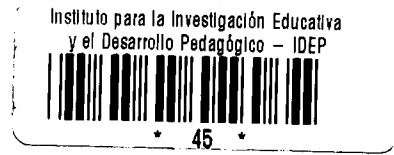


370.7  
B37a

18



# Análisis de algunos factores asociables al desempeño en el área de lenguaje: el Distrito Capital frente al resultado nacional

30/1/02

*Visión  
Análisis maestros.*

602000

INVESTIGACIÓN FINANCIADA POR EL «INSTITUTO PARA LA  
INVESTIGACIÓN EDUCATIVA Y EL DESARROLLO PEDAGÓGICO»  
A LA «SOCIEDAD COLOMBIANA DE PEDAGOGÍA».  
REALIZADA ENTRE SEPTIEMBRE DE 1999 Y JULIO DE 2000

**Investigadores:**

GUILLERMO BUSTAMANTE ZAMUDIO  
*Universidad Pedagógica Nacional*  
[bustama@colnodo.apc.org](mailto:bustama@colnodo.apc.org)

LUIS GUILLERMO DÍAZ MONROY  
*Universidad Nacional de Colombia*  
[lgdiaz@matematicas.unal.edu.co](mailto:lgdiaz@matematicas.unal.edu.co)

**Asistente:**

Omar Garzón  
*Sociedad Colombiana de Pedagogía*

Inv. IDEP  
33

387

## TABLA DE CONTENIDO

0. PRESENTACIÓN.....	5
<b>CAPÍTULO I: INTRODUCCIÓN .....</b>	<b>6</b>
1. ALGUNOS ANTECEDENTES.....	7
1.1 <i>La política educativa</i> .....	8
1.2 <i>Evaluación de impacto del PUEBP</i> .....	17
1.3 <i>Niveles del desempeño en lenguaje</i> .....	19
2. PLAN DE TRABAJO .....	25
2.1 <i>La herramienta estadística usada aquí</i> .....	25
2.2 <i>Metodologías estadísticas alternas</i> .....	39
2.3 <i>Escenarios de trabajo</i> .....	45
<b>CAPÍTULO II: ALGUNOS FACTORES ASOCIABLES AL DESEMPEÑO EN LENGUAJE .....</b>	<b>51</b>
3. PARTICIPACIÓN DEL ALUMNO.....	52
3.1 <i>La participación como factor asociable al desempeño</i> .....	52
3.2 <i>Algunos factores-participación asociables al desempeño en lenguaje</i> .....	61
3.3 <i>Resumen parcial</i> .....	76
3.4 <i>Un análisis parcial</i> .....	79
4. LIBROS .....	89
4.1 <i>Los libros como factor asociable al desempeño</i> .....	89
4.2 <i>Algunos factores-libro asociables al desempeño en lenguaje</i> .....	95
4.3 <i>Resumen parcial</i> .....	125
4.4 <i>Un análisis parcial</i> .....	130
5. “EFICIENCIA INTERNA” .....	139
5.1 <i>La “eficiencia interna” como factor asociable al desempeño</i> .....	139
5.2 <i>Algunos factores-“eficiencia interna” asociables al desempeño en lenguaje</i> .....	151
5.3 <i>Resumen parcial</i> .....	172
5.4 <i>Un análisis parcial</i> .....	174
6. DISCIPLINAS .....	191
6.1 <i>Por qué ciencias y matemáticas</i> .....	191
6.2 <i>Las disciplinas como factor asociable</i> .....	195
6.3 <i>Resumen parcial</i> .....	211
6.4 <i>Un análisis parcial</i> .....	214
<b>CAPÍTULO III: PARA CONCLUIR... Y SEGUIR .....</b>	<b>220</b>
7. POLÍTICAS EDUCATIVAS Y “EVALUACIÓN DE LA CALIDAD”.....	221
7.1 <i>Algunos componentes de la evaluación masiva</i> .....	221
7.2 <i>Una propuesta</i> .....	263
8. A MANERA DE CONCLUSIONES .....	266
BIBLIOGRAFÍA.....	282
ANEXO I.....	290

