

## MÉTODOS ESTADÍSTICOS DE ESTIMACIÓN DE LOS EFECTOS DE LA ESCUELA Y SU APLICACIÓN AL ESTUDIO DE LAS ESCUELAS EFICACES<sup>1 2</sup>

Tabaré Fernández Aguerre

### 1. PRESENTACIÓN

Los estudios sobre “escuelas eficaces” (“*effective schools*”), la investigación sobre los “efectos de la escuela” (“*school effect research*”) y los programas de mejoramiento de la escuela (“*school improvement*”) tienen entre sus puntos en común, la identificación empírica de un conjunto de escuelas que según criterios sustantivos se destacan en los logros cognitivos de sus alumnos<sup>3</sup>. En la bibliografía especializada, tanto en los países anglosajones donde son originarias estas perspectivas como también en América Latina, suele ser de gran interés el discutir en profundidad el perfil organizacional y pedagógico de estas escuelas, las posibilidades de generalizar esas experiencias exitosas o las diferencias observadas en el análisis comparado (Richards, 1991; Báez de la Fe, 1994; Reynolds *et al.*, 1996; Reynolds *et al.*, 1998; Martin *et al.*, 2000; LLECE 2002; Zorrilla, 2001; Murillo, 2003). Por su parte, los trabajos críticos sobre el conocimiento aportado por este paradigma suelen apuntar reiteradamente sobre distintos tipos de “debilidades atribuidas”: i) el vínculo con las políticas implementadas sobre aquellos fundamentos; ii) los compromisos ideológicos subyacentes; iii) el marco de teoría social; o iv) las decisiones metodológicas implicadas en el diseño o el en análisis (Rowan, 1983; Gibson y Asthana, 1998; Coe y Fitz-Gibon, 1998; Willmott, 1999a, 1999b)<sup>4</sup>.

En otro lugar (Fernández, 1999a, 2001, 2004), he abordado con cierto detalle algunos problemas relativos a la “escasez de teoría general” en este campo que han conducido a un cierto agotamiento en la “fertilidad” de la perspectiva de las escuelas eficaces y propuse algunas posibles soluciones también de teoría general que podrían contribuir a una nueva generación de estudios. Por tanto, no volveré aquí sobre esos temas. En este artículo me interesa concentrarme en particular sobre uno de estos aspectos problemáticos: los métodos estadísticos adoptados para estimar los efectos que una escuela tiene sobre los aprendizajes cognitivos de sus alumnos. Dado que estos términos pueden resultar equívocos, será

---

<sup>1</sup> Este trabajo se inscribe en el marco de la investigación de tesis doctoral que el autor lleva adelante en el Programa de Doctorado en Ciencia Social con especialidad en Sociología impartido en El Colegio de México, bajo la tutoría del Dr. Fernando Cortés Cáceres.

<sup>2</sup> Una primera y más general versión de este trabajo conformó el capítulo 2 del informe de investigación presentado al Instituto Nacional para la Evaluación de la Educación (INEE) de México titulado “Escuelas primarias eficaces en México”.

<sup>3</sup> Algunos autores sostienen que estas tres perspectivas si bien se originaron en preocupaciones diferentes, han tendido en los últimos años a fusionarse en un mismo campo de preocupaciones, por ejemplo, Bert Creemers (1998). Otros autores, en cambio tienen la postura de que se trata de perspectivas distintas, en particular en lo que hace a las “escuelas eficaces” y “la investigación de los efectos de la escuela”; esta es la opinión por ejemplo, de Valerie E. Lee (2001). A los efectos de este trabajo y dado que el punto de interés es otro, las tres perspectivas serán consideradas bajo el mismo rótulo, haciéndose uso del término “paradigma” de investigación (lo cual no implica desconocer los problemas epistémicos que esta decisión representa).

<sup>4</sup> Es importante hacer notar que en los últimos tres o cuatro años ha ido creciendo una muy sólida crítica a estos estudios que parten desde los estudios experimentales y ya no desde la econometría.

necesario precisarlos. Entiendo que se trata de “métodos especiales” en la medida en que constituyen procedimientos para tratar un conjunto de problemas (Bunge, 2000:7). Los métodos aquí tratados suponen la adopción (y opción) de un concepto de partida (el de “escuela eficaz”), requieren de determinadas decisiones relativas a su observación, y finalmente se formalizan en modelos estadísticos apropiados para cuantificar la magnitud que tienen la diferencia de efectos entre grupos. Estos métodos los califico de “estadísticos” con el propósito de resaltar una clase particular de problemas: la formalización de una conceptualización teórica. Esta formalización recurre a modelos de estimación cuyas propiedades han sido bastante estudiadas y discutidas: los modelos de regresión lineal clásico; de regresión ponderada y jerárquico-lineales<sup>5</sup>.

Los debates metodológicos y estadísticos suelen por lo general ocupar espacios más reducidos; implican discusiones técnicas a veces excesivamente matematizadas y por lo tanto terminan por interesar fundamentalmente a especialistas. Mi propósito a acometer en este trabajo una discusión metodológica y estadística es resaltar la importancia que tienen en la investigación varios de los problemas técnicos discutidos. Estoy firmemente convencido de que una discusión sobre estos problemas permitirá que los investigadores que estamos dedicados al estudio de este tema podamos consolidar esa nueva generación de estudios sobre efectos de la escuela que está emergiendo en nuestros países y que se está planteando cuestiones realmente cruciales cuando se trata de afrontar los graves problemas de desigualdad educativa que heredamos.

El artículo se organiza en seis apartados, además de esta introducción, procurando circunscribir en cada . En el segundo apartado, se trabajan cinco nociones que al menos pueden ser identificadas en las investigaciones sobre las escuelas eficaces: i) “eficacia absoluta”; ii) “la eficacia incremental”; iii) “eficacia relativa”, iv) el “enfoque combinado de Edmonds” y iv) las “nociones longitudinales”. Se discuten algunas críticas que se han presentado a cada una. Si bien no existe *un único* concepto “pacíficamente aceptado” por todos los investigadores, es preciso reconocer que el que aquí se presenta bajo el rótulo de “eficacia relativa” goza de una mayor aceptación. A continuación, en el tercer apartado, se abordan dos decisiones metodológicas del diseño que repercuten fuertemente en la elección del método de estimación: i) la elección de la unidad de análisis apropiada al concepto sustentado; y ii) la selección de las variables sociofamiliares que fungirán como controles (en el contexto de los conceptos de eficacia numerados más arriba como iii a v). El cuarto y quinto apartados abordan problemas específicamente estadísticos implicados en la correcta identificación de las escuelas eficaces y con la estabilidad de las identificaciones realizadas a lo largo del tiempo. Aquí me concentraré en los modelos estimados sobre mínimos cuadrados ordinarios y ponderados y los modelos multinivel o jerárquico-lineales. Se señalan tanto las potencialidades como las principales limitaciones de cada uno, apoyándome (principal pero no exclusivamente) en la bibliografía que en entre los expertos del mundo anglosajón aparece como de referencia obligada: desde la ponencia seminal de Aitkin y Longford leída ante la *Royal Statistical Society* en noviembre de 1985 y publicada en 1986 hasta los textos más recientes de Stephen Raudenbush y Douglas Willms (1995); de Harvey Goldstein (1997, con Woodhouse 2000) y de Stephen Raudenbush y Anthony Bryk (2002). Dado que la ausencia de control experimental en los estudios sobre efectos de la

---

<sup>5</sup> El lenguaje de la metodología no es necesariamente coincidente con el lenguaje de la estadística, una cuestión que puede molestar a quienes gustan del purismo terminológico. Por ejemplo, en la bibliografía econométrica el término de modelo tiene una extensión mayor que el término de “método”. Por el primero, se entiende tanto el “modelo clásico de regresión lineal” como el “modelo keynesiano de determinación del ingreso”. Por el segundo término, en cambio se refiere al “método de mínimos cuadrados ordinarios”, el “método de mínimos cuadrados ponderados”, el “método de máximo verosímil”. Véase al respecto, los índices analíticos en Bunge (2000), Greene (1999), Gujarati (2004) y García Ferrando, Ibáñez y Alvira (2000).

escuela suele invocarse para fundamentar el rechazo a usar estos métodos como instrumentos de una política de “responsabilización” de las escuelas (“*accountability*”), se dedica al apartado sexto al tratamiento de este problema, pasando revista a cuáles son los requisitos de diseño que se deberían cumplir a los efectos de aceptar las inferencias que se realicen con este tipo de investigaciones. En las conclusiones se revisan los principales puntos revisados y se plantean algunas sugerencias para la investigación futura que podría realizarse en la región de aceptarse los puntos aquí tratados.

Como corolario necesario dejar asentado que todos estos métodos tienen su base en modelos estadísticos del análisis de evaluaciones externas del aprendizaje cognitivo de los alumnos, aunque no existe ningún impedimento para extenderlos a otros indicadores de resultado tales como el comportamiento, la asistencia, la repetición o las trayectorias educacionales. Un buen ejemplo de esta extensión se puede encontrar en Mortimore et al (1988).

## 2. CONCEPTOS DE ESCUELA EFICAZ.

Los métodos estadísticos de estimación de los efectos de la escuela están condicionados por las conceptualizaciones de escuela eficaz de la que se partan. Suponiendo que se acepta la premisa de que los niveles teóricos, metodológicos y estadísticos están sub-ordinados en una jerarquía epistemológica, donde las decisiones en el nivel de la teoría supra-ordenan las modelizaciones y otras decisiones estadísticas (Bayce, 1983; Cortés y Rubalcaba, 1987; García, 2002), es necesario reconocer que el primer paso en la revisión de los métodos de estimación deberá hacerse mediante el análisis de la congruencia entre la noción teórica que se asienta y los requisitos que se le plantean a la observación.

Ahora bien, el problema es que una revisión de la extensa bibliografía sobre eficacia escolar muestra que no existe una única conceptualización sobre qué es una “escuela eficaz” ni menos aún una operacionalización que haya ganado el total consenso de la academia. Sin perjuicio de ello, se puede afirmar que algunas definiciones teórico-metodológicas han sido progresivamente desechadas en tanto que otras cuentan con sólidas argumentaciones y, por tanto, se han ido imponiendo en la investigación más reciente. Aquí mostraremos cinco distintas: a) la eficacia absoluta; b) las nociones de incrementalidad; c) la eficacia relativa; d) el enfoque combinado de la calidad y la equidad; y e) la noción de eficacia fundada en datos longitudinales.

### 2.a. La noción absoluta de eficacia.

La primera noción que fuera utilizada tanto en la investigación como en la discusión pública indicaba sencillamente que una escuela eficaz era aquella que obtenía los más altos resultados académicos promedio entre sus alumnos. Si nombramos como  $E$  al conjunto integrado por toda  $j$ -ésima escuela cuyo promedio de aprendizajes  $\mu_j$  es mayor que cierto valor crítico, denominado  $Z^*$ , formalmente tendremos:

$$[1] \quad J \in E \leftrightarrow \mu_J > Z^*$$

El problema básico de estimación para este concepto de eficacia es identificar el valor empírico de  $Z^*$ . Este valor entendido como umbral crítico, no requiere en principio de ningún modelo estadístico.

Teóricamente, un método de estimación podría consistir en la definición política de un estándar mínimo de conocimientos que los alumnos deberían alcanzar en determinado grado de su escolarización<sup>6</sup>.

Sin embargo, el método más comúnmente empleado puede catalogarse como contrario al concepto mismo. Parte por confeccionar (y por lo general publicar) un “listado” de escuelas según el puntaje promedio que sus alumnos obtuvieron en una o varias pruebas y a veces ordenadas también según un promedio de promedios de pruebas. Luego se establece un punto de corte definiendo que aquellas escuelas que se encuentran en los primeros 5, 10 o 50 lugares serán definidas como “eficaces”, “destacables”, “excepcionales”.

