

Un análisis de sus resultados en las evaluaciones estandarizadas de 1996 y 1999

Calidad y equidad en las escuelas de tiempo completo

por Manuel Cardoso

La equidad y sus “enemigos”

En el ámbito de las políticas educativas existe un largo debate entre autores y diseñadores de políticas que privilegian el criterio de la equidad frente a quienes priorizan la calidad y/o la eficiencia. También están quienes intentan realizar síntesis armónicas o relativizar estas dicotomías.

Una de las mayores controversias plantea un dilema entre el logro de una elevada calidad educativa con el logro de un grado satisfactorio de equidad. Este debate podría remitirnos a la discusión sobre las funciones del sistema educativo formal, pero ello sobrepasaría ampliamente las ambiciones de este artículo. El dilema *calidad o equidad* no es trivial, desde luego, ya que influye directamente en el diseño de políticas de asignación de recursos dentro de los sistemas educativos. Los docentes —razonablemente preocupados por lo que ocurre en los salones de clase con relación a este dilema— a menudo manifiestan su temor a la tendencia a “igualar para abajo”, inclinación que observan a veces en algunos de sus colegas o en las medidas de política educativa que se implementan desde los niveles superiores del sistema. No es necesario abundar sobre los efectos nocivos de sacrificar la calidad de la educación en pos de la equidad.

El autor. Máster en Educación por la Universidad de Harvard. Licenciado en Sociología por la Universidad de la República. Docente universitario. Consultor de ANEP.

La postura inversa, privilegiar la calidad en desmedro de la equidad, tiene varios correlatos empíricos: la exclusión de sectores de la población que no acceden a la educación formal (baja cobertura), que son expulsados (alta deserción), o que son retenidos en niveles inferiores del sistema (alta repetición). Estos tres factores hacen que determinados sistemas educativos muestren resultados muy favorables llegado el final del ciclo primario, ya que los estudiantes a los que se considera menos capaces de cumplir con los requerimientos de la carrera académica han sido excluidos de —o retrasados en— su pasaje por el sistema. Una forma más sutil de privilegiar la calidad sobre la equidad consiste en ignorar la dispersión de los resultados y la incidencia de los factores sociales cuando se analizan los desempeños académicos, limitándose simplemente a reportar promedios generales, sin mostrar las diferencias existentes entre las distintas escuelas, o entre distintos estudiantes de la misma escuela, ni las relaciones de estas diferencias con los orígenes culturales y económicos de los estudiantes.

Otro concepto que a menudo se contrapone al de equidad es el de *eficiencia*. Este es un punto de debate en el ámbito de la economía, aplicado a distintos sectores, no solo al de la educación. Sin embargo, para algunos autores esta discusión toma otro cariz cuando se refiere a la opción que los sistemas educativos enfrentan, entre asegurar el acceso igualitario a la educación para todos o maximizar los beneficios obtenidos a través de focalizar los esfuerzos de los sistemas en aquellos sectores que están mejor preparados para recibir educación formal, que son también, paradójicamente, los que menos la necesitan.

Según Pérez Morgan:

La equidad en materia educativa no está emparentada deliberadamente con la eficiencia. Si el criterio es el de buscar eficiencia en las asignaciones de recursos, los individuos involucrados en determinadas circunstancias serían dejados de lado. El criterio de asignar más fondos de un presupuesto fijo y por definición escaso, a los menos capacitados, disminuye la producción total, ya que el rendimiento marginal de los menos capacitados, es menor que el de los más capacitados. Por lo que la equidad supone en ciertos casos, no relacionarla necesariamente con la eficiencia porque se establece una prioridad normativa de la primera sobre la segunda.¹

En esta visión, valorar la eficiencia probablemente implicaría dejar de lado la equidad, ya que se supone que estos dos objetivos no pueden ser alcanzados simultáneamente.

¹ B. Pérez Morgan: Evaluación del Programa de Escuelas de Tiempo Completo propuesto por el CODICEN de la ANEP en 1995 desde el punto de vista de la equidad, CLAEH, Serie Investigaciones, Montevideo, 2001, p. 9.

Sin embargo, otros autores, como Okun, plantean visiones diferentes:

Lo que es bueno para la igualdad puede ser bueno para la eficiencia [...] Cuando podemos tener más justicia y más PBI, la sociedad debería aprovecharlo al máximo.²

Para Okun el ámbito de la educación es precisamente aquel donde el dilema equidad-eficiencia parece más fácil de resolver:

Igualar las oportunidades financieras para la educación superior es una de las formas por las cuales la nación puede obtener mayor eficiencia y mayor igualdad, sin sacrificar una por la otra.³

Si bien Okun habla específicamente de la educación universitaria (que en su país, Estados Unidos, representa el clivaje educativo más importante), este razonamiento es igualmente aplicable a otros países y otros niveles del sistema educativo formal. La idea no tiene nada de novedoso: se trata nada menos que de la vieja noción de igualdad de oportunidades, de la cual la reforma educativa vareliana estaba impregnada —por ejemplo, cuando hablaba de la importancia de que los individuos no se distinguieran más que por sus “virtudes y aptitudes”—. En definitiva, *si el sistema educativo logra que las oportunidades estén disponibles para los individuos en función de sus virtudes y aptitudes, y no de sus orígenes sociales, se estará contribuyendo al mismo tiempo a la eficiencia y a la equidad; ya que el acceso de los individuos a las distintas posiciones sociales se dará en función de la capacidad de éstos para cumplir los roles correspondientes* (lo que redundará en una mayor eficiencia) *y al mismo tiempo no se estará excluyendo a individuos a partir de características que no interfieren en el cumplimiento de estos roles* (lo que redundará en una mayor equidad).

Por otra parte, uno de los mayores problemas al aplicar el concepto de eficiencia al ámbito de la educación es que no resulta fácil cuantificar los beneficios que esta genera, ya que muchas veces tales beneficios trascienden a los receptores del servicio, lo que redundará en una mejora en la calidad de vida de la sociedad en su conjunto.