### **Siglas que se usan a lo largo del informe**

- Competencias-C: Desempeño en la parte *cerrada* de la prueba de lenguaje, analizado desde las categorías del equipo.
- Competencias-A: Desempeño en la parte *abierta* de la prueba de lenguaje, analizado desde las categorías del equipo.
- DC: Distrito Capital
- DNP: Departamento Nacional de Planeación
- ICFES: Instituto Colombiano para el Fomento de la Educación Superior
- MEN: Ministerio de Educación Nacional
- PUEBP: Plan de Universalización de la Educación Básica Primaria
- PEI: Proyecto Educativo Institucional
- SED: Secretaría de Educación Distrital
- SINECE: Sistema Nacional de Evaluación de la Calidad de la Educación (también conocido como “SABER”)
- SNP: Servicio Nacional de Pruebas, del ICFES
- TIMSS: Third International Mathematic and Science Study [Tercer estudio internacional en ciencias y matemáticas]
- Ley 115: Ley general de educación, Ley 115 de Febrero 8 de 1994

La estadística cuantifica la incertidumbre

Said Infante Gil y Guillermo P. Zárate de Lara

*Métodos estadísticos*

¿Cómo saber si el modelo está construido según la realidad si hace falta el modelo para informar sobre ella?

Gloria Benedito

«El problema de la medida en psicología»

Hay que reflexionar para medir y no medir para reflexionar

Gaston Bachelard

*La formación del espíritu científico*

Lo simbólico introduce un modo de operación que escapa a todo lo que podríamos hacer surgir a partir de una deducción de los hechos en lo real.

Jacques Lacan

*Seminario 3*



## 0. Presentación

Inicialmente, hacemos un reconocimiento al «Instituto para la Investigación Educativa y el Desarrollo Pedagógico» - IDEP que, siendo un organismo oficial, financió esta investigación que, en gran medida, es una objeción argumentada de un punto importante de las actuales políticas educativas oficiales.

El presente informe final de investigación consta de tres capítulos.

En el primero, se contextualiza la llamada evaluación de la calidad, en la cual se inscribe la evaluación de impacto del PUEBP, cuya información constituye la principal fuente utilizada en el presente estudio; se describe de manera esquemática el proceso de dicha evaluación de impacto y, específicamente, las maneras de analizar el desempeño en la prueba de lenguaje aplicada; se explican tanto la manera utilizada para buscar estadísticamente las correlaciones entre desempeño en el área de lenguaje y factores asociables a ese desempeño, como otras técnicas posibles; y se describe el recorrido del trabajo.

En el segundo capítulo, se buscan las sustentaciones de por qué suelen considerarse como asociados al desempeño, la participación del alumno, los libros, la “eficiencia interna” y las disciplinas (ciencias y matemáticas); en cada caso se buscan las correlaciones y se concluye de forma específica.

Y, en el tercer capítulo, hacemos un análisis aplicado en general a las políticas educativas y su relación con la evaluación masiva, componente por componente; concluimos de manera general y hacemos una propuesta que relanza nuestro trabajo hacia una línea de investigación.

En un CD aparte incluimos las tablas completas de toda la información utilizada, más otras correlaciones que no se tuvieron en cuenta en el presente estudio (como los resultados por región).

## **CAPÍTULO I: INTRODUCCIÓN**

## 1. Algunos antecedentes

La creciente inclusión de la evaluación entre los mecanismos de gestión de los sistemas educativos ha generado cambios estructurales\*

En el panorama internacional, los resultados del proceso educativo se relacionan cada vez más con los sistemas de evaluación de la calidad, particularmente con los llamados “factores asociados”; bajo la mirada de los actuales modelos [Peña, 1999:25], parecen insuficientes las maneras como se habían definido hasta hace poco los indicadores educativos. Colombia no es la excepción a esto: para los planes de desarrollo en educación, desde el gobierno nacional hasta el distrital, los términos son similares: para el gobierno nacional, uno de los tres propósitos fundamentales del sistema nacional de evaluación de la educación es «obtener, procesar y divulgar información pertinente, oportuna e inteligible según los destinatarios sobre el estado de la educación en el país, que apoye la toma de decisiones en los diferentes niveles del sistema educativo, siendo especialmente útil en la *definición o reorientación de políticas*\*\*» [DNP, 1999:52-53]... aunque, como dice el mismo documento, «La investigación sobre factores asociados a la calidad es incipiente en el país» [:11]\*\*\*. Por su parte, la SED [1998:4], con la aplicación de pruebas en competencias básicas y factores asociados se propone «establecer instrumentos para que la Secretaría oriente programas de apoyo».

