

POLÍTICAS EDUCATIVAS Y “EVALUACIÓN DE LA CALIDAD”*

Guillermo Bustamante Zamudio**
Luis Guillermo Díaz Monroy***

Algunos antecedentes

En decenios anteriores, los principales referentes de la política educativa eran algunos indicadores de gestión y eficiencia. Documentos del Departamento Nacional de Planeación [cf DNP, 1981] sobre educación en esa época se refieren a Cobertura, Escolarización, Número de docentes y Eficiencia interna (entendida ésta como: *retención*, escolaridad, tiempo promedio de duración de los estudios, alumnos por profesor y costos). Las cifras de unos y otros factores mostraban cierta independencia entre sí (dicho en jerga estadística: no se “cruzaban”), tal vez en el entendido de que cada uno tenía su propia justificación. En estos estudios, un parámetro importante para “medir” la calidad de la educación, en relación con el desempeño de los estudiantes, era el examen de Estado.

En la década del 90, este panorama cambió en Colombia. Desde comienzos de los 90, aparece cada vez más visible una política educativa que se diseña mencionando un apoyo en los resultados de la “investigación” que relaciona (que “*cruza*”) *factores asociables y desempeños de los estudiantes*:

Mejorar la calidad de la educación exige no sólo medir los resultados del logro escolar de los estudiantes sino recopilar información en torno a los procesos y recursos que contribuyen a generar dichos resultados. El conocimiento de la relación entre estos dos tipos de aspectos facilita la planificación y ejecución de acciones tendientes a la solución de los problemas en la calidad educativa [MEN, 1992:10].

Además de los factores asociables y de los desempeños de los estudiantes, también se ha querido introducir una dimensión “no académica” del trabajo escolar, mediante pruebas denominadas “de desarrollo personal y social”¹⁵ que intentan “aproximarse a aquellos factores que el individuo requiere para su proceso de adaptación al contexto social, cultural y político del país, no incluidos en los programas curriculares de un área específica” [MEN, 1992:24]. Esta dimensión —según reconoce el mismo Ministerio de Educación Nacional— ha tenido poco éxito. En todo caso, la definición de tal dimensión parece plantear, de un lado, que los programas curriculares poco tendrían que ver con los factores necesarios para la “adaptación al contexto social, cultural y político” (con lo que se hace posible dudar de su pertinencia) y, de otro, que el ser humano no produce tal contexto, sino que se adapta a él (lo que pondría en duda la especificidad del conocimiento humano).

* Este artículo hace parte de la investigación “Análisis de algunos factores asociables al desempeño en el área de lenguaje: el Distrito Capital frente al resultado nacional”, financiada por el IDEP para la vigencia 1999-2000.

** Universidad Pedagógica Nacional; bustama @ colnodo .apc.org

*** Universidad Nacional de Colombia; lgdiaz @ matematicas.unal.edu.co

¹⁵ La evaluación de impacto del Plan de Universalización de la educación básica primaria “que nos ha servido de base para la investigación” introdujo esta dimensión.

Esta “nueva” manera de hacer las cosas, conocida como evaluación de la calidad educativa, es una dinámica internacional, responde a requerimientos internacionales, con el aval de quienes toman las decisiones en política educativa; lo ejemplifican los siguientes hechos:

- Al argumentar a favor de la creación de un sistema de evaluación, el Ministerio de Educación Nacional (*MEN*, 1991:6) enumeró una serie de documentos que venían sugiriendo la evaluación de la calidad (algunos desde hacía 10 años); entre ellos se sitúan recomendaciones de la *UNESCO* y las *condiciones* de un contrato con el Banco Mundial;
- Ocho años después, la justificación sigue siendo igual de eufemística; Peña (1999:26) plantea que, de un lado, habría una necesidad de unificar información a escala nacional, dada la autonomía de los establecimientos educativos; y, de otro lado, habría “nuevas demandas” de información y de rendición de cuentas que varios actores formulan: “el usuario directo, que quiere saber qué pasa con sus hijos en la escuela” (como si siempre no hubiera sido así, como si a esa demanda siempre se hubiera respondido con un Sistema de Evaluación); “el sector productivo y la sociedad en su conjunto, que necesitan información sobre lo que sucede en el sistema educativo, dados los vínculos entre la educación y el desarrollo social” (algo que tampoco parece nuevo, aunque se hace difícil saber quién puede hacer demandas a nombre de “la sociedad en su conjunto”); “y el sector de los financiadores de la educación, cuyo interés tiene que ver con sus expectativas sobre los rendimientos de los recursos que se invierten en ésta” (esta es la única demanda que sí tiene demandante con nombre propio).
- Idénticas medidas de evaluación abundan en América Latina, con propósitos, instrumentos e incluso nombres muy parecidos: *SINECE* en Colombia¹⁶, *SINEC* en Argentina, *SIMCE* en Chile, Sistema Nacional de Evaluación en México... (idéntico nombre que en Estados Unidos, Italia e Inglaterra) [cf. Peña, 1999].

Las Cumbres Iberoamericanas de Jefes de Estado y de Gobierno (“escenario natural de debate de los problemas de la región”¹⁷) se realizan desde 1991 con temas específicos¹⁸. Pues bien, la quinta cumbre (San Carlos de Bariloche, Argentina, octubre de 1995), que tuvo como tema central “La educación como factor esencial del desarrollo económico y social”, creó el *Programa de evaluación de la calidad de la educación* para todos los países de Iberoamérica. Recientemente (Buenos Aires, marzo de 1997) el Comité Ejecutivo de dicho programa tuvo su primera reunión. Con esto queda bendecido como iniciativa regional lo que fue una imposición consentida país por país.

- El primer instrumento que se desarrolló en Colombia para evaluar la calidad de la educación en el área de lenguaje fue abandonado a favor de otro tipo de instrumento, pero la decisión no la tomaron los grupos de trabajo en el país, sino una comisión del Banco Mundial [cf. Bustamante y Jurado, 1993].

¹⁶ *SINECE* es la sigla del “Sistema Nacional de Evaluación de la Calidad de la Educación”, también conocido como *SABER*.

¹⁷ Declaración del presidente Pastrana al diario EL TIEMPO, noviembre 14 de 1999, pág. I5A.

¹⁸ Hace poco se realizó en La Habana, Cuba, la IX cumbre, con el tema de la Economía global.

