

Determinación del rendimiento escolar bajo una prisma de equidad: la evaluación del valor agregado

CHILE

1. Introducción desde la Historia de la Educación en la Segunda Mitad del Siglo XX.

La Educación en los últimos cincuenta años, muestra tres etapas claramente definidas en el énfasis de las orientaciones de las políticas educativas y en el tratamiento de los problemas de **equidad**. La referencia temporal señalada es propia de la realidad de América Latina ya que, en regiones del mundo desarrollado, estas etapas se presentaron con otras secuencias y duraciones.

Una primera etapa -entre fines de los cincuenta y comienzos de los setenta- estuvo caracterizada por **la igualdad de oportunidades de acceso a la Educación**. Esta consistió en que todos los alumnos en edad escolar tuvieran la posibilidad de asistir a la escuela y que los sistemas educativos, fundamentalmente en el nivel Básico, garantizaran un lugar para cada niño. Estuvo marcada por grandes inversiones en infraestructura escolar y en algunos casos por una reducción en la duración de la jornada escolar diaria, de manera de dar cabida a un grupo en la mañana y a otro en la tarde. La

evaluación educativa durante esta etapa, tuvo una connotación terminal o de producto y su rol primario fue de certificación. La medición se refirió a normas y la distribución del rendimiento se concibió como normal.

Un segundo período - entre los comienzos de los Setenta y fines de los Ochenta- estuvo centrada en **la igualdad de oportunidades de aprendizaje** y buscó no sólo garantizar a cada niño en edad escolar un espacio en la escuela, sino que a través de un enfoque denominado de enseñanza individualizada o de educación personalizada, hacer el aprendizaje accesible a todos los alumnos. Esta orientación tuvo su origen en los trabajos de Tyler, de Carroll, de Gagné y de Bloom. La evaluación se focalizó en el aula en un rol formativo. Las diferencias individuales hicieron centrar la atención en los prerrequisitos y la distribución de los resultados de aprendizaje se concibieron con un sesgo estadístico negativo.

Una tercera etapa se abrió a comienzos de la Década de los Noventa y continúa vigente y podría denominarse la etapa de **la igualdad**

* Director de la Oficina de Medición del Rendimiento Escolar del Instituto de Evaluación del Estado de Qatar - Golfo Árabe.

¹ Una versión anterior del presente trabajo fue publicada bajo el título *Evaluación de la calidad de la educación con equidad: el modelo de valor agregado*, en la Revista *Persona y Sociedad*, Universidad Alberto Hurtado-ILADES, Santiago de Chile. Vol. XVII, N° 1, abril 2003, pp. 165-178.

² Juan Enrique Froemel es Ingeniero, obtuvo un Master of Arts en Ciencias Sociales (1979) y un Doctorado en Educación (1980), en la Universidad de Chicago. Su área de especialidad es la medición y la evaluación del rendimiento en Educación. Ha sido Subsecretario de Educación Pública en Chile, Rector de la Pontificia Universidad Católica de Valparaíso también en su país, Gerente General de la Fundación Andes en Chile y Coordinador Técnico del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, operado por UNESCO-Santiago-OREALC. Es profesor en varias universidades, entre ellas en el Programa de Maestría en Política Educativa de la Universidad Alberto Hurtado, es consultor internacional y Socio Principal de EvaluArte, Froemel & Asociados. Se desempeña en la actualidad como Director de la Student Assessment Office, del Evaluation Institute, bajo el Supreme Education Council, del Estado de Qatar, en el Golfo Árabe.

de progresos de aprendizaje. Ya no sólo se debe procurar tomar en cuenta las diferencias individuales de los estudiantes, sino que debe procurar compensarlas e identificar las variables que afectan los resultados y de explicar como esas variables y el rendimiento se conjugan. Nuevos modelos teóricos y herramientas estadísticas, han posibilitado la evaluación de progresos o de “valor agregado”, ajustándolos por aquellos factores ligados al Contexto Familiar de los estudiantes y focalizándolos simultáneamente en varios niveles de agregación, aunque centrándose en el nivel de la escuela. El énfasis está puesto en lograr una equidad real y efectiva.

2. El condicionamiento de la idea de Equidad desde el contexto social y educativo.

Como se señalara con anterioridad, el tema de la incorporación de la equidad en los modelos que sustentan la política educativa es una característica central de la última década. Si se verifican las menciones al tema que, con relación a las reformas educativas en varios países, son reseñadas en una reciente publicación (Townsend, Clarke y Ainscow, 1999), ellas son abundantes, aunque los contextos y la terminología varían de país en país.

Primeramente, se ha estimado pertinente buscar la razón que genera un **condicionamiento** del término tal y como se señalara en el título de la presente sección. El caso es que el significado que el término equidad puede adquirir en distintos contextos, varía grandemente. Un buen ejemplo de esto último se encuentra en el Primer Informe del Primer Estudio Internacional Comparativo sobre Lenguaje, Matemática y Factores Asociados en Tercero y Cuarto Grado, (1998), el cual muestra desempeños desiguales entre escuelas públicas y privadas y entre las urbanas y las de mega-ciudad.

Sin embargo, si se presta atención al caso de Cuba, se puede observar que no se muestra una diferencia apreciable entre el comportamiento de las escuelas rurales con el de las urbanas

y el de las de mega-ciudad, en rendimientos en Lenguaje y Matemática como en los demás países. Además, las escuelas públicas cubanas están por sobre todas las escuelas privadas de la Región. Más aún, cuando se efectuó el ajuste de los puntajes de Rendimiento por nivel socioeconómico, en el caso de Cuba el efecto fue menor que en los demás países, debido a la alta homogeneidad en el país en esta variable. Aún más, al recurrirse a una variable alternativa para tal ajuste, el nivel socio cultural, básicamente conformado por la escolaridad de los padres, una vez más, el alto grado de homogeneidad existente en Cuba, hizo que el ajuste mostrara efectos menores.