La evaluación de impacto del PUEBP —en cuya ejecución participamos, uno de los autores desde el área de lenguaje, y el otro, desde el análisis multivariado— es una de las medidas que entra en esta nueva dinámica. El presente estudio pretende continuar la investigación realizada durante dicha evaluación de impacto [Jurado et al., 1998], con una mirada crítica sobre los límites y posibilidades de los estudios sobre educación que dicen apoyarse en investigaciones estadísticas (utilizaremos incluso su propia lógica).

---

\* Triana, 1996:39.

\*\* Cuando no se indique lo contrario, los destacados (con itálicas) en las citas son nuestros.

\*\*\* Cuando la cita se refiera al texto referenciado inmediatamente antes, sólo se pondrán dos puntos y los números de página o el aparte que se trae a colación.

El presente capítulo expone algunos elementos de la forma como se relacionan evaluación masiva y política educativa, contextualiza la evaluación de impacto del PUEBP y explica las tres maneras como se categoriza el desempeño de los estudiantes en la prueba de lenguaje que se utilizó durante la mencionada evaluación.

## 1.1 La política educativa

En decenios anteriores, los principales referentes de la política educativa eran algunos indicadores de gestión y eficiencia. Documentos del DNP [cf., 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<sup>1</sup>, era el examen de Estado.

En la década del 90, este panorama parece haber cambiado: los indicadores tradicionalmente empleados se tornan «absolutamente insuficientes para evaluar el desempeño de los sistemas educativos [...]. Dichos indicadores respondían al supuesto implícito de que dentro de la escuela se aprendía, supuesto que ha dejado de tener vigencia» [Toranzos, 1996: 65]. Por su parte, la UNESCO [1992:art.4] dice: «[...] la educación básica debe centrarse en las adquisiciones y los resultados efectivos del aprendizaje, en vez de prestar exclusivamente atención al hecho de matricularse, de participar de forma continuada en los programas de instrucción y de obtener el certificado final».

Así, desde comienzos de los 90, en Colombia aparece cada vez más visible una política educativa que se diseña *mencionando* un apoyo en los resultados de la “investigación” que relaciona estadísticamente (que “cruza”) dos tipos de información: i) sobre el grado de acercamiento entre los desempeños de los estudiantes y los objetivos del sistema educativo<sup>2</sup>, y ii) sobre los factores que pueden asociarse<sup>3</sup> a ese de-

---

<sup>1</sup> Hemos escogido como genérico el término “desempeño”, aunque sabemos de la proliferación terminológica al respecto (rendimiento, logros, competencias, etc.), que porta una discusión importante. Sobre esto, ver §7.1.1.

<sup>2</sup> Nótese la poca diferencia con el concepto tradicional de evaluación (particularmente, evaluación del rendimiento): confrontar objetivos y resultados. Para Triana [1996:51], aunque «ese nivel de generalidad en la respuesta resulta insuficiente», más adelante [:53] cae en él: «la valoración de tales resultados es tan importante como difícil. Es importante porque constituye la medida fundamental del éxito alcanzado por un sistema educativo determinado en relación con el logro de sus objetivos centrales».

sempañ [MEN, 1992:9]. Que sea “estadísticamente” es algo propio de los paradigmas que rigen nuestra “cultura”: Michel [1996:23], hablando de la “conducción” de los sistemas educativos, sitúa cuatro grandes funciones informativas que la evaluación masiva satisfaría, entre las que está «la función estadística o de conocimiento», ¡donde decir “conocimiento” parece equivalente a decir “estadística”! Esta posición quedó más o menos bendecida por la Ley 115, que establece el Sistema Nacional de Evaluación de la Educación [art. 80] en coordinación con el SNP del ICFES, que opera, en lo que respecta a evaluación y medición, con base en la estadística; así mismo, el artículo que asigna las funciones de la Junta Nacional de Educación, dice: «Presentar periódicamente al gobierno nacional, con base en los resultados de investigaciones *estadísticas*, informes de evaluación sobre el cumplimiento de los objetivos del servicio público de la educación» [art. 157, literal E.]. Es curioso pero de significar “relativo al Estado”, la estadística pasó a ser “datos relativos al Estado” y después, sencillamente, “datos” [cf. Desrosières, 1995].

