en cursos numerosos utilizan distintas aproximaciones pedagógicas que permiten monitorear el comportamiento y mantener a los alumnos ocupados: enseñanza directa, dictado o trabajo en el puesto. Si bien los alumnos prestan más atención, el desempeño no se ve afectado mayormente. Sin embargo los docentes que utilizan otras estrategias como *resolución de problemas, discusión, trabajo en grupo y escritura extensa* ven los beneficios de clases reducidas por cuanto este tipo de prácticas son más efectivas en cursos pequeños, sobre todo en los niveles iniciales donde estas estrategias toman la forma de trabajo en pequeños grupos y actividades prácticas sustentadas en una relación personal con el docente. Asimismo, los docentes de cursos reducidos dedican *más tiempo a alumnos con mayores dificultades* y los que se quedan atrás, lo que puede explicar el escaso impacto general de la intervención y el mayor beneficio para alumnos desaventajados.

Frecuentemente las investigaciones y estudios han encontrado que reducciones de alumnos por curso benefician en mayor medida a alumnos desaventajados, lo que se explicaría - como señala Ehrenberg- por una mayor disponibilidad de tiempo del docente para apoyar a estos alumnos de bajo desempeño. Sin embargo nuevas voces señalan que la razón radica en cambios de *motivación, actitudes y comportamiento* de los alumnos. El estudio de Babcock & Betts (2009) aporta al debate en cuanto entrega luces de un mayor impacto relativo de la capacidad del docente para motivar e impactar las actitudes y comportamiento de los estudiantes al implementarse reducciones de alumnos por curso. A partir del análisis de los datos académicos y de fuentes exógenas de variación del tamaño de los cursos en el distrito de San Diego en USA, concluyen que los cursos reducidos permiten a los profesores *motivar a alumnos desvinculados* mas efectivamente o que alumnos desmotivados pueden vincularse mejor al quehacer escolar en estos contextos, lo que explicaría en mayor medida el beneficio que perciben alumnos desaventajados de programas de esta naturaleza. Estos resultados relevan la importancia de la incorporación de aspectos *no cognitivos* en el estudio del impacto de intervenciones de reducción de alumnos por curso en tanto que las actitudes y conducta de los alumnos pueden moldear la manera en que los recursos educativos se transforman en capital humano. Reafirmando esta idea aparece el recientemente publicado informe TALIS 2014 que señala que en países con una distribución de profesores sub óptima, y en pos de asignar y aprovechar más eficiente los recursos, políticas como reducciones de alumnos por curso para docentes que se desempeñan en contextos difíciles son aconsejables (OCDE, 2014).

Adicionalmente a las investigaciones que estudian las relaciones entre los procesos educativos y la medida en cuestión, la experiencia que se recoge de la implementación de programas de reducción de alumnos por curso constituye un aporte fundamental en la comprensión de los procesos de aula que se ven impactados por esta intervención.

⁸⁴ Además describe al menos 5 problemas al estimar sus efectos: sesgos de clasificación, factores de distorsión, baja variación de la variable independiente, interacciones entre niveles y latencia.

El programa que redujo a 20 alumnos por curso la Educación Inicial Básica en la provincia de Ontario, **Canadá**, buscaba que cada estudiante, independiente de sus circunstancias sociales, tuviera acceso a una educación de calidad que le permita ser exitoso en el mundo actual (Bascia, 2010). Mejorar los logros de aprendizajes de todos, a la vez que disminuir la brecha entre los más desaventajados y aquellos más avanzados, constituía la principal meta. En términos de los procesos de aula intervinientes, los profesores participantes declararon que el grado de *apoyo individual* había aumentado significativamente con la reducción de alumnos por curso, lo que les permitía circular por la sala para ir ayudando a alumnos individualmente de acuerdo a sus necesidades y agrupar a los alumnos en múltiples configuraciones. En comunicación oral, donde los alumnos narran experiencias personales, las habilidades de *respetar turnos y cortesía*, se vieron reforzadas por la menor cantidad de estudiantes. Algunos docentes señalaron su interés por clases centradas en el estudiante, sin embargo no las implementaban ya que temían perder el control del curso sin la estructura de una aproximación centrada en el profesor, lo que se observó más frecuente en escuelas y cursos con alumnos vulnerables, de grupos minoritarios e inmigrantes. Cursos reducidos en general permitieron el establecimiento de *ambientes de trabajo* donde los alumnos están *relajados y son desafiados intelectualmente*, descubriendo y creando soluciones por sí mismos, en vez de trabajando para encontrar una solución predeterminada por el docente. Los profesores manifestaron tener un *mejor conocimiento de los alumnos*, lo que facilitaba la evaluación del aprendizaje y la adaptación de la instrucción a necesidades particulares de cada niño. Los alumnos pudieron desarrollar proyectos abiertos (*open-ended*) de su invención, de forma colaborativa con compañeros y profesores. Los alumnos con necesidades educativas especiales recibieron más *instrucción individualizada* producto de la reducción de alumnos por curso⁸⁵. Los docentes también declararon la *ocurrencia de menos interrupciones*, un mejoramiento de las relaciones entre compañeros y un *mayor compromiso* de los alumnos hacia la escuela. Producto de la implementación del programa, pudieron aprovechar el mayor espacio libre de la sala, utilizándolo flexiblemente para distintas actividades. También recibieron *más apoyo de profesionales especialistas* en aula. La mayoría de los profesores manifestaron sentirse más exitosos en el ejercicio de la profesión docente, al poder responder más efectivamente a las necesidades de sus estudiantes, lo que se refleja en un mayor bienestar emocional, físico y social de los alumnos. Los apoderados también manifestaron un mayor compromiso parental respecto de sus hijos y declararon una mayor percepción de efectividad en su rol de co-educadores con la escuela, producto de la medida.

6.1. El Docente en el Aula

La investigación en torno al comportamiento del docente en contextos de cursos reducidos es extensa pero poco concluyente. Ha sido complejo verificar el anhelado cambio en las prácticas pedagógicas de los profesores y más difícil aún vincularlas a mejoras en los aprendizajes. Una posible explicación es que los *instrumentos* con que se registran los cambios de comportamiento *no sean los suficientemente sensibles* para captar las variaciones que si ocurren.

⁸⁵ Señala el reporte que en contextos acomodados los alumnos con NEE representan 1/5 del curso, pero en contextos vulnerables esta proporción es significativamente más alta, constituyendo una dificultad adicional para entregar enseñanza individualizada de acuerdo a las necesidades de los alumnos.

Esta hipótesis surge a partir de los testimonios de profesores participantes en STAR, tanto de cursos reducidos como de regulares –con apoyo– que manifestaron haber contado con más tiempo de instrucción, haber enseñado de forma individualizada más frecuentemente y haber sufrido menos interrupciones. Sin embargo, como se describió anteriormente, sólo los estudiantes de cursos reducidos obtuvieron mayores logros de aprendizaje. Esto podría significar que los instrumentos de medición no son los adecuados o que estos beneficios reportados no son la clave exclusiva del éxito de este tipo de intervenciones y que relaciones más complejas, difíciles de registrar, son las que en su conjunto favorecen mejores resultados en contextos de cursos reducidos.

Como queda de manifiesto, en esta línea de investigación varios estudios utilizan los testimonios e impresiones de profesores como fuente de información, lo que de alguna manera los ha deslegitimado. En tanto que la medida impacta las condiciones laborales de los profesores, los docentes tienden a atribuirle importantes beneficios a la intervención, los que sin embargo no siempre se reflejan en la observación de clases.

Hay quienes señalan que sólo profesores efectivos logran aprovechar estos nuevos contextos y transformarlos en mejores oportunidades de aprendizaje, por lo que el tema se ha enfocado en vincular la calidad docente con el tamaño del curso. En esa línea, algunos académicos proponen, en vez de reducir la cantidad de alumnos, perfeccionar la formación docente y aumentar la calidad del profesorado, por ser una medida más costo-efectiva. Otros postulan que, independiente de las estrategias pedagógicas utilizadas, el *facilitar el acceso al profesor*, por el sólo hecho de que haya menos alumnos en el curso, pueda ser la *causa de mayores logros de aprendizaje* (Toppo, 2008 en Flessa, 2012, p. 15).

Más allá de estas posturas más bien extremas, corresponde analizar este nuevo contexto que enfrenta el docente, donde tiene a su disposición un abanico mayor de estrategias pedagógicas efectivas y mayor tiempo para implementarlas. Así, las principales hipótesis que se esgrimen respecto de los mecanismos que favorecen la obtención de mejores resultados vía el impacto del docente en cursos pequeños son:

- a) Mayor tiempo de instrucción (menos interrupción y trabajo administrativo)
- b) Más enseñanza individualizada
- c) Prácticas pedagógicas más efectivas
- d) Más retroalimentación oportuna y sostenida
- e) Mayor profundización y cobertura de contenidos
- f) Motivación e involucramiento de alumnos en actividades académicas

El trabajo de Ehrenberg et al., (2001) contrasta estas prácticas declaradas con la observación de clase y concluye que la evidencia *no arroja cambios sustanciales en las prácticas de enseñanza*. Plantea que esta inercia es esperable, ya que la enseñanza es concebida como una práctica institucionalizada sobre la cual los profesores tienen ideas bastante fijas y en consecuencia, la mantienen relativamente invariable. Plantea que esta resistencia al cambio se nutre en el trabajo individual de profesores, en las escasas instancias para compartir prácticas,

infrecuentes visitas y observaciones de clase, insuficiente disponibilidad de prácticas docentes innovadoras para profesores en formación, un currículum obligatorio y una multiplicidad de pruebas estandarizadas.

Más específico es el trabajo de Betts & Shkolnik (1999) que tiene como propósito comprender cómo los profesores del subsector de matemática cambian –o no– sus métodos de enseñanza al enfrentar cursos reducidos, en relación al grado de individualización de la enseñanza, al uso del tiempo y al porcentaje del libro de texto abordado. El estudio analiza una muestra de 2.170 profesores de matemática desde 7° de Educación Básica a IV° de Educación Media de escuelas públicas en EEUU. La aproximación teórica supone que el profesor desea maximizar los logros de aprendizaje promedio de sus estudiantes, sin embargo la evidencia empírica también señala que los profesores brindan más apoyo a los alumnos más desaventajados en cursos reducidos. Los resultados de las estimaciones como función del tamaño del curso muestran escasa diferencia respecto del grado de cobertura del libro de texto. Mucho más significativo que el tamaño resulta ser el *nivel de habilidad o composición del curso* para explicar este resultado. Respecto de la *individualización de la enseñanza*, la evidencia es estadísticamente significativa, al comprobar un giro desde la enseñanza grupal a la individual frente a reducciones de estudiantes por curso. Se predice que un aumento de 20 a 30 estudiantes, extiende el tiempo de enseñanza grupal en 4.8 minutos mientras que reduce la enseñanza individual en 7.8 minutos por estudiante. Este cambio es más pronunciado en docentes que enseñan cursos de baja habilidad. Por último, un aumento de 10 estudiantes reduce las actividades de enseñanza en un 2%, aumentando las no instruccionales (administrativas y control de disciplina). Esta investigación revela que reducciones en el tamaño de los cursos induce a muy pequeños cambios en el uso del tiempo por parte de los docentes respecto de actividades como revisión de nuevos contenidos, corrección, disciplina, tareas de rutina y evaluación. Un patrón interesante es que al enfrentar cursos pequeños los profesores en vez de revisar mas contenidos, *aumentan la corrección y retroalimentación* de actividades específicas.

Desde una aproximación cualitativa, las investigadoras Graue, Rauscher, & Sherfinski (2009) analizaron los datos de 3 años de trabajo de campo de 27 cursos de niveles iniciales en 9 colegios que implementaron reducciones de alumnos por curso en el marco de SAGE. Basados en observaciones etnográficas, entrevistas, sistema de evaluación de procesos de aula, resultados de pruebas estandarizadas, las autoras -mediante exámenes rigurosos de los procesos de aula que revelan los mecanismos que producen patrones de calidad- concluyen que este tipo de programa provee oportunidades que pueden ser activadas al organizar e implementar prácticas de aula de alta calidad. Los mayores avances se observaron en cursos donde la organización del espacio y el apoyo instruccional y emocional a los alumnos permitió desarrollar *estrategias de instrucción a grupos pequeños e individualizada* -respondiendo a necesidades de instrucción particulares y no curriculares-, generando mayores oportunidades de interacción con el profesor y entre los alumnos. En contraste, cursos donde el profesor no practica una enseñanza diferenciada, retroalimenta escasamente e interrumpe las interacciones por temas administrativos, pierden la oportunidad que brinda la reducción de alumnos por curso. Otras prácticas exitosas que se

observaron en cursos con configuraciones de co-enseñanza son la vinculación del contenido a la experiencia de los estudiantes, constante retroalimentación y conocimiento de cada uno de los estudiantes.

Basado en el estudio de aprendizaje del idioma inglés en aulas de Hong Kong anteriormente caracterizado, Harfitt (2013) indaga en las prácticas de aula de docentes de secundaria que enseñan tanto a cursos reducidos como a cursos numerosos, lo que permite controlar por calidad del profesor. A partir de datos de tipo cualitativos el estudio encuentra que los profesores si modifican sus prácticas con importantes diferencias en los *patrones de interacción profesor-alumno*, la *organización de la clase*, establecimiento de *reglas de curso* y el *uso del humor*. Se observó más trabajo en grupos y en parejas, más preguntas abiertas y desafiantes con mayor tiempo de retroalimentación. Sin embargo los profesores no siempre fueron capaces de teorizar o argumentar sus decisiones pedagógicas en este contexto. La falta de reflexión e intencionalidad en el desarrollo de estrategias que implican reformas como la aquí analizada, evidencian la ausencia de un marco conceptual que sustente dichas intervenciones y que guíe los procesos de preparación e implementación⁸⁶.

La percepción de *autoeficacia* de los profesores es otro factor relevante que incide en la eficacia docente y por consiguiente en los logros de aprendizaje de los estudiantes. El estudio de caso cualitativo conducido en una escuela suburbana de un estado del sureste de EEUU busca indagar en las actitudes de autoeficacia de cuatro profesoras de nivel inicial en cursos de distinto tamaño. Dos de ellas enseñan cursos reducidos (19-20 alumnos) en Kindergarten y dos en cursos numerosos (24 alumnos), en 1º de Educación Básica (Sullivan, 2012). La autora encuentra ciertas tendencias en las actitudes de las docentes relativas a su quehacer y el desarrollo de la comunidad educativa⁸⁷. Las dos profesoras de cursos reducidos mostraban sistemáticamente altos niveles de autoeficacia, y de satisfacción con su trabajo y logros alcanzados, a diferencia de las dos profesoras de cursos numerosos que se mostraban frustradas por problemas disciplinares y bajos niveles de logros de sus estudiantes. Talleres de manejo disciplinar, incorporación de profesores asistentes, estrategias para implementar el currículo y calificar pueden ser útiles para apoyar el aprendizaje de todos los alumnos. Involucrar algún apoderado voluntariamente es otra alternativa útil que puede ayudar a cumplir las demandas de cursos de nivel inicial numeroso.

Resumen de los estudios descritos en la sección 6.1. El Docente en el Aula						
Título	Autor/Año /País	Publicación	Alcance	Especificidad	Enfoque	Hallazgo Clave
<i>Why should reduced class size lead to increased student achievement?</i>	Anderson, L. 2002, EEUU	<i>Temple University</i>	Regional: usa evidencia empírica de la región para conceptualizar el modelo. Impacto:	Específico: vincula hipótesis con evidencia empírica que sustente las relaciones	Cualitativo: modelo conceptual que permite entender cómo CSR contribuye a mayores logros de aprendizaje.	Variables clave: disciplina, tiempo de instrucción; conocimiento de alumnos, satisfacción del profesor, compromiso de estudiantes,

⁸⁶ Anderson (2000) alega falta de discusión y desarrollo de modelos conceptuales en torno al tema.

⁸⁷ Apreciaciones respecto de su habilidad para trabajar con alumnos individualmente o en pequeños grupos, hacer libretas de notas/evaluaciones/ reportes, manejar la disciplina, involucrar a los apoderados, reflexionar, entretener a sus alumnos, trabajar con profesores asistentes, retener docentes, controlar el éxito académico de sus estudiantes, revisar actitudes/pensamientos/sentimientos al final de la jornada, confiar en estudiantes/apoderados/equipo directivo, nivelar grupos en matemática, etc.

			internacional.	propuestas.		individualización, etc.
<i>Class size and student achievement.</i>	Ehrenberg, R. G., Brewer, D. J., Gamoran, A., & Willms, J. D. 2001, EEUU	<i>Psychological Science and the Public Interest</i>	Nacional: Revisión de estudios sobre CSR. Síntesis de hallazgos.	General: impacto en logros de aprendizaje y efectos de largo plazo. Procesos que desencadena CSR.	Mixto: explica modelos para estimar efectos. Caracteriza hallazgos por tipo de estudio. Validez interna/externa.	Evidencia empírica muestra impacto favorable de CSR pero hace falta mayor contrastación con otras medidas posiblemente más costo efectivas.
<i>Reduced-class distinctions: Effort, ability, and the education production function</i>	Babcock, P., & Betts, J. 2009, EEUU	<i>Journal of Urban Economics</i>	Local: datos del distrito escolar de San Diego, año 1998-99 a 2001-02. Transición de 3° a 4°, curso reducido a regular	Específico: distingue esfuerzo de habilidad y correlaciona ambas con CSR.	Cuantitativo: parametrización de variable esfuerzo y habilidad. Correlación con CS.	Impacto de CSR sobre alumnos desmotivados o desvinculados es mayor que el impacto en estudiantes de bajo desempeño.
<i>Ontario's Primary Class Size Initiative: Report on Early Implementation.</i>	Bascia, N. 2010, Canadá	<i>Canadian Education Association</i>	Nacional: datos del programa de CSR a 20 alumnos en Ontario, Canadá, 2008-2009.	General: evaluación general del programa (beneficios y consecuencias)	Mixto: análisis de datos y entrevistas.	Exitoso programa, pero debe mejorar en desarrollo docente, instrucción combinada, agrupar estudiantes y asignar docentes; equilibrio entre ámbito local y regional.
<i>The Synergy of Class Size Reduction and Classroom Quality.</i>	Graue, E., Rauscher, E., & Sherfinski, M. 2009, EEUU	<i>Elementary School Journal</i>	Local: análisis de la experiencia del programa SAGE en Wisconsin.	General: como interactúa CSR con prácticas de calidad para favorecer el aprendizaje.	Mixto: observaciones etnográficas y entrevistas, documentos y análisis de puntajes de test.	CSR ofrece oportunidades que combinadas con prácticas docentes de alta calidad favorecen el aprendizaje.
<i>Why 'small' can be better: an exploration of the relationships between class size and pedagogical practices.</i>	Harfitt, G. J. 2013, Hong Kong	<i>Research Papers in Education</i>	Local: datos sobre 4 docentes de inglés de secundaria en Hong Kong, cursos reducidos y numerosos.	Específico: examina cambios de conducta y prácticas docentes al enseñar en cursos reducidos y numerosos.	Cualitativo: observación de clase, transcripciones, conducción de entrevistas	Se advierten cambios en prácticas docentes al enseñar cursos reducidos relativos a patrones de interacción, organización, reglas de aula y uso del humor.
<i>A Case Study of the Impact of Class Size on Teacher Efficacy.</i>	Sullivan, C. C. 2012, EEUU	<i>North Carolina State University.</i>	Local: estado del sureste de EEUU, caso de 4 docentes, 2 con cursos reducidos y 2 numerosos, K-1.	Específico: examina el impacto de CSR sobre autoeficacia docente.	Cualitativo: entrevistas, observación y bitácoras.	Ambas docentes de curso reducido exhiben niveles de eficacia altos, a diferencia de las docentes en cursos numerosos.