El posible concluir que el tema de la equidad depende del contexto, no siendo sólo un tema conceptual sino que un aspecto metodológico que surge del efecto de la homogeneidad de la distribución de la variable de ajuste. Una manera de salvar la dificultad consiste en abordar la equidad desde otras variables tales como el nivel sociocultural, el género, o las variables étnicas. Por lo señalado parece ya no ser más válido plantear comparaciones o ajustes para determinar equidad, sin antes establecer el rango de dispersión de las variables de ajuste.

La variedad de perspectivas mostradas por Townsend, Clarke y Ainscow (1999), señalan que en Malasia la equidad se concibe como la necesidad de que la educación constituya un instrumento para alcanzar la solidaridad nacional y se considera una interacción entre la educación y la sociedad y viceversa. En Sudáfrica se describe que entre los principios que guían la Reforma Educativa, se encuentran la equidad y el compromiso de todos los sudafricanos. En este caso particular la equidad es un logro deseable más que una variable puente. Se agrega que “quizás las inequidades son aquellas relacionadas con clase, género y religión, aunque podría pensarse que en este contexto las más gravitantes deberían ser aquellas ligadas a temas étnicos”. (Mkhatshwa, en Townsend, Clarke, y Ainscow, 1999). En Belarus, se señala que la Educación Secundaria requiere producir un equilibrio entre

los objetivos conjuntos de calidad y equidad: la igualdad de oportunidades para recibir una educación de calidad debería ser la base de las decisiones y políticas educacionales y se agrega que la socialización de los estudiantes y sus relaciones con el mundo circundante, deben estar basadas en un equilibrio entre los mejores intereses de los estudiantes y la utilidad de aquellos intereses para la sociedad. Como puede apreciarse el énfasis de la equidad, en este caso se sitúa en una acción concreta en el aula. Para Francia, la reciente y creciente independencia de las escuelas ha producido un gradual convencimiento de que la uniformidad tradicional del sistema francés estaba lejos de constituir una garantía para la equidad, ni para la eficacia de las escuelas. En Noruega el tema de la equidad se enfoca desde la perspectiva del género y su vigencia en las escuelas. En Chile la visión de la equidad es más bien tradicional, en cuanto a que los profesores tienden a atribuir al bajo nivel socioeconómico de los alumnos las dificultades halladas al intentar poner en marcha el enfoque de escuelas eficaces en el país.

En los Estados Unidos de Norteamérica, aunque el tema de la equidad es de larga data, en la actualidad existe una especial preocupación por él, evidente en la investigación, para relacionarla con el proceso escolar. El modelo más clásico es el desarrollado por Edmonds, en 1979 con sus cinco factores: fuerte liderazgo instruccional del Director, focalización de la instrucción, un ambiente de aprendizaje ordenado y seguro, altas expectativas, monitoreo frecuente y uso de la información evaluativa para orientar el mejoramiento. Otro aspecto importante es el del reconocimiento de los efectos negativos demostrados por las condiciones inadecuadas de entrada de los estudiantes sobre sus oportunidades educacionales (Kozol, 1991, citado en Townsend, Clarke y Ainscow, 1999) Otros aspectos son el impacto de la descentralización, la posibilidad de elección de la escuela y el movimiento de los denominados “charter schools”, en su relación con la equidad y eficacia de las escuelas y un elemento que ha

adquirido una dimensión nueva y que hoy es reconocido como una herramienta de aplicación múltiple en la Educación y central al tema de la equidad. Se trata de que, como señalan algunos autores (Townsend, Clarke y Ainscow, 1999) “el desarrollo de instrumentos válidos y confiables para evaluar la eficacia es aún un tema no zanjado. Sin embargo, existen esfuerzos en proceso para desarrollar un conjunto de tareas de evaluación, aunque los temas de confiabilidad siguen siendo su principal obstáculo”. Al respecto hay varias preguntas aún en espera de respuesta: ¿Cuáles son los sistemas justos y auténticos de evaluar el progreso de los estudiantes que son aceptables para el público y que a la vez sirven a las necesidades de los políticos, estudiantes y ciudadanos? ¿Cómo es posible determinar si el mejoramiento de las escuelas está agregando valor a los estudiantes?”.

3. Antecedentes de la idea de evaluación de “valor agregado” en relación con Educación.

La Evaluación en el campo educativo ha centrado su atención durante casi toda su historia en la determinación del rendimiento escolar. Este último corresponde a la expresión cuantitativa de lo que un alumno ha aprendido y es medido en base del resultado obtenido al responder una prueba.

Aunque menos difundida, la determinación estadística de los “progresos” en rendimiento es de larga data en Educación. Los enfoques y consecuentes procedimientos anteriores a la década de los Noventa, consistieron principalmente en la transferencia directa al campo evaluativo de la determinación estadística de las diferencias denominadas “progresos”, fuese entre muestras dependientes o independientes, utilizada en modelos de investigación. En la práctica, lo que se determinaba era la diferencia entre los resultados obtenidos por los alumnos, pertenecientes a cada grupo, en aplicaciones sucesivas de los mismos instrumentos o de pruebas equivalentes y seguidamente, se

calculaba el nivel de significación estadística de tales diferencias.

Sin perjuicio de lo señalado, la idea central de determinar y expresar cambios diferenciales en el rendimiento de los estudiantes, como se puede comprobar por lo recién señalado no es una idea nueva, sino que como señala un autor (Saunders, 1999): “Podría decirse que el valor agregado (progreso) era una idea esperando por su momento (... propicio)”.

La expresión “progresos” como se la maneja hoy en día en Educación, es asociada al término “valor agregado”, ha pasado por varias etapas evolutivas.

Su origen primero se ubica fuera del campo de la Educación, de modo que para algunos autores “El término valor agregado se origina en la Economía cuando se hace referencia a la diferencia entre insumos, tales como materias primas y la energía, versus los productos finales representados por el valor de venta... En años recientes el concepto económico de “valor agregado” ha adquirido un significado adicional, cuando se le asocia con la necesidad de las Economías Occidentales de focalizarse en la manufactura de alta calidad, como una forma de competir exitosamente con economías de rápido crecimiento. El término “valor agregado” en este contexto se aplica a productos elaborados cuya calidad y valor son aumentados por altos niveles de tecnología y destreza en el proceso de manufactura.” (Spours y Hodgson, 1996, citados en Saunders, 1999).