Los beneficios para los participantes del programa son los beneficios directos del programa. Los beneficios para los contribuyentes son los beneficios in-

² A. Okun: *Equality and Efficiency: The Big Tradeoff*, Brookings, Washington, 1975, p 79. Texto original inglés: “*what is good for equality may be good for efficiency [...] When we can have more justice and more real GNP, society should make the most of it*”.

³ *Ibidem*, p. 82. Texto original inglés: “*equalizing financing opportunities for higher education is one of the ways by which the nation can obtain more efficiency and more equality - without sacrificing one for the other*”.

directos, o externalidades, producidos como un subproducto de los efectos directos sobre los participantes.⁴

Al distinguir los beneficios directos de las externalidades, e incluir ambos en el análisis, estamos considerando la totalidad de los efectos favorables que la educación puede llegar a tener sobre la sociedad. Cuando se emprenda un estudio de costo-beneficio de las escuelas de tiempo completo, por ejemplo, será necesario tener en cuenta no solamente los beneficios directos para los niños participantes. Las escuelas de tiempo completo podrían constituir un apoyo a los padres (y en particular a las madres) que facilitara su inserción laboral, garantizando al mismo tiempo que los niños permanecen en un ambiente seguro y protegido de una serie de factores de marginalización.

Finalmente, algunos autores han intentado armonizar los criterios de calidad, equidad y eficiencia en modelos que guíen las opciones de política educativa. Hallak sugiere que el papel del Estado en esta materia debe estar guiado por un criterio de equidad en cuanto a la financiación, un criterio de costo/efectividad (o eficiencia) en lo referente a la provisión de servicios educativos, y de calidad en lo que se refiere a la regulación. En otras palabras, el Estado debería asignar recursos para asegurar la educación de aquellos que no pueden pagarla, procurar que en la prestación de los servicios no se desperdicien recursos, y asegurar la calidad mediante el establecimiento de regulaciones.

En el marco de nuestra región, Uruguay tiene una fuerte tradición de equidad (tanto en la teoría como en la práctica) en su sistema educativo, que a menudo es entendida como uno de los pilares de su relativa equidad económica y social. Sin embargo, una serie de cambios que se han producido en el país en las últimas décadas, que abarcan desde aspectos demográficos hasta políticos y económicos, han puesto seriamente en cuestión los fundamentos de esta presunta equidad de nuestro sistema educativo. En 1996, la realización de la primera Evaluación Nacional de Aprendizajes permitió cuantificar la magnitud de esta inequidad, mostrando claras diferencias entre la proporción de estudiantes que alcanzaban rendimientos satisfactorios en lenguaje y matemáticas en familias de distintos niveles culturales y económicos, y en centros de distintos "contextos socioculturales".⁵ En 1996, el 66,4% de los

⁴ W. Barnett: *Lives in the Balance: Age-27 Benefit-Cost Analysis of the High/Scope Perry Preschool Program Ypsilanti*, High/Scope, Michigan, 1996, p. 66. Texto original inglés: "The benefits to program participants are the direct benefits of the program. The benefits to taxpayers are the indirect benefits, or externalities, produced as a byproduct of the direct effects on the participants".

⁵ El contexto sociocultural es una medida agregada del nivel sociocultural de las familias que envían a sus hijos a una escuela. Se calcula teniendo en cuenta el nivel educativo de las madres de los niños evaluados y el nivel de equipamiento de sus hogares.

niños del contexto muy favorable obtuvieron la suficiencia en matemática, frente al 16,7% de los niños del contexto muy desfavorable.

Equidad y políticas educativas

¿Qué medidas pueden tomarse, en materia de política educativa, para reducir estas inequidades? Basil Bernstein sugiere:

Es posible crear una pedagogía visible que reduzca la relación entre clase social y resultados educativos. Esto puede exigir una estructura preescolar de apoyo, la relajación del enmarcamiento respecto a las reglas de ritmo y secuencia y la relajación del enmarcamiento que regula el flujo de la comunicación entre la clase escolar y la comunidad o comunidades sobre las que opera la escuela.⁶

Sin entrar en detalles acerca de lo que Bernstein llama una *pedagogía visible*, resulta llamativo observar la coincidencia entre sus propuestas y algunas de las medidas fundamentales en materia de educación primaria emprendidas en el marco de la reforma del período 1995-1999. Llamativamente, estas fueron también algunas de las políticas más controvertidas de dicha administración. La *estructura preescolar de apoyo* fue objeto de una política de universalización al nivel de 5 años —lo cual implicó una reducción de la cobertura en otros niveles—, precisamente en el entendido de que esta era una forma de lograr una mayor equidad de la escuela primaria, al contribuir a igualar los puntos de partida y disminuir el choque experimentado por los niños de sectores bajos al ingresar a la escuela. Lo que Bernstein llama “la relajación del enmarcamiento que regula el flujo de la comunicación entre la clase escolar y la comunidad o comunidades sobre las que opera la escuela” podría abarcar, por ejemplo, al programa “Todos los niños pueden aprender”, que intenta incorporar a los padres al proceso de aprendizaje de sus hijos a través, entre otros mecanismos, de la edición de un periódico (*Dame una Mano*). Pero, ¿qué quiere decir Bernstein con “la relajación del enmarcamiento respecto a las reglas de ritmo y secuencia”? Por *ritmo* entiende “la tasa esperada de adquisición, es decir, la tasa a la que se espera que se produzca el aprendizaje [...] Las reglas de ritmo regulan el ritmo de la transmisión, y la velocidad de este ritmo puede variar”.⁷ Por lo tanto, la relajación del ritmo implica la permanencia del niño en la escuela por un período mayor, lo cual obviamente nos lleva a las escuelas de tiempo completo, que experimentaron

⁶ B. Bernstein: *La estructura del discurso pedagógico*, Morata, Madrid, 1993, p. 88.

⁷ *Ibíd.*, p. 85

un importante crecimiento a lo largo de este período; también en esta línea podríamos mencionar la política de eliminación de la repetición en primer año.