6.1.1. Calidad Docente y Eficacia

Para que estos programas sean eficaces y las oportunidades que los cursos reducidos brindan sean bien aprovechadas, parece necesario entender cómo se vincula esta intervención y sus efectos a la *calidad docente*, es decir, es necesario indagar en la forma que interactúa la calidad de la enseñanza y la reducción del número de alumnos por curso en la generación de oportunidades de aprendizaje. Se han realizado estudios que intentan verificar si el grado de eficacia de los profesores promueve que las reducciones de alumnos por curso reporten mayores beneficios.

Coherente con dicho propósito Mueller (2013) analiza el efecto *conjunto* de las dos variables; cantidad de alumnos por curso y calidad del profesor. En este estudio alemán que trabaja con datos de STAR, el autor conceptualiza el efecto del profesor en términos de su experiencia docente, y lo mide a través del número de años de experiencia en aula. Concluye que sólo aquellos profesores con tres o más años de experiencia logran que sus estudiantes obtengan mayores logros de aprendizaje en forma significativa. Estos resultados serían consistentes con la

hipótesis de que los *profesores experimentados amplifican los efectos de la reducción del número de alumnos por curso*, vía alzas en la calidad de la instrucción. Ésta se manifiesta en una enseñanza más individualizada, mayor interacción profesor-alumno y un mejor conocimiento y abordaje de las necesidades individuales de los estudiantes. Señalan que los profesores principiantes estarían demasiado preocupados en cumplir con las tareas básicas de enseñanza y que por consiguiente no logran aprovechar las oportunidades que ofrecen cursos más pequeños. Ratifica lo anteriormente señalado el informe TALIS 2008 que señala que es más probable que docentes menos experimentados dediquen más tiempo al establecimiento del orden y la disciplina en el aula (OCDE, 2012).

Sin embargo en el contexto del estudio comparativo mundial por región de Wössmann & West (2006a) se concluye que los cursos reducidos tienen efectos beneficiosos observables en los logros de aprendizaje de sus estudiantes, sólo en aquellos países cuyo *cuerpo docente es relativamente mal remunerado y poco efectivo*, argumentando que los profesores efectivos logran aprendizajes en sus alumnos independiente del tamaño del curso. Lo reafirman Altinok & Kingdon (2012) al concluir que sólo países en vías de desarrollo -en donde la calidad del cuerpo docente es relativamente baja - muestran tamaños del efecto negativos y estadísticamente significativos.

Lo anterior podría parecer contradictorio respecto de los hallazgos anteriores, sin embargo resulta razonable analizar estos efectos en su contexto; allí donde hay más oportunidades para mejorar y avanzar en logros de aprendizaje, es donde más significativas resultarán este tipo de intervenciones, potenciadas por aquellos docentes más capaces, a diferencia de países o contextos altamente efectivos que ya han aprovechado ésta y otras estrategias para mejorar su calidad educativa y que requerirán de reformas de otra naturaleza para alcanzar un nivel superior.

Sistemas educativos de baja calidad de su cuerpo docente, con un reducido gasto en educación, con sueldos de profesores por debajo del promedio y con fuertes señales de falta de efectividad, ven favorablemente que intervenciones de esta naturaleza puedan generar mejoras sustantivas en los logros de aprendizaje. Sin embargo surge la pregunta por la efectividad de la medida respecto de otras alternativas⁸⁸ como *mejorar la calidad docente directamente*.

Dado el gran impacto económico que implican reformas educativas de reducción del número de alumnos por curso, algunos investigadores han centrado sus estudios en análisis costo-beneficio respecto de otras intervenciones que pudieran producir iguales o mayores aumentos en resultados académicos. Rivkin, Hanushek & Kain (2005) indagan en la importancia de la efectividad del docente en los efectos sobre desempeño en lenguaje y matemática. Señalan que un *aumento en la calidad docente de una desviación estándar produce mayores beneficios*⁸⁹ que una *reducción de 10 estudiantes por curso*, comparando directamente con programas ya implementados de reducciones del número de alumnos por curso en algunos estados de EEUU. Estos beneficios son más notorios en niveles iniciales. Sin embargo estas diferencias en calidad y

⁸⁸ Muchos investigadores rechazan la idea de que una estrategia pueda sustituir a la otra, y proponen su implementación en conjunto.

⁸⁹ 0.11σ en matemática y 0.095σ en lectura. No especifica el costo de esta reforma.

efectividad docente, no son capturadas por características observables (experiencia⁹⁰, certificación, resultados de pruebas, etc.) La inexperiencia de profesores nóveles respecto de docentes experimentados explica diferencias en la efectividad docente de un 10% en matemática y entre un 5% y un 20% en lectura. Por otra parte, los análisis sugieren que la tasa de rotación de profesores se correlaciona negativamente con resultados en logros de aprendizaje. Mientras menor es la rotación mejores son los resultados promedio y menor la varianza; al contrario -a mayor tasa de rotación docente- el desempeño promedio es menor y la varianza mayor. La evidencia muestra que a alumnos de contextos vulnerables y de grupos minoritarios les son asignados con más frecuencia profesores inexpertos y están afectos a mayor rotación de docentes, por lo que las políticas podrían intencionar la asignación de profesores experimentados y disminuir la rotación a la población más desaventajada para mejorar aspectos de calidad y equidad educativa. Este estudio revela la importancia de los procesos de contratación, despido, orientación y desarrollo profesional efectivos.

Por su parte Brühwiler & Blatchford (2011) encuentran que el *tamaño del curso* y la *calidad el profesor* son *factores independientemente importantes* en la consecución de mayores avances en logros de aprendizaje. Los autores se propusieron evaluar la efectividad docente en el progreso del aprendizaje, a la que llamaron *competencia de enseñanza adaptativa ATC*, derivada de la estrategia de educación adaptativa⁹¹. Señalan que su investigación constituye un aporte en cuanto captura las dinámicas de los procesos de aula, más allá de la observación, contemplando aspectos como *participación, interés y calidad docente*.

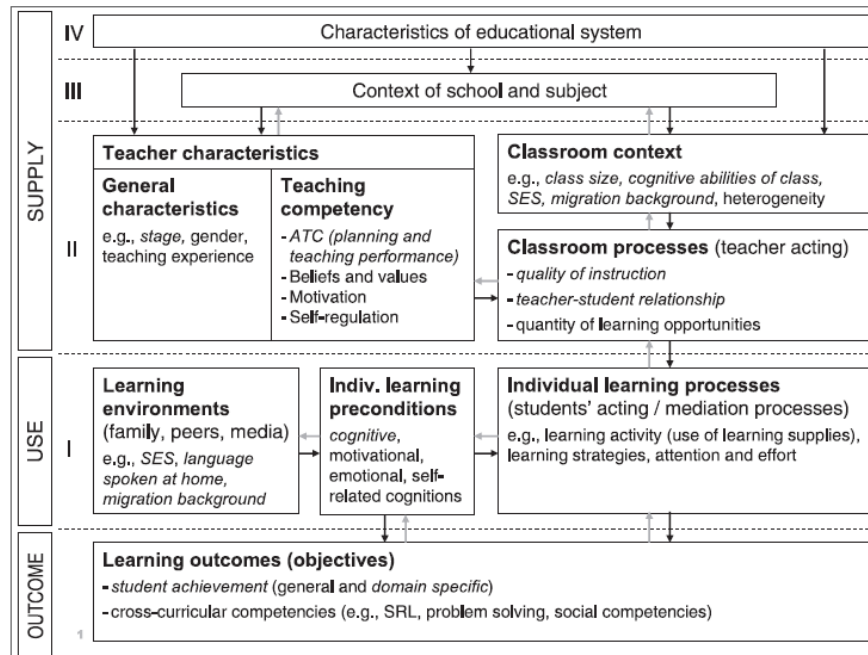
El modelo desarrollado y testeado en esta investigación no postula un efecto directo de las características del profesor ni de los procesos de aula sobre los resultados de aprendizaje. Al contrario, los efectos están mediados por los procesos de aprendizaje de los alumnos (estrategias de aprendizaje, atención, esfuerzo, etc.) y moderados por el contexto del aula (por ejemplo tamaño del curso) y condiciones de aprendizaje preexistentes (aspectos cognitivos y motivacionales). El modelo distingue 4 niveles: sistema, escuela, aula e individual, y se conceptualiza con influencias multidireccionales, como se observa en la figura 2.

Las consideraciones teóricas del modelo ATC enfatizan que la enseñanza debe satisfacer las diversas necesidades de estudiantes individuales en mayor grado que la configuración de enseñanza tradicional de grupo-curso. La enseñanza debe tomar en cuenta factores situacionales y ajustar su planificación a partir de la percepción de las condiciones de aprendizaje de sus estudiantes, comportamiento de aprendizaje y resolución de problemas, diagnosticando los procesos individuales de aprendizaje y de dificultades, como falta de comprensión o concentración. La muestra se tomó de la base de datos del proyecto de investigación *competencia de enseñanza adaptativa ATC* compuesta por 26 cursos de 4° y 5° de Educación Básica, y 23 cursos de 7° y 8° de Educación Básica de escuelas públicas en Suiza.

⁹⁰ Posterior a los dos primeros años, período de significativas mejoras en calidad y efectividad docente.

⁹¹ Definida como el uso de estrategias y recursos de enseñanza alternativos para responder a las necesidades de aprendizaje de estudiantes individuales.

Figura 2: Modelo de aprendizaje multinivel. Las variables observadas en el estudio están impresas en *italica*.



Como se esperaba, se encontró que las *clases reducidas muestran un efecto positivo en el progreso del aprendizaje de los estudiantes*. Una reducción de 5 estudiantes por curso resultó en un aumento promedio de un cuarto de desviación estándar basado en la distribución de los puntajes de los pre test, sólo para los niveles iniciales -4° y 5° de Educación Básica-, lo que confirma hallazgos de investigaciones previas donde se señala que los mayores beneficios se logran al iniciar la etapa escolar⁹². Los estudiantes percibieron una mejor calidad de sus profesores y también juzgaron la relación profesor alumno como mejor. *Se evidencia que las competencias de los profesores no se transmiten directamente a los alumnos sino que están mediadas por procesos de aula*.

Tomando en cuenta toda la gama de tamaños de cursos de los distintos sistemas educativos, algunas investigaciones han intentado caracterizar las restricciones que enfrentan los docentes tanto en aulas reducidas como numerosas respecto de las estrategias pedagógicas disponibles y su factibilidad de implementación, buscando identificar cuáles son más efectivas en uno y otro caso.

Es el tema del estudio de Pedder (2006) que propone el análisis de tres aspectos que impactan los aprendizajes vía variaciones del tamaño del curso. Esta revisión crítica es una de las más importantes investigaciones Inglesas que alegan tempranamente la ausencia de vigilancia en los procesos de aula en investigaciones precedentes, destacando que son estos mecanismos los que afectan el desempeño y logros de aprendizaje de los estudiantes vía las disminuciones de

⁹² La teoría sugiere que los niños pequeños al iniciar esta etapa, se enfrentan a muchos cambios y viven un período de confusión. Procesos como aprender a cooperar con otros, aprender a aprender, adaptarse a los hábitos de la escuela y adoptar el rol de estudiante se ven favorecidos por contextos de cursos reducidos.

alumnos por curso. Enfatiza la necesidad de investigaciones que sean capaces de identificar relaciones sistemáticas entre estrategias que son efectivas en cursos reducidos y no tanto en clases numerosas, y que a su vez tengan impactos en el desempeño de los estudiantes⁹³. Señala que también es relevante considerar *el tiempo que precisan los alumnos en desarrollar las habilidades o actitudes* que permiten aprovechar estas mejores oportunidades de aprendizaje, lo que ha estado ausente del debate. El autor presenta tres modelos teóricos que explican cómo el efecto del tamaño del curso sobre el desempeño de los alumnos está mediado por aquello que ocurre en el aula. El primero enfatiza los procesos que hacen que el docente maximice la *cantidad* de oportunidades de aprendizaje para sus estudiantes; el segundo maximiza la *calidad* de estas oportunidades de aprendizaje y el tercer modelo teoriza sobre los *efectos negativos diferenciados* relativos a la *cantidad y calidad* de las oportunidades de aprendizaje en *diferentes tipos de alumno para tamaños de curso distintos*.

- a) El primer modelo propone que a medida que aumenta el número de alumnos por curso, independiente de la calidad y habilidad del profesor, los docentes *perderán libertad para elegir de su repertorio de estrategias pedagógicas*, ya que más tiempo será requerido para actividades administrativas, de organización y de control conductual. En cursos numerosos este aspecto se aborda en la práctica mediante clasificación de alumnos de mejor conducta y desempeño a este tipo de cursos; a costa de la energía, estrés y ánimo del docente para responder a múltiples demandas simultáneamente y mediante contratos didácticos más estrictos y ambientes más controlados. En aulas reducidas puede suceder que ambientes más relajados generen un menor compromiso académico por parte de los alumnos o que conductas disruptivas de algunos alumnos ganen visibilidad y reduzcan los aprendizajes.
- b) El segundo modelo postula que el tamaño del curso puede impactar la *calidad de las oportunidades de aprendizaje* al influir en la libertad con que docentes y estudiantes optan de un conjunto de estrategias que buscan maximizar la calidad del aprendizaje. En cursos numerosos los profesores disminuyen las interacciones curriculares testeando menos frecuentemente las respuestas de los alumnos, o retroalimentando de forma más concisa, lo que se compensa con mayor tiempo recapitulando al inicio de la clase o revisando contenidos al final. Estrategias como preguntas desafiantes y aproximaciones formativas de retroalimentación son especialmente sensibles al tiempo disponible. Estructurar de forma clara y ejemplificada los nuevos conceptos y contenidos puede ayudar a evitar la redundancia y aprovechar mejor el tiempo en cursos numerosos, maximizando la calidad de las oportunidades de aprendizaje. Sin embargo esta estructura puede provocar mayor dificultad en la integración de conceptos previos con los nuevos, debilitando el compromiso con el aprendizaje de los alumnos. El tamaño del curso también afecta la calidad del aprendizaje, determinando el tipo de actividad que el profesor se siente capaz de desarrollar, sobre todo donde hay limitaciones de espacio o cuando los estudiantes presentan mala conducta. Una

⁹³ Se considera un curso reducido de entre 16-20 alumnos, y uno numeroso de entre 26-31 estudiantes.

menor frecuencia en las interacciones profesor-alumno dificulta el establecimiento y continuidad de la retroalimentación, desarrollo del conocimiento de las necesidades, intereses, contexto y avances de los alumnos, y la entrega de apoyo y orientación personalizada. Restricciones de tiempo limitan las oportunidades para desarrollar la habilidad y disposición para interactuar más largamente con el profesor, lo que conduce a una insuficiente reflexión y articulación sobre problemas y al debilitamiento del aprendizaje con otros. Sin embargo no existen beneficios automáticos de cursos reducidos; las mejores oportunidades de aprendizaje dependen de la capacidad que tenga el docente para comprometer a sus alumnos con el aprendizaje. Cabe mencionar que en cursos reducidos, mayores interacciones alumno-profesor pueden fomentar las expectativas de alumnos por una orientación individual, que debilita las instrucciones y actividades grupales.

- c) El tercer modelo postula que al aumentar el tamaño del curso los profesores destinan mayor tiempo a interacciones individuales con alumnos de bajo rendimiento, práctica que al aumentar el tamaño del curso se ve restringida. Lo anterior responde a una forma de compensación que se produce en relación a los alumnos que muestran mayores logros y respecto de los cuales los docentes desarrollan actitudes positivas, que pueden reflejarse en sus calificaciones y trato directo. Adicionalmente las clases numerosas exacerban el problema de la desafección de ciertos alumnos. Los docentes de cursos numerosos adoptan la evasión de estrategias de provocación con alumnos descontentos, lo que conlleva a menores oportunidades de aprendizaje.