- El país viene participando en procesos evaluativos internacionales como el Third International Mathematic and Science Study - Timss [cf. Díaz, 1998], que pretenden estandarizar los procesos evaluativos a escala internacional; por ejemplo, en el asunto del tiempo de estudio: considerándolo como asociado positivamente al desempeño (a más tiempo, mejores desempeños), se compara el tiempo que en los distintos países los niños permanecen en la escuela, cotejándolo a su vez con los desempeños promedio que cada país logra en las pruebas internacionales [cf. MEN, 1994 y 1996; UPN, 1999].

¿Ha habido “Sistema”?

El *Plan de apertura educativa* del presidente Gaviria, se proponía evaluar la calidad mediante pruebas aplicadas anualmente, gracias a las cuales se tendría información sobre los mismos cursos, con lo que se podrían hacer comparaciones para verificar qué había pasado con los desempeños por curso después de las inversiones; y se tendría información sobre los mismos niños —una cohorte— a lo largo de su educación, con lo que se podrían hacer comparaciones para verificar si sus desempeños habían o no mejorado. De ahí que se hicieran pruebas periódicas en 3^o, 5^o, 7^o y 9^o: “En la primera etapa, el sistema se concentrará en desarrollar pruebas que evalúen los logros cognitivos de los estudiantes de primaria en las áreas de matemáticas y lenguaje. Luego, evaluará a los estudiantes de secundaria en las mismas áreas. Posteriormente, abarcará todas las áreas académicas en ambos niveles” [DNP, 1991:36]. Más adelante, ante el evidente vacío que el grado 11 representaba en esta serie, se pensó en por qué el examen de Estado no culminaba todo el sistema; cosa que la *Misión de educación ciencia y desarrollo* planteó como factor de mejoramiento de la calidad y que hoy vemos realizado en el publicitado cambio de la prueba de Estado¹⁹.

Se creó, entonces, el *SINECE* en Colombia, desde comienzos de los 90, y ha venido llevando a cabo la evaluación de la calidad (léase: correlación entre desempeño y factores asociables) que sirve para ser mencionada a la hora de justificar la política educativa. El plan de trabajo que se trazó para las aplicaciones cada dos años, en ambos calendarios, puede verse en la primera publicación de resultados que se hizo [MEN, 1992:33]. Con la idea de *sistema*, también se pretendía “recopilar, documentar, clasificar y almacenar instrumentos de evaluación que hayan sido utilizados por el propio Sistema de Evaluación o por otros sistemas nacionales o internacionales [...] mantendrá un banco de ítems utilizados en diferentes pruebas y se encargará de su clasificación y análisis” [MEN, 1991:24].

No obstante, en Colombia es posible interrogar la sistematicidad del proceso, tal como la palabra “sistema” promete, en los nombres “Sistema Nacional de Evaluación de la Calidad de la Educación”, antes de la Ley 115/94, y “Sistema Nacional de Evaluación de la Educación”, a partir del artículo 80 de dicha Ley. A continuación enunciamos algunos argumentos a favor de la idea de que *sistema*, como tal, es muy difícil aceptar que haya habido, cosa que no ha impedido proceder *como si* hubiera. Esto no quiere decir que aspiremos a que haya sistema, con todo lo que, como se verá, tal cosa implicaría.

¹⁹ El *ICFES* viene haciendo público en los diarios nacionales lo que llama el “nuevo examen del *ICFES*”, que se empezó a aplicar en marzo del 2000. En el marco de esta divulgación, el 22 de noviembre de 1999 hubo una reunión técnica con jefes de admisión de las universidades.

Diseño de instrumentos

Los instrumentos, tanto para establecer el desempeño de los estudiantes como para establecer factores asociables, han sido elaborados por grupos e, incluso, por personas, que han operado independientemente unos de otros:

En el desarrollo de este trabajo, que se inicia antes de la expedición de la Ley General de Educación, participaron paralelamente tres entidades: el *MEN*, el Instituto ser de Investigación y el *SNP*, con la condición de llevarlo a cabo en muy corto tiempo. Esto último originó que las instituciones trabajaran separadamente y que, incluso en su interior, hubiera poca comunicación [ICFES, 1996:5].

Pero si, como creemos, en los instrumentos de desempeño hay teoría —así sea implícita— sobre la educación y, por lo tanto sobre los factores que inciden en su calidad; y si, de la misma manera, en los instrumentos de factores hay teoría —así sea implícita— sobre desempeño, entonces a la hora de correlacionar las informaciones que vienen de instrumentos aislados, se pueden estar cruzando medidas realizadas sobre objetos que pertenecen a clases diferentes, es decir, incompatibles, y pueden verse asociaciones debidas más a la accidentalidad o, en el mejor de los casos, al azar²⁰. De donde las correlaciones encontradas, en cualquier magnitud y sentido, pueden ser tan azarosas como las decisiones tomadas en política educativa.

Veamos un caso en relación con el factor *textos*. En la prueba de lenguaje para la evaluación de impacto del Plan de Universalización se concebía la competencia frente a los textos en función de la naturaleza de las prácticas que con ellos se realizan, en la escuela o fuera de ella, lo cual implica la diversidad del material impreso; de ahí que la base para realizar la prueba fuera una narrativa icónica [cf. Anexo I en: Jurado, Bustamante y Pérez, 1998] que enfatiza sobre la diversidad y sobre los usos. Pero mientras así se lo concebía en la prueba de desempeño en lenguaje, la concepción que se materializa al respecto en los instrumentos de factores asociables, en la misma evaluación de impacto, gira en torno al concepto de “dotación”, frente al cual resulta relevante la cantidad y resultan irrelevantes la diversidad y el uso. ¿Qué tan confiables pueden ser las conclusiones en relación con el factor en mención, cuando se cruzan informaciones recogidas con ideas opuestas sobre el mismo objeto?

Pensamos que esto sería contraproducente si la política se diseñara efectivamente con base en estos resultados. Pero incongruencias como estas resultarían secundarias si a la hora de definir la política educativa fueran determinantes otros asuntos...

Desempeño

En este proceso, el desempeño de los estudiantes ha tenido varias denominaciones: “logro cognitivo” o “académico” se decía en la época de Gaviria, “competencias” se dice actualmente; pese a que estas denominaciones puedan tener algún correlato conceptual y, en consecuencia, diferencias teóricas entre sí, han terminado utilizándose como sinónimos, pero con la connotación que esté de moda.

²⁰ No discutimos la validez estadística (o psicométrica) que puedan exhibir tales instrumentos.