Sin embargo, es imprescindible tener en cuenta que las ambigüedades asociadas con el término “valor agregado”, que no son pocas, han sido heredadas precisamente del contexto económico. Es por esto último que el hecho de que el concepto parezca originarse en tal entorno no garantiza en absoluto, ni la precisión ni la claridad conceptual. Por otra parte la definición de Spours y Hodgson, recién citada, puede llevar a pensar en que no sólo el contexto económico pueda haber afectado al educativo, sino que existiría razón para anticipar que

podiese haber también un efecto en la dirección opuesta, dado que “en el discurso dentro del cual se sitúa esta definición evolucionada del valor agregado (ella) está conectada con la relación entre la economía (de nivel nacional) y la entrega de educación/entrenamiento; y en qué medida puede ser demostrado o sólo supuesto que la educación y el entrenamiento sean elementos integrales de la salud económica de la nación.” ... “es posible señalar al respecto que recientemente ha empezado a surgir la sombra de cierta duda acerca de si en realidad la calidad de la educación/entrenamiento constituye efectivamente una variable clave para el éxito de una economía”.

Algunas consideraciones adicionales parecen ser necesarias (Saunders, 1999) en cuanto a completar esta visión del “valor agregado” desde la perspectiva de sus orígenes económicos. El primer aspecto consiste en la exploración de respuestas a las preguntas ¿valor que se agrega a qué, por parte de quién, y para el beneficio de quién? Es así como estas preguntas llevan a tomar cabal conciencia de la antes mencionada ambigüedad del concepto, si además se agregan a la búsqueda de respuestas a otras tales como ¿es el “valor agregado” un concepto absoluto o relativo? ¿Tiene un rol descriptivo o uno evaluativo? En cuanto a esto último, el tema de la ventaja o beneficio es un aspecto muy importante de captar ya que el “valor agregado”, incluso en un contexto económico, tiene un significado evaluativo, aunque éste pueda no estar presente en todas las acepciones del término. En este contexto educativo, sería posible argumentar que el *potencial* evaluativo, presente al interior del término se ha convertido más bien en un *propósito* evaluativo, en todas sus aplicaciones.

La introducción del término al campo de la Educación, fue por la vía de su segmento más avanzado, esto es por el de la Educación Superior. “Como *término* explícito” señala la autora mencionada (Saunders, 1999) “su primera aparición parece haber ocurrido en el sector de la educación avanzada/superior en el contexto de los indicadores institucionales de desarrollo

o rendimiento. Un *tema* crítico para el “valor agregado”, sin embargo, se originó a partir de la agenda de investigación en escuelas eficaces, la cual estuvo centrada en la Educación Media obligatoria. En ambos casos ha existido una investigación metodológica asociada, la que ha adquirido por derecho propio una dimensión educativa.

Otra de las primeras alusiones a éste término, define “valor agregado” como “una medida del rendimiento de los estudiantes, el que toma en cuentas los efectos de diferentes insumos propios de ellos (por ejemplo: calificaciones de ingreso) El rendimiento de los estudiantes así concebido, puede ser usado para contribuir a la evaluación de la calidad institucional, dentro las funciones de enseñanza y aprendizaje” (McGeevor y colaboradores, 1990, citados en Saunders, 1999).

Agrega la autora mencionada (Saunders, 1999) aludiendo a que el término surgió en la Educación Superior, que “Esencialmente el sector de la educación avanzada/superior parece haber estado particularmente interesado en la evaluación de la eficiencia y eficacia dentro de las instituciones educativas, de manera de justificar las asignaciones de fondos; mientras que el sector escolar estaba más preocupado con las formas de efectuar comparaciones más justas *entre* las instituciones educativas, particularmente en forma de explicar la aparente falta de coherencia entre resultados distintos para similares niveles de financiamiento, en diferentes instituciones o para niveles educativos locales.” Sin embargo, agrega la autora, “ambas conceptualizaciones tienen en común el hecho que responden a la misma pregunta ¿Cómo puede medirse el progreso de los estudiantes/alumnos en forma de arrojar luz acerca del desempeño de las instituciones?”.

Como señala una autora (Saunders, 1999) “el cálculo de medidas de “valor agregado” para la educación obligatoria (Primaria y Secundaria) estuvieron inicial y primariamente conectadas

con (el propósito de) hacer cada vez más objetivas y válidas, las comparaciones entre los desempeños de las escuelas”. Esta es una connotación que persiste en la actualidad, aunque las metodologías estadísticas denominadas como de “niveles múltiples” hoy permiten efectuar la determinación del “valor agregado” en varios niveles de agregación, en forma simultánea.

Uno de los trabajos pioneros en este aspecto se llevó a cabo en el Reino Unido y éste fue el de las comparaciones entre los resultados en el Certificado General de Enseñanza Media y el Nivel A (universitario), probadas y desarrolladas a través del sistema denominado ALIS (Sistema de Información del Nivel A), desde los comienzos de los Ochenta, por Fitz - Gibbon, quién en aquella época se desempeñaba en el Centro de Currículum, Evaluación y Gestión de la Universidad de Newcastle. Los resultados de tal estudio señalaron -por ejemplo “que las diferencias entre escuelas parecen no constituir una fuente muy importante de las diferencias alcanzadas por los estudiantes individuales, de distintos niveles de habilidad/rendimiento que asisten a distintas escuelas; y (también) que existe “evidencia contrapuesta en cuanto a la pregunta de si ciertas escuelas muestran una eficacia diferente para estudiantes cuyas habilidades / logros (también) diferían al ingreso”- también (de ese modo) prefigurando mucha de la evidencia acumulada, desde ese entonces, respecto de la eficacia relativa de tales mediciones. (Gray, Jesson y Jones, 1986, citados en Saunders, 1999).