Resumen de los estudios descritos en la sección 6.1.1. Calidad Docente y Eficacia						
Título	Autor/Año /País	Publicación	Alcance	Especificidad	Enfoque	Hallazgo Clave
<i>Teacher experience and the class size effect — Experimental evidence.</i>	Mueller, S. 2013, Alemania	<i>Journal of Public Economics</i>	Local: datos de STAR en función de calidad docente.	Específico: efecto de CSR mediado por calidad docente definida como experiencia.	Cuantitativa: caracterización de docente según años experiencia, dummy.	Solo docentes experimentados logran un efecto de CSR. Beneficia a todos, pero menos a deciles bajos.
<i>Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS.</i>	Wössmann, L., & West, M. 2006, Alemania	<i>European Economic Review</i>	Internacional: datos de TIMSS ciencias y matemática, 11 sistemas educativos 1994-95, nivel 7° y 8°.	Específico: impacto de CSR en resultados prueba TIMSS a los 13 años en 11 países.	Cuantitativa: usa variación de tamaño de cursos en 7° y 8° , ambas asignaturas. Utiliza efecto fijo de escuela y IV.	2 de los 11 países muestran efectos negativos significativos de CSR; países con cuerpo docente mal remunerado, menos efectivo.
<i>New evidence on class size effects: A pupil fixed effects approach</i>	Altinok, N., and, & Kingdon, G. 2012, UK	<i>Oxford Bulletin of Economics and Statistics</i>	Internacional: análisis de puntajes TIMSS 2003 para 8° nivel en 33 países.	Específico: efecto de CSR medido en puntajes de matemática y ciencias y tamaño de curso.	Cuantitativa: efecto fijo del alumno (puntajes matemática vs. ciencias) en vez de a través del tiempo.	Sólo en 1/3 de los países CSR es negativo pero pequeño. Países en vías de desarrollo, con docentes menos efectivos son más sensibles a CSR.
<i>Teachers, Schools, and Academic Achievement</i>	Rivkin, S., Hanushek, E., & Kain 2005, EEUU	<i>Econometrica</i>	Local: panel de datos de Texas EEUU, 3° y 7° básico, 3 cohortes de los años 90	Específico: comparación entre calidad del profesor (experiencia) y CSR	Cuantitativa: puntaje de test entre 3° y 7° básico, para 3 cohortes de los años 90.	Calidad profesor tiene mayor efecto sobre logros de aprendizaje que CSR.
<i>Effects of class size and adaptive</i>	Brühwiler, C., &	<i>Learning and</i>	Nacional: 26 cursos de 4° y	General: impacto en resultados de	Mixto: relación entre CS, calidad	Cursos reducidos conducen a mayores

<i>teaching competency on classroom processes and academic outcome.</i>	Blatchford, P. 2011, Suiza	<i>Instruction</i>	5°, y 29 de 7° y 8°, en 2004.	aprendizaje y de conducta producto de CSR	docente, procesos de aula y resultados de aprendizaje. Aplica cuestionarios a estudiantes.	resultados educativos, mejor conocimiento de los alumnos y procesos de aula. Calidad docente es importante independientemente.
<i>Are small classes better? Understanding relationships between class size, classroom processes and pupils' learning.</i>	Pedder, D. 2006, UK	<i>Oxford Review of Education</i>	Nacional: revisión de 30 años de literatura sobre CSR.	General: propone tres modelos conceptuales que establecen relaciones entre tamaño del curso, procesos de aula y aprendizaje.	Cualitativo: impacto de CSR en cantidad y calidad de oportunidades de aprendizaje; efectos diferenciales de CSR para distintos tamaños de curso.	Se requiere métodos más sofisticados de investigación cualitativa para entender las condiciones que favorecen el aprendizaje en cursos de distinto tamaño, en primaria y secundaria.
<i>How class size makes a difference: What the research says. The impact of class-size reduction (CSR)</i>	Achilles, C. M. 2003, EEUU	<i>Paper presented at the SERVE Research and Policy Symposium</i>	Nacional: revisión de la literatura nacional en torno a CSR.	General: posturas políticas sobre CSR, financiamiento, síntesis general de beneficios cognitivos y no cognitivos de CSR, retención docente.	Mixto: revisión del impacto de CSR en resultados cognitivos como no cognitivos.	CSR impacta resultados educativos, disciplina dentro y fuera de la escuela, compromiso, participación cívica, mejora productividad, contribuye a la sociedad.

6.1.2. Condiciones Laborales

La satisfacción laboral es un factor que incide en el desempeño del ejercicio de la profesión, a la vez que atenúa las altas tasas de rotación docente y deserción laboral, que impactan negativamente la continuidad de los procesos de aprendizaje de los alumnos afectados. El abandono de la carrera es mucho mayor en profesores nóveles que en docentes más experimentados, pero para todas las edades la deserción laboral de docentes es mayor que en otras profesiones similares. La falta de valoración social, la sobrecarga de trabajo, la estandarización de la enseñanza, la ausencia de apoyo, los problemas de disciplina, los bajos salarios y el escaso financiamiento de la educación, la baja influencia y posibilidades de desarrollo, entre otras razones lleva a los profesores a abandonar la profesión. Informes de la OCDE revelan que las “malas condiciones de trabajo suelen ser la razón que los docentes alegan para dejar la profesión. Por lo común éstas se asocian con preocupaciones acerca de una pesada carga de trabajo, la falta de recursos y apoyo y el trato con alumnos difíciles y, cada vez más, con padres de familia difíciles” (OCDE, 2009, p. 228).

Una medida que mejora la satisfacción laboral docente y que alivia la pesada carga de trabajo, es entre otras, la disminución de alumnos por curso. Por un lado, demanda menor tiempo de preparación de material y de clase en tanto se reduce la proporción de alumnos con dificultades de aprendizaje y necesidades particulares que atender; favorece climas de aula menos disruptivos haciendo más grato el ejercicio diario de la profesión y disminuye el tiempo de corrección fuera del aula que permite a profesores hacer uso de su tiempo libre, mejorando su salud y bienestar psicológico. Estas condiciones favorecen mejores resultados académicos, lo que se traduce en una mayor sensación de autoeficacia del docente, que a su vez se convierte en un factor de satisfacción laboral y retención docente (Achilles, 2003). Este círculo virtuoso hace además más atractiva la carrera docente, atrayendo a mejores estudiantes y profesionales más preparados a la docencia.

Respecto del concepto de satisfacción profesional, el estudio SCT *Small Class Teaching* de Galton & Pell (2009) realizado en Hong Kong reflejó una mayor comodidad, entusiasmo y relajo de los profesores al enseñar en cursos reducidos. Confirma esta impresión el resultado de la Encuesta *Metlife* de los Profesores Americanos de 2011, encontrándose que los docentes con baja satisfacción laboral, declaraban en mayor proporción (70% vs. 53%) que la cantidad de alumnos por curso había aumentado (Metlife, 2012). Asimismo y de acuerdo a la encuesta *Teach Plus* (2012) sólo el 5% de los docentes estaría dispuesto a aumentar el tamaño de los cursos como mecanismo para financiar mejores salarios.

La literatura ha caracterizado las difíciles condiciones de trabajo de profesores en particular. Un ejemplo es el libro *Teachers under Pressure* de Galton & Macbeath (2008) . Allí el tema del tamaño del curso es un aspecto central en las demandas docentes y surge como una de las maneras más efectivas de disminuir la carga de trabajo excesiva a la que se ven enfrentados los profesores. El comportamiento inapropiado de los alumnos es el principal problema que declaran enfrentar los maestros, lo que se comprobó mediante el trabajo de campo que los autores llevaron a cabo en colegios, verificando la relación intrínseca que existe entre comportamiento y tamaño del curso.

La capacidad de retener docentes experimentados es un aspecto esencial en la calidad de la educación que brinda un sistema educativo y la cantidad de alumnos por curso pareciera ser un factor determinante. Un estudio americano que indaga sobre el impacto de la cantidad de estudiantes por curso en la retención docente, analiza la correlación entre el tamaño del curso y la probabilidad de abandono del docente, encontrando una correlación positiva y estadísticamente significativa. Una reducción de 23 a 20 estudiantes por curso, disminuye la probabilidad de abandono del docente en un 4,2%. Sin embargo, y al igual que lo sucedido con el programa de reducción de alumnos por curso en California, las iniciativas de reducción focalizadas en NY tuvieron efectos adversos en la capacidad de contratar docentes experimentados de las escuelas con una mayor proporción de alumnos desaventajados (Pas, 2007).

Desde la perspectiva del bienestar del profesor, vinculado a calidad docente, el estudio correlacional realizado por el profesor Cornejo (2009) analiza el bienestar/malestar de docentes de enseñanza media de establecimientos de dependencia municipal, particular subvencionada y pagada del gran Santiago, para caracterizar las condiciones materiales y psicosociales de trabajo de los profesores. Utiliza un cuestionario de auto reporte y una lista de chequeo sobre condiciones materiales, que miden las siguientes variables:

1. Variables de identificación del docente y del establecimiento.
2. Indicadores de bienestar/malestar docente.
3. Condiciones materiales de trabajo.
4. Condiciones psicosociales de trabajo: las que están compuestas entre otras por *cantidad de alumnos a cargo* y existencia o no de selección de estudiantes.
5. Variables personales.

Respecto de los alumnos a cargo, en promedio los docentes enseñan en 8 cursos diferentes y tienen una jefatura. Cerca del 40% debe atender a más de 39 alumnos por sala, un

54% trabaja con cursos que fluctúan entre los 30 y 39 estudiantes y sólo un 6% enseña en aulas con menos de 30 alumnos. Al menos un tercio de los profesores presenta niveles altos en las escalas de “burnout” en relación al nivel de bienestar/malestar con la profesión docente, el que varía conjuntamente con el *nivel de significatividad* y el *nivel de demandas en el trabajo*. Por un lado la capacidad de otorgar un sentido no instrumental -propósito u objetivo moral trascendente- al trabajo beneficia el bienestar y la salud, y se malogra por el alto nivel de carga de trabajo que perciben los docentes, el que queda determinado entre otras variables por la excesiva cantidad de alumnos por curso.

Resumen de los estudios descritos en la sección 6.1.2. Condiciones Laborales						
Título	Autor/ Año/País	Publicación	Alcance	Especificidad	Enfoque	Hallazgo Clave
<i>Study on Small Class Teaching in Primary Schools in Hong Kong: Final Report.</i>	Galton, M., & Pell, T. 2009, Hong Kong	<i>Hong Kong: Education Bureau and Cambridge University.</i>	Local: datos de CSR de estudio a gran escala SCT conducido entre 2004-09 en Hong Kong, niveles 1-3.	General: desempeño y apreciaciones de CSR y desarrollo profesional combinado.	Mixto: análisis de datos, observación de clases, y entrevistas a alumnos y profesores.	Se constatan ciertas prácticas pedagógicas que se ven favorecidas en SC: individualización, desafío, trabajo en grupo y parejas, más interacción docente-alumno.
<i>The Effect of Class Size Reduction on Teacher Attrition and Recruitment: Evidence from Class Size Reduction Policies in New York State.</i>	Pas, E. 2007, EEUU	<i>Syracuse University.</i>	Local: datos de CSR en NY en niveles 1°- 3°, años 1997-98 al 2000-01.	Específico: examina el impacto de CSR en abandono docente; calidad de contrataciones y atractivo de trabajar en escuelas con CSR.	Cuantitativo: análisis de tasas de deserción laboral asociadas a CSR.	Una reducción de 23 a 20 estudiantes por curso, disminuye la probabilidad de abandono del docente en un 4,2%. Escuelas con CSR contratan menos docentes con gran experiencia.
<i>Condiciones de trabajo y bienestar/malestar docente en profesores de enseñanza media en Santiago de Chile.</i>	Cornejo, C. R. 2009, Chile	<i>Educación y Sociedad</i>	Local: respuestas de profesores de educación media del Gran Santiago	Específico: caracterizar las condiciones materiales y psicosociales de trabajo de los profesores	Cualitativo: cuestionario de auto reporte y lista de chequeo sobre condiciones materiales	Una gran proporción de docentes presenta estrés, asociado entre otras cosas a excesiva carga de trabajo donde un factor es el gran tamaño de los cursos.

6.2. Conducta en el Aula

Uno de los consensos que hay en este ámbito es la indiscutible relación que existe entre la cantidad de estudiantes de un curso y el comportamiento de éstos en el aula; a menor cantidad de estudiantes mejor comportamiento individual y viceversa. Una mejor disciplina en la sala, no sólo mejora el clima de aprendizaje sino que además permite aprovechar al máximo el tiempo de instrucción, al minimizar la cantidad de tiempo que los profesores destinan al manejo de la conducta y al control de interrupciones en el aula. Cada episodio de indisciplina mayor obliga al profesor a interrumpir la instrucción para sancionar, escribir reportes y contactar a inspectores, entre otros, lo que impacta negativamente la continuidad del proceso de enseñanza aprendizaje. Así, los cursos reducidos, mejoran el índice de tiempo dedicado a la instrucción, que -en países occidentales- fluctúa entre un 60% y 70%.

El primer intento serio por modelar el impacto del tamaño del curso en la conducta del curso e influir los resultados académicos lo desarrolló Lazear (2001). Reconociendo el aprendizaje en una configuración de aula como un bien público⁹⁴ y en el marco de una función de producción

⁹⁴ Bien público concebido como aquel que puede ser aprovechado por varias personas a la vez. En consecuencia, si un estudiante es disruptivo el aprendizaje de todos los demás también se ve afectado.

educativa, el autor provee un modelo analítico que vincula el tamaño del curso a interrupciones conductuales en el aula, mostrando cómo el tamaño de curso óptimo varía en función del grado de interrupción de los estudiantes que lo componen. Sostiene que el aprendizaje en aula presenta *efectos de congestión* que se traducen en externalidades. Si un alumno es disruptivo en clase el aprendizaje de todos los demás disminuye. El efecto par, reconocido como crucial en educación provee un sustento empírico a este postulado.

Particularmente, la función de interrupción que propone Lazear establece que el alumno puede estar en uno de dos estados posibles: un estado de *interrupción* y un estado de *no interrupción*. El aprendizaje se asume posible sólo cuando no existe interrupción (propensión a la no interrupción ρ), ya que las interrupciones inhiben el aprendizaje. El modelo supone igual propensión a la interrupción de todos los alumnos de un curso $(1 - \rho)^{95}$. Como el comportamiento se asume independiente entre los estudiantes, la probabilidad de no interrupción corresponde a la multiplicación de las probabilidades de cada uno de ellos de estar en un estado no disruptivo, es decir ρ^n , donde n corresponde al tamaño del curso; consecuentemente la proporción de tiempo que existe interrupción corresponde a $1 - \rho^n$.

Como se observa en el gráfico 4, si $\rho = 0.95$, es decir que la propensión a la interrupción es de un 5% para cada alumno, se obtiene sólo un 21% de tiempo de aprendizaje donde no existe interrupción. Se desprende que el tamaño óptimo del curso será mayor para grupos curso con una menor propensión a conductas disruptivas y menor en el caso contrario, lo que hace a la *conducta un sustituto del tamaño del curso* de acuerdo al autor. Este hallazgo ayuda a explicar las dificultades para encontrar evidencia del efecto del tamaño del curso en los datos, descrito anteriormente como problema o sesgo de clasificación (alumnos con mejor comportamiento y/o mejores calificaciones son ubicados en cursos más numerosos).

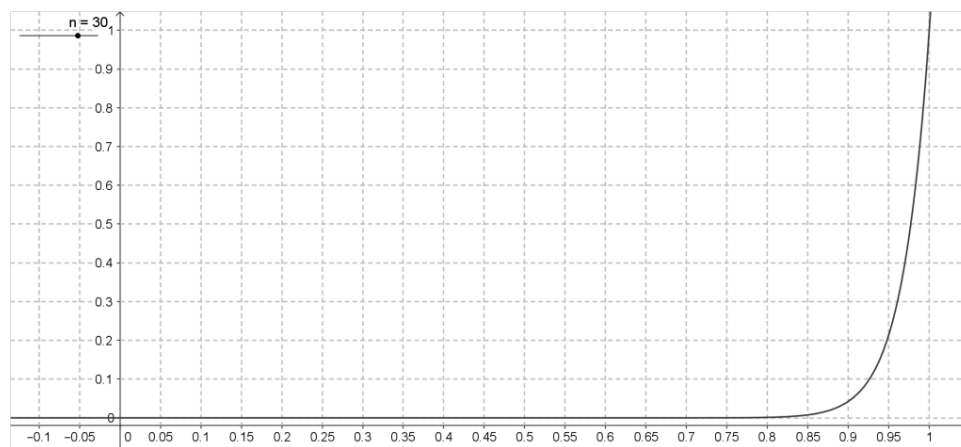


Gráfico 4: función de interrupción $f(x) = \rho^n$, donde $\rho \in [0,1] \wedge n = 30$. Proporción de tiempo para el aprendizaje en función de la propensión de interrupción de los alumnos de un curso. (*promedio Chile*)

⁹⁵ La conducta en el aula es una función tanto de las características de los alumnos como de la calidad y habilidad del profesor, lo que incorpora calidad docente a la formulación del problema.

Al igual que Lazear, los investigadores Cortes, Moussa, & Weinstein (2013) conceptualizan un modelo donde los cursos se componen de alumnos *disruptivos* y *no disruptivos*. El estudio utiliza datos de las escuelas de Chicago para analizar el mecanismo mediante el cual el tamaño del curso -vía la componente conductual de la composición del curso- afecta el desempeño de los estudiantes. Los estudiantes *disruptivos* se definen como aquellos que están ausentes de clase por alguna de las siguientes razones: *legalmente se le asignó a un correccional; perdido; excesivas inasistencias o violación del código de disciplina del uniforme*. En este trabajo el indicador de desempeño es la probabilidad de aprobar los cursos de inglés y/o álgebra en 1° de Educación Media. De acuerdo al estudio, la relación entre el promedio de alumnos *no disruptivos* y el momento del día en que se dicta una clase⁹⁶ presenta una forma de U invertida, con un máximo en la mitad del día. Los autores observan que la probabilidad de aprobar estas asignaturas aumenta al incrementarse la cantidad de alumnos *no disruptivos* de la clase en cuestión. Los tamaños de los efectos se comparan con los reportados para el proyecto STAR y para el experimento en Israel de Angrist & Lavy, siendo éstos bastante similares. Lo que difiere es la conceptualización de la variable de interés. En este estudio corresponde a la cantidad de alumnos *no disruptivos* de un curso (al aumentar mejora el clima de aula, ya que la proporción de alumnos disruptivos decae), a diferencia de la cantidad absoluta de alumnos por curso, que es la medida utilizada en la mayoría de los estudios sobre los efectos del tamaño del curso. Los resultados de esta investigación relevan la importancia de la componente conductual de la composición del curso en secundaria.

Como evidencia empírica del modelo teórico que propone Lazear, el profesor Terry Haydn de la Universidad de East Anglia, UK, publicó recientemente⁹⁷ un estudio que da cuenta del *deterioro del clima de aula* producido por las intensas disrupciones de alumnos conflictivos. De acuerdo a su investigación que incluye cuatro estudios conducidos por un período de 10 años, concluye que el aprendizaje de 700.000 estudiantes en UK se ve impedido por el mal comportamiento de algunos. Explora las dificultades que enfrentan escuelas y profesores que intentan reconciliar la inclusión en la educación y el derecho que tienen todos los alumnos a desarrollarse en un ambiente apropiado para el aprendizaje.