Para conocer el desempeño, se diseñan pruebas específicas; no se utiliza el examen de Estado, de un lado, porque se pretende hacer aplicaciones a lo largo de toda la educación básica, grados 3^o, 5^o, 7^o y 9^o, cosa que dicho examen no permite, pues se aplica solamente en grado 11 (en una época hubo un examen del *ICFES* para grado 5^o en Bogotá). Pero, entonces, ¿por qué no ampliar el examen de Estado a estos cursos?; porque,

los sistemas de evaluación de la calidad de la educación no tienen una relación directa con los sistemas de exámenes nacionales, como los que aplica el Servicio Nacional de Pruebas del *ICFES* en Colombia. Estos últimos son pruebas individuales de selección y, como tales, sus propósitos son distintos; esto es, seleccionar a los individuos y certificar, con base en un puntaje, si adquirieron unos conocimientos que les permitan ingresar, por ejemplo, a determinadas carreras universitarias. Por su parte, los sistemas de evaluación se realizan con base en pruebas diagnósticas que operan sobre una selección de indicadores que dan cuenta de los resultados que la sociedad le exige al sistema educativo [Peña, 1999: 26].

No obstante, no hay una diferencia nítida entre, de un lado, “certificar si se adquirieron unos conocimientos”, es decir, aquello que —según la cita de Peña— hace el examen de Estado, y, de otro lado, “dar cuenta de los resultados que la sociedad le exige al sistema educativo”, es decir, lo que —según Peña— harían las pruebas de calidad. Cabrían preguntas como: los conocimientos “adquiridos” por los bachilleres, ¿no son exigidos por la sociedad al sistema educativo?; los resultados que la sociedad le exige al sistema educativo, ¿no son parte de los conocimientos adquiridos por los bachilleres? Si las respuestas fueran afirmativas, la anterior justificación no sería admisible.

Según se dijo con el nacimiento del *SINECE*, el examen de Estado establece la posición de quien presenta una prueba frente al resto de personas que también la presenta (es el análisis *con arreglo a la Norma* que regularmente se ha hecho en el Servicio Nacional de Pruebas del *ICFES*); mientras que evaluar la calidad educativa requiere establecer la posición de quien presenta la prueba en relación con un saber estratificado (es el análisis denominado *con arreglo a criterio*, sobre el cual el *ICFES* reconocía, a principios de la década, no tener experiencia alguna).

La idea de que el examen de Estado no servía para medir la calidad de la educación y, en consecuencia, que no servía como base para tomar las decisiones en política educativa, parece no haber trascendido más allá de los equipos técnicos que desarrollaron las evaluaciones; es decir, no ha habido una justificación pública de por qué no se lo tomó para este propósito. Esto ha facilitado que, en cualquier momento, se lo retome para ponerlo en función de la calidad, tal como se ve en los siguientes tres casos:

En primer lugar, mientras el recién creado *SINECE*, en consecuencia con la diferencia entre interpretación con arreglo a la Norma e interpretación con arreglo a Criterio, invertía tiempo y dinero en la producción de pruebas que ubicaran a los estudiantes en relación con unos niveles de logro cognitivo —establecidos mediante teorías explícitas sobre el desarrollo del conocimiento en las áreas básicas—, mientras esto sucedía en una oficina del Ministerio de Educación Nacional, en otra se diseñaba el Programa de Mejoramiento Institucional²¹, donde el examen de Estado era el instrumento con el que se establecían los desempeños

²¹ Formato para aspirar a recursos del “Programa de Ampliación de la Cobertura en Educación Secundaria” - *PACES*.

con los que se mediría la incidencia del programa en la calidad de la educación secundaria donde se invirtieran sus recursos.

El segundo caso es el de la Ley General de Educación, que en su artículo 99 dice:

A los 50 estudiantes del último grado de educación media que anualmente obtengan los más altos puntajes en los exámenes de Estado [...] se les garantizará el ingreso a programas de educación superior en instituciones del Estado. De igual beneficio gozarán los estudiantes de último grado de educación media que ocupen los dos primeros lugares en cada uno de los departamentos, según las mismas pruebas.

Independientemente de que estos estímulos sean demandables jurídicamente por no ser equitativos, su instauración indica que el legislador cree en el examen de Estado como un instrumento que efectivamente mide el desempeño de los estudiantes y, entonces, habla de la calidad de la educación.

El tercer caso es el del gobierno de Pastrana: mientras bajo su gestión se están cambiando las pruebas de Estado, de otro lado, su Plan de desarrollo para el sector educativo, el *Plan educativo para la paz* dice: “Los resultados en materia educativa para la educación media son igualmente preocupantes. Al analizar los resultados obtenidos por los bachilleres colombianos en los exámenes de Estado del *ICFES* se puede apreciar un deterioro progresivo de la calidad de la educación” [UPN, 1999:9]; las cifras del examen de Estado comparadas son las que van de 1986 a 1995 [:10].

Cabría entonces preguntarse: si el examen de Estado sirve para medir calidad, ¿por qué se lo está transformando? Resulta paradójico que se lo transforme justamente a nombre de sus limitaciones para hablar de los desempeños, y bajo el modelo de la evaluación de la calidad, es decir, hacia el concepto de *competencias*, que fue el que se impuso como objeto de evaluación en el SINECE desde el comienzo. Que la propiedad, que a veces se le reconoce, de medir la calidad sea independiente del hecho de que esta información sirva para dar acceso a la universidad o para “diagnosticar” la educación, muestra que al modelo en boga lo único que le interesa es una clasificación del desempeño, independientemente de la manera que se use para obtenerla.

En atención a que las pruebas de desempeño en la evaluación de la calidad han sido un objeto de investigación, no han mantenido una misma conceptualización a lo largo del proceso, de manera que sus resultados son hasta cierto punto incomparables. Todo esto conduce a que a la fecha no se disponga realmente de una “línea base” de desempeños de los estudiantes en el país, pese a lo cual no sólo se utilizan los resultados cada vez para hablar supuestamente de la misma educación, de la misma población, sino que a veces se comparan entre sí...²² Otra cosa es pensarlo desde el ángulo de la investigación del desempeño: no ha habido permanencia de las pruebas, pues ello sería equivalente a decir que ya se sabe todo al respecto y que nada podría venir a cambiarlo. Esto plantea una tensión: o permanencia de las pruebas y, en consecuencia, comparabilidad de las cifras (es decir, *sistema*), o transformación permanente de los mecanismos de indagación de lo que los

²² Es como medir la estatura de una persona cada vez con unidades distintas y, sin tablas de convergencia entre sistemas de medición, comparar los resultados: “rebajó la estatura, porque hace un minuto medía 180 y ahora mide 70.8”. Claro: en un caso se usan centímetros y en el otro pulgadas.

niños saben y, en consecuencia, cifras hasta cierto punto incomparables (es decir, investigación). En otras palabras, o se tiene certeza y se habla de la educación y de los niños, y se toman medidas; o se tiene incertidumbre y se habla del aprendizaje que hemos ido teniendo, de la distancia insalvable, tanto con los procesos cognitivos de los niños, como con los fenómenos educativos, y se toman medidas *modestas*, que explicitan sus condiciones de posibilidad, que hacen historia y que se reconocen como limitadas.

