

La evaluación de la calidad

Un análisis del discurso oficial*

GABRIELA DIKER**
SILVINA FEENEY***

Revista del Instituto de
Investigaciones en Ciencias
de la Educación

Año VII, N° 12, Agosto de 1998

**Este material se utiliza con fines
exclusivamente didácticos**

* Este artículo se enmarca en el Proyecto “El Sistema Nacional de Evaluación : evaluación de su calidad”. Ubacyt F-215, bajo la dirección la Prof. Alicia W. De Camilloni.

** Licenciada en Ciencias de la Educación (UBA). Cursa estudios de Maestría en Educación y Sociedad en la FLACSO. Docente de las cátedras de Historia General de la Educación y Sociología de la Educación del Departamento de Ciencias de la Educación de la Facultad de Filosofía y Letras (UBA).

*** Licenciada en Ciencias de la Educación (UBA). Master en Tecnología Educativa de la Universidad de Salamanca. Cursa estudios de Maestría en Didáctica en la Facultad de Filosofía y Letras (UBA). Docente de la cátedra de Didáctica I del Departamento de Ciencias de la Educación de la Facultad de Filosofía y Letras (UBA).

El discurso de la evaluación de la calidad: líneas de análisis

El contexto histórico-político de emergencia de la preocupación por la calidad:

La preocupación por la calidad de la educación nace en la década del 80, en los Estados Unidos y en algunos países europeos, bajo la forma de un cuestionamiento a la productividad de la inversión en educación. En efecto, en el marco del proceso más global de crisis fiscal y retracción del gasto destinado a políticas públicas, los gobiernos comenzaron a someter a evaluación el volumen y la inversión misma en los sistemas educativos: “¿Hasta qué punto eran eficaces? ¿proporcionaban algo valioso a cambio del dinero?” (OCDE, 1991:22).

Estas preguntas no son menores. Desde la conformación de los sistemas educativos nacionales, la convicción de que a más educación seguía más progreso (ya sea porque produjera crecimiento económico, condujera a la integración política o garantizara el disciplinamiento social necesario para el desarrollo de las formas industrializadas de producción), no había sido puesta en duda en los ámbitos de decisión política.

Sin embargo, a pesar de que en la mayor parte de los países centrales se habían alcanzado ya las metas de universalización de los niveles obligatorios del sistema educativo, y de que había realizado una fuerte inversión en recursos materiales y humanos con el objeto de absorber sectores cada vez más amplios de la población, se llega a los 80 en el marco de un quiebre del desarrollo económico interrumpido que se había vaticinado en los 60, y de niveles sin precedentes de desempleo juvenil, que inspiraron la causación de que los jóvenes estaban siendo preparados inadecuadamente para el mundo del trabajo. A estos factores hay que sumarles, además, la amenaza que comenzó a construir —especialmente para los Estados Unidos— la creciente participación en el mercado mundial del Japón y los nuevos países industrializados de la zona del Pacífico.

Ante esta situación los diagnósticos críticos acerca de la calidad de los sistemas educativos comienzan a acumularse. Sólo para los Estados Unidos podemos citar: *A Nation at Risk* (1983); *American Competitive Challenge* (1983); *Academic Preparation for College* (1983); *Involvement in Learning* (1984); *To Reclaim a Legacy* (1984), *America 2000 and the politics and of reasure: democracy and cultural difference under siege* (1992), entre otros (cfr. Taylor, 1994).

En general todos estos diagnósticos coinciden en tres cuestiones: en primer lugar, en una crítica al sistema educativo por su inadecuación a las necesidades de los sectores de la producción y los servicios; en segundo lugar, en la baja rentabilidad individual (subempleo o desempleo de mano de obra calificada) y social de la inversión educativa; en tercer lugar, en los bajos niveles de logros educativos en comparación con los niveles de logros alcanzados por otros países. Respecto de este último punto resultan de sumo interés los resultados que arroja un estudio comparativo llevado adelante por el Educational Testing Service (corporación privada no lucrativa subvencionada por el Departamento de Educación de los Estados Unidos), que coloca a los estudiantes norteamericanos por debajo de sus pares coreanos, en todas las áreas evaluadas (matemáticas y ciencias) (Lapointe, A. et. al., 1989).

Para el caso de América Latina, los diagnósticos coinciden con esta descripción de la crisis del sistema, con el agregado de la falta de universalización de la educación en sus niveles obligatorios, atribuidos curiosamente también como producto de la crisis de calidad (Fernández Lamarra, 1991; Álvarez Tostado, 1987). En la Reunión del Comité Regional Intergubernamental del Proyecto Principal de Educación de la UNESCO para América Latina y el Caribe (PROMEDLAC), los ministros de educación de la región coincidieron en afirmar que “el mejoramiento de la calidad de la educación constituye una prioridad impostergable, aun cuando todavía no estén superados los problemas cuantitativos que presentan nuestros sistemas educativos” (Fernández Lamarra, 1991:1).

En otras palabras, si bien no se abandona la preocupación por la cobertura, el propósito no es ya sólo seguir expandiendo los sistemas educativos bajo las mismas condiciones sino proceder a una reestructuración compleja de éstos, bajo la premisa del mejoramiento de la calidad educativa.¹

¹ Tal reestructuración se enmarca en la propuesta más global de reforma del Estado. Estas reformas persiguen dos objetivos: 1) reducir el déficit fiscal a través de la disminución del gasto estatal y 2) consolidar un Estado “menos burocrático y más eficiente”, cuyo funcionamiento se rija más por la lógica de la productividad empresarial que por la lógica distributiva que había caracterizado al Estado benefactor, dado que, tal y como se venían desarrollando, “las actividades del sector público o estatal son vistas como ineficientes, improductivas, antieconómicas y como un desperdicio social, en tanto el sector privado es visto como eficiente, efectivo, productivo, pudiendo responder por su naturaleza menos burocrática con mayor rapidez y presteza a las transformaciones que tienen lugar en el mundo

Ahora bien, ¿qué se entiende por calidad de la educación?