Al preguntarle directamente a alumnos de varios países por su percepción respecto del clima de aula en el que están insertos y vincularlo a los resultados de pruebas estandarizadas internacionales, los investigadores Sortkær & Reimer (2014) buscaban testear la existencia de una relación estadística significativa entre clima de aprendizaje auto reportado y la habilidad matemática de éstos medida por los resultados en PISA. Los análisis muestran una correlación positiva y significativa -no trivial- para Canadá, Dinamarca y Noruega entre mejores conductas y clima de aprendizaje reportadas por estudiantes de 15 años que cursan el mismo grado y los mayores logros en la prueba internacional.

⁹⁶ En el sistema público educacional de Chicago los cursos en cuestión son cursos que constituyen un prerrequisito para tomar otros posteriormente y se dictan más de una vez a día a distintos horarios. Los alumnos seleccionan los cursos que toman administrando su horario de forma autónoma.

⁹⁷ Abril, 2014.

Importante evidencia al respecto entrega el estudio longitudinal natural a gran escala realizado en Reino Unido CSPAR. Los autores Blatchford, Edmonds, & Martin (2003) seleccionan un subgrupo de 235 niños de 4/5 años, distribuidos en 21 cursos pequeños (19 alumnos promedio) y 18 cursos numerosos (media de 33 estudiantes) para conducir observaciones de aula y poder establecer patrones de comportamiento al interior de aulas reducidas y compararlos con los de aulas regulares. Además de esta observación, se aplicó un cuestionario a profesores PBS *Pupil Behaviour Rating* que recogió datos de más de 5.000 estudiantes referentes al funcionamiento social en categorías como distracción, asocial, excluido, ansioso/temeroso, comportamiento agresivo y comportamiento pro-social. Los resultados de la observación sistemática muestran que estudiantes de cursos numerosos presentan una mayor probabilidad de estar desvinculados de la clase (*off-task*), particularmente no atendiendo al docente ni realizando su trabajo cuando éste es individual. Es más probable que estudiantes en cursos numerosos interactúen frecuentemente con sus compañeros y menos con el profesor. Participan más en contactos vinculados a tareas, contactos sociales y comportamiento desatento y desvinculado de la clase como molestar y burlarse de otros. Sin embargo los resultados del cuestionario PBS indican una débil relación entre el tamaño del curso y los factores medidos, emergiendo contra intuitivamente una leve tendencia de peores relaciones entre pares en cursos reducidos.

Años más tarde el académico Peter Blatchford lidera nuevamente un estudio, incorporando esta vez alumnos de educación secundaria. La investigación de este equipo de trabajo (Blatchford et al., 2011) indaga en la relación existente entre el nivel de compromiso de los alumnos, las interacciones profesor-alumno y el tamaño del curso. Contrasta ambas variables, y las mide respecto del desempeño previo de los estudiantes -agrupando a estudiantes de alto, medio y bajo desempeño previo-, evalúa los efectos en educación básica y secundaria, indaga sobre la existencia de umbrales para los cuales los efectos se intensifican o moderan y es metodológicamente riguroso⁹⁸. Además se controlaron posibles factores de confusión como género, alumnos con necesidades educativas especiales, asignatura, cantidad de profesores, y profesionales de apoyo. Sus resultados se basan en observaciones de 648 alumnos de 49 escuelas diferentes⁹⁹. Los autores señalan que estudios naturalísticos como este presentan ventajas respecto de experimentos aleatorios como STAR en cuanto a su autenticidad, ya que se basan en observaciones directas¹⁰⁰ de los procesos tal como suceden en aulas reales, y para distintas cantidades de alumnos, en contraposición a comparaciones algo artificiales de aulas de tamaños preestablecidos.

Los principales resultados de este estudio son:

- a) El tamaño del curso se relaciona positivamente con la cantidad de instrucción recibida. Este resultado resulta contradictorio y se interpreta como que en cursos numerosos la

⁹⁸ Utiliza modelos estadísticos multinivel.

⁹⁹ Las observaciones corresponde al proyecto de gran escala llevado a cabo en Inglaterra y Gales en los años 2005/06 DISS (*Deployment and Impact of Support Staff*).

¹⁰⁰ Observaciones cada 10 segundos realizadas por dos observadores independientes (confiabilidad de las observaciones /Coeficiente de Kappa).

mayor parte de la enseñanza es al grupo curso, donde el alumno tiene un rol pasivo de oyente, en contraposición a cursos reducidos donde los estudiantes reciben más atención individual.

- b) El tamaño del curso se relaciona negativamente con la frecuencia de contacto individual por alumno (Gráfico 5).
- c) El tamaño del curso se relaciona negativamente con el tiempo que el alumno permanece enfocado en el trabajo, participando y atento al profesor, y positivamente con el tiempo que el alumno presenta conductas disruptivas y desvinculadas a las actividades de aprendizaje. Sin embargo este resultado difiere para alumnos de distinto desempeño previo y nivel educativo. “Encontramos que para el grupo de menor desempeño en educación secundaria, se registró un comportamiento desvinculado y disruptivo más del doble de veces en cursos numerosos de 30 alumnos que en cursos reducidos de 15” (p. 13), gráfico 6.

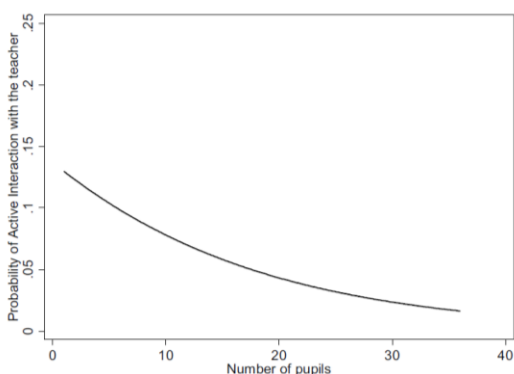


Gráfico 5: cantidad de alumnos por curso y probabilidad de interacción activa con el profesor.

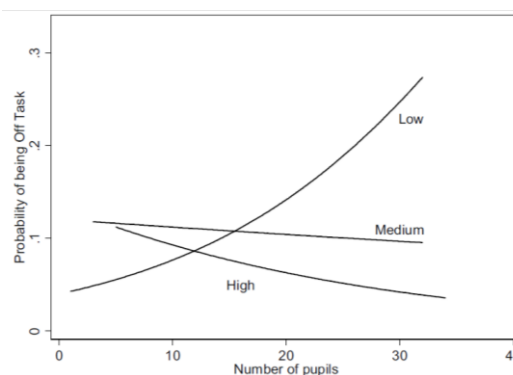


Gráfico 6: cantidad de alumnos por curso y probabilidad de estar desvinculado de las tareas de aprendizaje, para grupos de desempeño inicial alto, mediano y bajo

El hallazgo anterior se relaciona con efectividad educativa, en cuanto evidencia que el problema de cursos numerosos, sobre todo en educación secundaria, son los alumnos más desaventajados, cuyo comportamiento gatilla la necesidad de los profesores por intervenir en dicha conducta, lo que se exagera y redunda en un círculo vicioso.

Por su parte la extensa revisión de los autores Finn, Pannozzo, & Achilles (2003) indaga en las razones que explican los beneficios de cursos reducidos, centrando la atención en el comportamiento de los estudiantes¹⁰¹. Señalan que los alumnos *se comprometen más académicamente y socialmente en cursos reducidos*, y que este creciente mayor compromiso explicaría los mejores resultados de aprendizaje. El estudio del comportamiento de los alumnos en el aula se ha centrado en aspectos como el tiempo que el alumno permanece enfocado en el

¹⁰¹ Desestiman una mayor individualización de la enseñanza y no encuentran cambios significativos de estrategias pedagógicas; reconocen una disminución del tiempo dedicado a labores administrativas y un correspondiente aumento del tiempo de instrucción.

trabajo, participando y atento al profesor. Por el otro lado, se observa el tiempo que el estudiante está desatento y desvinculado de las actividades de aprendizaje o que presenta una conducta disruptiva en el aula. Además señalan, es necesario distinguir el *compromiso académico del social*, que describe la naturaleza de las interacciones tanto entre alumno profesor como entre estudiantes, y que puede ser un comportamiento *pro social* o *antisocial*. Se han encontrado correlaciones negativas y significativas entre la desvinculación de los alumnos y su desempeño académico, que van desde -0.34 a un -0.52. Otros hallazgos indican que el *compromiso en las actividades escolares disminuye a medida que los alumnos crecen* y que la mayor parte de la variabilidad en el compromiso entre estudiantes se explica por sus características personales y no por el colegio o el curso. Sin embargo, *influencias a nivel de aula pueden modificar el nivel de compromiso de los alumnos*.

Preocupados del factor conductual del aula, los grandes experimentos tales como el de Toronto en Canadá, STAR en Tennessee EEUU y Prime Time, entre otros, han llevado a cabo estudios paralelos del comportamiento de los alumnos¹⁰² en contextos de cursos reducidos. Sin embargo difieren mucho en su metodología (de gran y pequeña escala, cuantitativos y cualitativos), definición de variables (test de significancia, medidas del tamaño del efecto ausentes) y variadas fuentes de información (observación, entrevistas y cuestionarios a profesores, reportes, etc.). Ningún estudio revisado toma en cuenta el tiempo que el alumno fue expuesto a un curso reducido, aspecto esencial de la intervención en estudio. Principalmente las investigaciones sobre el número de alumnos por curso y la conducta en el aula se han vinculado a tres aspectos: *comportamiento relativo al aprendizaje*, *comportamiento social* y los *estilos de interacción de profesores*. Respecto de las investigaciones que estudian el comportamiento académico de los alumnos, independientemente de las diversas aproximaciones metodológicas y medidas utilizadas, todos muestran un *impacto positivo sobre el comportamiento hacia el aprendizaje cuando los cursos son reducidos*. Asimismo, los estudios que indagan en el comportamiento social de los alumnos en función de la cantidad de estudiantes por curso, muestran un positivo impacto de reducciones de alumnos por curso, al generar claras disminuciones del comportamiento antisocial y promover, aunque menos claramente, el comportamiento pro social. Por último, el modelo conceptual propone que los profesores también se ven impactados por reducciones de alumnos por curso, al afectar su ánimo, goce y satisfacción de enseñar, lo que a su vez afecta el nivel de compromiso de sus alumnos. Esto se refleja en una mejor actitud del profesor, donde los alumnos perciben que el proceso de enseñanza aprendizaje es más grato; que el docente es menos punitivo y permite que los alumnos exhiban una amplia gama de conductas en clase¹⁰³. Por último, un profesor que conoce a cada alumno mejor, favorece un sentido de comunidad en el aula. Los escasos estudios disponibles avalan los dos primeros

¹⁰² Ver Apéndice A de (Jeremy D. Finn et al., 2003) para detalle de los estudios revisados y el número de alumnos considerado en cursos pequeños y numerosos.

¹⁰³ Los problemas conductuales disminuyen no sólo porque las conductas inapropiadas se advierten instantáneamente y la respuesta se entrega sin demora, ni exclusivamente porque la proximidad profesor-alumno reduce su ocurrencia, sino porque las conductas inadecuadas se redefinen, ya que los alumnos tienen más libertad y una mayor comprensión de las actividades, lo que hace que se comporten mejor (A. Molnar, P. Smith, & J. Zahorik, 1999).

mecanismos, sin embargo, los autores señalan que no hay estudios que aborden el impacto en el ánimo y agrado por la docencia en profesores de cursos reducidos.

Para entender cómo el tamaño del curso afecta el comportamiento académico y social, los autores buscaron vincular estas investigaciones con teorías sobre tamaño de grupo y participación de la psicología y sociología, de donde surgen dos principios centrales: *visibilidad del individuo* y *sentido de pertenencia*. El principio de visibilidad del individuo señala que alumnos en cursos reducidos experimentan una creciente presión de participar; es difícil pasar inadvertido lo que genera un alto compromiso con las actividades escolares en estos contextos. La antítesis de la visibilidad es el anonimato, el que se ha estudiado en términos de la *difusión de la responsabilidad* y la *holgazanería social*. La primera establece que a medida que aumenta el tamaño del grupo la sensación de responsabilidad de actuar disminuye; la holgazanería social por su parte es la disminución del esfuerzo por parte de individuos que actúan en grupo. El esfuerzo disminuye más cuando la persona pasa desapercibida, cuando cree que su contribución tendrá poco impacto o cuando cree que su participación no será evaluada o recompensada. Todo lo anterior es más frecuente en cursos numerosos que en cursos reducidos.

Desde la perspectiva del *sentido de pertenencia*, la reducción de alumnos por curso fomenta la participación ya que grupos más pequeños tienen propósitos más unificados y porque cada miembro se siente vinculado al grupo, donde recibe apoyo y ayuda de otros. La cohesión de grupo ha sido definida como *el valor que le reporta a un miembro la participación en un grupo*. En la escuela ocurre si los alumnos se animan mutuamente a involucrarse en las actividades de aprendizaje, centrándose en la tarea, dándose apoyo y estimulando el esfuerzo mutuo.

Resumen de los estudios descritos en la sección 6.2. Conducta en el Aula						
Título	Autor/ Año/País	Publicación	Alcance	Especificidad	Enfoque	Hallazgo Clave
<i>Educational production.</i>	Lazear, E. P. 2001, EEUU	<i>Quarterly Journal of Economics</i>	Nacional: modelo que explica efectos de congestión en la educación escolar. Impacto: internacional.	Específico: modelo analítico que vincula el CS a disrupciones conductuales en el aula.	Mixto: teoría de congestión, modelo exponencial, cálculo de óptimos.	El tamaño óptimo del curso se obtiene en función de las características de los alumnos Conducta sustituta de tamaño.
<i>Making the Grade: The impact on classroom behaviour on academic achievement.</i>	Cortes, K., Moussa, W., & Weinstein, J. 2013, EEUU	<i>Texas A&M University, Syracuse University.</i>	Local: datos de escuelas públicas de Chicago, nivel I medio, desde 1993-94 hasta 2005-06	Específico: impacto de la componente conductual de la composición del curso sobre logros de aprendizaje.	Cuantitativa: caracterización de alumnos disruptivos y no disruptivos. Correlación con logros de aprendizaje.	La proporción de alumnos no disruptivos respecto de los disruptivos, impacta las tasas de aprobación de Álgebra e Inglés.
<i>Disciplinary climate and student achievement: Evidence from schools and classrooms.</i>	Sortkær, B., & Reimer, D. 2014, Dinamarca	<i>Social stratification and inequality – changing social dynamics in the 21st century</i>	Regional: utiliza datos de prueba PISA 2003 de habilidades matemáticas, alumnos de 15 años en seis países occidentales.	Específico: examina el impacto de la conducta en aula auto reportada en los resultados de la prueba.	Mixto: puntajes del test, tamaño del curso y cuestionarios a alumnos y directores.	Correlación positiva y significativa -no trivial- para Canadá, Dinamarca y Noruega entre mejor conducta y clima de aprendizaje en los resultados de PISA.
<i>Examining the effect of class size on classroom engagement and teacher-pupil interaction: Differences in</i>	Blatchford, P., Bassett, P., & Brown, P. 2011, UK	<i>Learning and Instruction</i>	Nacional: la muestra evaluada pertenece al proyecto a gran escala DISS sobre despliegue e impacto de	General: impacto de CSR en procesos de aula (interacción profesor-alumno y compromiso escolar) para	Cuantitativa: estudio naturalístico observacional (cada 10 segundos) con modelos	Cursos reducidos benefician a todos (enseñanza individualizada y atención del profesor) pero sobre todo al grupo de

<i>relation to pupil prior attainment and primary vs. secondary schools.</i>			personal de apoyo en escuelas de Inglaterra y Gales, año 2005/6.	distintos niveles /por desempeño inicial (bajo-medio-alto)	estadísticos multinivel.	desempeño inicial bajo en secundaria.
<i>The "Why's" of Class Size: Student Behavior in Small Classes</i>	Finn, J. D., Pannozzo, G. M., & Achilles, C. M. 2003, EEUU	<i>Review of Educational Research</i>	Internacional: Revisión de estudios sobre CSR y compromiso del alumno (EEUU, Australia, UK, Canadá)	General: Revisión bibliográfica y vinculación con teorías psicológicas y sociológicas.	Cualitativa. Proceso de aula que explica efectos positivos de CSR.	CSR impacta compromiso de los alumnos. Se vincula a teorías de comportamiento de grupo.

6.3. Habilidades No Cognitivas y Efectos de Largo Plazo

En general se considera que los recursos invertidos en educación deben generar retribuciones de mediano y largo plazo en la calidad de vida de aquellos que la reciben. Los economistas por su parte postulan que un rol fundamental de la escuela es preparar a los individuos para participar en el mercado del trabajo. Muchas veces las mejores oportunidades que se le presentan a un individuo y que impactan su calidad de vida positivamente, se vinculan a las llamadas *habilidades blandas*. A pesar del creciente valor que han adquirido, su medición es poco frecuente y en consecuencia, raramente son consideradas en los análisis a la par con las habilidades cognitivas.

Las *habilidades no cognitivas* se refieren a un vasto rango multidimensional de hábitos de trabajo y rasgos de comportamiento que no son medidos en evaluaciones convencionales y que abarcan una amplia gama de constructos tales como la disciplina, la motivación, el locus de control y la autoestima. Indicadores tan diversos como esfuerzo, auto control, confianza y estabilidad emocional aparecen vinculadas al éxito académico y a otros resultados de largo plazo. Estudios sobre habilidades no cognitivas y su relación con el tamaño del curso se sustentan en la línea de investigación que psicólogos educacionales han desarrollado en torno al concepto de *school engagement* o compromiso hacia la escuela y sus efectos en el largo plazo. Los principios de visibilidad y sentido de pertenencia recientemente descritos están intrínsecamente ligados a las habilidades no cognitivas, que como factor de mejoramiento de las condiciones laborales¹⁰⁴, últimamente han ganado una creciente importancia y reconocimiento.

A diferencia del coeficiente intelectual del alumno, que mayormente se estabiliza durante el periodo escolar primario, las *habilidades no cognitivas* parecieran ser más moldeables a edades más avanzadas (Carneiro & Heckman, 2003; Heckman, 2000). A lo anterior se suma lo señalado por (Boozer & Rouse, 2001) que propone la naturaleza acumulativa de la escolarización como una posible razón que explicaría menores magnitudes para efectos de reducciones del tamaño del curso en niveles tardíos de educación secundaria.