Este método es bastante “popular” en la medida en que a una familia le permite directamente establecer cuál es la mejor escuela y en consecuencia, enviar a sus hijos a ella. Un esquema de “información para el consumidor en un mercado educativo” está vigente en Chile y ha sido aplicado para informar los resultados del Sistema de Medición de la Calidad de la Educación (SIMCE) al menos desde 1995, y de la Prueba de Aptitud Académica (PAA). Recientemente, el CENEVAL publicó una separata con los resultados promedio de los alumnos en el EXANI- II para las escuelas de la zona metropolitana de Ciudad de México. En el año 2000 y con base en el TIMSS, la IEA (*International Association for the Evaluation of Educational Achievement*) publicó un reporte sobre “escuelas eficaces en matemática y ciencias” para el cual utilizó una noción de eficacia absoluta para agrupar a las escuelas “con alto y bajo logro académico” (Martin *et al.*, 2000: 8 y anexo A).

Esta noción de eficacia ha recibido amplia atención en la bibliografía estadística. En particular, Raudenbush y Willms (1995) la denominaron “efectos de tipo A”. Los autores sostienen que su atractivo reside en que representa el punto de vista de las familias al elegir una escuela. También señalar que desde el punto de vista de las políticas educativas tal concepción resulta, al menos, discutible su utilidad y su validez.

Una de las críticas más severas a este concepto se dirige a señalar que no logra distinguir o separar con exactitud si el resultado académico es un producto genuino de la organización escolar o un mero resultado de los procesos de reproducción del capital cultural de las familias<sup>7</sup>. Si fuera este el caso, el juicio estaría atribuyendo a la escuela un “valor agregado” que no existiría, generando señales incorrectas al magisterio y creando una distorsión gigantesca en la información pública sobre el estado de las escuelas.

La falta de controles apropiados sobre los indicadores sociofamiliares ha llevado a afirmar que este tipo de medidas no puede satisfacer los requerimientos básicos para una investigación válida sobre eficacia escolar (Goldstein y Woodhouse, 2000: 358). En el mismo sentido acuerdan dos destacados críticos británicos:

*“Cuando ninguna provisión se hace sobre las características antecedentes (por ejemplo, se usan los puntajes brutos), es generalmente consensuado que esto no mide la eficacia de una escuela: la mejor escuela con desventajosos antecedentes [sociales en su alumnado] no podrá desempeñarse*

---

<sup>6</sup> Por ejemplo, ha sido esta la idea detrás de la política curricular centrada en estándares de aprendizaje que se ha desarrollado bajo los programas *America 2000* y *Goals 2000* en Estados Unidos.

<sup>7</sup> Vía efecto composicional o “efecto de pares”. Por ejemplo, ver Pong (1998).

*mejor que una mediocre escuela con la más destacada población*” (Coe y Fitz-Gibon, 1998:424. Traducción del autor).

## 2.b. La noción de eficacia “incremental”

Una variante del anterior concepto de eficacia absoluta ha sido utilizada en algunos estudios pero añadiéndose un requisito temporal. Una escuela es eficaz si logra obtener una diferencia positiva entre los puntajes que obtienen sus alumnos a lo largo del tiempo. Lo distintivo de esta noción es que la comparación de la escuela es “contra ella misma” y no contra un ranking de escuelas. La estrategia de identificación consiste en la medición de los logros de los alumnos de una o varias escuelas y su comparación en el tiempo y entre las escuelas estudiadas. Formalmente:

$$[2] \quad J \in E \leftrightarrow \mu_{j(t+1)} - \mu_{j(t)} > \Delta^*$$

La noción tiene un atractivo importante: contrasta una escuela consigo misma a lo largo del tiempo, permitiendo observar su evolución o involución y relacionarlo por ejemplo, con políticas emprendidas durante ese período. En consecuencia, podría resultar muy práctico tanto a quienes trabajan en la gestión de los centros educativos como a los hacedores de políticas que procuran mostrar los efectos de las reformas. Desde el punto de vista del método también resulta relativamente sencillo dado que su definición formal sólo implica una operación algebraica básica (una resta). Sin embargo, el método de estimación encierra un problema de identificación bastante complejo: determinar cuándo un incremento o disminución en el aprendizaje se constituye en evidencia válida de que tal cambio es por un lado, estadísticamente significativo y por el otro, teórica o políticamente relevante. Es evidente por lo tanto, que de la misma manera que con la anterior noción, el método requiere que se defina el umbral o valor crítico que a diferencia de  $Z^*$  involucra un *cambio* ( $\Delta^*$ ).

El uso de la noción incremental ha sido bastante reducido en la región. En el año 2000 la Dirección General de Evaluación (DGE), dependiente de la Secretaría de Educación de México (SEP), aplicó este método a la evaluación promedio que las escuelas primarias y secundarias públicas urbanas habían en el Programa de Carrera Magisterial (Véase SEP 2000). En el año 2003, entre mayo y octubre, publicó en su página web (<http://snee.sep.gob.mx>) un listado con las 50 escuelas primarias y las 50 escuelas secundarias con resultados incrementales destacables.

Ahora bien, más allá de su atractivo y aparente simpleza, esta noción de eficacia es pasible de varias críticas. Una primera es que no se resuelve el problema *conceptual* de establecer un umbral a partir del cual se define que la diferencia es estadísticamente significativa, esto es que el cambio observado no es producto de errores de medición o de muestreo y de que además es sustantivamente relevante. Dicho umbral podría ser definido al igual que en la anterior noción a través de estándares curriculares o incluso operacionalizando objetivos de política. Por el contrario, las aplicaciones prácticas vuelven a recurrir a los “listados”.

Otra crítica es que la definición y el método que se emplea, tienden a confundirse el mejoramiento escolar con otros factores, tal como ha sido señalado clásicamente desde el trabajo seminal de Campbell y Stanley (1982). Estos autores señalaron que cuando se analiza una variable a través del tiempo en diseños pretest-posttest de un solo grupo (o de varios pero sin otros controles estadísticos o experimentales apropiados), aparecen confundidos muchos efectos, algunos de ellos denominados “factores de invalidez”, entre ellos: i) historia; ii) maduración; iii) administración de test; iv) instrumentación; v) regresión a la

media<sup>8</sup>.

Finalmente, es de notarse que todas las críticas referidas a la noción de eficacia absoluta se mantienen aquí, particularmente la ausencia de control sobre los factores socioculturales que inciden en los aprendizajes.

## 2.c. La noción de eficacia “relativa”

Tal como Shaffer *et al.* (2000: 239) lo establecen, una escuela eficaz puede ser definida como aquella que sobrepasa una diferencia estadísticamente significativa entre el promedio *observado* de aprendizajes de sus alumnos y el promedio *estimado* sobre la base de un conjunto de escuelas con similares características en el alumnado. Esta diferencia da lugar empíricamente a la construcción de un índice de efectos de la escuela cuyo valor ahora es condicional o relativo.

En esta noción aparentemente sencilla requieren para su implementación de dos vectores de características. Primero, delimita teóricamente un vector de indicadores sociofamiliares, actitudinales e intelectuales (tal como el nivel de conocimiento previo) del alumnado,  $S_{iJ}$ . Mediante diferentes procedimientos alternativos, se resume esta información para cada escuela  $S_J$ , donde el punto indica un promedio por ejemplo) a los efectos de obtener una medida del contexto sociocultural, que puede ser complementada con otras medidas socioeconómicas o institucionales de la localidad,  $L_j$ . Esto da lugar a un vector característico de la escuela,  $C_J$ . Se especifica para toda la población de escuelas, una función que vincula los resultados académicos con el contexto sociocultural y permite estimar cuál es el resultado académico esperado o promedio condicionado a aquellas características socioculturales,  $\mu^*$ . Si el promedio de la *j*-ésima escuela,  $\mu_j$ , supera el valor estimado,  $\mu^*$ , entonces la escuela será considerada eficaz. Debe suponerse además que tales medidas incluyen un error de medición y de estimación,  $\varepsilon$ , que cumple con las propiedades comúnmente asignadas ( $E(\varepsilon) = 0$ ;  $\text{Var}(\varepsilon) = \sigma^2$ ). Es precisamente esa diferencia,  $\varepsilon_j$ , la que define y cuantifica si los de la *j*-ésima escuela satisfacen algún valor o umbral  $Z^*$  que es preciso determinar esta vez y a diferencia de las nociones anteriores, siguiendo un criterio estadístico más que político. Formalmente:

$$[3] \quad C_j = S_j + L_j$$

$$[4] \quad \mu^* = C_j \beta$$

$$[5] \quad J \in E \leftrightarrow \mu_j - \mu_j^* \geq Z^*$$

Este enfoque se concreta directamente con uno de los métodos de estimación que más abajo se detallan. Un punto de discusión es cuál es el valor de  $Z^*$  que se adopta como umbral mínimo para establecer que el resultado de la escuela realmente indica eficacia. Nuevamente tal umbral puede ser fruto de una discusión teórica, una decisión de política educativa o una decisión puramente estadística. Por lo general, la solución más generalmente adoptada consiste en estandarizar la variable  $\varepsilon_j$ . Ravela *et al.*

---

<sup>8</sup> Es claro que estos “factores de invalidez”, también interfieren con las inferencias causales que se puedan hacer cuando se trabaja con datos longitudinales y se aplican otras nociones de eficacia.

(1999) utilizó para el análisis cuantitativo de las escuelas, un valor de más un desvío estándar para identificar los grupos efectivos de sexto grado de primaria, así como menos un desvío para identificar los “grupos bloqueados”. El LLECE (2002) por su parte adoptó el criterio de que  $\epsilon_j$  fuera mayor que 2 desvíos pero menor a tres desvíos<sup>9 10</sup>.

Una crítica importante que puede dirigirse hacia esta noción es que se supone que una escuela tiene efectos constantes sobre todos los alumnos, sin considerar cuáles son las características especiales de cada uno. Este es un aspecto que luego tiene directa consecuencia en los métodos de estimación. Por ahora, su reconocimiento conlleva a introducir lo que en otro lugar se denominó el “enfoque combinado de calidad y equidad” (Fernández, 1999) cuyo origen se remonta al artículo de Edmonds de 1979.

Sin embargo, es importante resaltar que las críticas se dirigen al método pero no al concepto que ha sido enunciado. Específicamente, exigen que el método tenga la misma estructura lógica que el concepto que se propone medir, lo cual es un principio de isomorfía bastante reiterado como exigencia epistemológica en la metodología cuantitativa (Cortés y Rubalcaba, 1993).

## 2.d. El enfoque combinado de Edmonds

Desde la publicación del trabajo de Edmonds (1979), se ha introducido en el debate una conceptualización restrictiva de eficacia que articula calidad con equidad en educación. Esta noción parte señalando que podría ser perfectamente posible que una escuela tuviera un rendimiento muy por encima de lo esperado pero que *la distribución social intra-escuela del conocimiento* conserve la misma desigualdad que se observa en otras escuelas. En consecuencia, no es suficiente con el requisito de la eficacia relativa; es necesario combinarlo con un requisito de igualdad de aprendizajes. Al respecto, sostenía:

*“Nosotros también utilizamos una definición particular de una escuela eficaz: es aquella en la que la proporción de niños de bajo nivel socioeconómico que demuestran maestría en las competencias básicas es virtualmente idéntica a la proporción de los alumnos de clase media que lo hacen. Si una escuela no cumple con este test, ninguna otra cosa podrá cualificarla como efectiva. Supongamos que una escuela tiene el 50% de alumnos de clase media y un 50% de alumnos de clase baja, que el 96% de los alumnos de clase media demuestran maestría anualmente: esa escuela podría ser nominada como efectiva sólo si el 96% de los alumnos de clase baja demuestra anualmente maestría académica. Esto no significa que las escuelas necesariamente lleven a los niños a desempeñarse iguales. Incluso en las escuelas efectivas, los niños de clase media como grupo todavía sobrepasan a los niños de clases bajas como grupo. Esto no viola nuestro estándar porque nosotros requerimos solamente que la proporción de aquellos que exceden el mínimo deben ser aproximadamente la misma”.* (Edmonds, 1979:95. Traducción propia)

Una escuela puede ser denominada como eficaz en el caso en que obtenga “buenos resultados académicos” en las áreas básicas de la educación escolar (matemática y lengua), a pesar de estar trabajando con una población estudiantil proveniente de clases sociales subalternas, obreras o urbano-

---

<sup>9</sup> Sin embargo, no informa cuál es el método estadístico utilizado.