Lo recién señalado, estableció la necesidad de contar con una forma de medir distinta a la tradicional, por dos razones concretas. Por una parte se planteaba que “metodológicamente dos temas emergen de la intención de comparar el desempeño de las escuelas: primero, que debe introducirse algún concepto de “contextualización”, de modo de dar posibilidad de manifestarse al hecho de que las escuelas están tratando con poblaciones estudiantiles

que difieren considerablemente en términos de rendimiento previo y en otras variables claves (el tema de los insumos) Por otra parte, desde el momento en que las escuelas son “parte de un sistema educacional que posee jerarquías naturales o niveles de inclusión (por ejemplo, los estudiantes agrupados en secciones, dentro de cohortes, a su vez al interior de escuelas)”, el modelo estadístico adoptado para efectuar los análisis de los datos de desempeño, debe reflejar esta estructura.” (Gray, Jesson y Jones, 1986, citado en Saunders, 1999).

Las razones recién enunciadas tienen una relevancia especial, dado que inciden en varios aspectos que por largo tiempo se han considerado centrales en la actividad evaluativa en Educación. El primero de ellos consiste en que aquella debe proporcionar información que permita tomar decisiones que estén encaminadas a producir efectos en las decisiones de política educativa y que necesariamente tales decisiones tienen que ser pertinentes. Pertinentes, en cuanto a que la evaluación genere información que permita adoptar las medidas necesarias sobre las escuelas, eso sí cuidándose que ello ocurra en la medida en que éstas sean realmente responsables del desempeño de los alumnos.

En segundo lugar, es crucial poder tomar en cuenta cuál es la contribución de cada uno de los niveles jerárquicos del sistema educativo y de sus variables asociadas, sobre el rendimiento. En éste último sentido, el nuevo enfoque a cuya aplicación se refiere este documento permite, casi por primera vez, considerar en forma exhaustiva los abundantes hallazgos que demuestran, por una parte, “que los niños provenientes de entornos socialmente deprivados tienden a obtener puntajes en los exámenes y resultados en las pruebas, inferiores a los de aquellos niños de entornos más privilegiados”.

De similar modo y aún más relevante que los aspectos recién señalados, es el hecho de que este enfoque evaluativo permite, a diferencia de otros anteriores, tomar en cuenta que “el logro de los niños cuando entran por primera vez a la escuela, constituye el factor más determinante

del rendimiento posterior” (Goldstein y Cuttance, 1988, citado en Saunders, 1999).

Otro aspecto central que está incluido en el enfoque en cuestión es “... que la eficacia de las escuelas varía en términos del desempeño relativo de diferentes sub-grupos. Intentar resumir las diferencias entre escuelas, aún ajustando por características de ingreso, sexo y antecedentes étnicos de los alumnos y por las características fijas de las escuelas, en una sola cantidad, es engañoso... El concepto de eficacia total, no es útil” (Nuttall, Goldstein, Prosser y Rabash, 1989, citado en Saunders, 1999).

Un tema adicional que es abordado por el enfoque en cuestión es que el ordenamiento de las escuelas por jerarquía de resultados, “sean estos brutos o ajustados por estatus socioeconómico y/o rendimiento previo es, en cualquier caso, altamente problemático, ya que el propio método usado para arribar a esas estimaciones de desempeño de las escuelas, parcialmente determina el orden resultante. Es necesario, por lo tanto, asumir que el “método” en este caso abarca (e influye sobre) ambos, el tema de cuáles productos son usados, particularmente aquellos que son capaces de diferenciar entre diferentes grupos de alumnos y el tema de cuáles técnicas de modelación estadística son usadas. (Goldstein y Cuttance, 1988).

Además, estos mismos autores previenen en cuanto a que “mientras mayor es el número de factores de contexto que se incluyen en el análisis, más inestables y menos confiables resultan las comparaciones.” “Este es un principio importante de tener en cuentas, dado que han existido demandas, razonables desde el punto de vista de aquellos que buscan administrar y “mejorar” las escuelas, por un modelo que incluyese todo aquello, que pudiese hacer pensar que constituye factores que afectan el desempeño de las escuelas.

Es por estas razones que el principio subyacente en la mayor parte del trabajo de valor agregado, ha sido más restrictivo que inclusivo... Su relación con la eficacia de las escuelas ha

estado radicada no tanto en identificar un rango de variables correlacionadas con la eficacia, sino que en discernir entre diferentes clases de “ruido” o información exógena, presentes en los análisis de eficacia y (finalmente) en (procurar) desembarazarse de estos lo más posible.” (Saunders, 1999).

En términos históricos, la primera referencia al “valor agregado” en Educación publicada, fue el artículo de Nuttal, de 1991, aparecido en el Suplemento Educativo del Times de Londres. “Este constituía una solicitud (casi un ruego) redactada en Lengua común pero explícito, presionando por investigación en el tema de la eficacia escolar en su relación con la generación de política, en particular respecto de la necesidad urgente de nuevos sistemas de medición. ” (Saunders, 1999) Textualmente manifestaba que “Por casi 30 años, la investigación sobre la eficacia de las escuelas ha (procurado) la utilización del *progreso* alcanzado por los estudiantes desde su nivel de desempeño al ingreso (a la escuela) hasta el nivel al momento de egresar (de ella), más que sus resultados “brutos” al momento de dejar la escuela. Este enfoque, adecuadamente iluminado por la metáfora del “valor agregado”, hace sentido intuitivamente y es fácilmente comprensible. (Nuttal, 1991).