Posibles mecanismos mediante los cuales los cursos reducidos promueven el desarrollo de habilidades no cognitivas, son una mayor facilidad con que el docente puede limitar las conductas disruptivas, lo que a su vez facilita el desarrollo de la atención y el autocontrol. Por otra parte, en cursos más pequeños, la capacidad del docente de moldear e influir motivacionalmente en el grupo curso se amplía, lo que permite maximizar el esfuerzo de estudiantes y desarrollar su resiliencia ante desafíos educativos.

¹⁰⁴ Estabilidad laboral, tener un empleo tiempo completo, nivel de ingresos, etc.

Estrechamente vinculado a las habilidades no cognitivas surge el auto-concepto académico que desarrollan los estudiantes durante su vida escolar, como una evaluación de sus propios logros académicos, al compararse con sus compañeros de curso. Un auto-concepto académico *positivo* es clave en la *motivación y desempeño* de los alumnos. El efecto *Big Fish Little Pond*¹⁰⁵ supone que este constructo se correlaciona negativamente con la habilidad de los pares. Los autores Thijs, Verkuyten & Helmond (2010) estudian este constructo en pre adolescentes Holandeses. Los investigadores toman en consideración el tamaño del curso, entre otras variables, en el desarrollo de este concepto. Señalan que el tamaño del curso inhibe el efecto negativo en la construcción del auto-concepto académico en su dimensión comparativa, ya que en cursos numerosos es más difícil saber cuál es el desempeño de los demás, se manejan menos datos y se hace más difícil establecer una posición académica relativa. Sin embargo se esperaba encontrar que en cursos numerosos la importancia que se le otorga a la comparación individual con pares de mejor rendimiento fuera mayor ya que al haber más estudiantes, habría más oportunidades de encontrar compañeros de desempeño levemente superior que hicieran más atractiva y desafiante dicha comparación. Esta comparación individual, a diferencia de la comparación con todos los demás compañeros, motiva a mejorar y tiene efectos positivos en el desempeño. Los resultados muestran que la importancia relativa de la comparación general no cambia a pesar de que ésta se vuelve menos precisa. En consecuencia, el tamaño del curso tiene un impacto mixto en la conformación del auto-concepto académico.

Vinculados a las habilidades no cognitivas, el auto concepto académico y la motivación, son factores altamente influyentes en resultados de mediano y largo plazo. Así, la mayoría de los estudios que relacionan las *habilidades no cognitivas* y tamaño del curso buscan establecer correlaciones con resultados académicos de mediano plazo específicos, como la finalización de la educación secundaria, rendición de exámenes, ingreso y titulación de la educación terciaria; y de largo plazo relativo a aspectos laborales como sueldo e ingresos. Desde el punto de vista económico, la efectividad de los recursos destinados a la educación debiera ser medida tanto por el impacto en los ingresos de sus graduados como por sus puntajes en test.

Los autores Finn, Gerber & Boyd-Zaharias (2005), a partir de los datos del proyecto STAR y un seguimiento por 13 años de 5.000 de estos alumnos en una extensa base de datos, indagaron sobre la relación entre haber sido parte de un curso reducido en los niveles iniciales y *finalizar la educación secundaria*. La investigación ha establecido claramente la relación entre desempeño académico y deserción escolar, encontrando que bajos logros de aprendizaje en los niveles iniciales (K-3) la explican como el principal factor. De la muestra evaluada, el 77,5% se graduó de secundaria. Si se desagrega la información por cantidad de años en cursos reducidos y vulnerabilidad, los resultados aparecen en la tabla 3:

¹⁰⁵ Análogo al dicho “el tuerto en el país de los ciegos”, que sugiere que el auto-concepto mejora cuando el desempeño de los pares es relativamente peor.

Tabla 3: Tasa de graduación por cantidad de años de tratamiento y vulnerabilidad. STAR.

N° de años en curso reducido	Tasa de Graduación		
	Vulnerable ¹⁰⁶	No vulnerable	Todos
0 (curso numeroso)	70.2	83.7	76.3
1	68.1	78.3	72.8
2	70.1	85.2	76.8
3	79.6	82.8	81.1
4	88.2	87.0	87.8

Destaca que la mayor tasa de graduación registrada para alumnos vulnerables que estuvieron los 4 años en cursos reducidos, es decir, se confirma la tesis de que los más beneficiados por esta intervención son los alumnos provenientes de los hogares más pobres. Se observa también un crecimiento monotónico de la tasa de graduación a medida que aumentan los años en cursos reducidos. Lo anterior confirma el supuesto de que el *efecto es acumulativo e incremental* y desmiente la suficiencia de 1 año que postula Hanushek. Se requieren 3 o 4 años de intervención para que las tasas de graduación se vean afectas de forma significativa, lo que es consistente con los 3 o 4 años requeridos para mantener mejoras de largo plazo en logros de aprendizaje (J.D. Finn et al., 2001). Llama la atención la menor tasa de graduación para alumnos que participaron por un año de cursos reducidos. Los autores atribuyen este resultado a una posible mayor transitoriedad de estos estudiantes, que probablemente fueron cambiados de colegio durante los niveles iniciales.

Asimismo el estudio de Browning & Heinesen (2007) busca relacionar la *probabilidad de continuar estudios* después de la educación obligatoria¹⁰⁷ en **Dinamarca** con el número de alumnos por curso y la cantidad de estudiantes por hora-profesor semanal, en base a un diseño de regresión discontinua donde el promedio de alumnos por curso es 20 y el umbral máximo permitido corresponde a 24. La educación secundaria -no obligatoria- consta de 3 años adicionales, y culmina a los 18 años. Se utilizó una base de datos con información de las 8 cohortes que cursaron 8° básico desde 1985 hasta 1992, con seguimiento hasta 2003. Las estimaciones sugieren que una reducción del 5% del curso (1 alumno en un curso de 20 estudiantes en promedio) aumenta la probabilidad de completar la educación secundaria en un 0.4% y aumenta la cantidad de años de educación en 0.02. Asimismo, si la variable explicativa es la reducción de un 5% en la cantidad de estudiantes por hora-profesor semanal, las probabilidades de completar la educación secundaria aumentan en 1,1% y en 0.056 los años de educación adicionales. Estos tamaños del efecto son consistentes con resultados de estudios americanos previos.

Por su parte los autores Dynarski, Hyman, & Schanzenbach (2011) examinaron el efecto de cursos reducidos en niveles iniciales sobre la *chance de ingreso a la universidad*, el tipo de *carrera elegida* y la *completación* de ésta. En base a los datos del experimento STAR hallaron que la asignación a cursos reducidos en K-3 aumenta la probabilidad de ingreso a la universidad en más de un 2,7%. Este efecto se duplica si se mide sólo a la población afroamericana, y si se toma en

¹⁰⁶ En el estudio esta categoría corresponde a aquellos alumnos que califican para obtener alimentación gratuita.

¹⁰⁷ Hasta 9° grado, 15 años de edad.

consideración sólo el tercio de las escuelas más pobres, el efecto alcanza un 7,3%. En general, el haber sido parte de un curso reducido aumenta la probabilidad de completación del grado universitario y moviliza a los alumnos hacia actividades mejor remuneradas asociadas a las áreas de las ciencias, tecnología, ingeniería, matemática, negocios y economía.

También hay estudios que vinculan la variable de interés a aspectos más generales que impactan la calidad de vida como salud, comportamiento cívico y bienestar.

Un estudio sueco explora la posibilidad de que el tamaño del curso pueda afectar la *prevalencia de problemas psicológicos y el bienestar* de adolescentes en Suecia. En un intento por relevar otras variables de interés que afectan la calidad de vida –además de los resultados de aprendizaje– estudian la relación de la intervención en cuestión con aspectos vinculados al bienestar físico y psíquico. Los autores Jakobsson, Persson, & Svensson (2013) utilizan datos de corte transversal del año 2008 de 2.755 adolescentes que cursan 9° grado. En particular se mide la prevalencia de problemas de concentración, preocupación, ansiedad y tristeza, mareos, dolores de cabeza y de estómago en alumnos de cursos numerosos y reducidos. Éstas son de interés además porque indirectamente afectan el desempeño académico. Sin embargo, los resultados basados en diversas aproximaciones econométricas que abordan potenciales variaciones endógenas a nivel de escuela y curso, no muestran una prevalencia robusta de estos problemas psicológicos y demás indicadores de bienestar en cursos numerosos.

Una dimensión clave del bienestar lo constituye una convivencia sana. Se asume que una mejor calidad de la educación y que mejores oportunidades debieran reducir la incidencia de comportamiento antisocial y delictivo. Bajo esa premisa los autores Arum & LaFree (2008) estudian la relación entre las características de la escuela, como la proporción profesor alumno, y el *riesgo de reclusión* en la vida adulta en EEUU. El estudio contrapone la mirada de los escépticos del sistema escolar que postulan que la escuela tiene un muy bajo impacto en la criminalidad y el riesgo de encarcelamiento, ya que la propensión de criminalidad se fija a una edad muy temprana. Adicionalmente las ineficiencias organizacionales de las escuelas públicas impedirían la utilización de los recursos de manera efectiva para alterar los resultados. La postura se contrasta con defensores del sistema que señalan que las escuelas proveen cada vez más momentos de definición crítica en el curso de la vida y que mejorando las oportunidades económicas y facilitando el control social en las escuelas, mayores recursos pueden incidir en la disminución de la criminalidad y el riesgo de reclusión. Los logros educativos se relacionan cada vez más con dimensiones de vida como la empleabilidad, el matrimonio y la reclusión. Estudios previos han hallado que comportamientos delictivos y criminales se asocian fuertemente con una serie de variables educativas como calificaciones, disgusto por el estudio, comportamiento disruptivo y logros de aprendizaje¹⁰⁸. La base de datos utilizada en este estudio contempla 5 cohortes partiendo de 1910 y toma datos del censo y del estudio longitudinal nacional de la juventud NLSY. Para evitar resultados espurios, se

¹⁰⁸ Jóvenes insatisfechos con la escuela son más propensos a abusar del consumo de drogas, presentar comportamiento sexual riesgoso, crimen y suicidio. Lévy-Garboua, L, Lohéac, Y, Fayolle, B (2006).

estudia el riesgo que enfrenta una persona nacida y escolarizada en un determinado estado (con alta o baja proporción de estudiantes por profesor¹⁰⁹) de ser encarcelada en otro estado, ya que cambió de residencia. La evidencia muestra que las personas son recluidas tanto por su comportamiento como por prejuicios y sesgos de aquellos que toman las decisiones¹¹⁰. La pregunta que se propone responder este estudio es si aquellos estudiantes que fueron educados en contextos donde había más alumnos por profesor enfrentan un mayor riesgo de reclusión en la adultez. La pregunta subyace la interrogante sobre el impacto que pueden tener mayores recursos educativos en los resultados y trayectorias de vida de los estudiantes. Los defensores enfatizan la maleabilidad de los jóvenes y argumentan que la escuela puede cambiar su comportamiento permanentemente. Los datos revelan que el riesgo de reclusión disminuye al aumentar los recursos en educación, sobre todo al disminuir la proporción de alumnos por profesor. Aquellos estados que invierten más en mejorar este indicador también destinan más recursos a programas de bienestar social, en consecuencia ambos están asociados a un menor riesgo de reclusión. Los autores señalan la imposibilidad de especificar el mecanismo mediante el cual la disminución de alumnos por profesor reduce el riesgo de criminalidad, sea ésta por una mejor socialización o por el impacto que produciría en la toma de decisiones de los jueces.

Muchos de los trabajos que vinculan las características de la escuela y la utilización de los recursos educativos a resultados de largo plazo, buscan determinar la rentabilidad de la medida como un indicador objetivo que justifique la significativa inversión que implica el financiamiento de esta intervención. Si bien la tasa de retorno a la inversión obtenida es desigual para distintos países y regiones, es consistentemente positiva.

El trabajo de los autores Dee & West (2011) utiliza el estudio longitudinal nacional de educación de EEUU NELS para indagar en la relación existente entre las habilidades no cognitivas detectadas en estudiantes de 8° de Educación Básica y las *condiciones laborales* futuras de estos estudiantes. Las habilidades no cognitivas se conceptualizan en el compromiso académico, el que se desagrega en compromiso conductual y compromiso psicológico o emocional el que se mide aplicando encuestas a profesores y alumnos. El diseño de la investigación se basa en la diferencia del tamaño del curso para dos asignaturas comparadas¹¹¹, con un promedio de 24.5 estudiantes por curso y una desviación estándar de 5.9. Los docentes de las asignaturas proporcionaron información de cada alumno en la muestra, relativa a su comportamiento y desempeño. Además la encuesta de seguimiento a estudiantes entrega mayor información sobre esfuerzo y compromiso. Los resultados indican que las habilidades no cognitivas -en particular el compromiso psicológico o emocional- presumiblemente facilitadas por cursos reducidos, son potentes predictores de logros educativos posteriores, además de generar beneficios para adultos jóvenes en el mercado laboral. Los autores estiman que el costo de reducir la cantidad de alumnos por curso en una desviación estándar, que es proporcional al costo por alumno, corresponde a

¹⁰⁹ La proporción profesor alumno por estado es relativamente estable y difiere entre estados para el periodo en cuestión.

¹¹⁰ En EEUU la probabilidad de ser recluido si se es afroamericano o hispano es más alta que para blancos.

¹¹¹ En pares: Matemática o Ciencias combinado con Inglés o Historia.

aumentar la cantidad de cursos en un 31%. Considerando el impacto en los ingresos futuros de estudiantes en contextos de cursos reducidos que se ven beneficiados tanto en términos cognitivos como no cognitivos, con tamaños del efecto que van del 0.05 a 0.09, se obtiene una *tasa interna de retorno* de la intervención que alcanza al 4,6% en general y a un 7,9% en sectores urbanos en particular. Otro hallazgo importante es el impacto positivo que se observó entre el aumento de habilidades no cognitivas y una mayor participación cívica, manifiesta en el ejercicio del voto en elecciones y la realización de trabajo voluntario por parte de los estudiantes que fueron parte de cursos reducidos.

Los europeos por su parte también buscan juzgar la efectividad de la intervención mediante el cálculo de la tasa interna de retorno en su contexto. Así, los investigadores **suecos** Fredriksson et al., (2012) estudian el impacto de largo plazo de reducciones del tamaño del curso sobre la *finalización de estudios, los ingresos y salarios* que obtuvieron los estudiantes entre los 27 y 42 años de edad. Desestiman la validez de otros estudios¹¹² que miden los ingresos muy tempranamente para poder observar tendencias de largo plazo o que muestran dificultades en el acceso a datos directamente vinculados a resultados del mercado laboral como es el caso de Krueger (2003) que concluye que la tasa interna de retorno alcanza apenas a un 6%. En particular estudian el efecto de cursos reducidos en segundo ciclo de primaria¹¹³, haciendo seguimientos de sus logros cognitivos y no cognitivos a edades de 13, 16 y 18 años, así como de resultados académicos posteriores, y condiciones laborales a muestras de las cohortes nacidas en los años 1967, 72, 77 y 82. Los indicadores de habilidades no cognitivas a los 13 años de edad provienen de cuestionarios que indagan sobre la situación del alumno en el colegio: se reportan aspectos como esfuerzo, motivación, aspiraciones, autoconfianza, sociabilidad, ausentismo y ansiedad. A los 18 años de edad se cuenta con un puntaje general de habilidades no cognitivas para varones obtenido de las entrevistas que realizan a los reclutas del servicio militar, donde se observan aspectos como estabilidad emocional, persistencia, aptitud social, disposición a asumir responsabilidad y capacidad para tomar la iniciativa. Los resultados muestran beneficios cognitivos y no cognitivos desde los 13 años, además de aumentos en la tasa de finalización de estudios, mejores salarios e ingresos. Un análisis costo beneficio sugiere que una reducción de 25 a 20 estudiantes por curso tiene una *tasa interna de retorno de casi un 18%*. Basados en la cantidad máxima de alumnos por curso fijada en 30, los autores utilizan un modelo de regresión discontinua para estimar el efecto causal del tamaño del curso. Las estimaciones sugieren que una reducción de 7 alumnos -equivalente a STAR- mejora las habilidades cognitivas a la edad de 13 años en 0.23σ . La magnitud del efecto en las habilidades no cognitivas como el compromiso hacia

¹¹² Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan, "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR" *Quarterly Journal of Economics*, 126 (2011), 593–1660.

¹¹³ El estudio que estima el efecto del tamaño del curso para alumnos de 6° de Educación Básica en Estocolmo "Home versus School Learning: A new approach to Estimating the Effect of Class Size on Achievement" del investigador Mikael Lindahl (2005) obtiene resultados similares a los encontrados en STAR.

la escuela, autoconfianza y perseverancia es levemente menor, cuantificada en 0.026σ por cada alumno menos. A los 16 años el 70% del efecto permanece y su tamaño es de 0.16σ por cada unidad reducida. La intervención además aumenta la cantidad de años de escolaridad en 0.05 años, equivalente a dos tercios de mes, y la probabilidad de obtener un título universitario en 0.8%.

Previamente Krueger & Whitmore (2001) habían encontrado un efecto significativo del tamaño del curso en la plausibilidad de rendir exámenes de ingreso a la universidad. Un 43,7% de los estudiantes que pertenecía a cursos reducidos en STAR versus el 40% de estudiantes asignados a cursos regulares, rindieron la prueba. Sin embargo el impacto para los afroamericanos fue mayor: la tasa de rendición del examen de ingreso aumentó de un 31,7% a un 40,2% para este grupo, lo que produjo una reducción de la brecha en la tasa de rendición de esta prueba entre alumnos blancos y negros de un 54%.

Relativo a los ingresos, éstos aumentan en 1,2% respecto del promedio, y los salarios crecen en un 4,4% por cada estudiante menos en el curso. Para el cálculo del análisis costo beneficio se considera el efecto del salario, ya que es un mejor estimador del impacto de la intervención sobre la productividad de los individuos. Asumiendo que durante 3 años se reduce de 25 a 20 la cantidad de alumnos por curso, por cada 100 alumnos, se pasa de requerir 4 docentes a 5. Además se contempla un tercio del costo adicional del profesor para infraestructura y administración. Se compara entonces el valor presente de estos costos cuando el alumno tiene 10 años con el valor presente de los beneficios. La tasa interna de retorno que equipara costos y beneficios es de 0,178, por lo que para tasas de descuento menores a esta cifra el valor presente neto de la reducción de 5 alumnos por curso durante 3 años es positivo.