Veamos el caso de dos pruebas realizadas en el área de lenguaje para el grado tercero: la primera utilizaba, por ejemplo, ejercicios de “completación” (frases en las que el niño debe introducir la palabra faltante, cuyo lugar está indicado por una línea), ejercicios de “pluralización” (pasar a plural una oración [sic] que está en singular). En este caso, el objeto de evaluación era “el manejo de los aspectos formales de la lengua y la adquisición de las habilidades necesarias (hablar, escuchar, leer y escribir) para comunicarse” [MEN, 1993a: 10]. Las pruebas que se hicieron después para el mismo grado no tienen este tipo de tareas; en lugar de eso, preguntan por la representación conceptual de la información brindada, por lo que de ella se puede inferir y por las posiciones que es posible asumir frente a ella. En este caso, el objeto de evaluación fue la *competencia comunicativa*, algo que puede no equivaler a “las cuatro habilidades” (y menos al “manejo de los aspectos formales de la lengua”), pues en la enseñanza de la lengua “las cuatro habilidades”, a diferencia de la competencia comunicativa, pretenden volcarse al uso pero no lo hacen de manera *contextual* [cf Widdowson, 1983].

Vale la pena anotar que al mismo tiempo que se aplicaba la primera prueba para tercero con las características anotadas, se aplicaba una prueba para quinto, elaborada por otro equipo, que tenía como objeto la competencia comunicativa.

Distintas pruebas promueven a la existencia distintos objetos, perfiles distintos de estudiante, con lo que no podrían justificarse las comparaciones ni las políticas educativas que dicen basarse en una información que es obtenida cada vez de formas tan diferentes. Al menos, habría que decir —pues muchas veces la política educativa no puede esperar a la terminación de los debates— cuáles son las condiciones bajo las que se obtuvo la información y, en consecuencia, las limitaciones que tienen las conclusiones. No obstante suele usarse esta especificidad de la decisión en relación con la discusión, justamente para no dar lugar a que se discuta.

Factores asociables

Los factores asociables entran a jugar en tanto “Un sistema evaluativo que realmente pretenda tener impacto en la educación, no debe prescindir de la medición de tales factores ni de su relación con los resultados de logro académico, para poder así llevar acabo una búsqueda de soluciones pertinentes y diferenciadas” [MEN, 1992:25]. Obviamente, pues el condicionamiento de parte de las instancias que prestan el dinero para educación era claro: hay que justificar en qué aspectos se invertirá el dinero, y no hay otra manera de hacerlo más que la “investigación” estadística. Veamos a continuación cómo los Planes de desarrollo de los distintos gobiernos asumen el asunto.

El Plan de desarrollo en educación del gobierno del presidente Gaviria, *Plan de apertura educativa*, decía que los resultados de la evaluación de la calidad “soportarán los programas de mejoramiento de la calidad de la educación básica” [DNP, 1991:36]. Por su parte, el Plan

de desarrollo en educación del gobierno del presidente Samper, *El salto educativo*, que usa los resultados de la evaluación de la calidad que comenzaron en el anterior gobierno, hablaba de conocer el rendimiento en las áreas básicas del currículo, y agregaba: “Especial énfasis se otorgará al análisis de los factores que explican el rendimiento de los estudiantes [...]” [MEN, 1994:25]; y, en consecuencia, plantea que las bases de la asignación de recursos serán este juicio sobre la calidad, y el establecimiento de la eficiencia [:27 y ss]. Y el *Plan educativo para la paz* del gobierno de Pastrana dice: “Los estudios de factores asociados constituyen un elemento fundamental en el diseño y evaluación de la política educativa” [UPN, 1999:11].

Entre los factores asociables sobre los que se ha indagado en esta década, se han venido incluyendo aquellos indicadores tradicionales que se utilizaban en décadas anteriores, pero también una serie de aspectos que varía notablemente de una aplicación a otra. A lo largo del proceso, los instrumentos de factores asociables muestran características como las siguientes:

- número extremadamente variable de factores, algunas veces bajo la creencia de que a mayor tamaño más exhaustividad. No obstante, “las determinaciones numéricas en ningún caso deben sobrepasar en exactitud a los medios de investigación” [Bachelard, 1979:254]. Esta variabilidad numérica se ve, por ejemplo, en el cuestionario de factores asociables para estudiantes: el primero que aplicó el SINECE tenía 41 preguntas; el segundo tenía 72; el que aplicó la evaluación de Impacto del Plan de Universalización tenía 11... De esto no se ha dado justificación.
- ajuste de los instrumentos a los recursos económicos disponibles para el estudio, y a la política educativa: “De estas variables descartadas se exceptuaron aquellas que presentaban especial importancia en relación con el diseño de políticas educativas” [MEN, 1992:55];
- falta de jerarquía, de relaciones, entre los factores indagados (no nos referimos a la jerarquía que tiene lugar para implementar cierta política educativa); esto muestra la ausencia de una teoría que organice y, en consecuencia, lo que se presentan son listados de factores que presuponen que todas las variables están en interacción y que el fenómeno resulta sensibilizado en todas sus variaciones; sin embargo, Bachelard plantea [1979:257] que la sensibilidad no es recíproca, y que eso nos obliga a hacer de cada investigación un caso especial. Cualquiera puede listar los factores que “cree” que afectan el desempeño de los niños. Lo difícil, en cambio, es establecer relaciones...
- excluyentes entre sí, a la hora de indagar el mismo objeto (ponemos a continuación un ejemplo a propósito del material impreso), lo que confirma la carencia de una teoría que organice la indagación;
- copia de otras indagaciones, poco o nada explicitadas, cuyas condiciones de realización pueden diferir radicalmente: por ejemplo, ante la afirmación de que es necesario cambiar la composición femenina del magisterio a favor de una composición masculina —dado que en el magisterio en básica hay más mujeres y que la calidad de la educación es baja—,

unos investigadores en Antioquia²³ aclararon que el instrumento de factores que habían aplicado estaba validado por la *UNESCO*.