Del mejoramiento de la calidad a la evaluación de logros:

La literatura sobre el tema producida tanto por organismos nacionales o internacionales como en el ámbito académico, muestra con claridad la dificultad para arribar a una definición de la idea de la “calidad”: “la definición del concepto de calidad de la enseñanza es compleja y difícil” (Soler García, 1994); “calidad puede significar diferentes standards en diferentes situaciones” (Álvarez Tostado; 1993); “El concepto de calidad es tan escurridizo como persuasivo; parece imposible llegar a un acuerdo universal o a una definición final de este concepto” (Van Vught, 1991); “calidad significa cosas diferentes para distintos observadores y grupos de interés” (OCDE, 1991); y podríamos seguir...

A falta de definiciones precisas, debajo del concepto de calidad se ubica casi todo: la relación escuela-ámbito productivo (Rodríguez Fuenzalida, 1994); la formación y captación de los docentes (Wilson, 1992; OCDE, 1991); la respuesta a las diferencias regionales y socio-culturales (Wilson, 1992; Fernández Lamarra, 1991); los modelos de gestión institucional (OCDE, 1991; Colom Cañelas, 1988); el rendimiento de los alumnos (Lapointe, A. et. al., 1989); el gasto por alumno, los materiales de instrucción, la duración del año y de la jornada escolar, el origen social del profesor (Banco Mundial, 1986).²

Semejante diversidad de definiciones y de indicadores de la calidad de la educación, haría esperable una diversidad análoga en las estrategias de intervención sobre los sistemas educativos tendientes a mejorar la calidad. Sin embargo, las medidas de mejoramiento son atravesadas en los distintos países por un eje que se impone y subordina a cualquier otra medida: *la evaluación de la calidad educativa entendida exclusivamente como medición de logros a escala nacional*.

Más aún, algunas investigaciones (Mena, 1991:31) revelan que las políticas de medición de logros a escala nacional como estrategias de mejoramiento de la calidad suelen implementarse en ausencia de una definición explícita de calidad. Se podría suponer que tal definición está implícita en la construcción de los instrumentos de medición, así como también, que para cada operativo de evaluación se puede definir “calidad” de una manera particular; lo que sigue llamando la atención es que más allá de *qué* se entienda por calidad educativa las políticas de mejoramiento sean las mismas.

Sostendremos aquí que el carácter comparativo del concepto de calidad obliga a un desplazamiento de las políticas hacia los dispositivos de evaluación, es decir, que no se puede hablar de calidad si no se evalúa. Esta evaluación ya no tendrá la función de relevar información con carácter diagnóstico, tal como ocurría en el marco de los modelos de planeamiento educativo, sino que supone la instalación de un mecanismo de control sobre un sistema educativo que los niveles centrales ya no gestionan.³

La evaluación de la calidad en la Argentina: un análisis preliminar del discurso oficial

Uno de los ejes del modelo de Reforma Educativa que se está implementando actualmente en la Argentina, es la evaluación de calidad, establecida en el marco legal de dicha reforma que es la Ley Federal de Educación, sancionada en el año 1993⁴. A continuación intentaremos reconstruir el discurso oficial de evaluación de la calidad en torno a tres ejes: qué se evalúa, cómo se evalúa y para qué se evalúa.

moderno” (Torres, 1995: 115). Y, en el llamado sector privado, productividad y eficiencia son dos cuestiones directamente asociadas con la **calidad** de los procesos y de los productos.

² En la mayor parte de los casos se combinan todos o algunos de estos elementos entre sí. Lo que interesa es señalar la diversidad y amplitud de indicadores de calidad que es posible relevar en la literatura sobre el tema. Conviene aclarar también, que las referencias bibliográficas son indicativas de las obras que enfatizan unos aspectos sobre otros, pero de ninguna manera pretenden ser exhaustivas respecto de la bibliografía disponible ni en relación con el pensamiento del autor.

³ En efecto, la preocupación por la calidad es, en la empresa, una preocupación por la productividad y, consecuentemente por la rentabilidad. Aún cuando la definición de calidad en el ámbito productivo y las formas de organización del trabajo que le están asociadas, han cambiado a lo largo de la historia del capitalismo, “es que en esta trilogía que cobran coherencia y continuidad las viejas y las nuevas discusiones sobre calidad y excelencia. Ningún empresario puede ser tan incompetente (a riesgo de desaparecer rápidamente del mundo de los negocios) como para no considerar que la calidad es importante en la medida en que garantiza la productividad (o la incrementa), lo que le asegura mantener su rentabilidad (o bien aumentarla)” (Gentili, 1993: 22).

⁴ El SNEC tiene como marco legal la Ley Federal de Educación 24.195, y más específicamente los artículos 48, 49 y 50 que se transcriben a continuación:

¿Qué se evalúa?

En el documento Bases Conceptuales del Sistema Nacional de Evaluación de la Calidad de la Educación Básica y Media, se responde extensamente a esta pregunta:

“Un primer sentido del concepto es el que surge por oposición a los fenómenos de vaciamiento [de contenidos]. En ese sentido la calidad es entendida como “eficacia”: una educación de calidad es aquella que logra que los alumnos realmente aprendan lo que se supone deben aprender –aquello que está establecido en los planes y programas curriculares- al cabo de determinados ciclos o niveles. En esta perspectiva el énfasis está puesto en que, además de asistir, los niños y los adolescentes aprendan en su paso por el sistema. Esta dimensión del concepto pone en primer plano los resultados del aprendizaje efectivamente alcanzados por la acción educativa.

Una segunda dimensión del concepto de calidad, complementaria de la anterior, está referida a qué es lo que se aprende en el sistema y a su “relevancia” en términos individuales y sociales. En este sentido una educación de calidad es aquella cuyos contenidos responden adecuadamente a lo que el individuo necesita para desarrollarse como persona (...) y para desempeñarse adecuadamente en los diversos ámbitos de la sociedad...”