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 (otros, muchas veces sin relación con los anteriores), se procede a un análisis *global* que finalmente “cruza” estas dos variables (y que puede ser hecho por nuevos expertos). En otras palabras, con los factores y el desempeño se opera estadísticamente una búsqueda de *correlaciones*<sup>4</sup>. Para ello se requiere que la evaluación del desempeño haga una clasificación de los estudiantes; como veremos, 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 PUEBP, en cambio, hizo una estandarización basada en el número de respuestas acertadas, independientemente de los niveles a los que cada pregunta estuviera teóricamente asignada [cf. §1.3.1].

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 avancen hacia mejores desempeños que en aquellas escuelas en las que hubo menos dotación. En tal caso, se diría que hay una correlación positiva (a más de lo

---

<sup>3</sup> En estos estudios se acostumbra a decir factores “asociados”. En el presente texto diremos “asociables”, pues justamente la investigación está destinada a indagar si los factores se asocian; en otras palabras, antes de cualquier demostración, son apenas *asociables*.

<sup>4</sup> En este trabajo, usaremos indistintamente los términos “correlación”, “dependencia” y “asociación” entre variables.

uno, más de lo otro) y, en consecuencia, que la dotación de medios pedagógicos efectivamente coadyuva a mejorar los desempeños y, por lo tanto, la calidad de la educación. De donde se justificaría una política educativa que se proponga entregar medios pedagógicos a las escuelas. Es por eso que se usa también para evaluar el “impacto” de programas educativos.

Se buscan relaciones sistemáticas entre logro y factor, de tal manera que pueda hablarse de una asociación positiva o negativa<sup>5</sup>; por ejemplo: que, a más textos, más tránsito a mejores logros y que, a menos textos, menos tránsito a mejores logros (asociación positiva); o que por encima de la edad promedio del estudiante, menos tránsito a mejores logros y viceversa (asociación negativa).

Se argumenta la necesidad de tener evidencia de que los aspectos sobre los que se va a intervenir influyen consistentemente sobre el desempeño estudiantil; de lo contrario, por muy bien intencionada que sea, la inversión podría desperdiciarse. Por ejemplo, distribuir textos a las escuelas, capacitar a los profesores, dotar a las escuelas de infraestructura, disminuir el número de alumnos por docente, aumentar la jornada de estudio, disminuir la repitencia, capacitar a los padres, etc.<sup>6</sup>, son acciones cuya efectividad debe “probarse”, pues el asunto no es evidente: aquello que parece importante puede mostrarse —en las cifras— ineficiente, y lo supuestamente contraproducente puede mostrarse —en las cifras— eficiente; veamos un par de ejemplos:

- De un lado, hay factores que podrían considerarse importantes y que, no obstante, contra toda evidencia, ciertos estudios dicen que no aportan al desempeño; es el caso de la capacitación docente. A todos nos parece evidente que los estudiantes de un docente que ha tomado mayor número de cursos de capacitación tienen ventajas para obtener mejores resultados en las pruebas de desempeño que los estudiantes de un docente con menos cursos de capacitación. Sin embargo, el Banco Mundial [Heyneman et al., 1991:134], afirma que no; asunto que se “verifica”<sup>7</sup> en Colombia poco tiempo después: «el número de cursos de capacitación recibidos en servicio no aparece con asociaciones significativas» [MEN, 1992:89]; y reiteradamente: «lo correspondiente a la capacitación para el ejercicio docente no tiene prácticamente efecto en el sector oficial urbano» [MEN, 1997b:54].

---

<sup>5</sup> Este tipo de asociación se verifica mediante el análisis de tablas de contingencia de variables con modalidades ordinales [cf. Agresti, 1996:34-39]. Esto se desarrolla en §2.1.

<sup>6</sup> Cf., por ejemplo, el programa de “Acciones focalizadas” de la SED.

<sup>7</sup> Ya veremos que no se trata sencillamente de una demostración “científica”, entendida como exenta de intereses.

- Y, de otro lado, hay factores que podrían considerarse contraproducentes y que, no obstante, ciertos estudios los muestran como favorables: de un lado, hay investigaciones que hablan de «disminución de las habilidades verbales entre los niños que crecieron viendo televisión» [citado por Flórez, 1997:1], mientras los apocalípticos, como dice Abad [1997:73], piensan que los niños frente a la TV son como muñecos descerebrados y que por eso leen cada vez menos y lo poco que leen no lo entienden; y, de otro lado, el SINECE [MEN, 1992:92-93] encontró una «asociación positiva entre la cantidad de tiempo que el estudiante dedica a la televisión y su logro académico»; Ávila [1999:38] menciona también tal “efecto positivo”, aunque lo independiza del tiempo que se dedique a esa actividad.