De ahí que las características de los diversos listados difiera de manera notable. Nuevamente, esto trae como resultado que a la fecha no se disponga de una historia del proceso: ni una “línea base” de factores asociables de la educación en el país, ni una teoría para su investigación; sin embargo, los estudios existentes se usan para tomar decisiones... De nuevo, no se trataría de que hubiera una lista de factores constante o listas equivalentes; esto no ha sido posible, pero no por las razones por las que la indagación del desempeño ha requerido la investigación. De nuevo, aparece la tensión: sistema o investigación, aclarando que con ayuda de la investigación se podría soportar una modesta política educativa.

Veamos un ejemplo, a propósito del material impreso disponible en la casa, como factor asociable al desempeño de los niños. En la evaluación de impacto del Plan de Universalización [*MEN*, 1997B] esto se indaga de la siguiente manera²⁴:

5. ¿Cuántos libros hay en tu casa?
(No incluyas las revistas ni los periódicos)
- A. Solamente los textos escolares
 - B. Los textos escolares y menos de 10 libros
 - C. Los textos escolares y más de 10 y menos de 50
 - D. Los textos escolares y más de 50 libros

Y en la evaluación censal del Distrito Capital, correspondiente a 1998²⁵, en relación con el mismo aspecto se hace la siguiente pregunta:

20.	Para ayudarte a hacer las tareas en tu casa tienes:	
Enciclopedia	Sí	No
Libros de literatura o cuentos	Sí	No
Diccionario	Sí	No
Otros libros	Sí	No

Como puede verse, sin ir más allá de la literalidad de las preguntas, hay dos concepciones completamente distintas sobre el material impreso: i) En la evaluación de impacto, se hace

²³ En este momento no contamos con la referencia.

²⁴ El instrumento se denomina: “Evaluación de impacto, Plan de Universalización. Prueba de matemáticas y ciencias naturales”. La pregunta está en la página 3.

²⁵ El instrumento se denomina: “Evaluación de competencias y saberes básicos en lenguaje y matemática”. La pregunta está en la página 17.

una diferencia, que la evaluación censal no hace, entre “texto escolar” y “libros”; ii) la primera hace una jerarquización, que la segunda no hace, entre materiales impresos, pues solicita excluir revistas y periódicos; iii) la primera indaga por la cantidad de textos, mientras la segunda indaga por una diversidad de textos; y iv) la primera interroga por la tenencia de los libros, mientras la segunda indaga por una de las posibilidades de su uso.

Correlaciones

Después de tener el análisis del “desempeño”, hecho por los respectivos expertos, y el análisis de los “factores”, hecho a su vez por los respectivos expertos, se procede a un análisis *global* que finalmente “cruza” estas dos variables. Este análisis, que a veces se menciona a la hora de establecer la política educativa, puede ser desarrollado por una instancia que incluso sea ajena a aquellas que desarrollaron cada instrumento. Esto puede introducir otra concepción que no se explicita y que puede ser determinante para la interpretación. Es el caso del estudio sobre repitencia en Colombia: “la parte cuantitativa del estudio se centró en el análisis estadístico de las respuestas a las preguntas cerradas de las entrevistas y en los resultados de las pruebas de logro académico. Dicho análisis fue adelantado por el equipo de Harvard Institute for International Development” [MEN, 1993b:71. También es el caso de la evaluación de impacto del Plan de Universalización: hubo un equipo distinto para cada una de las siguientes pruebas: matemáticas-ciencias, valores-actitudes infantiles, lenguaje, gestión, costo-efectividad (incluso para el diseño de la muestra); y otro equipo, de estadísticos, hizo el análisis global. En otras palabras, es posible encontrar trabajos “multidisciplinarios”, pero no interdisciplinarios, es decir, que obedezcan a alguna articulación. Las decisiones que se toman al respecto parecen revelar, en el fondo, la idea de que los instrumentos están hechos a prueba de sujetos, a prueba de interpretación (... a pesar de los sujetos de la evaluación, a pesar de la teoría).

Con los factores y el desempeño se opera estadísticamente una búsqueda de correlaciones. Para ello se requiere que la evaluación del desempeño haga una clasificación de los estudiantes; puede ser cualquier clasificación: por ejemplo, mientras la evaluación de la calidad, en el caso del área de lenguaje, ha hecho clasificaciones con arreglo a categorías provenientes de las teorías sobre el desarrollo lingüístico o sobre las etapas de la lectura y la escritura; la evaluación de impacto del Plan de universalización, en cambio, hizo una estandarización basada en el número de respuestas acertadas, independientemente de los niveles en los que cada pregunta estuviera teóricamente asignada.

Hecha la clasificación, se procede a ver cómo están distribuidos los diversos desempeños de los estudiantes, en relación con un factor asociable cualquiera. Por ejemplo, los “medios pedagógicos”: se busca saber si los desempeños de los estudiantes cambian a medida que varía la dotación de dichos medios; un caso sería aquel en el que en las escuelas donde hubo más dotación, hay una tendencia a que los estudiantes tengan mejores desempeños que en aquellas escuelas en las que hubo menos dotación. En tal caso, se diría que hay una correlación positiva y, en consecuencia, que la dotación de medios pedagógicos efectivamente coadyuva a mejorar los desempeños y, en consecuencia, la calidad de la educación. De donde se justificaría una política educativa que se proponga entregar medios pedagógicos a las escuelas.

Toda iniciativa y política educativa se inscribe automáticamente en esta dinámica de cotejar desempeño y factores: son los casos de la evaluación de impacto del Plan de

Universalización de la Educación Básica, el de la evaluación censal en el Distrito Capital, el estudio sobre repitencia en Colombia, el Tercer Estudio Internacional de Matemáticas y Ciencias (TIMSS). Claro que el desconocimiento de lo hecho incluso puede ser explícito: después de casi diez años de estar gastando empréstitos internacionales²⁶ en este tipo de estudios, el *Plan educativo para la paz* de Pastrana dice: “La investigación sobre factores asociados a la calidad es incipiente en el país”; y agrega que eso “limita el análisis costo-beneficio de los resultados obtenidos, y hace difícil el diseño de una canasta educativa de calidad que permita maximizar la utilización de los recursos asignados a mejorar la calidad educativa” [UPN, 1999:111. Veamos cómo se usa este mecanismo en cada plan de desarrollo educativo y, en consecuencia, cómo se desdice el plan de Pastrana:

- el *Salto educativo*, de un lado, no sólo retoma la idea de que la calidad de la educación es deficiente (pues, según dice, los estudiantes fallan en las habilidades que requiere la continuación del aprendizaje) sino que agrega que ésta se ha deteriorado. Y, de otro lado, utiliza las correlaciones entre indicadores para justificar la política educativa: “Las características del plantel y las de los procesos pedagógicos son las variables que más pesan²⁷ en la explicación del rendimiento de los estudiantes. Esto indica que el sistema educativo tiene en sus manos las principales herramientas para mejorar la calidad de la educación. En efecto, docentes calificados, planteles bien dotados, estudiantes con textos, y una mayor jornada, están asociados positiva y significativamente con el logro de los estudiantes” [MEN, 1994: 11];
- el *Plan educativo para la paz* plantea que los resultados del programa SABER indican que la calidad de la educación es deficiente: dice que no se alcanzan los objetivos curriculares nacionales ni los estándares internacionales [UPN, 1999:8-9]. A partir del cruce con factores asociables, plantea que las variables escolares explican hasta el 30% del rendimiento escolar [:101 y que factores como textos, bibliotecas, medios pedagógicos, promoción de la lectura y tiempo efectivo de aprendizaje son factores decisivos para la calidad de la educación [:11].