<sup>10</sup> Debe notarse sin embargo que esta decisión genera el problema adicional de distinguir los “casos desviados” de los “casos extremos”, siendo éstos últimos generalmente tratados como “no confiables”.

marginales o también con minorías étnicas marginadas (como son los hispanos y los negros en Estados Unidos). La obtención de tales resultados más allá de la clase social de origen, convertiría a estas escuelas en “eficaces” en relación a los ideales de “igualdad” en las oportunidades de acceso a un mínimo nivel común de los conocimientos científicos y artísticos que la sociedad define como su patrimonio cultural.

Esas observaciones han sabido generar estrategias metodológicas novedosas que permitieran alternativamente satisfacer los requisitos. Por ejemplo, Ravela (1993) generó un índice de efectos de la escuela considerando únicamente los resultados académicos excepcionales para los alumnos que sólo contaban con bajo capital cultural (medido por el nivel educativo materno). A estos alumnos se los denominó como “mutantes” extendiendo así la tipología que Bourdieu iniciara con el término de “herederos”. Concretamente la tipología se expresó de la forma en como lo muestra el cuadro 1.

Cuadro 1. Tipología de alumnos de 4to. primaria según resultados y capital cultural

	Nivel educativo materno hasta primaria	Nivel educativo materno de Secundaria y más
Insuficientes	“Previsibles”	“Perdedores”
Suficientes	“Mutantes”	“Herederos”

Fuente: Adaptado de Ravela (1993) y Rama (1991).

Ahora bien, este concepto no está libre de críticas. El uso de un indicador tal como el porcentaje de alumnos mutantes en una escuela, presenta el problema de decidir respecto del “valor de corte” similares a los ya discutidos anteriormente: qué porcentaje de “mutantes” se considerará indicativo de una escuela eficaz?. A esto se agrega el problema del tamaño relativo de la subpoblaciones utilizadas para comparar: por ejemplo, si son cuatro los alumnos con bajo capital cultural en un grupo de 30, y de esos cuatro tres obtienen la suficiencia, ¿estaríamos frente a una escuela eficaz?.

## 2.e. Las nociones longitudinales de eficacia relativa.

La temporalidad ingresa a los estudios sobre escuelas eficaces de dos formas distintas: una por razones metodológicas y otra por razones teóricas. ¿Cómo saber si la cohorte evaluada no era ya excepcional al inicio de su escolarización? ¿La eficacia ha de observarse para una misma cohorte al cabo de su paso por la escuela o en distintas cohortes y en distintos grados? ¿Alcanza que en una cohorte escolar la escuela tenga resultados excepcionales para ser considerada “eficaz”? ¿La eficacia ha de observarse en todos los grados o sólo al final del período en que se espera que los alumnos permanezcan en la escuela?. Sólo la primera pregunta es de tipo metodológico y corresponde al ámbito de las variables de control.

Ahora bien, es razonable suponer que la noción de eficacia escolar resulta lo suficientemente grave como para que una sola observación sea suficiente para realizar tal juicio. De aquí que se haya introducido en la bibliografía especializada la idea de que una escuela eficaz es una escuela que genera “valor agregado” en sus alumnos. Por tal se entienden por lo general dos aspectos analíticamente distinguibles pero empíricamente relacionados: por un lado, la definición de cuál es el cuerpo de conocimientos que un alumno traía consigo cuando ingresó a la escuela; por el otro, el conocimiento que el alumno ha desarrollado por el hecho de haber estado en la escuela. Si el primer aspecto tiene por implicancia, la realización de una medición ex-ante, el segundo aspecto requiere de la separación de efectos y del diseño de grupos de control.

Esta noción es citada como *la* noción en un trabajo sobre Brasil, en el cual sus autores señalan que: “*Mais especificamente, na escola eficaz o nível do progresso do aluno após um certo período na escola, deve ser superior ao nível do seu progresso, caso o aluno estivesse fora da escola, e contasse apenas com o conhecimento adquirido na sua rede de relações sócio-culturais: família, amigos e escolas anteriores. Considerando alunos de características semelhantes distribuídos por diversas escolas, a escola eficaz agrega valor extra aos alunos comparativamente às restantes. Em contraste, a escola ineficaz é aquela onde os alunos têm progresso inferior ao esperado, dadas as suas características sócio-culturais.*” (Barbosa e Fernandes, 2001:3).

Desde el punto de vista del método, la forma más directa de especificar este concepto es en el marco de una lógica cuasi-experimental (Campbell y Stanley, 1982), cuestión que requiere de: i) una misma población estudiantil; ii) una medición del logro ex-ante (“t”) del ingreso a la escuela; iii) una ex-post (“t+1”), y iv) con al menos un grupo de control que haga posible distinguir qué conocimientos desarrolla un niño des-escolarizado durante el mismo tiempo en que otro está escolarizado. Se supone que de lo contrario no se sabrá realmente si la escuela había intervenido logrando un cambio en el nivel de logro del alumno a través del tiempo de escolarización.

El indicador de eficacia de la escuela pasa a ser ya no el resultado absoluto sino el nivel de ganancia o pérdida de cada alumno en la escuela (Lee, 2001). Nuevamente esta diferencia,  $\varepsilon_J$ , puede ser estandarizada y aplicarse algunas de las reglas anteriormente citadas para confrontarlo con un umbral esperado, al que se denominó más arriba  $Z^*$ . Formalmente podría expresarse como sigue, siempre y cuando se tenga presente que la población de alumnos evaluada es la misma en cualquier  $t$  considerado.

$$[6] \quad \mu_{j(t+1)} - \mu_{j(t)} = \varepsilon_J$$

$$[7] \quad J = E \leftrightarrow \varepsilon_J > Z^*$$

Sin embargo, esta formalización acompañada con un diseño longitudinal de panel, no es una única posible de derivar a partir de la noción más general. Otra alternativa radica en sostener que la estabilidad (o cambio) en los resultados son una propiedad de la organización escolar, sin referencia al panel de alumnos que se estuviera analizando. El argumento es metodológico. Si el concepto de eficacia refiere a la escuela, consistentemente habría de tenerse una definición operativa de estabilidad que refiera a la *misma unidad de análisis*, con independencia de las propiedades singulares de las unidades de registro. En consecuencia, debería contarse con observaciones sistemáticas tomadas a varias cohortes de alumnos ( $P_1 \neq P_2$ ) para aquel grado escolar en el cual se espera la máxima visualización de los resultados del proceso de enseñanza – aprendizaje (por ejemplo, 6to grado de básica). Un concepto como este permitiría recuperar dentro de un diseño longitudinal, una noción como la de incrementalidad para reformularla dentro de una concepción relativa de eficacia escolar. Una escuela podría ser considerada como eficaz si sus efectos superan el umbral crítico  $Z^*$  y se mantienen estables para al menos dos cohortes de alumnos.

### 3. DECISIONES DE DISEÑO

El segundo condicionamiento que se debe considerar en un método estadístico de estimación de los efectos de las escuelas radica en la metodología empleada para contar con una matriz de datos sobre la cual se procesará el análisis. Por lo general, en los trabajos econométricos, todos los analistas partimos de un conjunto de supuestos sobre la validez y confiabilidad alcanza en el proceso de producción de los datos

disponibles<sup>11</sup> (Greene, 1999: capítulo 6) y sólo se llegan a discutir aquellos si se obtienen estimaciones que están fuera de todo lo teóricamente esperable. Sin embargo, el diseño también restringe las decisiones del análisis estadístico por otros dos lugares que tienen que ver con la definición de las filas y de las columnas de matriz: esto es las unidades de análisis y de las variables.

En todas las nociones no absolutistas de escuela eficaz se establece un estándar de referencia dado por el conocimiento esperable del alumnado dada una matriz de características socioculturales. Cualquiera sea el modelo que se desee ajustar para responder a una determinada noción, resulta imprescindible tomar decisiones metodológicas en dos momentos distintos: a) seleccionar la unidad de análisis apropiada; y b) delimitar la matriz de características socioculturales. Podría pensarse que estos son problemas menores de diseño y eventualmente de análisis, que no están relacionados con las virtudes o limitaciones del método de estimación elegido. Sin embargo, son conocidos y estudiados tanto los sesgos de agregación (Raudenbush y Bryk, 2002: caps.1 y 5) como los sesgos que se generan en las estimaciones al omitir variables relevantes, es esperable que los investigadores trabajemos detenidamente en estos aspectos. Sorprenderá sin embargo encontrar la escasa atención que estos puntos reciben en los trabajos especializados, aún cuando por lo general aquí se presentan problemas originados en la definición de las muestras, disponibilidad de los datos desagregados, en la consistencia de la información o en las limitaciones conceptuales y metodológicas de los cuestionarios. Tales cuestiones suelen introducir sesgos sobre los que por lo general se hace un supuesto “grueso”, a saber, de que se cancelan mutuamente, pero lo cierto es que desconocemos en la mayoría de los casos la dirección y magnitud de tales efectos.

### 3.a. La unidad de análisis

El predicado “eficaz” señala una modalidad de ser de una entidad denominada escuela. Dado que existe poco debate en este aspecto, parecería claro de que la unidad de análisis está unívocamente determinada. Sin embargo, toda la anterior discusión teórico-metodológica supone por lo general tres unidades de *registro*, concepto metodológico distinguible del concepto de unidad de análisis (Cortés, Rubalcaba y Yoclevsky, 1991).

En toda investigación que parta una noción relativista sobre eficacia escolar se requiere registrar información mediante encuestas (principal pero no necesariamente): i) de los alumnos, ii) de los padres de los alumnos (generalmente la madre del alumno), iii) de los maestros de la escuela y del director de la escuela. Incluso podrá requerirse integrar información generada por otras fuentes, como por ejemplo, medidas censales territoriales de marginación o pobreza. La información de los maestros y del director se combina posteriormente en una única matriz de datos relativa a la escuela, en tanto que la información provista por la familia es útil tanto para construir variables individuales del alumno como para construir una o varias variables que describan el contexto sociocultural de la escuela. El analista por tanto, trabajará con (al menos) dos matrices de datos “anidadas”. El desconocimiento práctico de la estructura anidada de los datos en este tipo de investigaciones genera problemas que condicionan los métodos estadísticos de estimación.

Un primer caso se presenta cuando por distintas razones, los datos se hallan agrupados ya a nivel de escuelas por ejemplo o no existen directamente los datos individuales. Tal es el caso de los análisis

---

<sup>11</sup> Estos supuestos metodológicos sustentarán luego buena parte de los supuestos de los modelos de regresión y de los modelos jerárquico-lineales. Véase Greene (1999), capítulo 6 y Gujarati (2004) parte II.

realizados sobre las bases de datos del SIMCE en Chile previos a 1998<sup>12</sup> o el trabajo sobre escuelas eficaces en Honduras (Fernández, Silva y Trevignani, 2002). Un segundo caso se presenta cuando el analista opta por suprimir en sus modelos la estructura anidada y trabajar exclusivamente con datos agrupados a nivel de escuela.