Por razones como éstas, comienza a ser cada vez más común encontrar en los enunciados de las instancias que gobiernan la educación frases como la siguiente: «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]. Obsérvese la similitud de estos enunciados con la jerga a escala internacional (España, en el siguiente caso): «La primera de las funciones que desempeña la evaluación en relación con la mejora cualitativa de la enseñanza es precisamente la de proporcionar datos, análisis e interpretaciones válidas y fiables que permitan forjarse una idea precisa acerca del estado y situación del sistema educativo y de sus componentes» [Triana, 1996:46].

Y, más recientemente, el MEN [1997b:10] afirma: «invertir en sistemas de evaluación tiene sentido, sólo en la medida en que se publican sus resultados y se utilizan en la *toma de decisiones* y mejoramiento de la calidad [...]. Estas evaluaciones tratan de obtener información sobre el efecto de la acción del “Sistema Educativo” que sea *útil para la determinación de políticas educativas*». Y la similitud continúa: «además de aumentar la capacidad de comprensión de los fenómenos educativos, pretenden proporcionar una base lo más sólida posible para la toma de decisiones» [Triana, 1996:48]; o, como se afirma en la investigación hecha en Colombia sobre repitencia: «con este aporte investigativo, esperamos contribuir a la reflexión sobre el tema y al diseño de políticas y acciones [...]» [MEN, 1993c.:2].

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”<sup>8</sup> que intentan «aproximarse a

---

<sup>8</sup> La evaluación de impacto del PUEBP —que nos ha servido de base para la investigación— introdujo esta dimensión con el nombre de «actitudes y valores infantiles».

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]. La evaluación de esta dimensión —según reconoce el mismo MEN— ha tenido poco éxito, en un contexto en el que «Las pruebas de Desarrollo personal y social (no cognitivo) han tenido un carácter experimental y están en exploración en los países con sistemas de evaluación» [Peña, 1999:27]. ¿Por qué insistir, entonces, en este tipo de factor?: «para convertirse en instrumentos de conducción más eficaces, estas evaluaciones de masa y sobre muestra deberán centrarse progresivamente en los conocimientos *no cognitivos* adquiridos: competencias transversales, comportamientos, valores, aptitud para el trabajo en equipo, etc. En resumen, en lo que respecta al “currículum oculto”, sobre el que se centran las principales expectativas de los empleadores» [Michel, 1996:28, destacado del autor]. Con lo que queda explícita la razón de hablar en estos términos: formar para el trabajo que está en la expectativa de los empleadores.

En todo caso, la definición que hace el MEN de estas pruebas “de desarrollo personal y social” 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 postularía una comprensión premoderna de la especificidad del conocimiento humano).

En el Plan de desarrollo en educación del gobierno de Gaviria, *Plan de apertura educativa*, hemos visto con más claridad cómo se va explicitando esta tendencia que señalamos. Dicho plan 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, en función de dichas inversiones. De ahí que se hicieran pruebas periódicas en 3°, 5°, 7° y 9°: «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 un examen de Estado modificado no culminaba todo el sistema; cosa que la *Misión ciencia, educación y desarrollo* planteó como factor de



mejoramiento de la calidad y que hoy vemos realizado en el publicitado cambio de la prueba de Estado<sup>9</sup>.

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 estadística entre desempeño y factores asociables) que se menciona 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].

Veamos a continuación cómo los Planes de desarrollo en educación de los distintos gobiernos nacionales asumen el asunto desde entonces:

- el *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];
- 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 [...]» [DNP, 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» [DNP, 1999:11].

Por su parte, el gobierno distrital dice en su plan sectorial: «Recordemos que con la aplicación de pruebas en competencias básicas y factores asociados el gobierno distrital persigue establecer instrumentos *para orientar programas de apoyo*» [SED, 1988:4].

---

<sup>9</sup> El ICFCES viene haciendo público en los diarios nacionales lo que llama el “nuevo examen del ICFCES”, que se empezó a aplicar en el 2000. En el marco de esta divulgación, el 22 de Noviembre de 1999 hubo una reunión técnica con jefes de admisión de las universidades.