Un segundo hito en la presencia explícita del término y del concepto en la literatura especializada, es el trabajo de McPherson de 1992, denominado *Midiendo valor agregado en las escuelas*. Su argumento central es que un “mal sistema de indicadores como aquel solamente basado en puntajes brutos de pruebas o resultados de exámenes, lleva implícitos los costos ocultos de juicios errados, ansiedades innecesarias e inútiles investigaciones adicionales originadas en falsas señales”. “La medición del valor agregado, explícitamente definida como el cálculo de la contribución que las escuelas hacen al progreso de los estudiantes, es central en el desarrollo de un buen sistema de indicadores. Tal medición

debe estar basada en “una teoría explícita sólidamente sustentada”, que esté abierta al escrutinio y refinación y que incluya elementos tales como:

- el rendimiento previo de los alumnos.
- la naturaleza longitudinal del progreso.
- la naturaleza multinivel de las escuelas.
- la naturaleza multivariada de los factores involucrados, especialmente “factores ajenos a la escuela que pueden potenciar o retardar el progreso”, tales como los antecedentes socioeconómicos de los estudiantes.
- la eficacia diferencial de escuelas específicas para distintos grupos de estudiantes.

4. Aspectos operacionales del “valor agregado”.

Para efectos de este documento se ha adoptado la definición de “valor agregado” que señala que tal concepto es (definido en el ámbito específico del Reino Unido) “una indicación de la medida en que una escuela en particular ha procurado el progreso de *todos sus estudiantes*, en un ámbito de asignatura y durante un específico período de tiempo, por ejemplo, desde el ingreso a la escuela hasta los exámenes públicos, en el caso de las escuelas secundarias o sobre años específicos en Primaria, en comparación con los efectos de otras escuelas de la misma muestra.” (Sammons, Thomas y Mortimore, 1997).

Cómo se puede advertir, aunque se tiene conciencia plena de la existencia y trascendencia de otros resultados del aprendizaje, tales como valores y actitudes, en el presente caso, la referencia a “progreso” de los estudiantes se ha restringido exclusivamente al ámbito del aprendizaje académico.

La obtención del “progreso” a través del enfoque del “valor agregado” implica que la información de insumo cumpla con tres

condiciones a saber: que cuente con una “línea de base”; que la evaluación se realice a nivel “censal” y no “muestral” y; que los alumnos que serán parte de la evaluación, estén plena e inequívocamente identificados.

Respecto de la primera de las condiciones, ello es contar con una “línea de base”, que permita sustentar la estimación estadística por regresión, existen varias alternativas de acuerdo a su diferente poder predictivo. La opción por una u otra alternativa configura, a la larga, distintos modelos que pueden, a su vez, ser clasificados de acuerdo a ese variado poder predictivo y explicativo, respecto de la variación en rendimiento. Tales modelos serán reseñados más adelante.

Sin perjuicio de la abundancia de alternativas, la opción más conveniente y ortodoxa para constituir la “línea de base”, se apoya en el empleo, como Primera Medición, de resultados obtenidos por los alumnos evaluados en pruebas estandarizadas de rendimiento, aplicadas idealmente con anterioridad al Grado en el cual se encuentren al momento de la Segunda Medición, lo que permitirá la determinación del “progreso”, a través del “valor agregado”.

Sin embargo, la opción más conveniente, en la mayoría de los casos no es la más factible. Lo recién señalado, por cuanto en la mayoría de las situaciones los sistemas de medición de amplia cobertura tienen carácter muestral y por la misma razón no siempre registran de manera utilizable la identificación de los sujetos que incluyen en sus evaluaciones.

Por otra parte, el hecho de que el enfoque de “valor agregado” constituya una forma innovadora de evaluación en el campo educativo, requiere que sus aplicaciones puedan exhibir en tiempo relativamente breve, resultados que sean utilizables con fines de generar decisiones en política educativa. Estas dos últimas razones llevan, la mayoría de las veces a emplear como “línea de base” alternativas

diferentes, menos eficaces y eficientes, que las mediciones estandarizadas.

Respecto de la condición relativa al carácter “censal” de las mediciones, tampoco ésta es una condición habitualmente encontrada en los sistemas de medición de amplia cobertura. Lo más común es que se estructuren de manera “muestral” y en el mejor de los casos que, si bien pueden en algunos, alcanzar dimensiones que se aproximan al universo de su población objetivo, dejan fuera segmentos de éste, aunque no de manera sistemática. En otras situaciones aunque existe cierta sistematicidad, las exclusiones involucran marginar segmentos de la población que por sus características son significativos en el sistema educativo. Esto último tiene por resultado que existan, en esos casos, pérdidas de información importantes que afectan el ámbito de generalización de la respectiva evaluación.

En cuanto a la identificación plena y unívoca de los sujetos participantes en el proceso evaluativo, la tendencia general ha sido, por una parte, que los resultados de las mediciones y evaluaciones no entreguen datos por estudiante, sino que agregados por cursos o por establecimientos. Por otra parte, situaciones aún más extremas han implicado que no existan mecanismos habilitados dentro de los sistemas de evaluación educativa y en algunos, ni siquiera dentro de los sistemas educativos como tales, que permitan identificar a los estudiantes. Ambas situaciones y en particular la segunda de las nombradas conspira en contra de la posibilidad de estructurar sistemas bajo el enfoque de “valor agregado”. Paradójicamente, aunque el establecimiento de mecanismos de identificación de los sujetos pareciera constituir el problema más difícil de resolver por lo largo y trabajoso, la decisión política de no hacer pública la información respecto de los sujetos individuales, es aún más difícil de remontar que aquellos, especialmente para fines evaluativos.

5. Aspectos metodológicos del “valor agregado”.

Como un primer paso antes de entrar de lleno en el tema de la metodología, es importante tener presente que la necesidad de este enfoque surge del hecho de que las evaluaciones que publican resultados brutos, organizados en ordenamientos jerárquicos (league tables) entregan una información imprecisa y que además, puede hasta ser peligrosa.

Tal peligro radica, por un lado, en el hecho de que puede inducir a la complacencia entre las escuelas que atienden a poblaciones más aventajadas. En el hecho, el progreso que alcanzan los estudiantes en algunos de tales establecimientos puede, en realidad, ser menor que el de otras escuelas que sirven a alumnos de menores recursos. En la práctica, sus resultados serían en tal caso, en realidad más bajos que lo pudiesen ser. Al contrario, hay escuelas de comunidades más pobres en las cuales los alumnos alcanzan un mayor progreso que los de otras escuelas con alumnos de niveles socioeconómicos más aventajados y sin embargo, en términos de resultados brutos sus logros aparecen como mediocres. En los primeros casos, erróneamente puede existir la percepción de que no se requiere mejoramiento cuando en los segundos, la situación puede llevar a serias e injustificadas desmoralizaciones.