En ambas situaciones, se conoce que el ajuste de modelos de regresión basados en mínimos cuadrados ordinarios o ponderados tenderán a: i) entregar estimaciones menos eficientes de los parámetros; y ii) obtenerse mejoras en la bondad de ajuste, que en algunos casos puede ser sustancialmente mayor al que se hubiera obtenido en casos contrarios (Greene, 1999:374-376).

Un tercer caso de desconocimiento práctico de la estructura anidada de los datos se presenta en forma particular cuando se adopta una noción de eficacia longitudinal. En este caso, el tiempo constituye uno de los niveles de análisis para cuyo tratamiento existen diferentes modelizaciones, una entre ellas, pueden ser los modelos jerárquico-lineales. Sin embargo, lo importante a retener aquí es que la estimación de los efectos de una escuela que se derivan de un método fundado en mediciones hechas sobre los resultados *de una sola cohorte de alumnos* podrían ser cuestionadas en su validez.

### 3.b. La teoría, los indicadores y los problemas de especificación.

Las restricciones en los datos están presentes también por el lado de las columnas. El problema aquí es delimitar un vector con  $k$  indicadores teóricamente válidos para cada uno de los  $i$ -ésimos *alumnos evaluados* que permitan “igualar” estadísticamente las poblaciones de alumnos (en ausencia de una asignación aleatoria). Por lo general, existe un consenso extenso sobre un núcleo de indicadores mínimos que se derivan de la discusión conceptual antes presentada, en particular fundada en la noción de “eficacia relativa”. Esto incluye tres bloques de variables: las sociofamiliares, las actitudinales y las de género.

Ahora bien, en los estudios sobre escuelas eficaces, el objeto de interés no está en analizar los efectos de estas variables; de aquí su papel de control estadístico. En el cuadro 2 se incluye un ejemplo reseñando los indicadores utilizados en estudios sobre Argentina, Uruguay, Chile, México y Bolivia. Tal como se puede observar existe cierta variedad tanto en la cantidad de variables así como la complejidad de su construcción. Esto conlleva dos problemas para cualquiera de los métodos de estimación: i) exogeneidad y ii) sesgo de selección. Por un lado, una discusión de importancia para el objetivo particular de identificar las escuelas eficaces, de cuáles de estas variables se puede afirmar que en sí misma no incluye “efectos de las escuelas”. Evidentemente los modelos de Chile y de Bolivia no presentan estos problemas, pero sí claramente los de México y de Uruguay y en menor medida podría discutirse el caso del modelo presentado para Argentina. Sin embargo, si se optara por excluir estas variables del modelo bajo el argumento de que no son en puridad “exógenas”, si este se demostrara falso, se cometería un error de especificación que ocasionaría sesgos en las estimaciones de los parámetros por excluir variables relevantes que además se sabe que están relacionadas con las variables socioeconómicas.

Cuadro 2. Variables individuales consideradas en estudios sobre efectos de la escuela en América Latina

Argentina	Bolivia	Chile	México	Uruguay
Nivel socioeconómico del	Idioma materno	Índice socioeconómico	Índice de capital familiar	Saldo educativo

<sup>12</sup> Véanse las observaciones que Mizala y Romaguerra 2000 respecto de los trabajos anteriores que para la educación media no contaban con datos individuales.

hogar		del hogar ( y su cuadrado)	(económico y cultural) del hogar
Índice de recursos culturales	Educación de padres	Género del alumno	Hábito de lectura
Edad	El niño vive con sus padres		Aspiraciones educativas
Sexo	Nº de libros en su hogar		Valoración del rendimiento académico
Repetición	La madre ayuda en las tareas		Repetición
Cambio de escuela en los 2 últimos años			Sexo (e interacciones con las anteriores variables)
Tiempo de traslado a la escuela			
			Tasa de niños con reducida o nula pre-escolarización
			Saldo equipamiento
			Tasa de hogares con hacinamiento
			Tasa de alumnos en viviendas precarias
			Tasa de alumnos en hogares con 5 y más hijos
			Tasa de niños con malestar personal

Fuente: Para Argentina, Cervini (1999); para Bolivia, Mizala, Romaguerra y Reinaga (2000); para Chile, Mizala y Romaguerra (2000); para México, Fernández (2003 b, 2003c); para Uruguay, Ravela *et al.* (1999). Para los análisis de Argentina y México se han incluido también variables contextuales que promedian estas propiedades del nivel individual.

Es claro que los argumentos que se puedan aportar para discutir exogeneidad y selección tienen un fundamento teórico y no estadístico. El punto es que los indicadores a incluir deberían estar conectados con las teorías disponibles y con los antecedentes de la investigación acumulada. Estos indicadores deberían estar representando, al menos, la teoría de las prácticas (Pierre Bourdieu), con sus distintos conceptos de capital, formas del capital y de habitus. En segundo lugar, habría introducir indicadores de la teoría de las elecciones educacionales (Raymond Boudon, John Goldthorpe), la teoría del capital social (James Coleman) y las estrategias que los estudiantes desarrollan para regular su propio aprendizaje (OECD-PISA, 2003). Es de advertir que en estas tres últimas dimensiones no está clara la precedencia temporal requerida por el control estadístico: es tan legítimo sostener que pueden ser desarrolladas o al menos alterada su incidencia mediante la escuela como suponer que constituyen variables derivadas del capital familiar. En tercer lugar y dado que ninguna de estas teorías fue formulada originalmente estableciendo una distinción entre las prácticas de varones y mujeres, es necesario incorporar aunque no sea más que por pura razón de control, al menos un indicador sobre el sexo del alumno. Lo más conveniente en todo caso es que los mismos indicadores incluidos en los cuestionarios incluyan ya una distinción de género; por ejemplo, en la cuestión del trabajo infantil.

Un sesgo de selección importante que se verificaría razonablemente en países pluriétnicos o con una significativa presencia de inmigrantes es no identificar estas propiedades en los alumnos, o en las escuelas si éstas por algún motivo, son escuelas segregadas<sup>13</sup>. Por ejemplo, se conoce desde el primer reporte Coleman que las minorías étnicas negras e hispanicas aprenden menos que sus pares blancos; se ha mostrado que tal es el caso para los niños indígenas en Bolivia (Mizala, Romaguerra y Reinaga, 2000), para el Perú (Cueto y Secada, 2003) y de las escuelas indígenas en México (Fernández, 2003a). Llama la atención, sin embargo, que fuera de estos casos, en América Latina estas variables sólo hayan sido objeto de atención en forma muy marginal.

<sup>13</sup> Por ejemplo, escuelas para indígenas como ocurren entre otros lados, en Bolivia, Ecuador, México y Perú. Sin embargo, las evaluaciones de aprendizaje de México no incluían hasta el presente año (2003) ningún indicador que permitiera identificar la pertenencia étnica de los alumnos.

La noción de eficacia longitudinal ha movido a los investigadores a procurar ampliar el registro de información sobre la trayectoria anterior del alumno, en temas tales como la asistencia a la educación pre-escolar, los cambios de escuela, los cambios de domicilio, los cambios de maestros durante el mismo año escolar, etc.

Finalmente, una mención especial deberá hacerse respecto a los conocimientos previos que un alumno contaba al inicio de su educación en una determinada escuela; variable que ya formaba parte de las discusiones al inicio de los años ochenta (Goldberber y Cain, 1982) pero que aún no se incorporado en América Latina. Es razonable pensar que el sector privado está en condiciones de aplicar políticas de selección de su alumnado y que éstas consideran específicamente los conocimientos previos. Si esto permite a estas escuelas contar con los alumnos más hábiles pero si a la vez no se incluyen en los modelos, los efectos de las escuelas privadas serán sistemáticamente sobre-estimados. Una correcta especificación debería controlar por conocimientos previos o utilizar como variable dependiente el cambio en el nivel de conocimientos del alumno. La importancia de este indicador es lo suficientemente grande como para que algunos autores sólo acepten la idea de “efectos de la escuela” como “valor agregado”. Excluirlo implica introducir sesgos en la estimación y atribuir a la escuela efectos producidos por la autoselección de los alumnos. Por ejemplo, Stevans y Sessions (2000: 170) refiriéndose a este problema en el campo de las comparaciones público /privado señalan las graves implicancias que tienen estos sesgos:

*“Tomando en cuenta el sesgo de selección, Noel (1981,1982) no encontró diferencias significativas en el desempeño entre los alumnos de las escuelas católicas y públicas en los Estados Unidos. Esto sugiere que los mejores estudiantes están eligiendo asistir a escuelas privadas, mientras que los estudiantes con menores aptitudes se encuentran en las escuelas públicas”* [traducción propia].

#### **4. EL MÉTODO DE ESTIMACIÓN POR REGRESIÓN LINEAL MÚLTIPLE.**

La estimación de los efectos utilizando un modelo de regresión lineal múltiple (generalmente por mínimos cuadrados), está entre las decisiones de método más frecuentemente hechas en la investigación sobre eficacia educativa. El método por lo general parte de una noción relativa de eficacia, agrega los datos a nivel de la escuela, desconociendo la estructura anidada de los datos, y calcula un valor promedio esperado de aprendizaje para la muestra dada una matriz de características sociofamiliares promedio.

Los parámetros de la ecuación pueden estimarse siguiendo diferentes alternativas. Aitkin y Longford (1986) señalaron cuatro alternativas, de las cuales tres eran modelos que utilizan estimación mínimos cuadrática ordinaria y hace el ajuste con datos individuales (modelos 1,2 y 4) y uno mediante estimación mínimo-cuadrática ponderada que opera con datos agregados (modelo 3).

En términos generales, las principales variantes para este método que destaca la bibliografía refieren: a) el uso datos individuales en lugar de agrupados; b) el ajuste mediante mínimos cuadrados ponderados; y c) el tratamiento de los casos desviados en la identificación del umbral  $Z^*$  de los efectos de la escuela.

##### **4.a. Mínimos cuadrados ordinarios con uso de los residuos promedio por escuela**

En esta primera alternativa, el método de estimación requiere trabajar en dos etapas, ajustando la ecuación de regresión con los datos individuales y estimando los parámetros mediante mínimos cuadrados

ordinarios. En la forma simple, se adopta la decisión de operar al nivel del *i-ésimo* niño con sus resultados de aprendizaje ( $Y_i$ ) y con un vector de características sociofamiliares ( $X_i$ ), el modelo se formaliza así:

$$[8] \quad Y_i = \beta_0 + X_i \beta + \varepsilon_i$$

Obsérvese que no se identifica la escuela de pertenencia del alumno. Lo cual supone la esfericidad del término de error individual (Greene, 1999:203,469-529). Se opera con todos los datos y por lo tanto  $\beta$  es el efecto “*pooled*” de las variables sociofamiliares sobre el aprendizaje. Para estimar el efecto de la escuela es necesario promediar para cada escuela los residuos individuales ( $\varepsilon_i$ ) para obtener el residuo propio de la escuela ( $\varepsilon_j$ ).

$$[9] \quad \varepsilon_i = Y_i - \beta_0 - X_i \beta$$

$$[10] \quad \varepsilon_i = Y_i - Y_i^*$$

$$[11] \quad \varepsilon_j = Y_j - Y_j^*$$

Es de observar que en ese modelo,  $\varepsilon_j$  representa un término residual compuesto por tanto por las variables individuales no contempladas en la parte no sistemática del modelo, como por las variables organizacionales sobre las cuales se inferirá que son causa de la eficacia. El residuo por tanto podrá ser de gran magnitud tanto si: i) se omiten variables individuales importantes; ii) si se desconocen efectos composicionales; como si iii) existen verdaderos efectos de la escuela de tal magnitud. Ahora se comprende los problemas teóricos que conlleva una correcta especificación del modelo.