Finalmente, una tercera dimensión es la que refiere a la calidad de los procesos y medios que el sistema brinda a los alumnos para el desarrollo de su experiencia educativa. Desde esta perspectiva una educación de calidad es aquella que ofrece a niños y a adolescentes un adecuado contexto físico para el aprendizaje, un cuerpo docente adecuadamente preparado para la tarea de enseñar, buenos materiales de estudio y trabajo, estrategias didácticas adecuadas, etc.”.(Doc. N° 12, 1994: 8-9)

Por otra parte el Documento N° 4 (1993) establece como competencias al SNE, las siguientes:

- “1. Establecer criterios para la medición de la calidad educativa del sistema.
2. Establecer un sistema de evaluación que permita medir la calidad de la enseñanza impartida por el Sistema Educativo Nacional en todas las jurisdicciones, ciclos, niveles y regímenes.
3. Evaluar el funcionamiento del Sistema Educativo Nacional en todas las jurisdicciones, ciclos, niveles y regímenes.
4. Evaluar la formulación de los programas nacionales de mejoramiento de la calidad de la enseñanza.
5. Implementar operativos de evaluación de calidad a fin de generar información sobre el rendimiento de los alumnos, para la elaboración de políticas de mejora.”(Doc. N° 4, 1993: 2)

A pesar de la complejidad con que el discurso oficial presenta la definición e implementación de los operativos de calidad de la educación, un análisis de las políticas efectivamente implementadas muestra una doble reducción: a) se reduce la definición de la calidad a la única dimensión “medible” esto es, al rendimiento de los alumnos (eficacia) y b) a pesar de definir como competencia del SNE las cinco arriba mencionadas, se termina reduciendo a estas competencias sólo a la implementación de operativos de evaluación de la calidad.

“Art. 48: el Ministerio de Cultura y Educación de la Nación, las provincias y la Municipalidad de la Ciudad de Buenos Aires, deberán garantizar la calidad de la formación impartida en los distintos ciclos, niveles y regímenes especiales mediante la evaluación permanente del sistema educativo, controlando su adecuación a lo establecido en esta Ley, a las necesidades de la comunidad, a la política educativa nacional, de cada provincia y de la Municipalidad de la ciudad de Buenos Aires y a las concertadas en el seno del Consejo Federal de Cultura y Educación.

A ese fin deberá convocar junto con el Consejo Federal de Cultura y Educación a especialistas de reconocida idoneidad e independencia de criterio para desarrollar las investigaciones pertinentes por medio de técnicas objetivas aceptadas y actualizadas.

El Ministerio de Cultura y Educación deberá enviar un informe anual a la Comisión de Educación de ambas Cámaras del Congreso de la Nación donde se detallan los objetivos que se establecen en la presente Ley.

Art. 49: la evaluación de la calidad en el sistema educativo verificará la adecuación de los contenidos curriculares de los distintos ciclos, niveles y regímenes especiales a las necesidades sociales y a los requerimientos educativos de la comunidad, así como el nivel de aprendizaje de los alumnos/as y la calidad de la formación docente.

Art. 50: las autoridades educativas de las provincias y de la Municipalidad de la Ciudad de Buenos Aires evaluarán periódicamente la calidad y el funcionamiento del sistema educativo en el ámbito de su competencia”.

¿Cómo se evalúa?

Las primeras acciones de la Secretaría de Programación y Evaluación Educativa tendientes a la consolidación de un SNE consistieron en una serie de *Jornadas de Discusión de la Metodología de la evaluación de la calidad de la educación*, que contaron con la participación de los Directores de Planeamiento Educativo de las distintas Jurisdicciones y representantes de las Direcciones Nacionales de Educación de los distintos niveles y ramas del sistema educativo, más el equipo de trabajo de la Dirección Nacional de Evaluación. Estas Jornadas de Discusión tuvieron lugar a lo largo del año 1991 y culminaron con la implementación del **Primer Operativo Nacional de Validación de Ítem** el día 4 de noviembre y el posterior procesamiento de la información y análisis de los resultados.

Las principales decisiones tomadas en relación a la consolidación de un SNE fueron las siguientes:

- a) La medición de logros del aprendizaje para evaluar la calidad de la educación.
- b) La construcción de instrumentos de evaluación.
- c) La consideración de las condiciones en que se dan los resultados de aprendizaje

Es nuestro interés en este apartado centrarnos en las decisiones que según la documentación oficial se han tomado alrededor de la construcción de los *instrumentos de la evaluación*⁵. Las acciones se centraron en la construcción de pruebas objetivas que respondieran al criterio de comparabilidad de los resultados: “El gran problema para cualquier tipo de acción que intente evaluar rendimientos a gran escala consiste en construir instrumentos que reúnan las características técnicas antes señaladas [pertinencia, confiabilidad y validez], así como logra que una sola y única prueba resulte apta para poblaciones diversas de estudiantes que han recibido enseñanza por parte de distintos docentes, con distintos materiales, en cursos con distintas orientaciones, y en distintos espacios geográficos” (Doc. N° 1, 1991, Cap. 1).

El problema de la validez se resuelve definiendo los objetivos que se evaluarán, “que abarcan tanto los amplios del proceso de enseñanza en su conjunto, como los particulares del curso o área que se quiere evaluar”, a partir de “una pauta ideal de lo que los alumnos deberán saber”. Esas pautas ideales fueron volcadas a una tabla de especificaciones, en la que se detallaron los alcances que se querían medir y el peso que tendría cada uno de ellos; la función de esta tabla de especificaciones era la de actuar como un contralor de la validez del instrumento que se aplicaría, en tanto permitía realizar una correspondencia precisa entre objetivos e ítem. Este modo se definieron los objetivos de la educación como los contenidos que se esperaba que los alumnos aprendieran al finalizar cada nivel del sistema educativo, para las áreas de lengua y matemática.