Desde esta perspectiva es ahora posible ingresar de lleno al tema de la determinación del “valor agregado”, el cual se presentará en base de un ejemplo sencillo, el cálculo del desempeño estimado de una escuela determinada, utilizando datos agregados al nivel de la escuela y calculando la relación estadística entre el logro de sus estudiantes al ingreso a esa etapa escolar determinada y aquel obtenido al finalizar la etapa.

Lo que en la práctica se obtiene es el gráfico de las medias del logro al ingreso, en el eje de las x, contra las medias del logro al fin de la etapa, en el eje de las y, ambas para cada escuela.

Cada punto en que se intersectan ambas medias para cada escuela constituye la relación que existe para cada establecimiento entre ambas variables. De modo estadístico, luego se ajusta una recta que se sitúe a las mínimas distancias de cada punto y su inclinación (pendiente) revela el grado de asociación entre ambas medidas, estando la magnitud de tal asociación expresada por el coeficiente de correlación. La correlación que como sabemos varía entre -1 y $+1$, si fuese 0 , lo que indicaría ausencia de relación entre ambas variables implicaría que la línea fuese más o menos horizontal. Una correlación de -1 o $+1$ que mostraría una recta inclinada en un ángulo de 45 grados, involucraría una alta correlación.

La pendiente de la recta es utilizada para determinar el puntaje estimado para cada escuela. Mientras más alejada la posición de una escuela de la recta implicará que su resultado bruto difiere del esperado y que la escuela tendrá mejores o peores resultados en términos de “valor agregado”. “La diferencia entre el puntaje final esperado y el puntaje final real de una escuela provee la medida del “valor agregado”, la cual es denominada en términos estadísticos como “residuos de la escuela”. Las escuelas ubicadas exactamente sobre la recta muestran total coincidencia entre su desempeño real y su desempeño esperado. Aquellas que están por sobre la recta, a su vez, muestran estar desempeñándose mejor que lo esperado y tienen residuos positivos, mientras que las escuelas bajo la recta están desempeñándose peor de lo esperado y los tienen negativos. “Claramente las escuelas con altos o bajos puntajes de rendimiento previo pueden obtener *cualquiera* de los dos tipos de resultados, mejores o peores. “ (Sammons, Thomas y Mortimore, 1997).

Si bien la forma más evidente y simple de calcular el valor agregado es utilizando el método de Regresión Simple por Mínimos Cuadrados, hay autores como Goldstein (1987, 1995) que han destacado la importancia de utilizar técnicas de niveles múltiples y datos detallados respecto del nivel de los estudiantes, para cada alumno

individual, en el cálculo del “valor agregado”. La ventaja de utilizar tales métodos es que ellos permiten explotar la estructura jerárquica de los datos educativos, los que están configurados por estudiantes agrupados en secciones, éstas dentro de escuelas y éstas últimas dentro de distritos, etc.

Del mismo modo en que la Regresión Múltiple es usada, esta metodología puede ser empleada para ubicar diferencias tales como aquellas existentes en los desempeños de niñas y niños, los cuales al haber tomado en cuenta sus respectivas situaciones al ingreso, permiten una comparación equitativa.

Sin embargo, aunque eficaz, la metodología reseñada no permite distinciones finas entre escuelas, tales como aquellas que emergen de las comparaciones de puntajes brutos. Más bien el “valor agregado” posibilita obtener indicadores de eficacia y de la medida en que algunas escuelas se están desempeñando mejor o peor que otras, una vez que se ha ajustado por las condiciones de entrada de sus estudiantes.

Dos aspectos adicionales deben ser mencionados antes de dar por terminada esta sección. El primero se refiere al hecho de que, si bien es posible entregar información a varios niveles de agregación, debe tenerse en cuenta que la comparación tanto de los alumnos de un aula determinada, como la de las aulas dentro de una escuela y las escuelas al interior de un distrito, deben llevarse a cabo con respecto de la Media de los alumnos del universo completo al cual se refirió la medición. Lo que se acaba de señalar implica no intentar comparaciones de alumnos en un aula respecto de la Media de esa aula o de aulas dentro de una escuela respecto de la Media de esa escuela o de las escuelas de un distrito con referencia a la Media de ese distrito. La razón es, ni más ni menos, que el hecho de que al hacer la comparación que se acaba de señalar, implica que la suma algebraica de todas las diferencias sería Cero.

El segundo aspecto se refiere a que se debe tener presente que un análisis inferencial por

regresión, utilizando el “valor agregado” como variable dependiente, daría en general por resultado que el porcentaje de varianza explicado fuese mínimo. La razón estriba en el hecho de que la determinación del “valor agregado”, de entrada implicó controlar por las variables de contexto y ya se hizo cargo de gran parte de la varianza. Es por lo señalado de que la mayoría de los estudios que pretenden relacionar “valor agregado” y Factores Asociados, por lo general utilizan como única manera de determinar el efecto de estos últimos sobre el primero, el hacer uso de estudios cualitativos que no requieren el uso de técnicas de estadística inferencial y en especial las de regresión.

6. Los diferentes modelos de “valor agregado”.

Ante la imposibilidad casi generalizada de acceder a “prima facie” a información de “línea de base” es preciso enfrentar el dilema respecto de cuál metodología emplear para obtener el “valor agregado”. Las alternativas de modelos diferentes emergen directamente del tipo, y número de las variables emplear y de las distintas posibilidades de sus combinaciones.