#### 4.b. Mínimos cuadrados ordinarios con información agregada al nivel de la escuela

Otra alternativa de ajuste por regresión lineal (modelo 3 de Aitkin y Longford, 1986) consiste en ajustar un modelo que relaciona los promedios de aprendizaje en la escuela ( $Y_j$ ) con los promedios de las variables sociofamiliares de control ( $X_j$ ) (datos agregados a nivel de la escuela). Formalmente:

$$[12] \quad Y_j = \beta_0 + X_j \beta + \varepsilon_j$$

$$[13] \quad \varepsilon_j = Y_j - \beta_0 - X_j \beta$$

#### 4.c. Mínimos cuadrados ponderados con información agregada al nivel de la escuela

Una variante de este método implica estimar los parámetros utilizando mínimos cuadrados ponderados. La propuesta fue reseñada por Aitkin y Longford (1986), quienes plantearon que la ecuación debe ajustarse ponderando por el número de alumnos evaluados en cada escuela ( $n_j$ ) y evitar así problemas de heterocedasticidad conocidos (Greene, 1999).

$$[14] \quad Y_j = \beta_0 + X_j \beta + \eta_j, \text{ tal que: } \eta_j \sim N(0; \phi).$$

Se supone que el residuo ( $\eta_j$ ) sigue una distribución normal con media cero pero con una varianza  $\phi = \sigma^2 / n_j$ . El argumento que la dispersión de los residuos del modelo es inversa al tamaño de las escuelas; esto es que los resultados académicos de las escuelas más grandes son más homogéneos respecto

de las características contextuales especificadas en los modelos. Para devolver el residuo a su magnitud, se ajusta la ecuación por mínimos ponderados (WLS). Sin embargo, es de notarse que actualmente este tipo de ajuste podría ser re-emplazado por un modelo que estime errores estándares robustos.

#### **4.d. Determinación del umbral $Z^*$**

Un problema que comúnmente se plantea con este método es establecer un intervalo dentro del cual se puede suponer que los valores observados en los residuos de la muestra de la escuela ( $\varepsilon_j$ ) pueden considerarse al azar y por tanto no podrían considerarse con cierta confianza como indicadores de eficacia escolar. Por lo general, se adopta el criterio de definir como umbral los valores  $Z^* = -2$  o  $Z^* = 2$  desvíos estándar. De esta forma, aquellos grupos cuyo valor residual promedio se encuentre por encima de  $Z^*$  pueden ser considerados como efectivos, ya que los logros obtenidos se encuentran razonablemente por fuera de la determinación de la recta de regresión.

Sin embargo, pueden existir valores de  $\varepsilon_j$  que superan en dos o tres veces el umbral establecido y para los cuales puede sospecharse que sus residuos no responden a un efecto sustantivo de la escuela, sino a problemas de medición y confiabilidad de los datos. Autores como Goldstein (1997) sostienen que estos casos debieran ser eliminados del análisis.

#### **4.e. Algunas aplicaciones**

El método de mínimos cuadrados en su segunda modelización ha sido el utilizado, entre otros, por Ravela *et al.* (1999) para los datos de la evaluación de 6to de Primaria de 1996. Este estudio tiene la particularidad de que analiza los efectos del grupo o sección de alumnos dentro de cada escuela, en lugar de analizar las escuelas. Para la estimación se agregó un importante número de variables sociofamiliares y un indicador de daño psicológico de los alumnos y luego se construyó un índice a través del análisis factorial. (Ver cuadro 2). Los puntajes promedio de matemática y español fueron regresados sobre esta variable, asignándosele a cada grupo el residuo. No se hizo corrección por heterocedasticidad. Para identificar los efectos de la escuela se definió para los residuos un intervalo de menos uno a más un desvío estándar.

#### **4.f. Limitaciones**

En todos los casos, el uso de este método implica desconocer las propiedades específicas de una de las dos unidades de análisis: si se trabaja con los alumnos se ignora el agrupamiento en las escuelas debido al supuesto de ausencia de autocorrelación (Greene, 1999: 501 y ss; Raudenbush y Bryk, 2002: capítulo 1); si se trabaja en el nivel de las escuelas, se ignora la heterogeneidad que pueda existir en el alumnado de una escuela. Ambas restricciones son fundamentales de tener presente.

### **5. EL MÉTODO DE ESTIMACIÓN POR MODELOS JERÁRQUICO-LINEALES.**

Luego de quince años de discusión sobre el método estadístico óptimo, hoy es generalmente aceptado que un abordaje satisfactorio para la identificación de las escuelas eficaces requiere modelos estadísticos multinivel o jerárquico-lineales (Goldstein, 1997; Shaffer *et al.*, 2000; Raudenbush y Bryk, 2002).

El uso de este método requiere de un proceso de decisiones de especificación fuertemente orientado por la teoría, en particular por la propia conceptualización de escuela eficaz que se tenga. De hecho, las diferencias más importantes entre los distintos los modelos que se pueden ajustar radican en el supuesto teórico que tienen por detrás, tal como se mostrará a continuación.

En términos generales, estos modelos trabajan con información del alumno y de la escuela sin fusionar o suprimir ambos niveles. Una modelización de este tipo permite así mantener la distinción teórica entre lo macro (la escuela y su entorno) y lo micro (los procesos individuales) sobre el nivel de aprendizaje de un alumno. Lo cual abre nuevas potencialidades al estudio de la eficacia escolar. Analíticamente se pueden ajustar ecuaciones de factores determinantes por separado, alcanzando una más elaborada comprensión de cómo y sobre qué tienen incidencia la organización escolar. Una escuela puede ser clasificada como “eficaz” en dimensiones que antes no eran posibles de hacerse; por ejemplo, introduciendo el requisito combinado de Edmonds.

### **5.a Razones para ajustar un modelo multinivel.**

Aunque su uso más destacado ha sido para identificar el objetivo más general de conocer cuáles son las características de la organización escolar que inciden sobre los aprendizajes, su utilización en este campo más específico tiene las siguientes virtudes. Entre otras razones señaladas por la bibliografía, se pueden señalar las siguientes tres.

En primer lugar, con un modelo multinivel es posible especificar de forma correcta y completa la compleja interdependencia entre las características sociofamiliares representadas en un vector  $S_{ij}$  del alumno y las características contextuales representadas en un vector composicional,  $C_j$ , y un vector localidad,  $L_j$  de la escuela que co-determinan el nivel de aprendizajes de cada alumno y por ende, también el proceso de reproducción cultural. Al hacer posible un modelo que considera ambas unidades de análisis, se evitan dos sesgos que se cometen al usar el método de mínimos cuadrados que son: a) desconocer los agrupamientos de alumno y escuela generados cuando se “aplanan los datos” (véase más arriba; y b) suprimir la heterogeneidad del alumnado como cuando se “agregan los datos” al nivel de escuela<sup>14</sup>.

En segundo lugar, distingue en la variación de los aprendizajes, la porción que corresponde a los atributos individuales del alumno y la porción atribuible a la escuela, cuantificación que responde a la pregunta más general sobre cuánto importa la escuela. Esta distinción levanta la restricción que impone el supuesto de ausencia de auto-correlación en el modelo de regresión lineal múltiple. Lo cual resulta congruente con la teoría; esto es que los niños aprenden en procesos grupales y que por lo tanto es esperable que los aprendizajes de los alumnos de un mismo grupo sean semejantes en alguna proporción.

En tercer lugar, la descomposición de la varianza en los aprendizajes permite representar más adecuadamente la parte no sistemática del modelo. En el método de mínimos cuadrados, se especifica una parte sistemática donde se incluyen todos los términos relativos a las variables sociofamiliares y una parte aleatoria con un único término de error, sea a nivel de la escuela, sea a nivel del alumno. Se mostró ya que este término de error no tiene directa una interpretación cuando se ajusta a nivel de los alumnos, porque es una mezcla no distinguible entre los residuos generados a nivel individual ( $\varepsilon_{ij}$ ) y los residuos generados a

---

<sup>14</sup> Este último aspecto será tratado más abajo con el nombre de prueba de homogeneidad de las regresiones.

nivel de las escuelas ( $u_j$ ). Este problema es resuelto en el modelo multinivel, incluyendo en su especificación mínima, representar dos términos de error en la ecuación.

En cuarto lugar, el uso del modelo permite establecer si la escuela tiene efectos diferenciales sobre sus alumnos dependiendo de cuáles sean sus características sociofamiliares, étnicas o de género. En comparación, el método de mínimos cuadrados hace el supuesto de que los efectos del capital cultural, por ejemplo, son homogéneos y constantes a través de todas las escuelas. Al levantarse tal restricción de forma como se detallará más abajo, se conquista una nueva dimensión para el análisis empírico que está presente en el enfoque “combinado de calidad y equidad”. Con esta modelización se está en condiciones de identificar si una escuela distribuye los aprendizajes entre sus estudiantes de forma más igualitaria, además de generar un “piso común” de aprendizajes más alto. Como se verá a continuación, no es frecuente que este tipo de consideraciones sea incluido en los análisis. Por esta razón se mostrarán las dos formas en que el método multinivel se ha utilizado.

### 5.b. La estimación multinivel bajo el supuesto de uniformidad

Los conceptos de “eficacia relativa” y de “eficacia longitudinal” no consideran como nota esencial de una escuela eficaz la modalidad con que los conocimientos se distribuyen socialmente dentro del alumnado de una escuela. En ese sentido, se puede asumir el supuesto de que los efectos de la escuela son constantes para todo su alumnado y que las únicas diferencias entre las escuelas radican en el promedio de aprendizajes o “piso común” que generan para sus estudiantes.

Si se adopta la notación clásica de los modelos multinivel propuesta por Raudenbush y Bryk (2002), lo cual implica escribir las ecuaciones respectivas para cada nivel, este modelo de efectos uniformes se expresaría así:

$$[15] \quad Y_{ij} = \beta_{0j} + S_{ij}\beta_{1j} + v_{ij}$$

$$[16] \quad \beta_{0j} = \gamma_{00} + C_j\gamma_{01} + u_{0j}$$

$$[17] \quad \beta_{1j} = \gamma_{10j}$$

Donde:

- $Y_{ij}$  es el nivel de aprendizaje del *i-ésimo* alumno en la *j-ésima* escuela
- $\beta_{0j}$  es el promedio de aprendizajes en la *j-ésima* escuela
- $\gamma_{00}$  es el nivel de aprendizajes promedio para todos los grupos de la muestra analizada
- $S_{ij}$  es un vector de características sociofamiliares del *i-ésimo* alumno
- $C_j$  es un vector de características del contexto de la escuela
- $\beta_{1j}$  es el efecto del vector S estimado para la *j-ésima* escuela.
- $\gamma_{10}$  es el efecto del vector S estimado a través de todas las escuelas de la muestra
- $\gamma_{01}$  es el efecto del vector C sobre el promedio de la escuela
- $v_{ij}$  es la diferencia entre el estimado y el observado para el estudiante

$u_{0j}$  es el efecto único de la escuela sobre los aprendizajes.