El paso siguiente fue la definición operativa de los objetivos en la construcción de ítems: “La estructura del ítem es una forma de control sobre el tipo de información que se quiere obtener, ya que involucra tanto al contenido como al procedimiento que se espera que se utilice a partir de ese contenido”. En el caso de esta primera experiencia de validación del instrumento de evaluación, se utilizaron ítems cerrados de selección múltiple y apareamiento y “una forma novedosa de ítem (...), llamado redes conceptuales”. Operando de este modo es posible la cuantificación –traducir a entidades mensurables- de los alcances del proceso de aprendizaje. La necesidad de cuantificar los resultados queda expresada del siguiente modo: “Esto no quiere decir que sólo así deba operarse, pero sí de que se trata de uno de los modos posibles, y el único en el caso de que se pretenda una evaluación abarcativa y amplia en cuanto a su extensión y que tienda a establecer criterios que permitan extraer conclusiones” (Doc. N°1, 1991, Cap.1).

Por último, los documentos oficiales del SNE expresan la necesidad de *considerar las condiciones en que se dan los resultados de aprendizaje*. La justificación de esta decisión aparece muy brevemente expresada en el Informe Preliminar (Doc. N°1, 1991) y gira en torno a la siguiente idea: “Para el análisis de resultados, en este caso de calidad de logro, no es suficiente considerar los puntajes obtenidos. Éstos sólo tendrán verdadero valor si se los redimensiona a partir de considerar las condiciones en que esos resultados aparecen (historia personal, características institucionales, etc.), porque no puede dejarse de considerar que el resultado es una forma de recorte arbitrario, que fuera de contexto, puede distorsionar las interpretaciones que a partir de él se hagan”. (Cap. 1, s/d). Para ello, se administraron cuestionarios a rectores donde se pedían datos del establecimiento y formularios a alumnos donde se recogían características personales de éstos; sin embargo, mucho de la información recogida en este sentido, es dejada de lado a la hora de analizar

⁵ El análisis técnico de los instrumentos de evaluación utilizados por el SNE, es otra de las líneas de indagación del proyecto Ubacyt F-215

el rendimiento de los alumnos y de proponer estrategias de mejoramiento, que, como veremos más adelante, se definen *exclusivamente* en términos de enseñar más y mejor los contenidos de las distintas áreas.

Finalmente, queda consolidado 1993, el **Sistema Nacional de Evaluación** de la calidad de la educación. A la fecha, se han llevado a cabo cuatro operativos nacionales de evaluación de la calidad, que tienen como punto de partida el año 1993. Sus notas características pueden resumirse como sigue:

1) Se mide el logro de los aprendizajes (rendimiento) de los alumnos de la enseñanza primaria y media, al finalizar cada nivel, en las áreas instrumentales (Lengua y Matemática). Esta selección de áreas de conocimiento se mantiene en los operativos 1993 y 1994; pero en 1995, por primera vez se administran -además- pruebas de Ciencias Naturales y Ciencias Sociales en 7º grado, y pruebas de Lengua y Matemática en 3º grado (EGB 1) y en 2º año (EGB3), previendo la implementación de la nueva estructura de niveles y ciclos.

2) Se consideran objetivos complementarios, aportar al conocimiento de la asociación del rendimiento escolar de los alumnos, con variables relacionadas con la conformación de sus hogares, la educación de los padres, los antecedentes escolares de los alumnos, las características de los docentes del curso y del director del establecimiento.

3) Para la aplicación de la pruebas se utilizaron técnicas de muestreo, ya que se decidió trabajar con una muestra y no con la totalidad (censo), atendiendo a “los objetivos que se persiguen, el presupuesto disponible y la oportunidad deseada en la presentación de los resultados”. En términos de diseño muestral, se utilizó la técnica del *muestreo aleatorio estratificado*⁶ para construir la muestra que comprendía a docentes (maestros y profesores) y directores de ambos niveles de enseñanza, y se aplicó la técnica de *diseño complejo*⁷ para construir la muestra de alumnos (ambos niveles) y familias (nivel primario) (Doc. N° 2, 1993).

4) A partir del operativo 1994 se incrementa la cobertura, alcanzando, con una muestra ampliada, a la totalidad de las provincias del país (Doc. N° 8, p.3). También se llevan a cabo dos experiencias de evaluación censal, en las Provincias de San Juan y Tierra del Fuego, que deciden aplicar las pruebas a la totalidad de los alumnos de 7º grado y 5º año. Esto marca una tendencia de aceptación por parte de algunos gobiernos provinciales a instalar una “cultura de la evaluación” como beneficiosa en sí misma.

5) El operativo 1995, incorpora otra novedad -acordada por el conjunto de las jurisdicciones en encuentros nacionales y regionales-, que consiste en confeccionar una tabla-base para cada uno de los cursos, que establece el punto de partida para construir el conjunto de las pruebas a aplicar. Esta tabla incluye los contenidos y las competencias básicas o mínimas que se tendrán en cuenta, así como también la cantidad de ítems que en cada caso se incorporarán. En este mismo orden de cosas, junto con los resultados del operativo 1995 se decide dar a publicidad la totalidad de las pruebas administradas en el operativo anterior (1994), “con la finalidad de que sean conocidas, analizadas y utilizadas por los docentes y así posibilitar que cada institución las pueda aplicar para confrontar resultados nacionales y jurisdiccionales con los alcanzados por sus propios cursos.” (Doc. N° 10, 1995: 5).

6) El tipo de resultados que se presentan son, en su mayoría, resultados globales nacionales de rendimiento y resultados globales jurisdiccionales de rendimiento, discriminados en: porcentajes de respuestas correctas por nivel, disciplina, ámbito y modalidad, contenidos que presentaron mayor dificultad en ambas áreas y niveles, rendimientos comparando regímenes estatal y privado en ambas áreas y niveles, porcentaje de rendimiento por competencia para cada una de las áreas en ambos niveles.