Llega a decirse, a propósito de la discusión sobre el concepto de “calidad de la educación”: «No obstante esta necesaria y conveniente controversia, en la presente década ha sido significativo el énfasis que ha vuelto a tomar el enfoque de que el punto de partida para mejorar la calidad de la educación, es evaluarla, mediante instrumentos que permitan obtener información sobre lo que los niños logran aprender y los factores que inciden de manera positiva o negativa en estos logros» [MEN, 1997b:9]. Es un enunciado que se postula como “más allá del bien y del mal”, como transcultural; veámoslo con más claridad en los siguientes casos en los que destacamos la misma palabra: «*Independientemente* de lo que entendamos sobre educación para propósitos particulares, ello no puede contradecir la concepción de educación de la Ley 115 [...]» [:13]; «asegurar la calidad de la educación es condición para mejorar la calidad de vida de un ser humano en cualquier contexto cultural, social en el que se halle, e *independientemente* de lo que en tales contextos se entienda por educación» [:14]. La discusión y la diferencia parecen quedar sobrepasadas por la trascendencia del propósito... y ya Zuleta [1994] comentaba sobre el riesgo que entraña sentirse dueño de una verdad tan absoluta.

Esta “nueva” manera de hacer las cosas, conocida como “evaluación de la calidad educativa”, no parece ser una iniciativa nacional, producto de la maduración de tendencias evaluativas, pedagógicas, disciplinares, etc. en nuestro país. Más bien es una dinámica internacional: «La demanda de respuestas objetivas y fiables a las cuestiones suscitadas acerca del sistema educativo estadounidense favoreció la canalización de notables recursos económicos hacia las actividades de evaluación, produciendo como consecuencia un gran impacto sobre su desarrollo académico y profesional. La evaluación educativa experimentaría así un apreciable desarrollo a partir de finales de la década de los sesenta, cuya influencia se haría sentir progresivamente en otros países» [Triana, 1996:37-38].

En el marco de esta “influencia” se hacen ciertos requerimientos que cuentan con el aval de quienes toman las decisiones en política educativa en Colombia. Esta relación de “nuestra” evaluación de la calidad con las exigencias y tendencias internacionales —que poco se explicita— la ejemplifican los siguientes hechos:

- Al argumentar a favor de la creación de un sistema de evaluación, el 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;
- 8 años después, la justificación utiliza similares eufemismos: Peña [1999:26] plantea que, de un lado, habría una necesidad de unificar información a escala

nacional, dada la autonomía de los establecimientos educativos<sup>10</sup>; y, de otro lado, plantea que hay “nuevas demandas” de información y de rendición de cuentas, formuladas por varios actores: «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”, a la vez que se hace contradictorio con el diagnóstico de que nuestra educación está divorciada del contexto); «[...] 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 verosímil, hecha por un demandante con nombre propio)<sup>11</sup>.

- Idénticas medidas de evaluación abundan en otros países, con propósitos, instrumentos e incluso nombres muy parecidos: SINECE en Colombia, SINEC en Argentina, SIMCE en Chile, Sistema Nacional de Evaluación en México... idéntico nombre que en Estados Unidos, Italia e Inglaterra [cf. Triana, 1996:39; Michel, 1996:29; Peña, 1999]. Y esto parece obvio cuando sabemos que, para la UNESCO [1992:art.4], «Que el incremento de las posibilidades de educación se traduzca en un desarrollo genuino del individuo y de la sociedad depende en definitiva de que los individuos aprendan verdaderamente como resultado de esas posibilidades, esto es, de que verdaderamente adquieran conocimientos útiles, capacidad de raciocinio, aptitudes y valores. [...] De ahí que sea necesario determinar niveles aceptables de adquisición de conocimientos mediante el aprendizaje en los planes de educación y *aplicar sistemas mejorados de evaluación de los resultados*».

---

<sup>10</sup> Lo mismo se dice en España: «Las escuelas pueden y deben hacer uso del creciente margen de autonomía de que disponen, organizando los procesos de enseñanza y aprendizaje en ese nuevo marco. Pero, a cambio, deben rendir cuentas de sus resultados a través de diversos mecanismos de evaluación. Es este un estilo de funcionamiento frecuente en las organizaciones productivas, que ha ido arraigando también entre las administraciones públicas. Parece ser que la educación no constituye una excepción a lo que va siendo *una regla cada vez más extendida en la gestión de los servicios públicos*» [Triana, 1996:42].