La uniformidad de los efectos de la escuela se traduce en “fijar” el coeficiente que representa cómo se distribuyen socialmente los aprendizajes *dentro de cada escuela* (ecuación 17). La variabilidad entre escuelas queda restringida a la diferencia entre los promedios o pisos comunes de aprendizaje que cada escuela proporciona a su alumnado. De acuerdo a esta especificación, el efecto de la escuela sobre los aprendizajes se identifica directamente despejando de la ecuación 15:

$$[18] \quad u_{0j} = \beta_{0j} - \gamma_{00} - C_j \gamma_{01}$$

Si se desea también se puede expresar este modelo de efectos uniformes mediante la ecuación combinada sustituyendo en 16 con las ecuaciones 17 y 18, tal como lo hacen Raudenbush y Willms (1995: 320):

$$[19] \quad Y_{ij} = \gamma_{00} + \gamma_{01} C_j + \gamma_{10j} S_{ij} + v_{ij} + u_{0j}$$

En de notarse que el modelo puede extenderse para que  $S_{ij}$  pueda representar un vector amplio de características individuales, tales como el sexo, la condición de actividad laboral, aspiraciones educacionales, disposiciones culturales, auto-valoraciones, etc. Del mismo modo, se pueden incluir dentro de las características contextuales todo el conjunto de características del entorno de la escuela sobre los que no se puede teóricamente suponer que sean manipulables por la gestión de la escuela.

### 5.c. El modelo de distribución social del conocimiento

Es razonable suponer y así se ha discutido mucho durante los años noventa, que las escuelas eficaces no son eficaces de la misma forma para cualquier tipo de estudiantado. Podría resultar que los varones aprovecharan más que las mujeres los efectos de una escuela eficaz; que los alumnos de alto capital más que los de bajo capital cultural; que los indígenas más que los blancos, y así sucesivamente. Esto ha sido designado en la literatura como “distribución social del conocimiento” (Lee y Bryk, 1989; Lee, 2001).

En consecuencia, junto con ajustar el promedio o “piso común” de aprendizajes para la *j-ésima* escuela, podría la investigación interesarse en cómo se distribuyen los aprendizajes dentro de la escuela y establecer por ejemplo, si las desigualdades de clase social se incrementan, se mantienen iguales o disminuyen. Tal preocupación se especifica permitiendo que el efecto de clase social sobre el aprendizaje varía (sea diferente) entre las escuelas. Formalmente se agrega un término a la ecuación 19 para obtener:

$$[20] \quad Y_{ij} = \gamma_{00} + C_j \gamma_{01} + S_{ij} \gamma_{10j} + v_{ij} + u_{0j} + u_{1j} * S_{ij}$$

Donde:

$u_{1j} * S_{ij}$  es el efecto variable entre escuelas de los antecedentes sociofamiliares.

Nótese en particular que la anterior especificación de los efectos de la escuela ha ampliado conceptualmente su rango de forma tal que ahora  $\epsilon_j$  ha sido representado por la adición de dos términos:

$$[21] \quad \epsilon_j = u_{0j} + u_{1j} * S_{ij}$$

### 5.d. El modelo de tres niveles.

Goldstein y Woodhouse (2000) han realizado la observación de que la estimación puede ser más compleja aún si se considera que las escuelas generalmente están agrupadas y dependen burocráticamente de una autoridad territorial. Este aspecto no está presente por lo general en la teoría y menos aún se sacan las consecuencias metodológicas esenciales.

*“Una de los rasgos más persistentes en la investigación sobre la eficacia escolar es la su conceptualización de las escuelas entidades discretas que no interactúan entre sí y que tienen características derivadas de sus alumnos y sus profesores así como otras estructuras permanentes. [...] hay un problema fundamental con el supuesto básico mencionado en el comienzo del párrafo, concretamente, que las escuelas son asumidas como entidades no interactivas. Un reflexión detallada sobre la formas en que las escuelas funcionan, especialmente en Inglaterra, revela que las acciones y las características de cualquier escuela están vinculadas con las características de otras escuelas. Mortimore y Whitty (1997) hicieron una puntualización similar sobre la imposibilidad de separar el funcionamiento de la escuela de las constricciones de la sociedad” (Goldstein y Woodhouse 2000: 256; traducción propia).*

Ahora bien, si las escuelas no pueden ser entendidas ignorando por ejemplo, las estructuras institucionales de gobierno, control y financiamiento del sistema educativo, la investigación sobre los efectos de las escuelas no puede ignorar que en realidad algunos efectos puedan ser más correctamente atribuibles a las sociedades nacionales, a los sistemas educativos o las entidades territoriales de los cuales son parte las escuelas. Las fuentes de variación en los aprendizajes se extienden para incorporar los territorios y reconocer que la estructura anidada o jerárquica de los datos tiene un nivel adicional<sup>15</sup>.

Si se designa ese tercer nivel, por ejemplo la jurisdicción territorial, como  $k$ , las ecuaciones 15 a 17 introducidas más arriba deben reformularse y ampliarse para incorporar esta nueva fuente de variación, conjuntamente con los “efectos diferenciales de la escuela” tratados en la sección anterior. Formalmente:

En el nivel 1 (alumnos) :

$$[22] \quad Y_{ijk} = \beta_{0jk} + S_{ijk}\beta_{1jk} + v_{ijk}$$

En el nivel 2 (escuelas):

$$[23] \quad \beta_{0jk} = \gamma_{00k} + C_{jk}\gamma_{01k} + u_{0jk}$$

$$[24] \quad \beta_{1jk} = \gamma_{10k} + u_{1jk}$$

En el nivel 3 (jurisdicciones):

$$[25] \quad \gamma_{00k} = \gamma_{000} + r_{00k}$$

---

<sup>15</sup> Uno de los comentaristas anónimos a este artículo señaló la importancia de señalar aquí el nivel “profesor” en los estudios de la escuela. Es posible que el argumento de Goldstein y Woodhouse podría extenderse para separa el efecto de la escuela respecto del efecto del profesor. Este es un tema en el cual aún no se ha avanzado lo suficiente (Reynolds et al 1994) a pesar de la evidencia disponible que señala que los

$$[26] \quad \gamma_{01k} = \gamma_{010}$$

$$[27] \quad \gamma_{10k} = \gamma_{100} + r_{10k}$$

La complejización del modelo no debe impresionar como innecesaria o atentadora de la parsimonia. Si se observa detalladamente, la consecuencia es clara. Si se sustituye en la ecuación 23<sup>16</sup>, con las respectivas ecuaciones 25, 26 y 27 para luego reordenar términos, se obtiene la estimación de los efectos de las escuelas:

$$[28] \quad \beta_{ojk} = \gamma_{000} + C_{jk}\gamma_{010} + r_{00k} + u_{ojk}$$

$$[29] \quad u_{ojk} = \beta_{ojk} - \gamma_{000} - \gamma_{010} C_{jk} - r_{00k}$$

Dado que se ha partido la varianza entre en su componente individual (alumno), organizacional (escuela), y territorial (autoridad educativa), se diluye el sesgo que ocasiona atribuirle a la escuela *también* el componente de eficacia (o ineficacia) originado en la administración educativa territorial.

Es de observarse también que este modelo de tres niveles de efectos muestra una aproximación particularmente interesante en la comparación internacional, de cuál puede ser el peso de la sociedad nacional y de su sistema educativo sobre los aprendizajes.

### 5.e. Estimadores bayesianos: ventajas y sesgos potenciales

Desde el punto de vista estadístico existe una ventaja adicional a favor de este método y de los distintos modelos: la introducción de los estimadores bayesianos. Esto permite corregir más eficientemente las formas peculiares que se observan en las distribuciones heterocedásticas de residuos cuando éstas son generadas por datos agregados con distintos tamaños de muestras de alumnos evaluados *por escuela*. Tal como lo señalan Raudenbush y Bryk (2002: 153):

*“Los residuos empíricos de Bayes estimados bajo un modelo jerárquico lineal proveen de un indicador estable para juzgar la performance de una escuela individual. Estos estimadores empíricos bayesianos tienen ventajas distintivas sobre los métodos previos. Aquellos a) toman en cuenta la pertenencia a los grupos [escuelas] aún cuando el número de grupos es muy grande, y b) producen estimadores relativamente estables aún cuando los tamaños muestrales por escuela son modestos”.*

Los paquetes estadísticos (HLM, MLwin) utilizan estimadores bayesianos para todos los parámetros y por ende,  $u_{0j}$ , aunque por ejemplo, se pueda disponer de estimaciones mínimo-cuadráticas para los primeros pasos del análisis y realizar comparaciones apropiadas.

La propiedad de este algoritmo es que opera reduciendo o restringiendo la varianza de  $u_{0j}$ , de aquí que se le denominen estimadores reducidos (*“shrunken residuals”*, Goldstein, 1997). La reducción se

---

efectos del profesor son incluso de mayor magnitud que los efectos de la escuela. Probablemente se deba a que las evaluaciones suelen no realizar censos de todos los grupos en cada escuela.

<sup>16</sup> Por razones de exposición se ha supuesto que el efecto de las características contextuales de la escuela es el mismo a través de las  $k$  entidades territoriales. Perfectamente podría hacerse variar también.

aplica al estimador mínimo cuadrático en forma inversamente proporcional a la falta de plausibilidad del valor observado en el residuo. Cuanto más confiable es el estimador OLS, menor es la reducción observada. En el extremo, los valores que han sido generados por muestras muy pequeñas de alumnos, son reducidos a cero. El supuesto por detrás es que los residuos de las escuelas siguen una distribución en el universo que debe ser estimada a partir de los estadísticos provistos por los promedios de las escuelas. Si dichos estadísticos provienen de muestras pequeñas, entonces es razonable que la estimación sea poco confiable<sup>17</sup>. Específicamente el estimador de los efectos de las escuelas bajo un modelo jerárquico-lineal es:

$$[30] \quad u^*_{0j} = \lambda_j u_{0j}$$

$$[31] \quad \lambda_j = \tau_{00} / (\tau_{00} + \sigma^2 / n_j)$$

Donde:

$\tau_{00}$  es la varianza entre escuelas en la variable dependiente

$\sigma^2$  es la varianza (homocedástica) a nivel de los alumnos

$n_j$  es el número de alumnos de la  $j$ -ésima escuela

Sin embargo, estas estimaciones no están libres de un potencial sesgo que puede denominarse “profecía autocumplida” (Raudenbush y Bryk, 2002: cap.5). Tal como se ha mostrado, por definición la falta de confiabilidad del efecto de la escuela provoca que el estimador tienda a acercarse al valor esperado con base a las características del alumnado. Si por ejemplo, una escuela inusualmente eficaz que atiende niños con muy bajo capital cultural, tendrá una estimación sesgada hacia la gran media, hacia el valor que típicamente muestran otras escuelas con un alumnado similar desde el punto de vista sociofamiliar. Si este fuera el caso, el procedimiento estadístico operaría una suerte de profecía estadística autocumplida, en la cual los efectos de la escuela serían sub-estimados para ciertas escuelas puntuales.

## 5.f. Aplicaciones

En Latinoamérica, donde los estudios sobre efectividad escolar son aún incipientes, existen pocas investigaciones que utilicen los modelos multinivel para estimar la eficacia escolar (Cervini, 1999, 2001, 2002a, 2002b; Fernández, 1999b, 2003b; LLECE, 2000; Barbosa y Fernandes, 2001; Cueto y Secada, 2003; Cueto *et al.*, 2003). Menos aún se conocen cuántas investigaciones que haya utilizado este método para el estudio específico de escuelas eficaces<sup>18</sup>.

Ahora bien, la mayor utilidad de tal técnica frente al método de mínimos cuadrados en el cual si hay experiencia de investigación, podría residir en el análisis de la distribución social del conocimiento. Podría pensarse que si se clasifican conjuntamente a las escuelas según su efecto sobre el aprendizaje ( $u_{0j}$ )

<sup>17</sup> Existe cierta discusión sobre este punto. Podría suceder que si la estimación de los promedios de las escuelas se hace con base en un censo de todos sus estudiantes, la ponderación de los efectos sea indebida. Esto tiene relación con el diseño de la muestra. Véase Raudenbush y Bryk 2002: capítulo 5.