Resumen de los estudios descritos en la sección 6.3. Habilidades No-Cognitivas y Efectos de Largo Plazo						
Título	Autor/ Año/País	Publicación	Alcance	Especificidad	Enfoque	Hallazgo Clave
<i>A Further Examination of the Big-Fish-Little-Pond Effect: Perceived Position in Class, Class Size, and Gender Comparison.</i>	Thijs, J., Verkuyten, M., & Helmond, P. 2010, Holanda	<i>Sociology of Education</i>	Nacional: estudio sobre pre adolescentes holandeses.	Específico: conformación de auto concepto académico en relación al tamaño del curso.	Mixto: cuestionario a alumnos último nivel de primaria. Desempeño categoría vocacional y notas	CS tiene efecto mixto en la construcción del auto concepto académico. CSR beneficia comparación grupal, pero no individual (mejora por proximidad).
<i>The Enduring Effects of Small Classes.</i>	Finn, J. D., Gerber, S. B., Achilles, C. M., & Boyd-Zaharias, J. 2001, EEUU	<i>Teachers College Record</i>	Local: utiliza los datos de STAR y seguimiento en 4°, 6° y 8°, Tennessee, EEUU. Lasting Benefits Study LBS.	Específico: cuantifica tamaño del efecto en función de años de intervención e impacto posterior.	Cuantitativo: se estima el tamaño del efecto cuantificado en meses de progreso. Usa modelo jerárquico lineal.	El año de inicio y la duración de la intervención impactan los beneficios ganados. Perdurabilidad en los efectos requiere Inicio temprano y mantención de al menos 3 años CSR.
<i>Class Size, Teacher Hours and Educational Attainment.</i>	Browning, M., & Heinesen, E. 2007, Dinamarca	<i>The Scandinavian Journal of Economics</i>	Nacional: base datos de Dinamarca, 8 cohortes, nivel 8°, desde 1985-92 seguimientos hasta 2003.	Específico: examina efecto de CSR en 8° sobre probabilidad de completar educación secundaria no obligatoria.	Cuantitativa: usa modelo de regresión discontinua para estimar efecto de CSR.	Efectos positivos sistemáticos de CSR. Aumenta probabilidad de terminar educación secundaria y aumentan años de estudios.
<i>Experimental Evidence on the Effect of Childhood Investments on</i>	Dynarski, S., Hyman, J., &	<i>National Bureau of Economic Research.</i>	Local: datos de STAR (termina ed. escolar en 1998) y los	Específico: estima impacto de CSR en chance de ingreso y término de ed.	Cuantitativo: estima el impacto de CSR en ingresar y completar ed.	CSR aumenta la probabilidad de ingresar y terminar una carrera universitaria, sobre todo

<i>Postsecondary Attainment and Degree Completion.</i>	Schanzenbach, D. W. 2011, EEUU		valida con NSC en admisión y término de college 06 -10.	postsecundaria.	postsecundaria, por raza y etnia.	a estudiantes de grupos minoritarios y vulnerables.
<i>Class-Size Effects on Adolescents' Mental Health and Well-Being in Swedish Schools.</i>	Jakobsson, N., Persson, M., & Svensson, M. 2013, Suecia	<i>Education Economics</i>	Nacional: datos de adolescentes de 1° medio, en escuelas de Suecia año 2008.	Específico: preguntas asociadas a bienestar y salud mental auto reportadas por alumnos.	Mixto: cuestionario mide: concentración, ansiedad, pena, mareo, dolor de cabeza y de estómago.	No se encuentra relación entre salud mental y bienestar con tamaño del curso.
<i>Educational Attainment, Teacher-Student Ratios, and the Risk of Adult Incarceration Among U.S. Birth Cohorts Since 1910</i>	Arum, R., & LaFree, G. 2008, EEUU	<i>Sociology of Education</i>	Nacional: datos del censo de EEUU para 5 cohortes desde 1910.	Específico: vincula proporción de alumnos por profesor con tasas de criminalidad y riesgo de reclusión en la vida adulta.	Cuantitativa: correlación entre tamaño del curso y criminalidad, método de control de efecto fijo.	Proporciones mas bajas de alumnos por docente se asocian a un menor riesgo de reclusión en la vida adulta.
<i>The Non-Cognitive Returns to Class Size</i>	Dee, T. S., & West, M. R. 2011, EEUU	<i>Educational Evaluation and Policy Analysis</i>	Nacional: panel de datos NELS:88, seguimiento 1990 y 2000.	General: impacto de CSR en 8° básico en resultados académicos y resultados de largo plazo.	Mixto: análisis cuasi experimental de NELS; entrevistas de seguimiento.	Impacto de CSR en las habilidades no cognitivas con implicancias educativas y en el mercado del trabajo. TIR entre 4,6% y 7,9%.
<i>Long Term effects of class size.</i>	Fredriksson, P., Öckert, B., & Oosterbeek, H. 2012, Suecia	<i>The Quarterly Journal of Economics</i>	Nacional: datos de estudiantes suecos de 13, 16 y 18 años, seguimiento entre los 27-42. Cohortes de 1967-82.	General: examina impacto en resultados académicos y de largo plazo.	Mixto: análisis de puntajes de test, salarios. Conducción de entrevistas para habilidades no cognitivas.	CSR en los 3 últimos años de primaria mejoran puntajes de test y habilidades no cognitivas, lo que impacta el desempeño y resultados de largo plazo (educación y salarios)
<i>The Effect of Attending a Small Class in the Early Grades on College-Test Taking and Middle School Test Results: Evidence from Project Star.</i>	Krueger, A. B., & Whitmore, D. M. 2001, EEUU	<i>Economic Journal</i>	Local: utiliza una muestra de STAR, seguimiento y cruce con información rendición de SAT y ACT en 1998.	Específico: estima impacto de CSR en K-3 sobre desempeño en cursos superiores y rendición de exámenes de ingreso college.	Cuantitativo: análisis de puntajes de test. Variable dicotómica rinde o no examen de ingreso.	Beneficio de CSR en K-3 cae a la mitad al pasar a cursos regulares, pero persiste; aumenta la probabilidad de rendir el ACT o SAT, cae la brecha entre blancos y negros a la mitad.

6.3.1. La Percepción de los Estudiantes

En el contexto actual de responsabilización, los académicos y la comunidad educativa han tendido a equiparar la noción de éxito educativo a desempeño académico y con esto han introducido reformas que si bien avanzan en logros de aprendizaje aumentan muchas veces el desagrado hacia la escuela.

Es cada vez más clara la relación que existe entre la satisfacción hacia la escuela¹¹⁴ y la vida en general con el bienestar. Conocer estos predictores de satisfacción debiese ayudar a los actores educativos a organizar las escuelas de tal forma que se maximicen ambos aspectos: aprendizaje y calidad de vida. La literatura ha estudiado estos indicadores que se clasifican en tres categorías:

1. Factores de escuela/cursos: posibilidad de participar en actividades y controlar su ambiente, recibir un trato justo, sentirse seguro, beneficiarse de un clima de aula positivo.
2. Factores sociales: nivel de apoyo que recibe del profesor y en menor medida de sus pares, nivel de aceptación que percibe y ausencia de victimización.

¹¹⁴ Ha llegado a constituir uno de los 5 factores críticos de la satisfacción de los niños en general.

3. Factores individuales: competencia académica y social percibida, compromiso hacia el aprendizaje.

Los autores Randolph, Kangas, & Ruokamo (2008) expanden esta línea de investigación identificando los predictores de satisfacción hacia la escuela en 331 escolares holandeses y finlandeses de entre 6 y 13 años de edad. Los resultados de este estudio, consistentes con la literatura precedente, arrojan como principales factores de satisfacción el género, la edad y el nivel de agrado o desagrado hacia el profesor o profesores. En general las mujeres y alumnos de niveles iniciales reportan mayor satisfacción hacia la escuela. Sin embargo y al contrario de lo que concluyeran Verkuyten & Thijs en 2002, resulta significativa la relación entre *satisfacción y tamaño del curso*, apareciendo un *máximo* en torno a los 20 estudiantes por curso. Relativo a esta variable, los estudiantes que reportaron menor satisfacción hacia la escuela, pertenecían a cursos de menos de 15 alumnos o a más de 25 estudiantes.

Asimismo, la encuesta a estudiantes Gallup de Marzo de 2009, que recoge las apreciaciones de más de 70.000 alumnos de entre 5° de Educación Básica y 4° de Educación Media de 18 estados de EEUU, es consistente con lo hallado anteriormente. La encuesta mide el bienestar *experimentado*, que revela cómo se siente el individuo respecto de su vida pasada. El tamaño del curso se asocia negativamente con este índice, es decir, *a mayor tamaño del curso menor bienestar* reportado (Lopez, 2009).

La satisfacción hacia la escuela también se vincula a la ansiedad generada producto de las altas expectativas en ambitos académicos específicos. En particular en el contexto asiático la ansiedad de aprender un idioma extranjero es muy alta y se observó que el tamaño de la clase ayuda significativamente a disminuir este tipo de tensión. El trabajo de Harfitt (2012b) indaga en el porqué y el cómo el tamaño de la clase puede hacer una diferencia en el proceso de enseñanza aprendizaje desde el punto de vista de los alumnos de secundaria. Este estudio de caso realizado en Hong Kong analiza datos tanto de entrevistas como de observación de clase. Conceptualiza el aprendizaje de los alumnos no como resultados de pruebas estandarizadas sino como un proceso: indaga en el grado de participación de los alumnos y otras dimensiones sociales, culturales y psicológicas relevantes; la visibilidad de los alumnos y el sentido de pertenencia a la clase. De acuerdo al autor, los alumnos reportaron mantener relaciones mas cercanas con sus pares, y las tutorías entre compañeros resultaron mas frecuentes en clases mas pequeñas. En cursos reducidos el apoyo entre compañeros, incluso en la preparación para exámenes, es mayor. Los mejores resultados se asocian al concepto de zona de desarrollo próximo planteada por Vigotsky.

Resumen de los estudios descritos en la sección 6.3.1. La Percepción de los Estudiantes						
Título	Autor/ Año/País	Publicación	Alcance	Especificidad	Enfoque	Hallazgo Clave
<i>Predictors of Dutch and Finnish Children's Satisfaction with Schooling.</i>	J., Randolph., Kangas, M., & Ruokamo, H. 2008, Holanda	<i>Journal of Happiness Studies</i>	Local: estudiantes holandeses y finlandeses de entre 6 y 13 años, 2008.	Específico: ratifica predictores conocidos de satisfacción con la escuela e incluye nuevas variables (tamaño curso*).	Mixto: utiliza la escala <i>The Children's Overall Satisfaction with Schooling</i>	* Alumnos de cursos de aprox. 20 alumnos reportaban mayor satisfacción con la escuela, que de cursos con menos o mas compañeros.

<i>How class size reduction mediates secondary students' learning: hearing the pupil voice.</i>	Harfitt, G. J. 2012, Hong Kong	<i>Asia Pacific Education Review</i>	Local: estudio de caso de 3 profesores de inglés de secundaria en Hong Kong, ejerciendo en cursos reducidos y numerosos.	Específico: perspectiva del alumno que experimenta CSR. Cómo y porqué hace una diferencia aprender en contextos de cursos reducidos.	Cualitativo: observación de clase, transcripciones, conducción de entrevistas	Los estudiante perciben CSR como capaz de realzar el sentido de pertenencia y cohesión con el curso; relación mas cercana con el docente, y mayor participación en clase.
---	--------------------------------	--------------------------------------	--	--	---	---

VII. CONCLUSIONES

Existe un conjunto amplio de investigaciones que buscan responder a la pregunta por el impacto de la cantidad de alumnos por curso en los aprendizajes de los estudiantes. Si bien el debate ha estado caracterizado por posturas fuertemente encontradas que representan visiones de la educación muy distintas, los estudios e investigaciones conducidos en los últimos 30 años han despejado en parte el confuso panorama. Se distinguen dos enfoques en la investigación que abordan este tema: uno que relaciona directamente el tamaño del curso con logros de aprendizaje medidos por puntajes de test; el otro examina cómo la reducción de alumnos por curso impacta los procesos de aula, y busca identificar cuáles y en qué medida estos mecanismos son los causantes de mejores resultados educativos.

La principal dificultad que presenta la estimación del impacto del tamaño del curso, es la aislación de su efecto respecto de las otras múltiples variables que intervienen en los aprendizajes de los alumnos. En consecuencia, la metodología utilizada, las técnicas estadísticas de análisis y las estrategias de especificación elegidas han concentrado gran parte de la discusión. Actualmente existe una definición clara del constructo de interés *número de alumnos por curso*; sus implicancias educativas están definidas y despejadas las diferencias con la *proporción de alumnos por docente* que introdujeron confusión al debate temprano. Adicionalmente, se ha recabado abundante evidencia sobre prácticas de *clasificación* de alumnos *entre escuelas y al interior* de éstas, que en conjunto con una creciente tendencia al uso de políticas educativas *compensatorias*, establecen ciertos requisitos metodológicos para una correcta estimación de la variable analizada, que corrija estos sesgos de clasificación. Por su parte la comparación del efecto del tamaño del curso entre distintos países y sistemas educativos disímiles, ha permitido desechar la hipótesis de un *efecto universal* absoluto, admitiendo la comprensión del fenómeno en forma contextual y dependiente de factores y características educativas particulares a cada sistema. Actualmente hay consenso en torno a la idea de que los cursos reducidos ofrecen mejores *oportunidades* de aprendizaje, pero *no causan* una mejor enseñanza. Cabe señalar que los experimentos aleatorios como STAR y los estudios naturalísticos como CSPAR presentan una mayor capacidad predictiva respecto de las otras estrategias de identificación utilizadas. Éstos han reportado impactos positivos y significativos, con un tamaño del efecto en logros de aprendizaje en torno a 0.25σ producto de una reducción de entre 7 y 10 alumnos por curso.

Hay una serie de dimensiones clave asociadas a reducciones del número de alumnos por curso como la existencia de umbrales, la etapa escolar más sensible de reportar beneficios, el efecto sobre grupos minoritarios, la perdurabilidad del efecto, su naturaleza acumulativa y su impacto en habilidades no cognitivas, que han sido abordadas en la literatura extensamente. La hipótesis que postula un *umbral* de 20 alumnos por curso –tamaño máximo que favorece el desencadenamiento de procesos educativos beneficiosos en cursos reducidos– queda parcialmente descartada, dada la abundante evidencia de mayores logros de aprendizaje a partir de disminuciones que no alcanzan ese tamaño. Sin embargo algunos estudios que establecen un *tamaño óptimo* del curso basados en la teoría del aprendizaje social, y que consideran criterios

académicos, de satisfacción o de desarrollo regional, han encontrado que este corresponde a 20 estudiantes por curso, o aproximaciones cercanas.

Gran parte de la investigación en este ámbito se ha concentrado en el *inicio de la etapa escolar*, por considerar más efectiva la medida en estos niveles que se caracterizan por una mayor plasticidad y adaptabilidad de los alumnos. La hipótesis sugiere que un inicio escolar más contenido y mejor fundado junto a incitativas remediales tempranas, sienta las bases para un mejor desempeño posterior. Al contrario, la naturaleza acumulativa de la escolarización, debilitaría el impacto de la medida en niveles educativos superiores. Sin embargo la investigación ha dado cuenta de que las reducciones de alumnos por curso traen beneficios de otra naturaleza en los niveles educativos posteriores, producto de la moldeabilidad de las *habilidades no cognitivas* a edades más avanzadas. Asimismo la intervención es efectiva en una etapa más tardía al abordar la desmotivación y desvinculación con el quehacer escolar, que se manifiestan más frecuentemente en niveles secundarios.

Respecto de la polémica capacidad de generar *mayores beneficios en alumnos desaventajados* y de *grupos minoritarios*, a pesar de los múltiples casos en que se verifica un mayor efecto para este grupo de alumnos y donde se observan disminuciones en la brecha de desempeño entre estudiantes de mejor y peor rendimiento, la evidencia no es concluyente. Si bien existen estudios que dan cuenta de beneficios diferenciados en favor de alumnos vulnerables, atribuyéndole a la medida una mayor equidad educativa, hay otros que encuentran que los más beneficiados son los estudiantes de mayor desempeño, lo que le da un carácter regresivo a la intervención. Asociado a lo anterior, la composición del curso resulta una variable fuertemente explicativa de los logros de aprendizaje cuando las reducciones en la cantidad de alumnos por curso no arrojan resultados concluyentes, lo que sugiere que el efecto par es significativo¹¹⁵.

En relación al efecto *acumulativo* de los beneficios y la *perdurabilidad* de éstos en el tiempo, las investigaciones evidencian la insuficiencia de un año de tratamiento, como originalmente fue planteado. Más bien hay un consenso en torno a que son necesarios tres o cuatro años de intervención para que estos mayores logros de aprendizaje incrementales, sean persistentes y mejoren los resultados de largo plazo como: finalizar la educación secundaria, iniciar y completar estudios terciarios, acceder a mejores condiciones laborales y obtener mayores ingresos y bienestar general. Es decir, la medida favorece la disminución de indicadores como ausentismo, deserción y violencia escolar y afecta las trayectorias de largo plazo en aspectos como estabilidad laboral, comportamiento cívico y calidad de vida.

La investigación más reciente ha permitido sobre todo entender que la relación entre el tamaño del curso y los logros de aprendizaje no es directa y está *mediada por procesos de aula intervinientes*. Los principales mecanismos mediadores dicen relación con la *maximización del*

¹¹⁵ No hay que confundir la identificación de variables que expliquen mayores logros de aprendizaje con medidas que los puedan favorecer a nivel global. El sistema educativo chileno ya exhibe un grado de segregación importante. Además, se ha podido comprobar a nivel internacional y en Chile en particular, que la diversidad social en el aula tiene un efecto positivo sobre el aprendizaje (Taut & Escobar, 2012).

tiempo de instrucción y la *individualización de la enseñanza*, al disponer el docente de más tiempo por estudiante. Producto de una reducción del tiempo dedicado a labores administrativas y de control de la disciplina, las interacciones profesor-alumno aumentan y con ello la frecuencia y calidad de la retroalimentación, facilitando la reflexión y la corrección del error. Una mejor conducta de los estudiantes favorece un clima de aprendizaje más adecuado, lo que junto a un mejor conocimiento y abordaje de las necesidades individuales de los estudiantes, y a una mayor influencia sobre alumnos desmotivados y disruptivos, explicarían los beneficios tanto cognitivos como no cognitivos asociados a la medida. Cursos reducidos favorecen la visibilidad de todos los alumnos y afianzan el sentido de pertenencia al grupo, lo que se traduce en un mayor compromiso hacia el aprendizaje de los estudiantes. Asimismo los docentes que trabajan con cursos más pequeños son menos rígidos y punitivos, permitiendo un amplio repertorio de conductas dentro del aula, lo que aumenta las estrategias pedagógicas disponibles, impactando tanto la satisfacción de alumnos como de profesores. Si bien la investigación no demuestra un cambio sistemático en las prácticas pedagógicas al disminuir el tamaño del curso, si da cuenta de que ciertas estrategias como la formulación de preguntas desafiantes y el trabajo en parejas y grupos pequeños, son más efectivas y viables de implementar en cursos reducidos, ya que éstas son altamente sensibles al tiempo disponible y a un clima de aprendizaje apropiado.