Si bien la preocupación central de Bernstein es claramente la equidad, una vez más la cuestión de la eficiencia aparece como problemática:

Tal relajación del enmarcamiento de una pedagogía visible eleva el coste de la transmisión y tiene consecuencias importantísimas en la preparación de los profesores y la gestión de los centros.⁸

Siguiendo a Bernstein, hipotetizaremos que las escuelas de tiempo completo, al posibilitar un relajamiento de las reglas de ritmo, permitirán no solo el logro de una mayor calidad, sino también de una mayor equidad.

Hipótesis

1. A mayor nivel socioeconómico promedio de una escuela, mejores resultados académicos promedio.
2. Dentro de una escuela determinada, a mayor nivel socioeconómico de un niño, mejores resultados académicos de dicho niño.
3. En la actualidad, las escuelas de tiempo completo muestran niveles más elevados de calidad que las escuelas urbanas comunes, una vez que se controla el efecto del nivel socioeconómico de los estudiantes. Esto implica que el promedio de los resultados académicos de las escuelas de tiempo completo deberá ser más alto que el promedio de los resultados académicos de las escuelas urbanas comunes, una vez que se controlan los efectos del nivel socioeconómico.
4. En la actualidad, las escuelas de tiempo completo muestran niveles más elevados de equidad que las escuelas urbanas comunes, una vez que se controla el efecto del nivel socioeconómico de los estudiantes. Esto implica que la asociación entre los niveles socioeconómicos de los estudiantes y sus resultados académicos en las escuelas de tiempo completo deberá ser más baja que la asociación entre los niveles socioeconómicos de los estudiantes y sus resultados académicos en las escuelas urbanas comunes, una vez que se controlan los efectos del nivel socioeconómico.
5. La diferencia en calidad a favor de las escuelas de tiempo completo (hipótesis 3) aumentó en el período 1996-1999.
6. La diferencia en equidad a favor de las escuelas de tiempo completo (hipótesis 4) aumentó en el período 1996-1999.

⁸ *Ibíd.*, p. 88.

El problema metodológico

La investigación social de corte cuantitativo ha evolucionado influida por los desarrollos de otras disciplinas científicas, cuyas problemáticas son diferentes. Como resultado de ello, en ocasiones las técnicas estadísticas aplicadas no resultan idóneas para el conocimiento de organizaciones tales como los centros educativos. En una escuela, los resultados académicos que un niño obtiene son producto de un conjunto de factores relativamente específicos de su situación particular (su herencia genética, el tipo de crianza que le ha dado su familia, los recursos de que dispone en el hogar para facilitar su aprendizaje, etc.) y de otro conjunto de experiencias que comparte con sus compañeros de clase (el tipo de interacción que tiene con su maestra, la infraestructura y el ambiente de la escuela, las características del barrio, etc.). Todo esto resulta altamente intuitivo para toda persona que conozca la realidad de una escuela. Sin embargo, cuando se trata de explicar los resultados de las escuelas utilizando modelos cuantitativos y trabajando con ambos tipos de factores (del niño y de la escuela) como predictores, surgen problemas estadísticos, dado que ciertos supuestos sobre los cuales trabajan estos modelos no se cumplen en el caso de las organizaciones educativas.

Según Bryk:

Hay dos procedimientos obvios que se han visto algo desacreditados. El primero es desagregar todas las variables de orden superior al nivel de los individuos. Las características del maestro, la clase y la escuela son todas ellas asignadas al individuo, y el análisis es realizado al nivel del individuo. El problema de este enfoque es que, si sabemos que los estudiantes están en la misma clase, también sabemos que tienen el mismo valor en cada una de las variables de la clase. Por lo tanto, no podemos usar el supuesto de independencia de las observaciones, que es básico para las técnicas estadísticas clásicas.⁹

El incumplimiento de este supuesto genera una serie de problemas computacionales que terminan invalidando las conclusiones de este tipo de estudios. Sin embargo, no es infrecuente encontrar estudios que trabajan con esta metodología.

⁹ A. Bryk y S. Raudenbusch: *Hierarchical Linear Models*, Sage, Newbury Park, 1992, p. xiv. Texto original inglés: "There are two obvious procedures that have been somewhat discredited. The first is to disaggregate all higher order variables to the individual level. Teacher, class, and school characteristics are all assigned to the individual, and the analysis is done on the individual level. The problem with this approach is that if we know that students are in the same class, we also know that they have the same value on each of the class variables. Thus we cannot use the assumption of independence of observations that is basic for the classical statistical techniques".

El otro procedimiento tiene consecuencias menos graves, y de hecho nosotros mismos lo hemos utilizado, pero también tiene, como apunta Bryk, algunas desventajas importantes:

La otra alternativa es agregar las variables del nivel del individuo al nivel superior y hacer el análisis en el nivel superior. Por lo tanto, agregamos las características de los estudiantes por clase y hacemos un análisis por clase, quizás ponderado por el tamaño de la clase. El principal problema es que descartamos toda la información intragrupos, que puede llegar a ser el 80 o 90% de la variación total antes de comenzar el análisis. Como consecuencia, las relaciones entre las variables agregadas son a menudo mucho más fuertes, y pueden ser muy diferentes de las relaciones entre las variables no agregadas. De este modo desperdiciamos información, y distorsionamos la interpretación si tratamos de interpretar el análisis agregado en el nivel individual. Es así que tanto agregar como desagregar resulta insatisfactorio.¹⁰

El *Estudio de los factores institucionales y pedagógicos que inciden en los aprendizajes en escuelas primarias de contextos sociales desfavorecidos en el Uruguay*, publicado por la Unidad de Medición de Resultados Educativos, se basa en esta metodología que, como mencionaban Bryk y Raudenbusch, no toma en cuenta una proporción importante de la información disponible y podría, teóricamente al menos, llevar a conclusiones erróneas. Digamos en nuestro descargo que, a diferencia de la alternativa anterior, esta metodología no implica ninguna violación de los supuestos de la estadística clásica y constituye una práctica habitual hasta el momento, criticada solamente por los defensores del (relativamente novedoso) análisis multinivel.