Junto con los resultados del operativo 1995 -correspondientes a las áreas de lengua y matemática- se publican unas tablas con los *valores de rendimiento por jurisdicción y nivel (medias provinciales)* ordenados de mayor a menor rendimiento y con información que incluye las muestras 1993, 1994 y 1995, en los casos que fue posible. La información incluye, también, la media nacional de rendimiento de los alumnos, que marca una línea divisoria entre las jurisdicciones que quedaron por encima y por debajo de esa media de rendimiento, en cada una de las áreas -lengua y matemática- y en cada uno de los cursos donde se administraron las pruebas -3º y 7º grado, y 2º y 5º año.

⁶ El muestreo es estratificado cuando previo a la selección de la muestra el universo es dividido en clases separadas (estratos), y la muestra se elige independientemente en cada estrato por el procedimiento conocido como muestreo aleatorio simple sin reposición.

⁷ Se denomina complejo a un diseño muestral que combina dos o más técnicas simples de muestreo, por ejemplo, la estratificación previa con el muestreo por conglomerados.

Nos parece importante destacar cómo van avanzando las acciones del SNE, en el sentido de prescribir lo que pasa dentro de cada una de las aulas que conforman el sistema educativo nacional, al punto de indicar contenidos, competencias y hasta actividades que luego serán evaluadas nuevamente desde y por el nivel central. Evidentemente este tipo de acciones tiene consecuencias directas en el denominado "currículum real" más allá de las decisiones que automáticamente puedan y deban haber tomado las jurisdicciones a la hora de elaborar sus diseños curriculares, de acuerdo con lo establecido en la Ley Federal de Educación.

¿Para qué se evalúa?

En todos los países las políticas de evaluación de la calidad se instalan bajo la convicción de que la calidad del sistema educativo es baja. Ante esto cabría preguntarse: ¿para qué invertir tantos recursos en la elaboración e implementación de costosos y sofisticados sistemas nacionales de evaluación de la calidad educativa, cuyos resultados -o más bien, cuyos malos resultados- se conocen a priori? Ante esta pregunta podemos reconocer dos líneas de respuesta, ampliamente utilizadas hoy en el contexto argentino:

1) *Se evalúa para generar una información sistematizada sobre el funcionamiento del sistema* que los organismos oficiales históricamente no han construido: "En lo que respecta a la evaluación de aprendizajes, aun cuando en nuestro país se han realizado estudios, se carece de información secuenciada sobre los logros de los alumnos, así como de investigaciones sistemáticas acerca de los factores que afectan esos logros. Los indicadores disponibles, tales como deserción y repitencia, tienen utilidad limitada, porque no proveen datos acerca de lo que aprenden los alumnos. Además, las mediciones de los aprendizajes con fines de promoción, no poseen la estandarización necesaria para que sean tomados como indicadores de calidad de educación". (Doc. N° 7, 1994: 1). "La puesta en primer plano del problema de la calidad de los aprendizajes torna absolutamente insuficientes los indicadores tradicionalmente empleados para evaluar el desempeño de los sistemas educativos: evolución de la matrícula, cobertura, repetición, deserción, etc. Dichos indicadores respondían al supuesto implícito de que dentro de la escuela se aprendía, supuesto que ha dejado de tener vigencia" (Doc. N° 8, 1994: 9).

Digamos que, realizados bajo ciertas condiciones, los operativos nacionales de evaluación de la calidad pueden suministrar información más precisa acerca de en qué áreas del currículum el rendimiento de los alumnos (que es lo que en definitiva miden los operativos de evaluación de la calidad) es más bajo, en qué escuelas, en qué provincias, en qué países. Desde ya, parece fundamental producir toda la información posible sobre el sistema, en particular en países como el nuestro, donde es llamativa la escasez y la irregularidad de la información de que se dispone sobre la situación educativa. Pero, si lo que se pretende es "*conocer lo que se aprende en las escuelas*" el tipo de resultados que proveen los operativos en la Argentina, no aportarían información relevante habida cuenta del nivel de agregación de los resultados para cada una de las áreas. El tipo de información que se presenta como resultados de los operativos tiene la siguiente forma: porcentaje de respuestas correctas, incorrectas y sin respuesta por ítem, por jurisdicción y a nivel nacional.

2) *tal información es imprescindible para intervenir pedagógicamente en el mejoramiento de los aprendizajes*: "...evaluar es obtener información sobre los logros de aprendizaje de los alumnos con el objeto de identificar los problemas y sus causas, para poder generar distintas estrategias que aporten soluciones específicas para cada una de las dificultades" (Doc. N° 7: 1). "La información acerca de resultados alcanzados es un insumo insustituible para rediseñar y actualizar contenidos y metodologías. Sólo a partir de la información provista por la evaluación se puede planificar y saber por qué rumbo está andando el sistema educativo en su conjunto y cada uno de los tramos o subsistemas que lo componen". (Doc. N° 1, 1991: Introducción).

Profundizando esta perspectiva, el SNE propone vincular los niveles de rendimiento de los alumnos con diferentes variables con el fin de aislar aquellos aspectos de la tarea escolar sobre los que es necesario intervenir. "Para el análisis de resultados, en este caso de calidad de logro, no es suficiente considerar los puntajes obtenidos. Estos sólo tendrán verdadero valor si se los redimensiona a partir de considerar las condiciones en que esos resultados aparecen (historia personal, características institucionales, etc.), porque no puede dejarse de considerar que el resultado es una forma de recorte arbitrario, que fuera de contexto, puede distorsionar las interpretaciones que a partir de él se hagan". (idem, Cap. I: s/d).