Central en la investigación del impacto del número de alumnos por curso ha sido la variable calidad docente, que al estar vinculada a muchos aspectos no observables, resulta compleja de caracterizar y medir. Algunos estudios plantean que independientemente de la calidad de los profesores la medida es beneficiosa en cuanto hace más asequible al profesor y desencadena los procesos anteriormente descritos. Otros sugieren que las habilidades y competencias docentes no se transmiten directamente a los alumnos, sino que son mediadas por los procesos pedagógicos y aspectos contextuales como el tamaño del curso, siendo factores independientemente importantes. Existen también trabajos que postulan que sólo docentes efectivos y experimentados son capaces de generar mayores logros de aprendizaje producto de reducciones del tamaño del curso. Por otro lado, investigaciones que evalúan la razón costo-beneficio de las reformas educativas, comparan la efectividad de mejorar la calidad docente con la reducción de alumnos por curso, concluyendo que la primera medida es más costo efectiva. Cabe señalar que estas comparaciones no contemplan aspectos no cognitivos que favorecen tanto mayores logros de aprendizaje como otros resultados educativos y de bienestar de largo plazo. Empíricamente se ha comprobado que países con un cuerpo docente más capacitado, con mayores remuneraciones y certificaciones, son menos sensibles a reformas de esta naturaleza, y que por el contrario, sistemas educativos menos efectivos se ven más beneficiados por reducciones del tamaño de los cursos.

Asimismo, la evidencia empírica que arroja la evaluación de programas de reducción de alumnos por curso implementados en distintos países del mundo ha entregado importantes aprendizajes respecto de su ejecución. Esta literatura advierte de potenciales externalidades que podrían surgir al implementar en forma masiva este tipo de programas, como el debilitamiento de la calidad docente al requerir nuevas contrataciones, que muchas veces no alcanzan la

certificación ni experiencia mínima necesaria; la redistribución de recursos en favor de los niveles beneficiados y en desmedro de los demás alumnos, ya sea en términos del tamaño de los cursos o del acceso a profesionales especialistas y recursos educativos.

En síntesis, gracias al avance de la investigación desde los distintos enfoques desarrollados, como al análisis del fenómeno desde una perspectiva cuantitativa y cualitativa, el impacto del número de alumnos por curso ha podido ser estimado e interpretado de forma cada vez más certera y multidimensional. La tendencia internacional es a un aumento de la implementación de programas de reducción del tamaño de los cursos, incluso en países asiáticos, lo que ha hecho posible recabar mayor evidencia de sus beneficios vía aquellos procesos de aula que impactan los aprendizajes de los alumnos. Así, este cúmulo de experiencias ha aportado a una mayor comprensión del impacto de esta medida en las habilidades de corto y largo plazo, tanto cognitivas como no cognitivas que favorecen los contextos de cursos reducidos.

7.1. Desafíos para la investigación en Chile

Como parte de las conclusiones expuestas se menciona la imposibilidad de establecer un efecto universal de una medida de esta naturaleza, altamente sensible al nivel de desarrollo de la nación que se analiza, a las características del sistema educativo donde se implementa, a la efectividad de sus docentes, a los logros educativos alcanzados previamente, a las características de los alumnos, al tamaño de la reducción y a la situación inicial, entre otros. Las escasas evaluaciones disponibles para Chile y su limitado alcance, no son suficientes para orientar la implementación de políticas de este tipo. Las investigaciones internacionales que cuentan con datos para Chile son alentadoras (Altinok & Kingdon, 2012), pero sin duda se requiere de una validación de sus hallazgos. De esta forma, Chile enfrenta enormes desafíos para el desarrollo de investigaciones a gran escala en esta materia al contar con un sistema educativo único, que como plantea Urquiola (2009), no cumple con los supuestos para utilizar estrategias de investigación basadas en diseños de regresión discontinua. Esto principalmente porque el modo de financiamiento incentiva a que los establecimientos retengan su matrícula en múltiplos del límite máximo de alumnos por curso, lo que imposibilita observar variaciones en el tamaño de los cursos, producto de cambios exógenos en la matrícula. Así, por el momento, el escenario se ve restringido a análisis de datos en contextos de variaciones naturales de la población y a experimentos aleatorios. Para el primer caso la confiabilidad y validez de sus resultados dependerá, entre otras cosas, de la disponibilidad de extensas y ricas bases de datos de panel que permitan controlar por distintas características tanto a nivel escuela, como de docentes y alumnos, aislando la variable de interés y estimando su efecto. Esta primera necesidad constituye un desafío para las autoridades del Ministerio de Educación y en particular a su Centro de Estudios, que están llamados a enriquecer y sistematizar la recolección de datos que sirva los propósitos de la investigación que se requiere para el desarrollo del país. Adicionalmente en el caso chileno, los sesgos de clasificación entre escuela e intra escuela se dan fuertemente, no sólo a nivel académico sino que principalmente a nivel socioeconómico y por medio (urbano y rural), lo que implica la incorporación de distintos enfoques metodológicos adicionales que corrijan la endogeneidad de la variación en el tamaño de los cursos. La educación compensatoria, que se manifiesta

principalmente mediante proyectos de integración educativos, es otro fenómeno presente en nuestro sistema, necesario de contemplar a la hora de estimar el efecto de reducciones de alumnos por curso.

Otra importante conclusión de este trabajo es que los cursos reducidos ofrecen mejores *oportunidades* de aprendizaje, pero son los docentes -impulsores de los cambios y transformaciones en el aula- los llamados a aprovecharlas. Sin embargo la responsabilidad de identificar, transmitir e incorporar en las didácticas aquellas estrategias diferenciadoras recae en la institución que los forma; de ahí la importancia de que las universidades se involucren desde sus programas de formación inicial docente. Igualmente importante para los resultados de aprendizaje mediados por intervenciones de esta naturaleza, resulta la distribución de los docentes a lo largo del sistema. La capacidad de generar incentivos y condiciones laborales atractivas que permitan retener a los profesores más preparados ejerciendo en los sectores más vulnerables de la sociedad es de suma importancia. Asumiendo que las universidades y el estado logran generar las condiciones descritas, será preciso -para poder captar el impacto de la medida en su totalidad- diseñar instrumentos más sofisticados que sean capaces de distinguir y recoger aquellos cambios de comportamiento, esfuerzo y participación que los cursos reducidos favorecen, para identificar las interrelaciones de efectos que se dan a nivel de aula y establecer eventuales vínculos con mayores logros de aprendizaje y otros resultados de largo plazo. No obstante, estas relaciones requieren de un marco conceptual que las modele, y como plantea Anderson hacia el año 2000, éste es un aspecto que precisa de un mayor desarrollo multidisciplinario que involucre aportes de educadores, sociólogos y psicólogos. Asimismo es importante nutrir la investigación con datos provenientes de diversas fuentes de información, y los estudios realizados demuestran que los alumnos, sobre todo de secundaria, son una fuente de información útil y confiable al aportar una mirada que refleja aspectos centrales de la educación como satisfacción hacia la escuela, compromiso y trayectoria educativa. Todo lo anteriormente planteado requiere de la acción coordinada de investigadores, universidades, centros de estudio, escuelas, profesores, apoderados y alumnos, lo que enriquece el proceso y consolida los hallazgos pero no está exento de dificultades. Por su parte, los estudios experimentales, recomendados para contextos como el chileno, conllevan importantes gastos y son de difícil implementación. Su ejecución debe abarcar un tiempo suficiente para desencadenar patrones y conductas propias de la condición experimental y debe garantizar la aleatoriedad del resto de las variables explicativas de logros de aprendizaje. Requiere de seguimientos y análisis sistemáticos que perduren en el tiempo. Estas características se alejan de los requisitos actualmente vigentes de los Fondos de Investigación y Desarrollo en Educación FONIDE que establecen una duración máxima de ejecución de proyectos insuficiente para esta realidad. Otras fuentes de financiamiento ligadas a Conicyt como Fondecyt y Fondap podrían ajustarse a las necesidades que plantea la investigación en torno al número de alumnos por curso, como área prioritaria a desarrollar.

Una alternativa que permitiría avanzar en comprender mejor el fenómeno contextualizado a la realidad chilena, y que se ajusta al proceso de reforma que vive el país, sería la implementación de programas de reducción del tamaño de los cursos como SAGE, con un diseño

cuasi-experimental a modo de experiencias piloto, representativas de las distintas realidades del país. Estas acciones permitirían recoger antecedentes e identificar dificultades tempranamente, con el fin de sistematizar procesos y establecer las condiciones ideales que garanticen el éxito de la medida, en un escenario en que ésta se amplía y generaliza al resto del sistema de forma gradual.

7.2. Recomendaciones para la implementación de reducciones de alumnos por curso

Recogiendo lo que describe la literatura y el aprendizaje de la implementación de programas de reducción de alumnos por curso en varios países del mundo, es recomendable para eventuales intervenciones tener en consideración las siguientes dimensiones del diseño y ejecución de la medida.

1. Implementación focalizada: la evidencia analizada sugiere que los beneficios de la medida son mayores para los estudiantes más vulnerables, por lo que las escuelas que atienden a una mayor proporción de alumnos prioritarios y de familias de escasos recursos debieran ser aquellas que adopten la medida primeramente. Contribuye así a una mayor equidad educativa que es un objetivo que esta medida persigue.

2. Nivel educativo: la intervención genera beneficios tanto a nivel inicial como secundario. A nivel inicial ayuda a sentar las bases para un desempeño escolar exitoso y permite advertir y remediar posibles dificultades de forma temprana. En una etapa más tardía los beneficios se asocian al desarrollo de habilidades no cognitivas en alumnos adolescentes, que además de favorecer mayores logros de aprendizaje, reduce la deserción, mejora aspectos de convivencia escolar, favorece el compromiso con la escuela e impacta resultados de largo plazo. En el caso de Chile podría facilitar un acceso más democrático y equitativo a la educación superior, tema altamente controvertido en la actualidad.

3. Docentes adicionales: dada la importancia del factor docente en los resultados de aprendizaje de los alumnos, es de suma importancia contar con profesores bien preparados y experimentados para hacerse cargo de los cursos reducidos. La medida ofrecería mejores condiciones laborales, haciendo más atractivo el ejercicio de la profesión en ese sector, sin embargo se requieren incentivos adicionales suficientes para que estos docentes permanezcan en estas escuelas. Es necesario prever y evitar que, producto del aumento de la demanda de profesores, se genere o profundice una asignación de los docentes más capaces y experimentados a sectores más acomodados, lo que deja a los alumnos más vulnerables con docentes principiantes, menos certificados y posiblemente afectos a una mayor rotación docente.

4. Infraestructura: considerando que una medida como esta requiere de mayor cantidad de salas de clase, se debe aprovechar al máximo la capacidad ociosa que exista en las escuelas públicas producto de la baja de matrícula sostenida que ha enfrentado el sector. Estos nuevos requerimientos deben contemplarse en conjunto con otros cambios estructurales que puedan suceder y que demanden modificaciones a la organización de los establecimientos. Para el caso chileno habría que contemplar el cambio de régimen a 6 años de Educación Básica y 6 años de Educación Media que establece Ley General de Educación (LGE) para 2018 y el eventual proceso

de desmunicipalización que requerirán reordenar y reestructurar la infraestructura disponible al nuevo contexto y planificar las nuevas instalaciones necesarias.

5. Desarrollo profesional: el desarrollo profesional docente, asociado a prácticas efectivas en cursos reducidos, es necesario para maximizar el impacto de la medida, tanto a nivel de formación del profesorado como para docentes en ejercicio. Las universidades a través de sus facultades de educación, están llamadas a incluir en sus programas de pedagogía estrategias de enseñanza para cursos de diferentes tamaños. Además contar con procesos de inducción de docentes principiantes más prolongados, en donde el aprendiz pueda observar y practicar diversas aproximaciones pedagógicas y prácticas de aula efectivas en distintos tipos de curso, favorece el impacto de la intervención.

6. Medidas complementarias: se deben fomentar medidas que tienen el potencial de favorecer programas de reducción de alumnos por curso como mayores sueldos, programas para alumnos prioritarios, proyectos de integración, escuela para padres y jornada escolar extendida, articulando sus metas y resguardando que su ejecución simultánea no entorpezca los objetivos de una y otra.

7. Consideraciones costo-beneficio: las decisiones deben basarse en análisis costo efectividad que comparen diversas medidas, pero esos análisis deben contemplar y valorar todos los beneficios de la intervención, esto es, aspectos cognitivos y no cognitivos, de corto y largo plazo que contribuyan a mejorar la calidad y equidad de la educación, y no sólo su impacto en logros de aprendizaje inmediatos. Valorar y cuantificar aspectos como disminuciones en la tasa de violencia escolar o una mayor participación ciudadana es un desafío, sin embargo son aspectos no cognitivos deseables que se asocian a reducciones de alumnos por curso.

8. Evaluación: como sugiere la evidencia, el éxito de esta medida va a depender en gran parte de las condiciones y circunstancias en que es aplicada. Documentar y evaluar extensamente la implementación y ejecución de programas de reducción de alumnos por curso en sistemas educativos nuevos a la intervención es vital para obtener la información de su funcionamiento en el contexto local, lo que permitirá identificar los aspectos más críticos, y así modificarlos y ajustarlos a la realidad de los establecimientos participantes, logrando obtener el mayor beneficio posible de la medida.

9. Seguimiento: para informar y estimar adecuadamente los beneficios de largo plazo de la medida se hace necesario realizar un seguimiento a los alumnos que fueron favorecidos con la intervención. Indicadores como rendimiento escolar y resultados educativos de mediano plazo además de aspectos no cognitivos como comportamiento cívico y otros resultados de largo plazo como estabilidad laboral, ingresos y bienestar, deben ser recogidos y analizados. Para el caso chileno habría que contemplar la finalización de la educación secundaria, rendición y puntaje PSU, ingreso y completación de educación terciaria, entre otros.

VIII. APÉNDICE: Dirección y Liderazgo Escolar

La orientación que se le dio a este trabajo dejó fuera un análisis de los aspectos relacionados a la dirección y al liderazgo escolar relativos a la reducción del número de alumnos por curso de forma directa. Sin embargo, y atendiendo a que la mención asociada al grado de magíster al que se busca optar mediante esta tesis corresponde a Dirección y Liderazgo Educacional, se ha incluido una pequeña sección que contextualiza su vínculo y que sintetiza algunos hallazgos. Además se presentan referencias bibliográficas y se plantean ciertas interrogantes que pueden ser relevantes para el contexto chileno.

La literatura disponible ha descrito las condiciones y circunstancias económicas y políticas bajo las cuales se diseñan e implementan programas de reducción de alumnos por curso. En general es una decisión que se toma a nivel país, producto de acuerdos políticos entre los diferentes actores involucrados, cuando otras necesidades más básicas ya se han superado.

Los directores de establecimientos educativos son agentes que comúnmente enfrentan estas medidas ya resueltas. Sin embargo juegan un rol clave en el éxito de la reforma, en cuanto inciden en su eficaz implementación y ejecución. El liderazgo del director, su nivel de involucramiento en la intervención y el desarrollo profesional del cuerpo docente aparecen como los aspectos más importantes con los que contribuye el líder en la consecución de los propósitos de la intervención¹¹⁶:

- a) Liderazgo directivo: los directores muestran patrones de conducta de acuerdo a nivel de logros de sus escuelas; algunos directores exhiben respuestas proactivas que responden a desafíos asociados a metas de desempeño, mientras que otros son reacios a iniciar cambios mayores o unificar programas, favoreciendo el status quo. Resulta esencial contar con un equipo directivo proactivo, con altas expectativas y con grados razonables de autonomía para implementar la medida eficazmente. Igualmente importante resulta la articulación de otras medidas en ejecución con la implementación de reducciones del tamaño del curso.
- b) Involucramiento: la visión y las creencias de los líderes afectan el modo de entender y actuar en torno al cambio escolar. Líderes afines a los propósitos y objetivos de la medida determina en parte el éxito de ésta. De esta forma los directores pueden beneficiarse de un desarrollo profesional que aborde las motivaciones subyacentes de su conducta, incluyendo sus actitudes hacia el cambio y su filosofía educativa.
- c) Desarrollo profesional: la necesidad de un perfeccionamiento profesional docente, que considere distintos niveles de habilidades y que intencione el desarrollo de estrategias pedagógicas efectivas en contextos de aulas reducidas debe ser liderado

¹¹⁶ Mayor detalle sobre dirección y liderazgo escolar en la implementación y ejecución de programas de reducción de alumnos por curso se encuentra en (Bascia, 2010; Burch, Theoharis & Rauscher, 2010; Flessa, 2012 y Galton & Pell, 2009).

por un director capacitado e informado, que apoye la adopción de nuevos métodos de enseñanza y fomente la colaboración entre colegas.

Se abren ciertos temas y preguntas para un futuro debate en relación a políticas y decisiones a nivel de dirección y liderazgo escolar:

- ¿Cuál es el nivel de participación que deben tener los líderes escolares en el diseño e implementación de una medida de esta naturaleza?
- ¿Cuáles son las competencias clave que requiere un director para liderar una intervención como esta de forma exitosa?
- ¿La visión y filosofía educativa de nuestros actuales directores es favorable a estas medidas?
- ¿De qué manera podrían contribuir los directores por medio del desarrollo profesional de sus docentes a la implementación exitosa de una medida como esta, dadas las actuales dificultades que experimenta la transferencia de competencias adquiridas en programas de perfeccionamiento y capacitación al aula?