A partir de los años ochenta, investigadores europeos y americanos han estado trabajando en el diseño y la aplicación de técnicas específicas para el tratamiento de objetos de estudio tan complejos como las organizaciones educativas. Estas técnicas (conocidas en las márgenes del Atlántico como *análisis multinivel* o *modelos lineales jerárquicos*) procuran tener en cuenta el hecho de que los factores explicativos de los rendimientos estudiantiles se ubican en distintos niveles: el nivel individual, la clase y la escuela. Esta nueva metodología, además de no estar sujeta a las limitaciones de las alternati-

¹⁰ *Ibidem*, p. xiv. Texto original inglés: "The other alternative is to aggregate the individual-level variables to the higher level and do the analysis on the higher level. Thus we aggregate student characteristics over classes and do a class analysis, perhaps weighted with class size. The main problem is that we throw away all the within-group information, which may be as much as 80% or 90% of the total variation before we start the analysis. As a consequence, relations between aggregated variables are often much stronger, and they can be very different from the relation between the nonaggregate variables. Thus we waste information, and we distort interpretation if we try to interpret the aggregate analysis on the individual level. Thus aggregating and disaggregating are both unsatisfactory".

vas previamente descartadas, incorpora algunos elementos que la hacen ideal para el tratamiento de temas en los que las cuestiones de la calidad y la equidad aparezcan como esenciales. ¿Cómo se manifiesta la calidad en un modelo jerárquico lineal aplicado a la investigación educativa? Los centros con mayor calidad educativa exhibirán resultados promedio más altos que los demás. ¿Cómo se manifiesta la equidad en dichos modelos? Los centros con mayor equidad educativa exhibirán asociaciones más leves entre el nivel socioeconómico y los resultados académicos que los demás centros. ¿Cómo se representa habitualmente la asociación entre dos variables como las mencionadas? Lo más habitual es calcular coeficientes de correlación y trazar rectas de regresión. Estas rectas tienen, cada una, una determinada pendiente que indica la fuerza de la relación entre las variables: cuanto más empinada la pendiente, mayor será la relación entre las variables. Por lo tanto, los centros con mayor equidad exhibirán menores pendientes que los demás. En definitiva, *para cada centro obtendremos un promedio de resultados académicos que será un indicador de su calidad, y una pendiente que será indicador de su equidad.*

Pero nosotros no deseamos extraer conclusiones para ninguna escuela en particular, sino para el conjunto de las escuelas de tiempo completo en relación con las escuelas urbanas públicas comunes; por lo tanto, compararemos el promedio de los promedios (por así decirlo) de las escuelas de tiempo completo con el equivalente de las escuelas urbanas públicas comunes, y haremos algo similar con las pendientes.

La variable dependiente: competencia matemática

En el momento de seleccionar la variable dependiente, se nos plantearon varias restricciones, y algunos dilemas. Por un lado, nos encontramos limitados por la inexistencia de algunos indicadores y la escasa confiabilidad de otros; por otro, disponíamos de resultados académicos de las escuelas que podían ser procesados de distintas formas, y debimos optar por una de estas alternativas.

En primer lugar, si bien se trata de evaluar los resultados del programa de tiempo completo, hay que tener en cuenta que no son muchos los indicadores disponibles en lo que se refiere al desempeño de las escuelas en nuestro país. Los centros pueden tener distintos tipos de impactos en sus estudiantes y, si bien la mayoría de quienes trabajan en educación coinciden en afirmar que las escuelas deberían apuntar a contribuir fundamentalmente al desarrollo cognitivo de sus estudiantes, muchos subrayan además la importancia de los aspectos sociales y morales. Sin embargo, la medición de estos aspectos resulta a menudo más compleja que la de los aspectos cognitivos, y

en nuestro país no existen ejemplos de este tipo de evaluaciones a gran escala. Por lo tanto, decidimos centrarnos fundamentalmente en aquellos indicadores que dan cuenta del rendimiento académico de los estudiantes.

Por otra parte, algunos indicadores que el propio sistema emplea regularmente presentan algunos problemas para su uso en investigación; un claro ejemplo de esto es la repetición, que, si bien es muy útil a los efectos de la gestión de los centros, a menudo no se aplica con criterio uniforme a través de estos. Mientras desde la perspectiva de los docentes esto puede responder a la necesidad de adaptar los criterios a los contextos específicos, o a cierta diversidad de opiniones en materia pedagógica, desde el punto de vista de la investigación acaba constituyéndose en un problema de confiabilidad de la información, que incrementa el error en la medición de la variable dependiente.

La deserción resulta, en primera instancia, un indicador más valioso que la repetición, pero tampoco está exento de problemas de confiabilidad (por ejemplo, a veces una deserción no es registrada como tal por el centro porque los padres piden un pase, que luego no retiran). En la medida en que nuestro estudio trata de medir el impacto de los centros en el estudiante sobre el final de su ciclo escolar en el nivel primario, se nos presentaba un problema adicional: en Uruguay la deserción es relativamente baja, lo cual redundaba en que este indicador tenga poca variabilidad, aproximándose a 0% en una proporción importante de las escuelas.

Por lo tanto, optamos por trabajar a partir de resultados académicos tal como son medidos por las evaluaciones estandarizadas de la Unidad de Medición de Resultados Educativos. En sexto año ha habido dos instancias de evaluación de este tipo: una con carácter censal en 1996 y otra basada en una muestra aleatoria en 1999. En esta última instancia, junto con la muestra de escuelas públicas y privadas, se aplicó la prueba a la totalidad de las escuelas de tiempo completo, lo cual implica que para este tipo de escuelas la prueba tuvo, nuevamente, carácter censal. Estas dos instancias tuvieron carácter equivalente en la medida en que se aplicaron pruebas de matemática y lenguaje compuestas cada una por 24 ítems, los cuales medían el mismo tipo de competencias con similares niveles de dificultad en una y otra instancia. Una vez tomada la decisión de trabajar sobre estos resultados, se nos planteó la disyuntiva de trabajar con el puntaje total de la prueba (de 0 a 48 puntos) o solamente con el puntaje de una de las dos áreas. En ese momento tuvimos contacto por primera vez con la *teoría de respuesta al ítem*, lo cual nos planteó un nuevo dilema, pero que en definitiva nos ayudó a resolver el anterior.