Ante la ausencia de datos acerca de rendimiento previo es posible utilizar en su reemplazo antecedentes socioeconómicos y otras características de los estudiantes. En este sentido, en algunos estudios han explorado en forma específica el uso de las notas asignadas por los profesores de cada asignatura al fin del curso previo al cual se trata, para la Segunda Medición. Los resultados de esta exploración, que realizó el autor, para Quinto Grado en la Provincia de Buenos Aires, indican que es más recomendable el uso de calificaciones como una forma de imputar los puntajes eventualmente faltantes y no perder sujetos, antes que utilizarlos como una forma de reemplazar la línea de base obtenida a través de pruebas estandarizadas.

Un ejemplo emblemático de un enfoque alternativo cuando no se dispone de datos de rendimiento previo obtenidos a través de

mediciones estandarizadas, es el empleado por la Oficina de Estándares en Educación (OFSTED) del Reino Unido, reportado por Sammons et al., en 1994, el cual propone el establecimiento de un procedimiento provisorio para evaluar el desempeño de las escuelas, ante la ausencia de datos nacionales disponibles para los alumnos a su ingreso a la escuela secundaria. La metodología desarrollada por la OFSTED, de acuerdo a lo señalado por Sammons, Thomas y Mortimore, 1997, “ha influido considerablemente en el establecimiento de estándares y en el planteamiento de metas en el tema”. “Sin embargo”, agregan los autores, “numerosos investigadores han abordado el tema de comparar diferentes modelos para separar y medir los efectos de la escuela de aquellas características de entrada de los estudiantes, tales como rendimiento previo y factores socioeconómicos”. Los hallazgos de Thomas y Mortimore (1996) y los de Sammons et al. (1994) señalan que “un número de escuelas (quizá un 20 por ciento del total) puede obtener diferentes resultados cuando se cuenta para el análisis con datos disponibles respecto de las condiciones de entrada, (en contraste) con los resultados basados solamente en datos del

contexto de los estudiantes -tales como edad, género, y derecho a alimentación escolar.

Más aún plantean que, “Es interesante constatar que cuando se han usado adecuadas “líneas de base”, la inclusión de información socioeconómica en los cálculos contribuye poco, en cuanto a explicar las diferencias entre los estudiantes. Sin embargo, ésta es útil como una manera de lograr una “sintonía fina” de los puntajes de “valor agregado” (Willms, 1992; Sammons, 1996; Tomas y Mortimore, 1996). Como manifiestan Sammons, Thomas y Mortimore (1997) “Por lo tanto, es esencial que, donde sea posible, los sistemas de “valor agregado” estén basados en un diseño longitudinal para medir el progreso de los estudiantes, donde ambos la línea de base y los resultados, estén medidos en términos de lo que es enseñado en las escuelas: el Currículum Nacional”.

Con el fin de comentar las diferentes posibilidades entregadas por el uso de distintas medidas de “línea de base” y de variables explicativas, se presentan a continuación los cuatro “modelos” distinguidos por Sammons, Thomas y Mortimore (1997).

Modelo I	Intercepto	No incluye variables explicativas.	Entrega puntajes brutos.
Modelo II	Sólo contexto.	Incluye por ejemplo: género, edad, etnia y derecho a alimentación en la escuela.	Entrega valor agregado.
Modelo III	Sólo rendimiento previo.	Incluye medidas tales como: resultados de pruebas estandarizadas anteriores o notas.	Entrega valor agregado. Entrega valor agregado.
Modelo IV	Completo.	Incluye por ejemplo: género, edad, etnia y derecho a alimentación escolar, además de medidas de rendimiento previo tales como: resultados de pruebas estandarizadas o notas.	

A continuación se muestran, de acuerdo a un estudio realizado por los autores antes nombrados, los porcentajes de varianza explicados por cada modelo, en Lenguaje, Matemática y en Puntaje Total.

	Lenguaje	Matemática	Total
Modelo II. Varianza total explicada.	9,5	6,9	11,6
Modelo II. Varianza de la escuela explicada.	52,7	37,1	43,8
Modelo III. Varianza total explicada.	36,5	33,7	40,4
Modelo III. Varianza de la escuela explicada.	57,3	48,1	57,4
Modelo IV. Varianza total explicada.	40,9	36,6	45,9
Modelo IV. Varianza de la escuela explicada.	68,2	53,9	70,0

Como es posible apreciar, los modelos, a medida en que aumentan en los elementos que incorporan, aumenta también su capacidad de explicación de la varianza en los resultados, tanto de la escuela como en el total. En este sentido el Modelo IV es el de mayor capacidad para tal propósito. Como señalan Sammons, Thomas y Mortimore (1997) “la bondad de ajuste de los distintos modelos está representada en función del porcentaje de varianza explicada en cada asignatura y en el Puntaje Total”.

Como puede verse, el uso de medidas previas de rendimiento solamente, como variables explicativas (Modelo III) produce resultados más pobres que el Modelo Completo IV. Esto indica que el hecho de controlar por medidas adicionales de contexto, tales como género, es necesario para obtener medidas precisas de eficacia de las escuelas (residuos) Es importante también notar que el sólo uso de factores de contexto (Modelo II), explica un porcentaje substancialmente menor de la varianza total, que los Modelos III y IV, sugiriendo en forma clara que el Modelo II -caso en el cual el análisis carece de rendimiento previo- es inadecuado en cuanto a proveer controles debidos de las condiciones de ingreso de los estudiantes.

7. Limitaciones y objeciones del enfoque de “valor agregado.”

Si bien las ventajas del enfoque tratado aquí por sobre los puntajes brutos, son evidentes, él adolece de algunas limitaciones, entre ellas los efectos del error de medición y la precisión de los datos.

Otro aspecto necesario de considerar es el hecho de que parte de los insumos para la determinación del “valor agregado” son de naturaleza retrospectiva, basándose en períodos previos del devenir de las escuelas. Este es un aspecto que es preciso tener en cuenta al momento de interpretar los resultados obtenidos. Es por ello que desde la perspectiva de quién interprete los resultados, sea esta de la esfera académica, política o docente, es imprescindible que previamente a extraer conclusiones se las ponga en el contexto temporal en que se sitúan ambas mediciones.