Como decíamos más arriba, se considerarán entonces objetivos complementarios "aportar al conocimiento de la asociación (correlación) del rendimiento escolar de los alumnos con variables relacionadas con la conformación de sus hogares, la educación de los padres, los antecedentes escolares de

los alumnos, las características de los docentes del curso y del director del establecimiento" (Doc. N° 3, 1993: 3).

En este sentido, el universo a *investigar* (el subrayado es nuestro) definido en el Documento N° 3, es el conjunto de alumnos que en la Argentina cursan, en 1993, el último grado de la enseñanza primaria común y de la enseñanza media en las modalidades Bachillerato, Comercial, Técnica (Industrial y Agraria). Pero las unidades de observación a investigar son el alumno, la familia del alumno, director del establecimiento donde cursa el alumno⁸.

Pero la presentación de estos resultados de las correlaciones realizada en mayo de 1994, contiene la distribución de frecuencia de algunas variable que fueron medidas en los diversos cuestionarios aplicados a docentes, alumnos y directores, y se excluyen los resultados obtenidos a través de los cuestionarios administrados a familias. Sólo se establecen correlaciones entre los rendimientos de matemática y lengua y algunos factores escogidos a modo de ejemplificación: recursos didácticos del establecimiento, tamaño del establecimiento, antigüedad en la docencia, que no proveen de información distinta o más relevante que las numerosas investigaciones tanto a nivel nacional como internacional, acerca de las variables que intervienen en la producción del éxito y el fracaso escolar. Un operativo de tal alto costo y complejidad, sólo estaría justificado si puede proveer de información diferente, sea porque determine nuevas variables asociadas al logro, cosa que no ocurre (además en ese caso, ¿por qué operativos de evaluación, y no investigaciones educativas?⁹; sea porque permite discriminar cada área del curriculum escolar, cuál es el factor que produce cierto nivel de logro, pero para ello, la agregación de información en indicadores nacionales y el establecimiento de rankings por provincias no tendría sentido.

Por otra parte, si se trata de evaluar para obtener información más precisa de las variables asociadas al logro en cada área, habría que determinar la unidad de intervención -¿el alumno que no rinde? ¿el maestro del alumno que no rinde? ¿la escuela?, el departamento, la provincia?-. Resulta evidente que resultados del tipo "el 60 % de los alumnos argentinos de 4º grado no supera más del 43% de los ítems de Matemática", no permiten establecer sobre qué factores intervenir, ni tampoco sobre quiénes intervenir. En ese caso, sólo se podrá concluir en términos generales, que hay que "enseñar más Matemática". A lo sumo se podrían determinar los temas que presentaron estadísticamente mayores niveles de fracaso, cuya enseñanza habría que intensificar. Semejante nivel de generalidad y ambigüedad no parece ser una buena base para el diseño de políticas de mejoramiento de la calidad. La reducción de la escala a niveles provinciales conserva el mismo problema.

Sin embargo, ésta parece ser la opción que ha seguido la Secretaría de Programación y Evaluación Educativa del MCE a través de la producción de las "Recomendaciones metodológicas para la enseñanza" y "Aportes para el mejoramiento de los resultados del aprendizaje".

En efecto, estos materiales tienen como objetivo plantear "algunas alternativas metodológicas para la enseñanza de los contenidos identificados que mostraron dificultad". Fueron elaborados para los niveles primario y medio en las dos áreas evaluadas y presentan, en términos generales la siguiente estructura: 1) presentación del ítem (se reproduce el ejercicio tal como fue presentado en la prueba); 2) identificación del "área de conocimiento en el que está incluido el tema evaluado"; 3) identificación de la operación requerida (ubica cuáles son los procedimientos y/o estrategias que el alumno debió utilizar para resolver el problema); 4) identificación de la dificultad que presentó el ítem a los alumnos (se describen las posibles causas que puedan haber llevado a la elección de algunas de las opciones que no constituyen la respuesta correcta); 5) consideraciones didácticas y 6) sugerencia de actividades para desarrollar en el aula que permitan ejercitar las competencias comprometidas en la resolución del tipo de problema planteado.

Estas recomendaciones están dirigidas al conjunto de los docentes que enseñan cada área en las escuelas medias del país y al conjunto de los docentes de nivel primario. El análisis de estos materiales permite, en principio, algunas observaciones: en primer lugar, en cuanto al recorte de la unidad de

⁸ Los instrumentos para recoger información en cada una de estas unidades de observación son:

- en el caso de alumnos: pruebas de matemática y lengua (ambos niveles) y cuestionario (distinguiendo alumnos rurales y urbanos),
- en el caso de docentes: cuestionario maestro de lengua y cuestionario maestro de matemática y cuestionario maestro rural (idem para profesores de nivel medio de lengua y matemática),
- en el caso de familias: cuestionario director nivel primario (distinguiendo rurales y urbanos) y cuestionario director nivel medio.

⁹ Recordemos que en el organigrama del MCE, la Dirección de Investigación y Desarrollo forma parte de la Subsecretaría de Programación y evaluación educativa, mientras que la Evaluación de la Calidad tiene subsecretaría y Dirección nacional paralelas.

intervención, se recurre finalmente a la consigna “hay que enseñar más Matemática”, medida cuya dudosa eficacia ya ha sido analizada; en segundo lugar, a pesar de las correlaciones con otros factores que incidirían en los bajos rendimientos, lo que está en la base de estas estrategias de mejoramiento es la hipótesis excluyente de que el bajo rendimiento de los alumnos es producto de una mala enseñanza; y en tercer lugar nos parece necesario destacar que el hecho de que desde un nivel central que no tiene ya la responsabilidad sobre la gestión de ningún establecimiento, se prescriban *actividades para realizar en el aula*, constituye un mecanismo de regulación que podría inscribirse en lo que hemos llamado un nuevo estilo de gestión política del sistema.