La teoría de respuesta al ítem (TRI) se desarrolló como alternativa a la llamada teoría clásica de los tests (TCT), frente a la cual presenta varias ven-

tajas, muchas de ellas irrelevantes a los efectos de nuestro análisis. Una ventaja, sin embargo, resulta digna de mencionar: la TRI permite un mayor grado de precisión en la medición de una determinada competencia, en tanto genera una medida continua en función de las respuestas a un conjunto de ítems dicotómicos.¹¹ Al evaluar a una población o muestra utilizando TCT y una prueba compuesta por 24 ítems dicotómicos, cada uno de los evaluados recibe generalmente una puntuación que expresa la cantidad de ítems que respondió correctamente; o sea que un determinado individuo puede obtener una puntuación de 0, 24 ó 12 en función de cuántos ítems “acertó”. Esto no presenta mayores problemas en el caso de quien responde correctamente a todos o a ninguno, pero ¿qué ocurre cuando dos individuos responden correctamente 12 ítems, pero uno de ellos acertó los 12 más fáciles y el otro los 12 más difíciles de la prueba? La TRI toma esto en cuenta y genera una medida continua, donde dos individuos que respondieron correctamente la misma cantidad de ítems pueden recibir puntajes diferentes en función del grado de dificultad de esos ítems. Para resumir, en el momento de calcular puntajes, la TCT solo toma en cuenta cuántos ítems fueron respondidos correctamente; en cambio, la TRI toma en cuenta, además de cuántos, cuáles.

Por lo tanto, la TRI parecía una alternativa útil para incrementar la precisión de la medición de nuestra variable dependiente, y su variabilidad. Como contrapartida, para aplicar un modelo de este tipo era necesario que se cumplieran dos supuestos:

1. El *supuesto de unidimensionalidad* exige que todos los ítems de una prueba midan la misma habilidad; por ejemplo, la competencia matemática o la competencia lingüística. Este supuesto nos obliga a descartar la posibilidad de trabajar con el puntaje total de la prueba (de 0 a 48) y nos lleva necesariamente a optar por una de las dos áreas de la prueba, o bien a trabajar con ambas, pero de forma independiente.
2. El *supuesto de independencia local* exige que las respuestas a los ítems sean estadísticamente independientes entre sí una vez que se controla el nivel de la competencia evaluada. Esto puede no ocurrir cuando la resolución de un ítem exige que se haya resuelto correcta-

¹¹ En evaluación cuantitativa se utiliza a menudo la expresión *ítems dicotómicos* para referirse a ítems cuyas respuestas son codificadas como “correctas” o “incorrectas”, con independencia de que originalmente las posibles respuestas al ítem fueran más de dos. Típicamente, estos son ítems de opción múltiple con una clave y tres o cuatro distractores, al dicotomizar el ítem se recodifica la clave como “respuesta correcta” y los tres o cuatro distractores como “respuesta incorrecta”. El uso de más de un distractor responde a varias razones, tanto estadísticas como pedagógicas: por un lado, el aumento en el número de distractores reduce las probabilidades de acierto al azar; además, distintos distractores pueden dar cuenta de distintos tipos de errores en el razonamiento de los evaluados.

mente un ítem anterior (por ejemplo, cuando el resultado del ítem 1 es un insumo necesario para poder resolver el ítem 2), o incluso podría llegar a ocurrir, aunque es menos probable, cuando varios ítemes tienen un mismo referente textual o gráfico. Este supuesto podría hacer problemático trabajar con la competencia lingüística, dado que nuestras pruebas de comprensión de texto y reflexiones sobre el lenguaje se basan en dos textos (uno argumentativo y uno narrativo) y se componen de un gran número de preguntas referidas a cada uno de ellos. Por lo tanto, resulta preferible en esta primera instancia trabajar con la competencia matemática, dado que, por la estructura de la prueba, ofrece mayores garantías de cumplimiento del supuesto de independencia local.

Fue así que finalmente optamos por utilizar la competencia matemática, calculada a partir de la teoría de respuesta al ítem, como variable dependiente de nuestro análisis. Esto no significa, ni mucho menos, que consideremos que este sea el único ni el más importante de los resultados de la actividad de las escuelas; por el contrario, podrían haberse utilizado otros indicadores de resultado, como la competencia lingüística, la deserción o la repetición, pero, como hemos mencionado, cada uno de ellos planteaba diversas dificultades metodológicas que no los hacían óptimos para este tipo de análisis.

La variable de política: tiempo completo

La variable independiente de nuestro análisis es la categoría de la escuela, variable dicotómica o *dummy* que asume valor 1 cuando la escuela es de tiempo completo y valor 0 cuando la escuela es urbana pública común. La decisión de limitar el análisis a las escuelas públicas urbanas (comunes o de tiempo completo) se basa en la necesidad de controlar otras fuentes de variación que pueden nublar nuestra percepción de las relaciones entre las variables estudiadas. Si incluyéramos escuelas privadas o rurales en el estudio, estaríamos involucrando, inevitablemente, otros factores que harían más compleja la interpretación de los resultados. No se trata de restar importancia al papel que tanto las escuelas rurales como el sector privado desempeñan en nuestro sistema educativo; simplemente se trata de comparar escuelas que comparten una mayoría de rasgos (las escuelas urbanas públicas) y se diferencian fundamentalmente en una variable, cuyos efectos están siendo evaluados (la categoría de escuela de tiempo completo).