<sup>11</sup> Lo de la “rendición” de cuentas y lo de las múltiples “demandas” sociales tampoco es un invento nuestro; es el mismo lenguaje en todos los países [cf. Triana, 1996:43-44, 57]. Esta estrategia de asignarle demandas a los estamentos sociales es de utilidad para la imposición de los sistemas de evaluación sin necesidad de sacar de los enunciados la diferencia y los contextos: «No se puede hablar de sistemas educativos mejores o peores en términos absolutos, sino mejor o peor adaptados a sus circunstancias específicas y *más o menos capaces de dar respuesta a las demandas que se le formulan*» [:59].

- Las Cumbres Iberoamericanas de Jefes de Estado y de Gobierno («escenario *natural* de debate de los problemas de la región»<sup>12</sup>) se realizan desde 1991 con temas específicos<sup>13</sup>. 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 considerado como iniciativa regional lo que ha sido una política internacional, consentida país por país. Incluso las asesorías ahora se prestan entre países (Colombia asesora a Bolivia, por ejemplo).
- 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].
- Colombia viene participando en procesos evaluativos internacionales: «El país no ha rehuído tampoco las evaluaciones internacionales. Fue así como participamos en una evaluación internacional, junto a otros 40 países, sobre la enseñanza y los aprendizajes en ciencias y en matemáticas. Sabíamos de antemano que no nos iría muy bien al principio, como se corroboró en los resultados. Pero era necesario afrontar los riesgos de compararse con países de alto desarrollo» [MEN, 1997b:7]. Se trataba del TIMSS [cf. Díaz, 1998], pruebas que pretenden estandarizar los procesos evaluativos a escala internacional<sup>14</sup>; por ejemplo, en el asunto del tiempo: considerándolo como asociado positivamente al desempeño (a más tiempo de estudio, 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. DNP, 1994 y 1999; MEN, 1996b; Fundación Social, 1998]. De ahí que el DC invierta \$7.100 millones en el programa de “Ampliación de la jornada educativa” [SED, 1998:46].

En esto no se ven diferencias entre gobiernos ni entre partidos (si las hubiere): según el Plan de Pastrana, «para mantener informada a la comunidad educativa y a la ciudadanía sobre el estado y los cambios en los logros de los estudiantes en áreas fun-

---

<sup>12</sup> Declaración del presidente Pastrana al diario EL TIEMPO, noviembre 14 de 1999, pág. 15A.

<sup>13</sup> A finales de 1999 se realizó en La Habana, Cuba, la IX cumbre, con el tema de la Economía global.

<sup>14</sup> La *International association for the evaluation of educational achievement* promueve y realiza estudios internacionales de evaluación educativa desde la década del 60; entre los últimos está el TIMSS [cf. Triana, 1996:38; UNESCO, 1998:61].

damentales de la educación básica y de los factores del entorno personal, familiar, escolar y social de los estudiantes que contribuyen a que ellos obtengan altos o bajos logros *se continuará* con los esfuerzos adelantados para la evaluación de la calidad de la educación» [DNP, 1999:53].

Se echa a andar una política y después se registra, como desde la barrera, que hay brotes aquí y allá que muestran una tendencia: la UNESCO que, como vimos más atrás, es una de las que recomienda e impulsa este tipo de políticas, lo “registra” así: «un interés creciente de las autoridades nacionales por el seguimiento, la medición y la evaluación de los procesos y el rendimiento escolares» [UNESCO, 1998:50]. Por su parte, Triana [1996:53] dice: «no hay país que pueda ignorar la relevancia de dicha valoración [...] puede afirmarse que la importancia de la tarea de evaluación de los resultados prevalece sobre su evidente dificultad. Como se mencionaba más arriba, son cada vez más los países que han puesto en marcha programas de esa naturaleza».

## **1.2 Evaluación de impacto del PUEBP**

Con los objetivos explícitos de ampliar la cobertura y de mejorar la calidad de la educación básica, el MEN implementó, entre 1989 y 1997, el llamado «Plan de Universalización de la Educación Básica Primaria». Este Plan se financió con un préstamo del Banco Mundial, que permitió invertir en obras civiles, mobiliario escolar, materiales didácticos, libros y capacitación docente [cf. MEN, 1988 y Banco Mundial, 1988].