Compárense los factores que en cada caso se califican como más relevantes y que, sin embargo, se obtienen de los mismos estudios o de estudios que, por principio, tendrían que ser equivalentes: se nota que sin problema alguno se hacen desaparecer y aparecer factores en el escenario de la política educativa, de una manera misteriosa (pero, ¿los conejos estaban en la chistera del mago?).

Algo en común que, a propósito de la política educativa, tienen estas referencias es el hecho de señalar crudamente que la educación ya no es un factor de ascenso social, que no puede solucionar los asuntos de calidad educativa debidos a las diferencias sociales (los estudios no dejan de encontrar todas las veces que estas diferencias tienden a coincidir con las diferencias de desempeño), pero que sí puede incidir sobre un porcentaje del desempeño, el debido a factores propiamente educativos. Claro que mientras para el *Salto educativo* esto indica que el sistema tiene “las principales herramientas para mejorar la calidad de la educación”, el *Plan educativo para la paz* las reduce a un 30%. Para sustentar

²⁶ El modelo para regular la política educativa calculaba que en las etapas iniciales las evaluaciones se harían con préstamos internacionales, pero que a medida que avanzara el sistema de evaluación, los recursos tendrían que ser del presupuesto nacional.

²⁷ Lo del “peso” se refiere a la significatividad estadística.

esto, que es *vox populi* en la jerga neoliberal en educación a nivel mundial, se habla de manera que ¡nuevos factores, nuevas cifras, nuevas correlaciones y nuevos ámbitos aparecen!:

La Misión Social del Departamento Nacional de Planeación ha realizado análisis con los datos de las diversas evaluaciones adelantadas, y una de sus principales conclusiones señala que los factores asociados al plantel educativo pueden afectar hasta en un 29% el logro.

Los resultados de todos estos análisis plantean que la política tiene un gran margen de acción y que la escuela, por consiguiente, merece mayor atención. Sugieren que es importante analizar aspectos como las prácticas pedagógicas, el desarrollo institucional, la cualificación de los métodos de gestión, la oportunidad de la información y la asignación eficiente de los recursos [Ávila: 1999:37].

Tenemos, entonces, basados en los mismos estudios, los siguientes factores que en cada caso se “escogen” para decir de qué depende el mejoramiento de la calidad y, como hemos venido diciendo, hacia dónde debe dirigirse la política educativa:

	SALTO EDUCATIVO	PLAN DUCATIVO PARA LA PAZ	DNP
Factores “Asociados”	Docentes calificados		
	Dotación de planteles	Medios pedagógicos	
	Textos	Textos	
		Bibliotecas	
	Mayor jornada	Tiempo de aprendizaje	
		Promoción de la lectura	
			Prácticas pedagógicas
			Desarrollo institucional
			Métodos de gestión
			Oportunidad de la información
		Asignación eficiente de los recursos	

Pese a que se supone que están basados en las mismas cifras (o equivalentes), se observa: en ningún caso concuerdan las tres maneras de concluir; sólo hay coincidencias parciales en dos casos (textos y tiempo de estudio); se da aproximación en un caso (dotación de planteles/medios pedagógicos); y hay muchas casillas vacías que hablan de inconsistencia. Las conclusiones del *DNP*, basadas según Ávila en “los datos de las diversas evaluaciones adelantadas” (es decir, incluyendo las que subyacerían a los planes de Samper y de Pastrana), no coinciden en ningún punto con las otras dos. De otro lado, es notable que su manera de concluir tenga la marca del modelo que impulsa para toda la política nacional: la administración. Es decir, para el *DNP*, el asunto educativo es más de índole administrativa; repasemos su lenguaje: “desarrollo institucional”, “gestión”, “oportunidad”, “eficiencia”, “recursos”.

Así mismo, en los diversos estudios se hacen afirmaciones que, si no son contrarias, por lo menos son parcialmente excluyentes. Veamos los siguientes casos:

- i) Sobre número de alumnos por docente, por ejemplo, en la evaluación de impacto del Plan de Universalización hemos encontrado —mediante la interpretación con arreglo a criterio

de los desempeños en el área de lenguaje— que allí donde hay menos estudiantes, los desempeños son superiores; es decir, una correlación *negativa*, contrario a lo que se plantea como fundamento de la política de aumento de alumnos por docente: a saber, que no es un factor que influya sobre el desempeño, sobre todo porque se plantea que “Una disminución de 2 alumnos por curso en promedio en el país, generaría la necesidad de nombrar cerca de 25.000 docentes para atender la misma población” [Rosas et al., 1996:10].

- ii) Sobre capacitación, el *SINECE* plantea que los cursos de capacitación docente recibidos en servicio no aparecen con asociaciones significativas [*MEN*, 1992:89]; sin embargo, el *Salto educativo* [*MEN*, 1994:11] afirma que la calificación de los docentes —entre cuatro factores— está asociada positiva y significativamente con el logro de los estudiantes.
- iii) Sobre textos escolares, el primer informe del *SINECE* recomienda su dotación, en tanto factor *decisorio* de la calidad [*MEN*, 1992:100]; mientras que uno de sus últimos informes plantea que, en el sector oficial urbano, la dotación *de la biblioteca* (tener enciclopedias, diccionarios y cuentos) se asocia positivamente con un mayor logro promedio [*MEN*, 1997a:45]. O sea, a diferencia del primer informe, el material impreso que estaría asociado al desempeño no serían los textos escolares, ni este hallazgo sería generalizable para toda la educación.