Finalmente, si se pretendiera reducir esta incertidumbre reduciendo más aún la escala, y se tomara como unidad de análisis una escuela, un grado dentro de la escuela o un alumno, nos encontraríamos con que cuanto más se particulariza la unidad de análisis -aún si esto permitiera asociar con mayor precisión factores y rendimiento- más se imposibilitaría la definición de una intervención en la escala política. La única política posible sería que cada escuela dispusiera de toda la información que provea el dispositivo de evaluación y elaborara sus propias estrategias de mejoramiento; pero, en primer lugar, las escuelas por su labor cotidiana ya disponen periódicamente de información sobre el rendimiento de todos sus alumnos, con lo cual no se estaría produciendo información nueva en esta escala -excepto por su comparabilidad a nivel nacional, que no sería particularmente útil a los fines del diseño de la intervención pedagógica-; y, en segundo lugar, si conocer con exactitud los guarismos fuera suficiente para diseñar estrategias de mejoramiento, la tan mentada crisis de la calidad educativa quedaría reducida a un problema de disponibilidad de información. La ingenuidad del razonamiento es evidente.

El discurso internacional de la evaluación ofrece una tercera línea de respuesta, por cierto menos técnica pero más contundente, a la pregunta ¿para qué evaluar la calidad?, que consiste en afirmar que *la evaluación misma produce una elevación de la calidad*: “la evaluación ejerce una doble función: 1) alimentar la reflexión de cada actor (enseñantes, administrativos) sobre el estado del sistema en el cual trabaja e, indirectamente, sobre su propia acción sobre el sistema, y 2) provocar, gracias a la comparación, una mejor comprensión, una apreciación más exacta del sistema, de los establecimientos escolares y de las unidades educativas, y también favorecer una cierta emulación entre ellos”. (Thelot, 1994: 8). En una formulación menos elegante, la Asociación Internacional para la Evaluación del Rendimiento Educativo (IEA) afirma que “los instintos competitivos de la mayoría de las sociedades modernas las empujan a ver los resultados de los estudios como retos para mejorar el rendimiento” (Lapointe, A., et. al., 1989:8); de hecho, “los problemas se agudizan debido a la falta de competencia entre las escuelas” (Banco Mundial, 1993). En efecto, “creo que si vamos a tener un sistema educativo inglés-británico necesitamos tener tantas pruebas como podamos para que nos lleven a los niveles de los franceses y alemanes” (declaraciones del Secretario de Estado para la Educación de Gran Bretaña, Mr. Patten, a la BBC, el 28/2/1993, citado en Graham, 1994:36).

Es decir, la evaluación provocaría en los actores -sea porque aumenta la reflexión sobre la tarea educativa, porque pone información a su disposición, o porque estimula la competitividad-, una acción que mejorará la calidad de aquello que se evalúa. En esta línea, el artículo 48° de la Ley Federal de Educación de la Argentina, afirma que las autoridades políticas “deberán garantizar la calidad de la formación impartida en los distintos ciclos, niveles y regímenes especiales *mediante la evaluación permanente del sistema educativo*”. Un año de implementación del Operativo Nacional de Evaluación de la Calidad parece confirmar la correlación entre evaluación y calidad establecida en la ley, a juzgar por los notables avances en los resultados del Segundo Operativo Nacional de la Calidad, por sobre los resultados del primero, con sólo un año de diferencia entre ambos.

Algunas ideas para seguir pensando: la evaluación de la calidad como dispositivo regulado

A partir de lo expuesto, los sistemas nacionales de evaluación de la calidad deberían ser entendidos, más que como un modo de intervención eficaz para mejorar el rendimiento de los alumnos, más que como un sistema de relevamiento de información, como un nuevo estilo de gestión político-educativa, que permite retener en los niveles centrales al capacidad de control sobre el sistema educativo, sin intervenir directamente en su gestión.

En efecto, la evaluación de la calidad a nivel nacional constituye uno de los pilares de un modelo de estado que, a nivel central, no presta directamente el servicio educativo, pero define las metas que deben ser alcanzadas y evalúa su cumplimiento.

El control que algunos gobiernos ejercen sobre el sistema educativo a través del dispositivo de evaluación de la calidad, se expresa de manera visible, a través de la distribución de recursos a las instituciones escolares o la determinación de los incrementos salariales de los docentes, según los niveles de rendimiento de los alumnos (Chile o Gran Bretaña, para citar dos claros ejemplos).

Ahora bien, más allá de estas medidas que pueden ser anotadas como mecanismos directos de control político sobre el sistema, interesa retener que la evaluación de la calidad produce otras modalidades de control -devenidas en autocontrol- en tanto se constituye como un dispositivo que se impone con fuerza de evidencia, como resultado de varias operaciones que se pueden resumir en las siguientes tesis:

- 1) La evaluación de la calidad refleja “lo que pasa en las escuelas”, el estado de la educación en su versión más fidedigna, y por lo tanto es una información que nadie se puede negar a conocer (¿qué puede tener de malo saber *lo que de verdad pasa* en el sistema educativo?).
- 2) La evaluación de la calidad se presenta como un procedimiento técnico, y, en consecuencia, exento de conflictividad política. En efecto, la sofisticación técnica que exigen los procedimientos de evaluación a escala nacional, produce un efecto de verdad que naturaliza y despolitiza el dispositivo todo. En palabras de Tadeu da Silva, se trataría de “outra das operações centrais do pensamento neoliberal em geral e, em particular, no campo educacional, consistente em transformar questões políticas e sociais em questões técnicas. Nessa operação, os problemas sociais -e educacionais- não são tratados como questões políticas, como resultado -e objeto- de lutas em torno da distribuição desigual de recursos materiais e simbólicos e de poder, mas como questões técnicas...” (da Silva, 1994:8).
- 3) A la vez, el discurso evaluador se presenta a sí mismo como fundante, como enteramente nuevo: “la preocupación por la evaluación del sistema educativo es reciente: ella data de cinco a diez años” (Thelot, 1994:143). Más aún, suele afirmarse que una de las razones por las que los sistemas educativos llegan a una situación tan crítica como la actual es que nunca antes habían sido evaluados. No hace falta argumentar demasiado sobre el efecto de naturalización y desactivación de cualquier clase de respuesta política, que produce la negación de la historia de un proceso social. Para Tadeu da Silva, este procedimiento forma parte de lo que él llama la “pedagogía de la política”: “quando um discurso desse tipo se torna hegemônico ele não apenas coloca ‘novas’ questões, introduz novos conceitos e categorias; ele, sobretudo, desloca e reprime outras categorias, obscurece a memória popular, ocupa o lugar de categorias que moveram nossas lutas no passado, redefinidas agora como anacrônicas e ultrapassadas” (Thelot, 1994: 21).