Un tercer aspecto a considerar es que el “valor agregado” en general, presenta inestabilidades en el tiempo, lo cual hace aconsejable, de acuerdo a la opinión de la mayoría de los autores, que para poder emitir y sustentar conclusiones respecto de una escuela, sección o individuo, se disponga de a lo menos de dos determinaciones secuenciales del “valor agregado” e idealmente de tres.

Finalmente, el otro aspecto que debe considerarse es el hecho de que el “valor agregado” no necesariamente muestra comportamientos consistentes a través de diversas asignaturas, por lo cual es preciso interpretar sus resultados en cada materia y ser muy cuidadoso en cuanto a extrapolar o comparar los resultados obtenidos en diferentes materias.

En la raíz de las objeciones existen preocupaciones que exceden a los aspectos exclusivamente operativos. Uno de ellos emerge

del planteamiento de Goldstein y Thomas (1995) en cuanto a que:

“Mientras nos sea posible estudiar los factores asociados con el desempeño de los estudiantes y arribar a conclusiones acerca de cuál de ellas parece estar relacionada con el ‘éxito’, aún así parece difícil identificar precisamente cuáles escuelas están haciéndolo bien o mal ...En otras palabras, la investigación acerca de la eficacia de las escuelas es una actividad útil en nuestros intentos por obtener conocimiento acerca del proceso de educación, pero es una herramienta muy pobre para hacer que las escuelas asuman su responsabilidad.”

Valga entonces finalizar la presentación de estas ideas con un planteamiento que si bien no muestra mayor optimismo, sí coincide con una visión realista y objetiva. Dice un autor (Byatt, 1997) que “Valor agregado, ciertamente en su contexto educacional, debe ser propiamente visto como uno de esas ‘metáforas y analogías,

de las cuales (el pensamiento científico) debe hacer uso y (a la vez) sospechar”

Es posible, sin embargo comentar que al igual que la manida discusión del pasado respecto de la bondad comparativa de las pruebas estructuradas y aquellas de respuesta abierta, en el caso del valor agregado cuando éste se compara con los logros brutos, cada uno tiene sus ventajas y desventajas. De este modo, tal como no hay mejor forma de medir la presencia de aprendizajes de alto nivel de complejidad que las pruebas abiertas, no hay forma más eficaz de determinar la contribución de una escuela a sus alumnos que el uso del valor agregado. Análogamente, como no hay mejor manera de determinar la presencia de aprendizajes de bajo nivel de complejidad que las pruebas estructuradas, la determinación del logro cumple un rol irremplazable en la evaluación como certificación y como variable dependiente para la determinación de relaciones de variables de producto con los Factores Asociados a la Educación.

Bibliografía

- Byatt, A. S. (1997). The inequality express. *The American Prospect*. 20, 81-93.
- Emonds, R. (1979). Some schools work and more can. *Social Policy*, 9, (4), 28-32.
- Goldstein, H. (1987). *Multilevel models in educational and social research*. London: Charles Griffin and Company. New York: Oxford University Press.
- Goldstein, H. (1995). *Multilevel Statistical Models* (2nd. Ed.). London: Edward Arnold. New York: Halsted Press.
- Goldstein, H. y Thomas, S. (1995). School effectiveness and value added analysis. *Forum*. 37, 36-38.
- Goldstein, H. y Cuttance, P. (1988). A note on national assessment and school comparisons. *Journal of educational policy*, 3, 197-202.
- Gray, J., Jesson, D., y Jones, B. (1986) The search for a fairer way of comparing schools examination results. *Research papers in education*, 10 (1), 45-68.
- Kozol, J. (1991). *Savage Inequalities: Children in America's schools*. New York: Harper Perennial.
- Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, UNESCO-SANTIAGO, Oficina Regional de Educación para América Latina y El Caribe (1998) *Primer Estudio Internacional Comparativo sobre Lenguaje, Matemática y Factores Asociados en Tercero y Cuarto Grado, Primer Informe*. Santiago de Chile: UNESCO.

- McGeevor, P., Giles, C., Little, B., Head, P. y Brennan, J. (1990) *The measurement of Value Added in higher education. A joint PCFC/CNAA project report*. London: Council for national academic awards.
- McPherson, A. (1992). *Measuring value added in schools* (NCE briefing 1). London: National Commission on Education.
- Mkhathshwa, S. en Townsend, T., Clarke, P., y Ainscow, M (Eds.) (1999) *Third Millenium Schools. A world of differences in effectiveness and improvement*. Lisse: Swets and Zeitlinger, 339-351.
- Nuttal, D. (1991) An instrument to be honed: Tables do not reflect schools' true performance. *Times educational supplement*, 22.
- Nuttal, D., Goldstein. H., Prosser, R y Rabash, J. (1989). Differential school effectiveness. *International journal of educational Research*, 1 769-776.
- Sammons, P. (1996). Complexities in the judgment of school effectiveness. *Educational research and evaluation*. 2, (2),113-149.
- Sammons, P., Thomas, S., Owen, C. y Pennell, H. (1994). *Asssesing school effectiveness: Developing measures to put school performance in context*. London: Office for Standards in Education (OFSTED).
- Sammons, P., Thomas, S. y Mortimore, P. (1997). *Forging Links. Effective schools, effective departments*. London: Paul Chapman.
- Saunders, L. (1999). A brief History of educational 'Educational Value Added: How did we get to where we are? *School efecctiveness and school improvement*.
- Spours, K. y Hogdson, A. (1996). *Value Added and raising attainment: A formative approach. A resource pack for practitioners*. Poole: BP Education Service.
- Thomas, S. y Mortimore, P. (1996) Comparison of value added models for secondary school efectiveness. *Research papers in education*, 1, (11), 5-33.
- Townsend, T., Clarke, P., y Ainscow, M (Eds.) (1999) *Third Millenium Schools. A world of difference in effectiveness and improvement*. Lisse: Swets and Zeitlinger.
- Willms, J. D. (1992). *Monitoring of school performance: A guide for educators*. London: Falmer.