Bibliografía

- Achilles, C. M. (2003). How class size makes a difference: What the research says. The impact of class-size reduction (CSR). New York. ERIC Document Reproduction Service No. ED475012.
- Akabayashi, H., & Nakamura, R. (2012). Can Small Class Policy close the gap? An empirical analysis of class size effects in Japan. Working Paper E(51), Tokyo Center for Economic Research.
- Altinok, N. , & Kingdon, G. (2012). New evidence on class size effects: A pupil fixed effects approach. *Oxford Bulletin of Economics and Statistics*, Department of Economics, University of Oxford, vol. 74(2), pages 203-234, 04.
- Anderson, L. (2000). Why should reduced class size lead to increased student achievement? In M.C. Wang and J.D. Finn. eds., *How small classes help teachers do their best* (pp. 3 – 24). Philadelphia, PA: Temple University Center for Research in Human Development and Education.
- Angrist, J., & Lavy, V. (1999). Using Maimonides' Rule To Estimate the Effect of Class Size on Scholastic Achievement. *The Quarterly Journal of Economics*, 114 (2): 533-575.
- Arum, R., & LaFree, G. (2008). Educational Attainment, Teacher–Student Ratios, and the Risk of Adult Incarceration Among U.S. Birth Cohorts Since 1910. *Sociology of Education* 81: 397.
- Averett, S., & McLennan, M. (2004). Exploring the Effect of Class Size on Student Achievement: What Have We Learned Over the Past Two Decades? *International Handbook on the Economics of Education*, Geraint Johnes and Jill Johnes, eds. Edward Elgar Publishers.
- Babcock, P., & Betts, J. (2009). Reduced-class distinctions: Effort, ability, and the education production function. *Journal of Urban Economics*, 65, 314–22
- Bascia, N. (2010). Ontario's Primary Class Size Initiative: Report on Early Implementation. Toronto: Canadian Education Association, 1-121.
- Betts, J., & Shkolnik, J. (1999). The behavioral effects of variations in class size: The case of Math Teachers. *Educational Evaluation and Policy Analysis*, Vol. 21, No. 2, pp. 193-213.
- Bilagher, M., & Holz, M. (2012). Class Size: Can Demographic Transition Help Quality Of Education In Latin America And The Caribbean? Evidence From Serce. UNESCO Regional Office for Education in Latin America and the Caribbean. Presentado en ICSEI 2013 "The 26th International Congress for School Effectiveness and Improvement".
- Blatchford, P. (2011). Three generations of research on class size effects. In K.R. Harris, S. Graham & T. Urdan (Eds.), *APA Educational psychology handbook (Vol. 2: Individual differences and cultural and contextual factors, pp. 529 - 554)*. Washington, DC: American Psychological Association.
- Blatchford, P., Bassett, P., & Brown, P. (2011). Examining the effect of class size on classroom engagement and teacher-pupil interaction: Differences in relation to pupil prior attainment and primary vs. secondary schools. *Learning and Instruction*, 21(6), 715-730.
- Blatchford, P., Edmonds, S., & Martin, C. (2003). Class size, pupil attentiveness and peer relations. *British Journal of Educational Psychology*, 73, 15–36.
- Bonesrønning, H. (2003) Class size effects on student achievement in Norway: Patterns and explanations. *Southern Economic Journal*. volum 69 (4).
- Booser, M., & Rouse, C. (2001). Intrascchool Variation in Class Size: Patterns and Implications. *Journal of Urban Economics*, 50(1), 163-189.
- Borland, M., Howsen, R., & Trawick, M. (2005). An investigation of the effect of class size on student academic achievement. *Education Economics* Vol. 13, No. 1, 73–83.
- Bosworth, R. (2011). Class size, class composition, and the distribution of student achievement. *Education Economics*, 1-25.

- Breton, T. R. (2012). Evidence that Class Size Matters in 4th Grade Mathematics: An Analysis of TIMSS 2007 Data for Colombia. *International Journal of Educational Development*, 34, 51-57.
- Brewer, D. J., & McEwan, P. J. (Eds.). (2010). Economics of education: *Tiebout Sorting and Competition*. Amsterdam: Elsevier.
- Browning, M., & Heinesen, E. (2007). Class Size, Teacher Hours and Educational Attainment. *The Scandinavian Journal of Economics*, 109(2), 415-438.
- Brühwiler, C., & Blatchford, P. (2011). Effects of class size and adaptive teaching competency on classroom processes and academic outcome. *Learning and Instruction*, 21(1), 95-108.
- Burch, P., Theoharis, G., & Rauscher, E. (2010). Class Size Reduction in Practice Investigating the Influence of the Elementary School Principal. *Educational Policy*, 24(2), 330-358.
- Case, A., & Deaton, A. (1999). School inputs and educational outcomes in South Africa. *Quarterly Journal of Economics* 114: 1047-1084.
- Chile de Todos. (2014) Programa de Gobierno Michelle Bachelet 2014-2018; recuperado de: <http://michellebachelet.cl/michelle-bachelet-presenta-programa-de-gobierno-con-enfasis-en-una-reforma-estructural-la-educacion/>.
- Chingos, M. M. (2012). The impact of a universal class size reduction policy: Evidence from Florida's statewide mandate. *Economics of Education Review*, 31 (5), 543-562.
- Cho, H., Glewwe, P., & Whitler, M. (2012). Do reductions in class size raise students' test scores? Evidence from population variation in Minnesota's elementary schools. *Economics of Education Review*, 31(3), 77-95.
- Cohen, G., Miller, C., Stonehill, R., & Geddes, C. (2000). The Class-Size Reduction Program: Boosting Student Achievement in Schools Across the Nation: U.S. Department of Education.
- Cornejo, C. R. (2009). Condiciones de trabajo y bienestar/malestar docente en profesores de enseñanza media en Santiago de Chile. *Educación y Sociedad*, vol. 30, n. 107, 409-426.
- Cortes, K., Moussa, W., & Weinstein, J. (2013). Making the Grade: The impact on classroom behaviour on academic achievement. Texas A&M University, Syracuse University.
- Dee, T. S., & West, M. R. (2011). The Non-Cognitive Returns to Class Size. *Educational Evaluation and Policy Analysis*, 33(1), 23-46.
- Ding, W., & Lehrer, S. (2011). Experimental estimates of the impacts of class size on test scores robustness and heterogeneity. *Education Economics*, 19(3), 229-252.
- Dobbelsteen, S., Levin, J., & Oosterbeek, H. (2002). The causal effect of class size on scholastic achievement: distinguishing the pure class size effect from the effect of changes in class composition. *Oxford Bulletin of Economics and Statistics*, 64(17), pp.17-38.
- Dynarski, S., Hyman, J., & Schanzenbach, D. W. (2011). Experimental Evidence on the Effect of Childhood Investments on Postsecondary Attainment and Degree Completion. Working Paper No. 17533, National Bureau of Economic Research.
- Educational Leaders. (2011). School support staff: collectively making resources count. Recuperado de <http://www.educationalleaders.govt.nz/Managing-your-school/School-support-staff>.
- Ehrenberg, R. G., Brewer, D. J., Gamoran, A., & Willms, J. D. (2001). Class size and student achievement. *Psychological Science and the Public Interest*, 2(1), 1-30.
- Finn, J. (2002). Class Size Reduction in Grades K-3. In A. Molnar (Ed.), *School Reform Proposals: The Research Evidence*. United States: School Improvement Programs.
- Finn, J., Gerber, S., & Boyd-Zaharias, J. (2005). Small Classes in the Early Grades, Academic Achievement, and Graduating From High School. *Journal of Educational Psychology*, 97, No. 2, 214-223.

- Finn, J. D., & Achilles, C. M. (1990). Answers and Questions about Class Size - A Statewide Experiment. *American Educational Research Journal*, 27(3), 557-577.
- Finn, J. D., & Achilles, C. M. (1999). Tennessee's class size study: findings, implication, misconceptions. *Educational Evaluation and Policy Analysis*, 21(2), 97-109.
- Finn, J. D., Gerber, S. B., Achilles, C. M., & Boyd-Zaharias, J. (2001). The Enduring Effects of Small Classes. *Teachers College Record*, 103(2), 145-183.
- Finn, J. D., Pannozzo, G. M., & Achilles, C. M. (2003). The "Why's" of Class Size: Student Behavior in Small Classes: *Review of Educational Research*, 73, 321-368.
- Flessa, J. J. (2012). Principals as Middle Managers: School Leadership During the Implementation of Primary Class Size Reduction Policy in Ontario. *Leadership and Policy in Schools*, 11, 325-343.
- Fredriksson, P., Öckert, B., & Oosterbeek, H. (2012). Long Term effects of class size. *The Quarterly Journal of Economics*, Oxford University Press, vol. 128(1), pages 249-285.
- Galton, M., & Macbeath, J. (2008). Teachers under pressure? London, England: SAGE Publications Ltd.
- Galton, M., & Pell, T. (2009). Study on Small Class Teaching in Primary Schools in Hong Kong: Final Report. Hong Kong: Education Bureau and Cambridge University.
- Glass, G. V., & Smith, M. L. (1978). Meta-Analysis of Research on the Relationship of Class-Size and Achievement. *Educational Evaluation and Policy Analysis*, 21(2), 143-163.
- Graue, E., Hatch, K., Rao, K., & Oen, D. (2007). The Wisdom of Class-Size Reduction. *American Educational Research Journal*, 44, 670-700.
- Graue, E., Rauscher, E., & Sherfinski, M. (2009). The Synergy of Class Size Reduction and Classroom Quality. *Elementary School Journal*, 110(2), 178-201.
- Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research*, 66(3), 361-396.
- Gutiérrez, N., Castro, A., & Segure, T. (2000). El número de alumnos por curso y su relación con la calidad de los procesos pedagógicos. *Docencia*, 11, 37-44.
- Hanushek, E. (1999a). Some findings from an independent investigation of the Tennessee STAR experiment and from other investigations of class size effects. *Educational Evaluation and Policy Analysis*, 21, No. 2, 143-163.
- Hanushek, E. (1999b). The evidence on class size. In *Earning and learning: How schools matter*. Susan E. Mayer y Paul E. Peterson (eds). Washington, DC: Brookings Institution: 131-168.
- Hanushek, E., & Woessmann, L. (2010). The Economics of International differences in Educational Achievement. NBER Working Paper 15949. Cambridge, Mass.: National Bureau of Economic Research.
- Harfitt, G. (2013). Why 'small' can be better: an exploration of the relationships between class size and pedagogical practices. *Research Papers in Education*, 28 n. 3, 330-345.
- Harfitt, G. J. (2012a). Class size and language learning in Hong Kong: the students' perspective. *Educational Research*, 54(3), 331-342.
- Harfitt, G. J. (2012b). How class size reduction mediates secondary students' learning: hearing the pupil voice. *Asia Pacific Education Review*, 13(2), 299-310.
- Hoxby, C. (2000). The effects of class size on student achievement: New evidence from population variation. *Quarterly Journal of Economics*, 115(4), 1239-1285.
- J., R., Kangas, M., & Ruokamo, H. (2008). Predictors of Dutch and Finnish Children's Satisfaction with Schooling. *Journal of Happiness Studies*, 11, 193-204.
- Jakobsson, N., Persson, M., & Svensson, M. (2013). Class-Size Effects on Adolescents' Mental Health and Well-Being in Swedish Schools. *Education Economics*, 21(3), 248-263.

- Jepsen, C., & Rivkin, S. (2009). Class Size Reduction and Student Achievement: The Potential Tradeoff between Teacher Quality and Class Size. *Journal of Human Resources*, 44, 1.
- Krueger, A. B. (2003). Economic considerations and class size. *Economic Journal*, 113, 34–63.
- Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*, 114(2), 497–532.
- Krueger, A. B., & Whitmore, D. M. (2001). The Effect of Attending a Small Class in the Early Grades on College-test Taking and Middle School Test Results: Evidence from Project Star. *Economic Journal*, 111, 1–28.
- Lazear, E. P. (2001). Educational production. *Quarterly Journal of Economics*, 116(3), 777–803.
- Leuven, E., Oosterbeek, H., & Rønning, M. (2008). Quasi-experimental estimates of the effect of class size on achievement in Norway. *Scandinavian Journal of Economics*, 110(4), 663–693.
- Levin, J. (2001). For whom the reductions count: A quantile regression analysis of class size and peer effects on scholastic achievement. *Empirical Economics* 26, 221–246.
- Lévy-Garboua, L, Lohéac, Y, Fayolle, B (2006) Preference formation, school dissatisfaction and risky behavior of adolescents. *Journal of Economic Psychology* 27(1): 165-183.
- Lopez, S. (2009). Well-Being, Success, and the Gallup Student Poll: Gallup, Inc.
- Medina, C., & Bríñez, O. (2007). The Effects of Class Size on Quality of Education in Colombia. Paper presented at the International conference *The Quality of Education in Latin America and the Caribbean*, Mexico.
- Metlife, I. (2012). The Metlife Survey of the American Teacher: Teachers, Parents and the Economy: Metlife, Inc.
- MINEDUC. (2012). Serie Evidencias. Centro de Estudios Mineduc (Ed.), Hacia la medición del costo de una educación de calidad: resultados preliminares.
- Minzyuk, L., & Russo, F. (2012). The Causal Effect of Class Size on Pupils' Performance: Evidence from Italian Primary Schools. Paper presentado en la 34th Conferencia de la Società italiana di economia pubblica, Pavia, 24-25 September.
- Mishel, L., & Rothstein, R. (eds). (2002). The Class Size Debate. Economic Policy Institute, Washington, DC.
- Molnar, A., Smith, P., & Zahorik, J. (1999). 1998–99 evaluation results of the Student Achievement Guarantee in Education (SAGE) program. Milwaukee: University of Wisconsin, School of Education.
- Molnar, A., Smith, P., Zahorik, J., Palmer, A., Halbach, A. & Ehrle, K. (1999). Evaluating the SAGE program: A pilot program in targeted pupil-teacher reduction in Wisconsin. *Educational Evaluation and Policy Analysis*, 21, 165–177.
- Mueller, S. (2013). Teacher experience and the class size effect — Experimental evidence. *Journal of Public Economics*, 98, 44–52.
- Muralidharan, K., & Sundararama, V. (2010). Contract Teachers: Experimental Evidence from India. NBER Working Paper No. 19440.
- Myers, J. (2004). La relación entre el número de alumnos por curso y el rendimiento en la prueba de Matemáticas del SIMCE 2000. *Docencia*, N°23, 14-19.
- NEA. (2013). Maryland Paraeducator Named NEA 2013 ESP of Year. Recuperado de <http://www.nea.org/home/54915.htm>
- Nye, B., Hedges, L., & Konstantopoulos, S. (2002). Do Low-Achieving Students Benefit More from Small Classes? Evidence from the Tennessee Class Size Experiment. *Educational Evaluation and Policy Analysis*, 24, No. 3, 201–217.
- Nye, B., Hedges, L. V., & Konstantopoulos, S. (2004). Do Minorities Experience Larger Lasting Benefits From Small Classes? Evidence from a five-year follow-up of the Tennessee class size experiment. *Journal of Educational Research*, 98, 94-100.

- OCDE. (2009). Los docentes son importantes. Política de educación y formación. Atraer, formar y conservar a los docentes eficientes: OCDE.
- OCDE. (2012). The Experience of New Teachers: Results from TALIS 2008, OECD Publishing.
- OCDE. (2013). Education at a Glance 2013: OECD Indicators: OECD Publishing.
- OCDE. (2014). Talis 2013 Results: An International Perspective on Teaching and Learning, OECD Publishing.
- Pas, E. (2007). The Effect of Class Size Reduction on Teacher Attrition and Recruitment: Evidence from Class Size Reduction Policies in New York State. Syracuse University.
- Pedder, D. (2006). Are small classes better? Understanding relationships between class size, classroom processes and pupils' learning. *Oxford Review of Education*, 32, No. 2, 213–234.
- Teach Plus. (2012). Great expectations: Teachers' views on elevating the teaching profession. Teach Plus Great Expectations.pdf. Recuperado de <http://www.teachplus.org>.
- Rivkin, S., Hanushek, E., & Kain, J. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, 73, No. 2, 417–458.
- Shapson, S., Wright, E., Eason, G., & Fitzgerald, J. (1980). An Experimental Study of the Effects of Class Size. *American Educational Research Journal*, 17(2), 141-152.
- Sharp, M. A. (2002). An analysis of pupil-teacher ratio and class size. EdD Dissertation. Eastern Michigan University.
- Shin, Y., & Raudenbush, S. (2011). The Causal Effect of Class Size on Academic Achievement: Multivariate Instrumental Variable Estimators With Data Missing at Random. *Journal of Educational and Behavioral Statistics*, Vol. 36, No. 2, pp. 154–185.
- Sims, D. R. (2009). Crowding Peter to educate Paul: Lessons from a class size reduction externality. *Economics of Education Review*, 28(4), 465-473.
- Sojourner, A. (2013). Identification of Peer Effects with Missing Peer Data: Evidence from Project STAR. *The Economic Journal*, 123: 574–605.
- Sortkær, B., & Reimer, D. (2014). Disciplinary climate and student achievement: Evidence from schools and classrooms. Paper presentado en la Conferencia de Psicología Danesa "Social stratification and inequality – changing social dynamics in the 21st century". Copenhagen 23-24 January, 2014.
- Sullivan, C. C. (2012). A Case Study of the Impact of Class Size on Teacher Efficacy. North Carolina State University.
- Taut, S., & Escobar, J. (2012). El efecto de las características de los pares en el aprendizaje de estudiantes chilenos de enseñanza media. Infomes Técnicos Mide UC. Centro de Medición MIDE UC, Pontificia Universidad Católica de Chile.
- Thijs, J., Verkuyten, M., & Helmond, P. (2010). A Further Examination of the Big-Fish-Little-Pond Effect: Perceived Position in Class, Class Size, and Gender Comparison: *Sociology of Education*, 83: 333.
- Urquiola, M. (2006). Identifying Class Size Effects in Developing Countries Evidence from Rural Bolivia. *Review of Economics and Statistics*, 88(1), 171-177.
- Urquiola, M., & Verhoogen, E. (2009). Class-Size Caps, Sorting, and the Regression-Discontinuity Design. *American Economic Review*, 99(1), 179-215.
- Wössmann, L., & West, M. (2006a). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review*, 50(3), 695-736.
- Wössmann, L., & West, M. (2006b). Which school systems sort weaker students into smaller classes? International evidence. *European Journal of Political Economy*, 22 (4), 944-968.
- Yeh, S. S. (2009). Class size reduction or rapid formative assessment? A comparison of cost-effectiveness. *Educational Research Review*, 4(1), 7-15.