Como producto de estas condiciones de instalación del discurso de la evaluación de la calidad, ella es visualizada cada vez más, no sólo como el mejor procedimiento para conocer la situación actual del sistema educativo, y así mejorar su calidad, sino como el *único medio posible*. Estamos entonces frente a un dispositivo que, en la medida en que es asumido como legítimo por los sujetos sobre los cuales se aplica, prefigura sus acciones y sus percepciones, convirtiéndose así en un medio de regulación, altamente eficaz, en la producción de nuevas relaciones sociales en la educación. Analicemos algunas de ellas:

- 1) En primer lugar, la publicidad de los resultados de los operativos nacionales de evaluación de la calidad -uno de los más valorados argumentos de justificación de esta política-, en general bajo la forma de rankings que ordenan las instituciones en una escala “bueno-malo”, altera la tradicional relación escuela-padres: en un sentido, los padres ahora disponen de una información que tradicionalmente era provista por la escuela, que les permite tomar decisiones (cambiar de escuela), así como demandar y exigir que se modifiquen las pautas de trabajo de la institución, bajo la expectativa de mejorar el rendimiento escolar de sus hijos, habilitándose una posibilidad de intervención mucho mayor de los padres en el trabajo escolar. Sin embargo, el tipo de intervención a que habilita el manejo de los resultados de los operativos de evaluación de la calidad coloca a los padres más en el rol de consumidores, que toman decisiones por intereses individuales, que en el de ciudadanos preocupados por la marcha de la educación en su conjunto. La situación se agrava si se considera además, que no todos los padres disponen de las mismas posibilidades de elección y demanda hacia las escuelas, y que, por otra parte, estas posibilidades en lugar de incrementarse tienden a reducirse, como consecuencia de las presiones fiscales del Estado neoliberal (el caso británico es paradigmático al respecto: en 1988 se implanta la política de “libre matriculación” sobre la base de la información que provee a los padres el sistema

nacional de evaluación, mientras que, esta misma información es utilizada por el Estado central para cerrar establecimientos o reducir puestos de trabajo) (Graham, 1994).

- 2) El Estado evaluador se coloca fuera del sistema educativo, con la única “responsabilidad de responsabilizar” a las escuelas, y en particular, a los docentes, del estado de la educación, midiendo cuánto el rendimiento de los alumnos, se acerca o se desvía de la norma impuesta por él. En otras palabras, el estado evaluador es a la vez “legislador y juez”¹⁰, mas nunca responsable de la calidad educativa. En el caso argentino el Estado Nacional, ahora sin escuelas, a lo sumo provee a los docentes de más “normas para enseñar bien, aquello que se enseñó mal”, y que, por tanto, produjo un bajo rendimiento en algunos ítems (véanse las “Recomendaciones metodológicas para la enseñanza”. Subsecretaría de Evaluación de la Calidad Educativa, MCE, 1995). Si el rendimiento en futuros operativos volviera a ser bajo en los mismos ítems, será, sin dudas, porque los docentes no siguieron las “recomendaciones para la enseñanza” correspondientes.
- 3) Esta hiper-responsabilización de los docentes del bajo rendimiento de los alumnos, y por ende de la baja calidad educativa, sumada a la publicidad de los resultados que mencionáramos más arriba, impulsa a maestros y profesores a “enseñar para las pruebas”, sea porque está en juego la materialización de una sanción, a través de su salario (Gran Bretaña), sea porque se ve frente a la obligación pública de rendir cuenta (accountability) de su trabajo. De este modo, aun en los contextos en los que se les da un alto margen de autonomía a las instituciones y a los docentes en la definición de lo que se suele llamar “el aspecto pedagógico” de la institución (con formulaciones que recuerdan con nostalgia la agenda progresista de años atrás), en la práctica se reemplazan las metas de aprendizaje definidas institucionalmente, por las definidas en las pruebas, en tanto son éstas últimas las que definen “la calidad”. De este modo “en lugar de que las pruebas sean utilizadas para sostener y validar los juicios de los maestros sobre las metas de aprendizaje definidas autónomamente, ahora son usadas para definir esas metas y para reemplazar al maestro como agente de evaluación” (Elliott, 1994:46). Efecto no menor, si se considera que tiene lugar sin intervención directa sobre el sistema, sino a través de un proceso de interiorización de las pautas de control.

Para finalizar, nos interesa rescatar otros modos de pensar la calidad de la educación: “es importante reafirmar que la retórica neoliberal sobre calidad en educación, a pesar de ser hegemónica en este momento no es la única manera de definirla. Aunque precise ser revisada y revitalizada hay también una tradición democrática de calidad en educación, desarrollada en años de lucha por educadores y educadoras involucrados en la teoría y en la práctica educacionales. Esta tradición está basada en una concepción sociológica y política de la educación y su noción de calidad está estrechamente vinculada al combate contra las desigualdades, las dominaciones y las injusticias de cualquier tipo. En esta perspectiva, la calidad es un concepto inevitablemente político”. (da Silva, 1996:170).

¹⁰ Tomamos esta expresión de Narodowski, M. (1995).