No analizaremos en detalle las características que diferencian a las escuelas de tiempo completo de las demás escuelas urbanas públicas, por limitaciones de espacio; pero es importante destacar que una caracterización

exhaustiva de estas escuelas sería un proyecto de investigación en sí mismo, que contribuiría, seguramente, a explicar los resultados que se observen en dichas escuelas.¹²

La variable de control: nivel socioeconómico

El rendimiento de los alumnos de sexto año se encuentra fuertemente afectado por sus orígenes sociales, y entre las características familiares se destacan el nivel educativo materno y el nivel económico del hogar. Estas dos variables se hallan altamente asociadas entre sí. Si bien desde el punto de vista teórico tiene más sentido proponer al nivel educativo materno como predictor de los resultados académicos de los niños, un índice de nivel socioeconómico (medido a través de un *proxy* de nivel interval, en este caso un índice que mide la disponibilidad de equipamiento en el hogar) puede ser utilizado como predictor en un modelo de regresión lineal o, en esta oportunidad, en un análisis multinivel utilizando modelos jerárquicos lineales.

La tabla de la página siguiente muestra los 11 ítemes de equipamiento que componen el índice, y el porcentaje de hogares que, en 1996 y 1999, disponían de cada uno de ellos.

El índice, que puede asumir valores entre 0 y 11, se calcula asignando un punto por cada ítem disponible en el hogar.¹³ Como vimos en la tabla, si bien muchos de los ítemes se encontraban prácticamente universalizados ya en 1996 (energía eléctrica, televisión color), otros resultan menos frecuentes y aparecen fuertemente relacionados con el contexto sociocultural (computadora, automóvil de uso exclusivo del hogar).

El índice de equipamiento del hogar (*proxy* del nivel socioeconómico) formará parte del modelo de análisis en sus dos niveles:

1. En el nivel 1, el del estudiante, se tomará el número de ítemes de equipamiento del hogar de *cada estudiante* como predictor de su rendimiento en matemática. El índice será utilizado como una variable interval *discreta* con un rango teórico de variación entre 0 y 11.
2. En el nivel 2, el de la escuela, se tomará el *promedio* de los índices de todos los estudiantes de sexto año de la escuela como predictor del rendimiento promedio de la escuela en matemática. El índice será uti-

¹² Uno de los aspectos que sería necesario investigar a este respecto es la conformación de los equipos docentes, que a menudo varía significativamente con respecto a las escuelas comunes.

¹³ El refrigerador con *freezer* vale dos puntos y el refrigerador sin *freezer* vale un punto, pero si el hogar tiene ambos, solo se contabilizan los dos puntos del refrigerador con *freezer*. Esta es una forma de asignarle mayor valor al ítem que indica un mayor estatus, sin alterar el valor máximo del índice.

Porcentaje de alumnos en hogares con:	Contexto sociocultural											
	Muy favorable		Favorable		Medio desfavorable		Desfavorable		Muy desfavorable		Total país	
	1999	1996	1999	1996	1999	1996	1999	1996	1999	1996	1999	1996
Energía eléctrica	99,6	99,3	99,3	99,0	98,6	98,7	97,7	97,3	93,4	96,3	97,3	97,9
Calefón o calentador de agua	97,8	96,1	90,8	92,8	79,8	84,2	66,0	76,1	46,1	64,6	72,8	80,4
Refrigerador sin freezer	50,8	44,2	69,7	61,3	73,2	66,5	72,8	70,8	65,1	68,5	66,9	63,7
Refrigerador con freezer	54,7	63,0	31,3	43,8	23,1	34,3	18,8	25,8	14,9	24,2	26,4	35,7
Lavarropa automático	85,4	88,8	65,1	80,1	50,2	64,2	37,4	53,1	21,9	41,2	48,2	62,0
Teléfono	91,3	92,2	67,6	80,7	47,6	70,8	32,7	60,5	17,2	48,1	46,9	67,3
Televisión color	98,5	98,6	95,7	97,3	91,5	96,2	86,5	93,2	75,2	89,7	88,0	94,3
Video casetero	82,9	76,5	59,6	55,4	43,8	43,0	32,9	38,5	21,0	27,0	44,4	44,8
Videogame o family game	76,0	63,0	64,4	58,0	54,8	53,4	45,1	41,6	27,8	37,4	50,7	48,7
Computadora	40,8	56,3	13,7	30,1	6,3	14,4	3,3	7,1	1,3	4,9	11,0	19,3
Automóvil de uso exclusivo del hogar	66,7	63,8	42,8	48,4	30,7	39,5	24,3	26,8	15,9	22,3	33,2	37,3

FUENTE: ANEP-Unidad de Medición de Resultados Educativos: Evaluación 6to. año Educación Primaria, 1999.

lizado como una variable interval *continua* con un rango teórico de variación entre 0 y 11.

Análisis de los resultados¹⁴

Nuestro planteamiento del problema de este estudio comenzó con un esbozo del debate acerca de la calidad y la equidad en el ámbito de las políticas educativas, en el que se enmarca la evaluación de las escuelas de tiempo completo. El análisis de los modelos lineales jerárquicos (HLM) arroja coeficientes que permiten, en este caso, establecer si una determinada variable incide en el logro de una *mayor calidad* (*mayores promedios*) y una *mayor equidad* (*menores pendientes* de las rectas de regresión que representan la asociación entre el nivel socioeconómico y los resultados académicos). Por lo tanto, en la medida en que hemos hipotetizado que las escuelas de tiempo completo son mejores en calidad y más equitativas que las escuelas urbanas comunes, deberíamos obtener:

- una *asociación positiva* entre la variable dicotómica *tiempo completo* y los *promedios* de resultados académicos por escuela (hipótesis 3);
- una *asociación negativa* entre la variable dicotómica *tiempo completo* y las *pendientes* de las rectas de regresión que representan la asociación entre el nivel socioeconómico y los resultados académicos (hipótesis 4)

A su vez, en función de la hipótesis 1, nuestra expectativa es que se observe una asociación positiva entre el nivel socioeconómico promedio de una escuela y el promedio de sus resultados académicos.