<sup>18</sup> Recientemente, el autor ha concluido un análisis de las escuelas primarias de México definiendo un modelo de tres niveles. (Fernández 2003 c)

y según su efecto sobre la distribución social del conocimiento entre sus alumnos ( $u_{lj} * S_{ij}$ ) se obtenga un panorama como el siguiente:

**Cuadro 3. Propuesta de tipología de posibles resultados de calidad y equidad**

	Bloqueada	Predecible	Eficaz
Disminuye la desigualdad	1	2	3
La conserva	4	5	6
Incrementa la desigualdad	7	8	9

Si bien toda la cuarta columna (celdas 3, 6 y 9) son de interés en la investigación sobre la eficacia escolar, una política educativa orientada por objetivos de equidad estará particularmente interesada en conocer el perfil de las escuelas que se clasifican en la celda 3.

## 6. REQUISITOS Y LIMITACIONES PARA UNA INFERENCIA VÁLIDA

Hasta aquí se han revisado los múltiples requisitos de diseño y estimación que deberían tener los estudios sobre las escuelas eficaces que usan métodos estadísticos como para aceptarse académicamente las conclusiones que se presenten. Sin embargo, es claro que las alternativas presentadas *solamente remiten a estudios no-experimentales*. Por fuera han quedado hasta ahora dos niveles de requisitos de diseño más generales que se introducen explícitamente cuando el objetivo es hallar una explicación causal sobre los efectos identificados mediante los métodos expuestos: la aleatorización y la identificación hipotética de los determinantes.

Sin entrar al debate sobre la causalidad en las ciencias sociales, debe reconocerse que actualmente existe una creciente tendencia a recomendar y discutir la necesidad de contar con diseños experimentales que permitan extraer conclusiones más certeras. En general se acepta que los estudios de las escuelas eficaces no pueden realizar inferencias causales debido a que el diseño no es apropiado para este objetivo. Raudenbush y Willms (1995) por ejemplo, señalan que las inferencias causales serían aceptables sólo bajo condiciones *experimentales* que aseguraran:

- i) asignación aleatoria de los alumnos a las escuelas
- ii) asignación aleatorias de las escuelas a los tratamientos (políticas) que teóricamente se suponen generan diferencias (por ejemplo, liderazgo, clima, enfoque didáctico, etc).

Cuando no es posible realizar una asignación aleatoria de alumnos y escuelas, una forma de sortear estos problemas se ha realizado a través de la ampliación de los controles estadísticos “*ex-ante*” y “*ex-post*” sobre aquellas variables que están relacionadas con el aprendizaje y con la pertenencia a determinada escuela. Sin embargo, este desplazamiento involucra de una forma muy fuerte a la teoría disponible, dado que sólo se está en condiciones de controlar aquellos factores determinantes o perturbadores sobre los cuales existen hipótesis (Kish, 1995). Esto abre nuevos problemas: en primer lugar, el consenso sociológico se reduce sobre un conjunto reducido (aunque muy relevante) de variables. Esto tiene que ver necesariamente con el ciclo teórico-empírico de construcción de teoría, tanto a nivel de los atributos individuales de los alumnos como a nivel organizacional o institucional de las escuelas. Y es de reconocer que la teoría de las escuelas eficaces se encuentra en un estadio embrionario de su desarrollo

donde la mayor parte de los factores que se reseñan como determinantes no están sólidamente operacionalizados. Como señalan Coe y Fitz-Gibon (1998), no está para nada claro qué es conceptualmente el liderazgo del cual se hablan en muchos estudios cualitativos ni menos aún cuales son sus dimensiones observables.

Pero, en segundo lugar, existen fuertes discrepancias entre los investigadores sobre las formas correctas de especificar los modelos de control estadístico, y se sabe que los resultados son sensibles a los modelos. El principio general se puede expresar de la siguiente forma:

*“Cuanto más dramáticamente difieren los grupos en las características antecedentes, más sensibles serán las inferencias a los diferentes métodos de ajuste y menos creíbles serán las inferencias resultantes”* (Raudenbush y Bryk, 2002: 155, traducción propia)

Aún suponiendo que se dispone de una amplia teoría que se pueda contrastar, existen otros requisitos metodológicos a completar si se desea una inferencia válida. Goldstein (1997) ha sintetizado cuatro requisitos mínimos que este tipo de investigaciones *cuasi-experimentales* debiera cumplir:

1. ser un estudio **longitudinal**, de tal forma que las diferencias pre-existentes en los estudiantes y los subsecuentes eventos en las escuelas puedan ser tomados en cuenta;
2. enfocar el análisis con un apropiado modelo **multinivel**, de tal forma que las inferencias estadísticas sean válidas y en particular, sean exploradas la existencia de “eficacias diferenciales” (esta noción será detallada más abajo);
3. disponer de cierta **replicación** del estudio en otros momentos y en otros lugares a los efectos de fundamentar la replicabilidad; y
4. disponer de cierta **explicación plausible** del proceso por el cual las escuelas devienen en eficaces.

Señala también Goldstein (1997) que de todos los estudios más frecuentemente citados en la bibliografía especializada, sólo el informe de Mortimore *et al.* (1988) cumple con estos requisitos; seguido parcialmente por Rutter *et al.* (1979). Esta recurrente falta de rigor metodológico está en la base de las agudas críticas que se han hecho a quienes trabajan en esta perspectiva en el mundo anglosajón (Coe y Fitz-Gibon, 1998).

Para finalizar es necesario sostener que los métodos estadísticos deben enfrentar un plano de cuestionamientos que resulta general a todos ellos y que se deriva de su carácter no-experimental. Una posición crítica más radical al respecto ha sido recientemente reiterada por Thomas Cook quien sostiene que:

*“La superioridad de la asignación aleatoria para la extracción de inferencias sobre las consecuencias de ensayos de cambios planificados es rutinariamente reconocida en la filosofía, la medicina, la salud pública, la agricultura, la estadística, la microeconomía, la psicología, la criminología, la investigación preventiva, la educación pre-escolar, el marketing y todas aquellas partes de la ciencia política y la sociología involucradas con el mejoramiento de las encuestas de opinión. [...] Sin embargo, la asignación aleatoria es relativamente rara en la investigación educativa, especialmente para la evaluación de intervenciones educativas de obvia relevancia en*

*las políticas. [...] Sin embargo, enunciados causales son rutinariamente hechos en este campo, usualmente a través de procesos que vinculan teorías sustantivas con varias prácticas cualitativas o cuantitativas no experimentales. Este artículo no argumenta que las conclusiones causales correctas provienen sólo de experimentos. Argumenta, sin embargo, que los experimentos provee de una mejor garantía para tales conclusiones que otros métodos. Por lo que si los experimentos pueden ser conducidos en la escuela, deben ser hechos. No usarlos requiere de una muy fuerte justificación". (Cook, 2003: 117).*

## 7. IMPLICANCIAS PARA FUTURAS INVESTIGACIONES

Este artículo ha delineado las alternativas más frecuentes que se encuentran en el plano de la conceptualización, en el plano del diseño y en el plano de la modelización estadística; dichas alternativas aparecen organizadas en la bibliografía como “métodos estadísticos de estimación”, en la medida en que se han enfatizado las discusiones relativas a las formas en cómo se estiman los parámetros de las ecuaciones que formalizan las conceptualizaciones. Aquí por el contrario, se ha enfatizado las relaciones entre aquellos niveles epistémicos, sosteniendo que la congruencia entre sus implicancias debiera ordenarse según el principio de isomorfía entre la teoría y las técnicas (Cortés y Rubalcaba, 1993).

En esta revisión también se han mostrado las debilidades implicadas en las nociones absolutistas de escuela eficaz que han derivado en “rankings” público; así como también los problemas de estimación que acarrea la falta de un diseño longitudinal y el desconocimiento de la estructura anidada de los datos. Respondiendo a esta lógica, una conclusión central del artículo reside en recomendar que los estudios de efectos de la escuela aceptan una noción de eficacia relativa o longitudinal, y se estimen los efectos utilizando modelos jerárquico-lineales, de preferencia de tres niveles (alumnos, escuelas, jurisdicciones). Esta conclusión ya es parte de las decisiones que distintos investigadores han adoptado y que han sido citados debidamente a lo largo de este artículo.

Este relativamente amplio consenso existente se contrasta con la investigación que se ha realizado en la región latinoamericana. Sin perjuicio de reconocer el valor académico y político que tienen la creciente masa crítica que se ha venido generando durante los últimos diez años en este campo de las escuelas eficaces, es preciso señalar que no toda la investigación realizada alcanza la misma calidad teórica y metodológica, precisamente por algunas de las razones aquí señaladas. Una observación crucial que tiene relación con la transformación de los hallazgos científicos en programas de política sobre la base de inferencias causales que en algunos casos carecen de sustento alguno.

Ahora bien, tanto el interés político por las escuelas eficaces, como la sospecha ideológica que han levantado y la crítica metodológica deben ser valorados como potentes motores para el desarrollo de nuevos estudios cada vez más rigurosos. Es necesario reconocer que se han hecho progresos gigantescos gracias a estos ingredientes y que estamos en condiciones de participar de una tercera generación de estudios en latinoamérica<sup>19</sup>. Pero, es necesario aceptar que deberemos ser más modestos y cautelosos en nuestras conclusiones y cada vez más reacios en aceptar “listados de factores”.

---

<sup>19</sup> Véase recientemente por ejemplo, Picaroni (2002) para Uruguay y Cueto, Ramírez, León y Pain (2003) para el Perú.

**REFERENCIAS BIBLIOGRÁFICAS**

- Aitkin, M. y Longford, N. (1986). Statistical Modelling Issues in School Effectiveness Studies. *Journal of the Royal Statistic Society* vo., 149. pp. 1-43.
- Ball, S. (1994). *Education Reform: A Critical and Post-Structural Approach*. Open University Press. Buckingham.
- Ball, S. (2003). *Class Strategies and the Education Market: The Middle Class and Social Advantage*. Routledge/Falmer. London
- Barbosa, M.E. y Fernandes, C. (2001).. *A escola brasileira faz diferenca? Uma investigacao dos efeitos da escola na proficiencia em Matemática dos alunos da 4ª série*. Laboratorio de Estadística Computacional. Departamento de Engenharia Eletrica. PUC. Río de Janeiro.
- Bryk, A. y Lee, V. (1992). “Is Politics the Problem and Markets the Answer?. An Essay Review of Politics, Markets and America ‘s Schools. *Economics of Education Review*, 11(4). pp. 439-451.
- Bryk, A.; Lee, V. y Holland, P. (1993). *Catholic Schools and the Common Good*. Cambridge University Press. MA.
- Bunge, M. (2000). *La investigación científica. Su estrategia y filosofía*. Primera Edición, corregida de la 2da. Edición española. Siglo XXI Editores de México. México, D.F.
- Campbell, D. y Stanley, J. (1982). *Diseños experimentales y cuasiexperimentales en investigación social*. Amorrortu. Buenos Aires.
- Cervini, R. (1999). *Calidad y equidad de la educación básica de Argentina*. Serie Factores asociados al logro escolar, informe n°5. Ministerio de Cultura y Educación de la Nación. Buenos Aires.
- Cervini, R. (2001). Efectos de la oportunidad de aprender sobre el logro en matemática en la educación básica argentina. *Revista Electrónica de Investigación Educativa*, 3(2). Consultado en: <http://redie.ens.uabc.mx/vol3no2/contenido-cervini.html>.
- Cervini, R. (2002a). Desigualdades en el logro académico y reproducción cultural en Argentina. Un modelo de tres niveles. *Revista Mexicana de Investigación Educativa*, VII (16).
- Cervini, R. (2002b). Desigualdades socioculturales en el aprendizaje de matemática y lengua de la educación secundaria en Argentina: un modelo de tres niveles. *Revista Electrónica de Investigación y Evaluación Educativa*, 8(2). Consultado en : [http://www.uv.es/RELIEVE/v8n1/RELIEVEv8n2\\_1.htm](http://www.uv.es/RELIEVE/v8n1/RELIEVEv8n2_1.htm).
- Coe, Robert y Fitz-Gibbon, Carol Taylor (1998). School effectiveness research: criticisms and recommendations. *Oxford Review of Education*, 24(4), pp. 421-438.
- Cook, Thomas (2003). Why Have Educational Evaluators Chosen Not to Do Randomized Experiments?. *Annals\_AAPSS* 589, pp. 114-149.
- Coleman, J.S; Hoffer, T. y Kilgore, S. (1982). *High School Achievement: Public, Catholic and Private Schools Compared*. Basic Books; New York.
- Cortés, F. y Rubalcaba, R.M. (1987). *Métodos estadísticos aplicados a la investigación en ciencias sociales. Análisis de asociación*. El Colegio de México. México, DF.
- Cortés, F. y Rubalcaba, R.M. (1993). Consideraciones sobre el uso de la estadística en las ciencias sociales. Estar a la moda o pensar un poco. En González Casanova (ed.), *Matemática y ciencias sociales*. Miguel Ángel Porrúa Editores. México DF.