Dadas estas diferencias, cabe preguntar: sobre los estudios en los que dice apoyarse la política educativa en cada caso; sobre la responsabilidad de los instrumentos en los hallazgos; y sobre el papel que realiza la interpretación de las cifras. No se explicitan las diferencias entre los diversos estudios (si los hay), siendo que llegan a afirmar cosas que podrían ser contrarias. En cada caso, la política educativa parecería basarse en un estudio que, dado que sirve para tomar decisiones fundamentales, ha de considerarse objetivo, independientemente de que otros estudios digan otra cosa, e incluso, de que el mismo gobierno tome más adelante decisiones en otra dirección. Aquí no hay más remedio que recordar el principio musulmán gracias al cual queda demostrado que, excepto *El Corán*, se pueden quemar todos los libros: si dicen lo mismo que éste, sobran, pues están repitiendo; si dicen cosa distinta, deben ser falsos, pues sólo aquél dice la verdad. Pero entonces queda claro que es el interés el que decide qué es y qué no es verdad, y no al contrario.

El uso de la estadística

Pese a que la estadística en su misma lógica se ofrece como una ayuda para procesar una información cuya naturaleza y sentido son establecidos por otros campos conceptuales (sin desdeñar la implicación recíproca), la manera como se la ha utilizado en todo este proceso de evaluación de la calidad educativa y de establecimiento de la política educativa ha implicado dos tipos de problema:

De un lado, que los criterios técnicos condicionen en gran parte los criterios teóricos, pedagógicos y educativos. Ejemplos: i) es la dificultad de procesar la información venida de pruebas abiertas (confiabilidad) la que restringe su uso, y no la naturaleza del estudio²⁸ u) se ha pretendido tener pruebas de desempeño estandarizadas, antes de tener suficiente investigación sobre el conocimiento y la evaluación en las diferentes áreas objeto de análisis;

²⁸ Y no obstante, en la investigación hemos encontrado que la prueba de desempeño que más se presta para encontrar correlación entre desempeño y material impreso es aquella de naturaleza abierta

iii) y se ha omitido la discusión sobre qué factores indagar, bajo la suposición de que, en este caso, la consistencia estadística hablaría de un amarre “ideacional” que no habría por qué buscar antes; de ahí que a las preguntas por el fundamento de los instrumentos se les pueda responder que éstos ya han sido “probados”.

Y, de otro lado, la manera como se ha utilizado la estadística durante este proceso ha implicado que incluso su coherencia resulte violentada,

—en primer lugar, por el uso indebido de algunos procedimientos en los cuales los datos no se corresponden con los supuestos teóricos que su teoría demanda (por ejemplo, hacer regresiones incluyendo factores que son asociables a unidades de observación jerarquizadas);

— en segundo lugar, por la selección de los resultados que, a juicio del funcionario de turno, parezcan acordes con las expectativas previa, con la continuidad de la política educativa; por ejemplo, después de que herramientas estadísticas se hace una selección de variables para un modelo de regresión, se descartaron “las que presentaran signos contrarios a la predicción teórica” [MEN, 1992:55]; pero si se descartan los que no, ¿por que no se descartan los que sí, si son resultado de los mismos instrumentos y las mismas metodologías? Una actitud como esta descuida el hecho de que la interpretación estadística puede ser un interlocutor que advierte sobre la validez de la teoría, sobre la confiabilidad de los instrumentos y sobre la pertinencia del estudio. Pero ahí no para la arbitrariedad: “De estas variables descartadas se exceptuaron aquellas que presentaban especial importancia en relación con el diseño de políticas educativas” [:55], es decir, que la política educativa en realidad no depende de la investigación de las cifras y, más bien, se la pone como justificación cuando conviene; parecen usar la estadística como los borrachos usan los faroles: para apoyarse, no para iluminarse [Silva, 1997:2];

— y, finalmente, por la asignación de causalidades en la realidad educativa a partir de correlaciones de cifras, sin establecer las mediaciones entre ambas; sabido es que dos eventos pueden mostrar asociación estadística no solamente porque sean uno causa del otro sino, por ejemplo, porque media un tercero que no está considerado. Es el caso del número de nacimientos y el número de cigüeñas: un estudio estadístico encontrará correlación entre ambos, con lo que podríamos concluir causalmente que a los niños los trae la cigüeña; no obstante, como las cigüeñas anidan en las chimeneas de las casas, ambos factores —niños y cigüeñas— dependen de un tercero:

el tamaño de la población, de donde, a más población, más casas y, por ende, más chimeneas y más cigüeñas.

Una propuesta

Los “hallazgos” —o lo que ocupa su lugar— han sido tomados por quien hace política educativa como totalizantes y concluyentes, incluso cuando se abandonan para tomar nuevos resultados o nuevas posiciones como totalizantes y concluyentes. Es decir, no se ha procedido a interrogar la naturaleza de la información contenida en los datos. Tampoco ha habido consecuencia lógica en el sentido de continuar analizando los hallazgos mediante investigaciones que pongan en juego elementos y relaciones más puntuales y más complejos. En menos palabras: se ha procedido de manera arrogante. No obstante, si estamos obligados a adoptar, ajustar e incluso producir modelos teóricos le interpretación, es porque aceptamos el hecho de estar limitados para describir la realidad educativa tal cual es.

De tal manera, ninguna propuesta puede postularse como la que va a mostrar efectivamente cómo es la realidad educativa; a lo sumo puede postularse como una iniciativa cuya dinámica está en la puesta en juego de sus presupuestos y en su disposición a la interlocución con miradas pares que pretenden propósitos similares.

En un contexto así constituido, ¿qué significaría una propuesta de mejoramiento de la calidad de la educación?, ¿pasaría por los procesos que se han tenido en cuenta en las evaluaciones anteriores?, ¿habría que continuar evaluando todo el país con los mismos instrumentos y métodos, haciendo caso omiso de las especificidades (que podrían hacer heteróclita nuestra educación) o, por lo menos, no sustentando la equivalencia entre regiones, modalidades, etc.?

Hay quienes piensan que si se quiere velar por la calidad de la educación, habría que consolidar una “voluntad política”, incluso de una sola persona: para la *Misión de ciencia, educación y desarrollo* [1997:191], “Las experiencias en la historia de Colombia y en la de otros países del continente muestran que una transformación educativa de fondo sólo puede ser exitosa si es asumida decididamente por el primer mandatario”. Ante esto, valdría la pena preguntar por qué la voluntad política invocada no ha tenido lugar por su propia dinámica (es decir, que hay otra voluntad, no invocada pero efectiva); por qué no se dio la “transformación educativa de fondo”, si los “sabios” le dijeron hace ya 5 años a quien los había convocado que entre ambos grupos constituían la solución: los únicos que piensan y el único que decide (“tú me convocas, yo te alabo”). Es, con más mediaciones, la separación que hemos estado comprobando para el diseño de la política educativa: el sabio —en nuestro caso el estadístico— susurra al oído del Príncipe —en nuestro caso quien toma las decisiones políticas— y éste dispone de manera impredecible: hace caso o, en un arrebato, hace lo contrario o, incluso, le manda a cortar la cabeza al sabio.