La tabla siguiente muestra los resultados de nuestro análisis para la Evaluación Censal de Aprendizajes de 1996. Como vemos, la tabla está dividida en dos sectores: en el primero se informa sobre el modelo que explica los promedios escolares; en el segundo se reportan los resultados del modelo que explica el comportamiento de las pendientes que, dentro de cada escuela, relacionan el nivel socioeconómico de cada niño con sus respectivos resultados académicos. (La información sobre los interceptos es incluida porque forma parte del modelo original, pero en este caso carece de valor analítico.)

¹⁴ Advertencia: El análisis de un conjunto de datos utilizando modelos lineales jerárquicos es habitualmente un proceso largo que involucra varias etapas. A efectos de este artículo, y por restricciones de espacio, decidimos limitarnos a presentar los resultados finales de dicho proceso, sin hacer referencia a las etapas previas.

En el primer sector de la tabla, en la fila referida al equipamiento, observamos un coeficiente positivo (0.218), que confirma nuestra hipótesis 1.¹⁵ Por lo tanto, concluimos que en 1996 se observaba una asociación positiva entre el nivel socioeconómico de una escuela y sus resultados académicos. Esta conclusión dista de ser alentadora —de hecho, constituye evidencia de las inequidades que estamos tratando de disminuir—, pero hemos de admitir que es altamente plausible y que no abundan los sistemas educativos que se encuentren exentos de esta condición. También en el primer sector de la tabla, pero en la fila referida a tiempo completo, observamos un coeficiente que indica una asociación positiva, aunque leve, entre las escuelas de tiempo completo y los promedios de resultados académicos. Esto muestra que las escuelas de tiempo completo obtenían resultados promedio ligeramente superiores a los de las escuelas urbanas públicas comunes.¹⁶

1996	Coeficiente	Error estándar	T	Valor de <i>p</i>
Modelo para los promedios escolares				
Intercepto	-1.280	0.057	-22.516	0.000
Equipamiento	0.218	0.010	21.43	0.000
Tiempo completo	0.080	0.067	1.193	0.233
Modelo para las pendientes NSE (equipamiento) – rendimientos académicos				
Intercepto	0.024	0.010	2.435	0.015
Equipamiento	0.005	0.002	3.244	0.002
Tiempo completo	0.006	0.013	0.504	0.614

En el segundo sector de la tabla podemos observar el modelo para las pendientes, es decir, el modelo que explica la asociación que dentro de cada escuela existe entre el nivel socioeconómico y los resultados académicos. En la fila referida al equipamiento observamos un coeficiente positivo de escasa magnitud (0.005), que indica una leve relación entre el nivel socioeconómico promedio de una escuela dada y la asociación, dentro de esa escuela, entre el nivel socioeconómico de cada uno de los niños y sus respectivos resulta-

¹⁵ En la medida en que los datos de 1996 son censales, es escasa la utilidad de los errores estándar, las pruebas T y el valor de *p*.

¹⁶ Es interesante observar que, si bien los datos obtenidos son censales, de haber sido extraídos de una muestra no habríamos podido considerar este coeficiente como indicador de una diferencia estadísticamente significativa a favor de las escuelas de tiempo completo.

dos académicos. Esto indicaría que, dentro de las escuelas urbanas públicas, aquellas de niveles socioeconómicos promedio más bajos tienden a producir resultados menos determinados por los orígenes sociales de los estudiantes. En la fila referida a “Tiempo completo” observamos un coeficiente positivo, de escasa magnitud, que indicaría que las escuelas de tiempo completo tendían, en 1996, a ser ligeramente más desigualitarias, es decir, que en ellas había una asociación ligeramente más fuerte entre orígenes sociales y resultados académicos que en las escuelas urbanas públicas comunes. Esta es una conclusión preocupante, dado que, aunque la magnitud del coeficiente es mínima, el signo es positivo, y escuelas más equitativas deberían llevarnos a obtener un signo negativo.

La siguiente tabla presenta los resultados de la evaluación muestral de aprendizajes en 1999. Nuevamente, el primer sector de la tabla presenta los resultados del modelo para los promedios de resultados académicos por escuela (*calidad*) y el segundo sector el modelo para las pendientes (*equidad*).

1999	Coeficiente	Error estándar	T	Valor de p
Modelo para los promedios escolares				
Intercepto	-1.354	0.153	-8.832	0.000
Equipamiento	0.206	0.023	9.007	0.000
Tiempo completo	0.176	0.066	2.649	0.008
Modelo para las pendientes NSE (equipamiento) – rendimientos académicos				
Intercepto	0.028	0.031	0.893	0.372
Equipamiento	0.004	0.005	0.882	0.378
Tiempo completo	-0.012	0.013	-0.925	0.355

En el primer sector, la fila correspondiente al equipamiento muestra un coeficiente positivo e importante (0.206), de valor muy similar a su equivalente de 1996 (0.218), aunque algo menor, lo cual indica —entre las escuelas urbanas públicas, tanto comunes como de tiempo completo— una cierta disminución de la asociación entre los niveles socioeconómicos promedio y sus respectivos promedios de resultados académicos. Sin embargo, como era de esperar, esta asociación es estadísticamente significativa, tal como lo indica el resto de la información (el reducido error estándar; el alto valor de T de Student; el valor de p, que indica significación estadística del coeficiente tanto al nivel del 95 como del 99% de confianza). La fila correspondiente a “Tiempo completo” muestra un coeficiente claramente mayor (0.176) que su equiva-

lente de 1996 (0.080). El dato de 1999, extraído de una muestra, es estadísticamente significativo al nivel del 95 y 99% de confianza, lo cual indica que *la calidad de las escuelas de tiempo completo parece haberse afianzado a lo largo de los años transcurridos entre ambas evaluaciones.*

En el segundo sector de la tabla, la fila correspondiente al equipamiento muestra un coeficiente de 0.004 (ligeramente inferior al valor de 1996) que no resulta estadísticamente significativo. Esto parecería indicar que ya no resulta válido el hallazgo de 1996 respecto a la (leve) asociación entre el mayor nivel socioeconómico promedio de una escuela y la mayor estratificación de los resultados académicos entre los niños de esa escuela. De todos modos, este resultado no constituye nuestra preocupación central.