- Cortés, F.; Rubalcaba, R.M. y Yocelovsky, Ri. (ed) (1990). Escalas básicas de medida En idem, op. cit. *Metodología volumen IV. Programa nacional de Formación de Profesores Universitarios en Ciencias Sociales*. Ed. SEP / Universidad de Guadalajara / Consejo Mexicano de Ciencias Sociales. México DF.
- Cueto, S.; Ramírez, C.; León, J. y Pain, O. (2003). *Oportunidades de aprendizaje y rendimiento en matemática en una muestra de estudiantes de sexto grado de primaria de Lima*. Documento de Trabajo. GRADE. Lima.
- Cueto, S. y Secada, W. (2003). Eficacia escolar en escuelas bilingües en puno, Perú. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en la Educación*, vol. 1 (1). Consultado en <http://www.rinace.org/reice/vol1n1/cuetoysesada.pdf>
- De Corte, E. (1994). Learning Theory and Instructional Science. En Reimann y Spada (eds.), *Learning in humans and machines. Towards an interdisciplinary learning science*.
- Edmonds, R. (1979). Effective Schools for the Urban Poor. *Educational Leadership*, 37, pp. 15-27.
- Fernández Aguerre, T. (1996). Escuelas eficaces: una perspectiva sociológica de la organización escolar. *Revista de Ciencias Sociales*, 11. Departamento de Sociología. Universidad de la República. Montevideo.
- Fernández Aguerre, T. (1999a) Efectividad en Educación. *Revista de Ciencias Sociales*, 16. Departamento de Sociología. Universidad de la República. Montevideo.
- Fernández Aguerre, T. (1999b). *Estudio de Caso en los grupos efectivos en Tercer Año de las Escuelas Públicas de Contextos Desfavorables del Uruguay*. Unidad de Medición de Resultados Educativos (UMRE). Inédito. Montevideo.
- Fernández Aguerre, T. (2001). *Contribución al análisis organizacional en educación*. Editado por Facultad de Ciencias Sociales / Plural Editores. Montevideo
- Fernández Aguerre, T. (2002). Determinantes sociales e institucionales de la desigualdad educativa en sexto año de educación primaria de Argentina y Uruguay, 1999. Una aproximación mediante un modelo de regresión logística. *Revista Mexicana de Investigación Educativa*, VII (16).
- Fernández Aguerre, T. (2003a). La desigualdad educativa en el Uruguay entre 1996 y 1999. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en la Educación*. 1(1). Consultado en <http://www.rinace.org/reice/vol1n1/fernandez.pdf>
- Fernández Aguerre, T. (2003b) *Determinantes sociales y organizacionales del aprendizaje en la Educación Primaria de México. Un análisis de tres niveles*. Informe de Investigación para el Instituto Nacional para la Evaluación de la Educación (INEE) de México. México , D.F.
- Fernández Aguerre, T. (2003c). *Escuelas primarias eficaces de México*. Informe de Investigación para el Instituto Nacional para la Evaluación de la Educación (INEE) de México. México , D.F.
- Fernández Aguerre, T. (en prensa). De las “escuelas eficaces” a las reformas educativas de “segunda generación. *Estudios Sociológicos*, 65, Centro de Estudios Sociológicos, El Colegio de México, México, D.F.
- García, R. (2001). Fundamentación de una epistemología en las ciencias sociales. *Estudios Sociológicos*, 57, Centro de Estudios Sociológicos, El Colegio de México. México, D.F.
- García Ferrando, M.; Ibañez, J. y Alvira, F. (eds.) (2000). *El análisis de la realidad social. Métodos y técnicas de investigación. Tercera edición*. Alianza Editorial. Madrid.

- Goldberger, A. y Cain, G. (1982). The Causal Analysis of Cognitive Outcomes in the Coleman, Hoffer and Kilgore Report. *Sociology of Education*, 55, pp. 103-122.
- Goldstein, H. (1997). Methods in School Effectiveness Research. *School Effectiveness and Improvement*. 8 pp. 369-395.
- Goldstein, H. y Woodhouse, G. (2000). School Effectiveness and Educational Policy. *Oxford Review of Education*, 26 (3-4), pp 353-363.
- Greene, W. (1999). *Análisis Económico. Tercera Edición*. Prentice Hall. México DF.
- Gujarati, D. (2004). *Econometría. Cuarta Edición*. McGrawHill. México, D.F.
- Kish, L. (1995). *Diseños estadísticos para la investigación social*. Editado por Centro de Investigaciones Sociológicas. Madrid.
- Lee, V. (2001). *Restructuring High Schools for Equity and Excellence. What Works*. Sociology of Education Series. Teacher's College Press. Columbia University. New York.
- Lee, V. y Bryk, A. (1989). A Multilevel Model of the Social Distribution of High School Achievement. *Sociology of Education*, 62(3), pp.172-192.
- Lee, V.; Bryk, A. y Smith, J. (1993). The Organization of Effective Secondary Schools. *Review of Research in Education*, 19, pp. 171-267.
- LLECE (2000). *Primer estudio internacional comparativo en Lenguaje, Matemática y factores asociados Segundo Informe*. UNESCO. Santiago.
- LLECE (2002). *Estudio cualitativo de escuelas con resultados destacables en siete países latinoamericanos*. UNESCO, Santiago de Chile.
- Martin, M; Mullis, I; Gregory, K.; Hoyle, C.; Shen, C. (2000). *Effective Schools in Science and Mathematics*. Third International Mathematics And Science Study (TIMSS) International Association for the Evaluation of Educational Achievement (IEA) / Boston College. MA.
- Mizala, A. y Romaguera, P. (2000). *Determinación de factores explicativos de los resultados escolares en educación media en Chile*. Serie Economía N° 85. Centro de Economía Aplicada, Departamento de Ingeniería Industrial Facultad de Ciencias Físicas y Matemáticas Universidad de Chile. Santiago de Chile.
- Mizala, A.; Romaguera, P. y Reinaga, T. (2000). *Factores que inciden en el rendimiento escolar en Bolivia*. Centro de Economía Aplicada de la Universidad de Chile. Santiago.
- Mortimore, P.; Sammons, P. ; Stoll, L.; Lewis, D y Ecob, R. (1988). *School Matters*. London: Open Books.
- Murillo, F.J. (2003) Una panorámica de la investigación iberoamericana sobre eficacia escolar. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en la Educación*, 1(1) En <http://www.ice.deusto.es/rinace/reice/vol1n1/Murillo.pdf>.
- Murillo, F.J.; Barrio, R. y Pérez-Albo, M.J. (1999). *La dirección escolar: análisis e investigación*. Madrid: Ed. Centro de Investigación y Desarrollo de la Educación (CIDE). Madrid.
- OECD - PISA (2001). *Knowledge and Skills for Life. First Results for the OECD Programme for International Student Assessment (PISA) 2000*. OECD. Paris.

- OECD - PISA (2003). *Learners for Life. Student approaches to learning. Results from PISA 2000*. OECD. Paris. France.
- Picaroni, B. (2002). *Cuando la escuela marca la diferencia*. Inédita. Tesis de maestría en Políticas Públicas. Universidad ORT del Uruguay. Montevideo.
- Rama, G. et al. (1991). *Quiénes y qué aprenden en las escuelas primarias del Uruguay*. Oficina de la CEPAL: Montevideo.
- Raudenbush, S. y Bryk, A. (2002). *Hierarchical Linear Models . Second Edition*. Ed. Sage. Thousand Oaks. CA.
- Raudenbush, S. y Willms, D. (1995). The Estimation of School Effects. *Journal of Educational and Behavioral Statistics*, 20(4), pp 307-335.
- Ravela, P. (1993). *Escuelas productoras de conocimientos en los contextos socioculturales más desfavorables*. Ed. CEPAL, Oficina de Montevideo.
- Ravela, P.; Picaroni, B.; Cardozo, M.; Fernández, T., Gonet, D.; Loureiro, G. y Luaces, O. (1999). *Factores institucionales y pedagógicos explicativos de los aprendizajes. Cuarto Informe de la Evaluación Nacional de Aprendizajes en Sextos Años de Educación Primaria*. Ed. UMRE-MECAEP-ANEP. Montevideo.
- Reynolds, D.; Bollen, R.; Creemes, B.; Hopkins, D.; Stoll, L. y Lagerweij, N. (1998). *Las escuelas eficaces. Claves para mejorar la enseñanza*. Editorial Santillana / Aula XXI. México DF.
- Reynolds, D.; Creemers, B.; Nesselrodt, P.; Schaffer, E.; Stringfield, S. y Teddie, C. (1994) *Advances in school effectiveness research and practices*. Oxford: Pergamon.
- Richards, C. (1991). The meaning and measure of school effectiveness”. En J. BLISS; ; W. Firestone y C. Richards, *Rethinking Effective Schools: Research and Practice*. Prentice Hall. NJ
- Rowan, B. (1983). Research on Effective Schools: a cautionary note@. *Educational Researcher* , 12 (4).
- Schafer, W.; YEN, S.J. y Rahman, T. (2000). School Effects Indices: Stability of One & Two Level formulations. *Journal of Experimental Education*, 68 (3).
- SEP –México (2000). Distribución de los planteles públicos de educación primaria y secundaria según el nivel de aciertos de sus alumnos en los exámenes de carrera magisterial. Ed. SEP, Subsecretaría de Planeación y Coordinación, Dirección General de Evaluación. México Df.
- SEP –México (2001). *¿Cómo transformar las escuelas? Lecciones desde la gestión escolar y la práctica pedagógica - Resumen*. Ed. SubSecretaría de Planeación y Coordinación. Dirección General de Evaluación. México DF.
- Slavin, R. (1996). *Salas de clase efectivas, escuelas efectivas: plataforma de investigación para una Reforma Educativa en América Latina*. PREAL. Santiago.
- Stevans, L.K y Sessions, D.N. (2000) Private/public school choice and student performance revisited. *Education Economics*, 8(2).
- Willmott, R. (1999a). Structure, agency and school effectiveness: researching a ‘failing school’. *Educational Studies*, 25(1), pp. 5-18.
- Willmott, R. (1999b). School effectiveness research: an ideological commitment?. *Journal of Philosophy of Education*, . 33, pp. 253-268.