Pero, en cualquier caso, el presidente y los “sabios” no son los que llevan a cabo el día a día de la escuela de millones de niños y de miles de maestros; que puedan y sepan, respectivamente, termina por no importar cuando la educación es lo que sus agentes del día a día son capaces de hacer, capacidad que ha sido producida en condiciones específicas. Sólo la producción de otras condiciones haría otra cosa posible. Pero esas condiciones no vienen del otro, llanamente; sólo se consolidarían en la interacción entre una posición ética de los agentes educativos y una posición social, política que construya-gane esas condiciones.

Que la política se deje orientar por los hallazgos investigativos que provengan de la construcción de una teoría compleja, exhaustiva en la medida de lo posible, dinámica (cambiante en el tiempo) y perfectible; que se pueda aprender de la historia del diseño de la política educativa en el país; que haya sustentos explícitos y debatibles en el campo académico, que permitan proponer unos factores asociables al desempeño, pero estructurados entre sí; que sea posible diseñar instrumentos al servicio de la investigación educativa, que se transformen a la luz de ésta, pero que también posibiliten su desarrollo; que se haga una evaluación del desempeño cuyo fundamento busque relación con el de los factores asociables, y que sea igualmente discutible y perfectible; que se expliciten las metodologías y que se proceda en consecuencia con sus requerimientos y alcances... Todo esto sólo puede ser un efecto, no es una condición del trabajo.

Todo el mundo está de acuerdo con mejorar la calidad de la educación; sin embargo, en medio de tanta diversidad de intereses, nos parece sospechoso el hecho de que haya

semejante acuerdo. Nuestra relación con la evaluación de la calidad de la educación nos ha mostrado que donde no hay homogeneidad es fácil encontrar diferencias (claro, ¡porque es una tautología!). En lugar de eso, hemos hecho investigaciones en las que no sólo aparecen datos que contradicen muchas de las certezas que circulan como verdades, sino que aparece evidencia empírica para señalar que es el punto de vista el que crea el objeto.

Cuando se pasa a discutir qué quiere decir “calidad de la educación”, resulta que es algo distinto, dependiendo de los propósitos de los que es portador quien responde. La propuesta es ir construyendo a cielo abierto una concepción de calidad ajustada a lo que vayamos encontrando... y hacer con esto la política educativa, lo que puede no requerir de las evaluaciones masivas.

BIBLIOGRAFÍA

ÁVILA, Rosa. “Factores asociados al logro educativo”. En: Revista *Alegría de enseñar* N° 38. Cali: FES, enero-abril de 1999.

BACHELARD, Gaston. *La formación del espíritu científico*. México: Siglo XXI, 1979.

BUSTAMANTE, Guillermo y JURADO, Fabio. “La evaluación oficial del área de lenguaje”. En: Revista *Opciones pedagógicas* N° 10. Bogotá: Universidad Distrital, 1993.

DEPARTAMENTO NACIONAL DE PLANEACIÓN. “Análisis de la educación primaria y secundaria”. En: *Revista de planeación y desarrollo*, Vol. XIII, N° 3 de diciembre de 1981.

___ *Plan de apertura educativa*. Bogotá: DNP, 1991.

DÍAZ, Carlos. “Diseño, metodología y resultados del Tercer Estudio Internacional de matemáticas y Ciencias -TIMSS- en Colombia”. En: Revista *Edu.Co*. Vol. 1, N°1, enero-marzo de 1998. Bogotá, Ministerio de Educación Nacional.

DUARTE, Patricia, et al. “Posición sobre los componentes del sistema de evaluación, a propósito del área de lenguaje». En: BUSTAMANTE, Guillermo (comp). *Evaluación y Lenguaje*. Bogotá: SOCOLPE, 1998.

ICFES. “Cruce entre la variable logro y los factores asociables correspondientes a plantel-rector, docentes y estudiantes”. Bogotá: ICFES s/p, 1996.

JURADO, Fabio; BUSTAMANTE, Guillermo y PÉREZ, Mauricio. *Juguemos a interpretar*. Bogotá: Plaza & Janés, 1998.

MINISTERIO DE EDUCACIÓN NACIONAL. “Ciclo de Foros sobre el Plan de Apertura Educativa.

Tema: La evaluación de la calidad de la educación” (Documento de trabajo). Bogotá,

___ *Saber Sistema nacional de evaluación de la calidad de la educación. Primeros resultados: matemáticas y lenguaje en la básica primaria*. Bogotá: MEN, 1992.

____ *Avances de Saber. N° 2* (Evaluación de primaria: los primeros resultados). Bogotá: MEN, 1993a.

____ *Avances de Saber. N° 3* (Informe sobre repitencia). Bogotá: MEN, 1993b.

____ *El salto educativo. La educación, eje del desarrollo del país*. Bogotá: MEN, 1994. ____ *La escuela: el lugar predilecto de las niñas y los niños. Por una mejor educación para la Colombia que soñamos. Programa nacional de incentivos*. Bogotá, MEN, 1996.

____ *Factores asociados al logro cognitivo de los estudiantes. Grados 3°. 5°. 1993-1995*. Bogotá: MEN, 1997a.

____ *Evaluación de impacto Plan de Universalización Educación básica primaria. Crédito 3010-CO BIRF. Resultados*. Bogotá: MEN, 1997b.

MISIÓN DE CIENCIA, EDUCACIÓN Y DESARROLLO. *Colombia al filo de la oportunidad*. Bogotá, IDEP: 1997.

PEÑA, Margarita. "Los sistemas de evaluación de la calidad: un vistazo al panorama internacional". En: Revista *Alegría de enseñar* N°38. Cali, FES, enero-abril de 1999.

Rosas, Gabriel, et al. «Racionalización del gasto público en educación». Bogotá: Departamento Nacional de Planeación, 1996.

SILVA, Luis Carlos. *Cultura estadística e investigación científica en el campo de la salud: una mirada crítica*. Madrid: Díaz de Santos, 1997.

UNIVERSIDAD PEDAGÓGICA NACIONAL. "Un plan educativo para la paz". Documentos Pedagógicos N° 4. Vicerrectoría Académica. Mayo de 1999.

WIDDOWSON, H.G. "Manejo y uso". En: *Lenguaje y Sociedad*. Cali: Univalle, 1983.