Lo más destacable de la tabla es el coeficiente referido a las escuelas de tiempo completo, que ha cambiado de signo con respecto a 1996. Mientras tres años antes las escuelas de tiempo completo mostraban una leve tendencia a una mayor inequidad de los resultados de sus estudiantes, en 1999 muestran lo contrario, aunque este resultado no resulte aún estadísticamente significativo.

Conclusiones

En función de los resultados presentados, podemos afirmar que las escuelas de tiempo completo han experimentado una clara evolución favorable entre 1996 y 1999, en comparación con las escuelas urbanas públicas comunes, tanto en lo que se refiere a la calidad como en lo que tiene que ver con la equidad. Queda, de todos modos, un largo camino por recorrer, que desde el punto de vista de la evaluación deberá pasar por análisis de costo/efectividad del programa.

Hemos visto cómo el incremento en la precisión al medir los resultados de un programa o la utilización de técnicas más sofisticadas para la recolección, el procesamiento y análisis de los datos reducen el error y aumentan el poder explicativo de los modelos. Sin embargo, estos progresos técnicos no bastan para solucionar uno de los problemas centrales de la evaluación: el establecimiento de los parámetros contra los cuales se compara aquello que está siendo evaluado. Es aquí, más que en ningún otro punto del proceso, donde el trabajo del técnico requiere de la acción del político en tanto intérprete del sentir de la comunidad. A menudo se sostiene que la evaluación de programas y políticas en los países en desarrollo está fuertemente limitada por la escasez de recursos. Esta apreciación es correcta, seguramente, pero el problema fundamental va más allá de la disponibilidad de recursos, o de la conveniencia de su asignación para estos fines. Tanto en países desarrolla-

dos como en los nuestros, el punto crítico en cualquier proceso evaluativo es definir los objetivos que se supone que estos programas deben cumplir, y esta es, como mencionamos, una competencia del político, *asesorado* por el técnico.

Por último, es importante destacar que, a medida que las metodologías orientadas a la evaluación de programas y políticas se hacen más sofisticadas, crece el desafío para los técnicos en cuanto a la comunicabilidad de los resultados, tanto hacia los políticos tomadores de decisiones como hacia el público en general. Esto, sin embargo, no debería ser óbice para que estas nuevas técnicas se apliquen; simplemente nos obliga a buscar formas más eficaces de transmitir ideas complejas de manera clara y concreta, pero al mismo tiempo rigurosa.

Bibliografía

- BARNETT, W.: *Lives in the Balance, Age-27 Benefit-Cost Analysis of the High/Scope Perry Preschool Program Ypsilanti*, High/Scope, Michigan, 1996.
- BERNSTEIN, B.: *La estructura del discurso pedagógico*, Morata, Madrid, 1993.
- BRYK, A., y S. RAUDENBUSCH: *Hierarchical Linear Models*, Sage, Newbury Park, 1992.
- HALLAK, J.: *Investing in the future*, UNESCO, París, 1990.
- KREEFT, I., y J. DE LEEUW: *Introducing Multilevel Modeling*, Sage, Londres, 1998.
- MARTÍNEZ ARIAS, R.: *Psicometría*, Síntesis, Madrid, 1995.
- MUÑIZ FERNÁNDEZ, J.: *Teoría de respuesta al ítem*, Pirámide, Madrid, 1990.
- OKUN, A.: *Equality and Efficiency, The Big Tradeoff*, Brookings, Washington, 1975.
- PÉREZ MORGAN, B.: *Evaluación del Programa de Escuelas de Tiempo Completo propuesto por el CODICEN de la ANEP en 1995 desde el punto de vista de la equidad* Montevideo, CLAEH, Serie Investigaciones, Montevideo, 2001.
- SNIJDERS, T., y R. BOSKER: *Multilevel Analysis*, Sage, Londres, 1999.
- UMRE: *Estudio de los factores institucionales y pedagógicos que inciden en los aprendizajes en escuelas primarias de contextos sociales desfavorecidos en el Uruguay* Montevideo, ANEP-BIRF, Montevideo, 1999.

Resumen

El presente estudio constituye una evaluación de las escuelas de tiempo completo, enmarcada en el debate sobre equidad, calidad y eficiencia de los sistemas educativos. El artículo busca además introducir en nuestra comunidad académica dos técnicas relativamente nuevas y sofisticadas para el análisis de datos cuantitativos: la *teoría de respuesta al ítem* (cuyo ámbito de aplicación probablemente se restrinja a la evaluación educativa y psicológica) y los *modelos lineales jerárquicos* (aplicables cuando se trata de explicar resultados individuales en función de características de los propios individuos, así como

de variables de las organizaciones en que se encuentran insertos: niños en escuelas, pacientes en policlínicas, copropietarios en cooperativas de vivienda, etcétera).

Palabras clave: reforma educativa, Uruguay, calidad de la educación, evaluación de la educación, análisis de datos, métodos cuantitativos.

Abstract

The present study constitutes an evaluation of full time schools, within the framework of the debate on equity, quality and efficiency of educational systems. The article also seeks to introduce in our academic community two relatively new and sophisticated techniques for the analysis of quantitative data: *the theory of response to the item* (which application range is probably restricted to educational and psychological evaluation) and the *hierarchic linear models* (applicable when it comes to explaining individual results in function of the characteristics of the individuals, as well as of variables of the organizations in which they are inserted: children in schools, patients in policlinics, co-owners in housing cooperatives, etcétera).

Key words: educational reform, Uruguay, quality of education, evaluation of education, data analysis, quantitative methods.

Copyright of Prisma is the property of Universidad Catolica del Uruguay Damaso Antonio Larranaga and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.