Para establecer su impacto, al término del programa<sup>15</sup> el MEN recogió información en 100 municipios en todo el país sobre las cifras educativas del municipio, los recursos de la escuela, las características de los profesores y los desempeños de 6447 niños de 3° y 4°<sup>16</sup> en matemáticas, ciencias, lenguaje, y valores y actitudes infantiles. La muestra a escala nacional fue diseñada y administrada por el MEN.

Las áreas de ciencias y matemáticas aparecen mezcladas en un solo cuestionario de 30 preguntas (15 para cada área) con respuesta verdadera única (que llamaremos “cerradas”); el aspecto de actitudes y valores infantiles se dividió en un cuestionario para estudiantes y otro para docentes (48 y 18 preguntas, respectivamente); los fac-

---

<sup>15</sup> El hecho de que hubiera sido “al término del programa” implica que el estudio fue de tipo transversal (o sea, en un instante del tiempo) y, en consecuencia, que las conclusiones no pueden presentarse longitudinalmente (para otros tiempos).

<sup>16</sup> Los niños de calendario B fueron encuestados empezando el año; o sea, que acababan de terminar 3° y 4°, respectivamente.

tores asociables se indagaron en formularios para estudiantes (9 preguntas), maestros (29 preguntas), director de escuela (32 preguntas) y municipio (44 preguntas); y la prueba de lenguaje —en cuyo diseño y análisis participamos— estuvo constituida por 20 preguntas.

Con todo esto, el equipo del MEN hizo un informe global [MEN, 1997], en el que se incluyeron algunas páginas del informe específico del área de lenguaje, preparado por la Asociación Colombiana de Semiótica<sup>17</sup>, encargada de esta prueba. Este último informe respondía al propósito de contestar a las preguntas que se hacía el equipo de investigación en lenguaje, desde la teoría que puso en juego; por ello,

- se ajustaron los resultados a las categorías y niveles definidos teóricamente por competencias (además de la determinación estandarizada que se hizo para los objetivos del MEN);
- se tuvieron en cuenta ambas partes constitutivas de la prueba (la parte cerrada y la parte abierta), ya que para el análisis del MEN se determinó considerar solamente lo relativo a la parte cerrada;
- y se establecieron las relaciones entre esta evaluación y las evaluaciones de la calidad de la educación en Colombia, en las cuales el equipo también había participado.

Una vez *estandarizados* los puntajes [cf. §1.3.1], y respondiendo a sus propios intereses y criterios de evaluación, el MEN los relacionó con la información proveniente de los instrumentos de factores asociables. Por nuestra parte, en el presente estudio investigamos la relación entre algunos factores asociables y el desempeño en el área de lenguaje, pero desde los objetivos y criterios de evaluación propios del equipo. Por tal razón, se hacen otras asociaciones:

- entre los desempeños de los niños en la parte abierta y en la parte cerrada;
- entre los niveles estandarizados definidos por el equipo de evaluación del MEN y los niveles por competencia definidos por el equipo de lenguaje;
- entre factores seleccionados del conjunto establecido en el instrumento aplicado por el MEN<sup>18</sup> y los desempeños (por competencias y estandarizados);

---

<sup>17</sup> Publicado con el nombre de *Juguemos a interpretar* [Jurado et al. 1998].

<sup>18</sup> Es decir, no se tomarán todos los factores indagados por el MEN (sería un trabajo demasiado extenso), sino aquellos que, a nuestro juicio, son más pertinentes para investigar la relación con el área de lenguaje.

- entre factores redefinidos<sup>19</sup> por el equipo y desempeños.

En consecuencia, pese a que el diseño muestral fue hecho pensando en una expansión de los resultados a la población nacional —con lo que se buscaba el perfil del niño impactado por el PUEBP en todo el país—, destacamos que hay una sucesión de aspectos previamente conceptualizados y que, como tales, son susceptibles de examinarse; son ellos los que nos interesa discutir, pues en la eventualidad de que fueran cuestionados, los “resultados” perderían validez, mientras que la discusión conceptual siempre será enriquecedora. Esa sería una forma de trascender ciertas maneras de investigar la educación, a favor de otras que den cuenta de más relaciones, de más complejidades...

### **1.3 Niveles del desempeño en lenguaje**

La prueba de lenguaje constó de 20 preguntas referidas a una narrativa icónica que incluía parlamentos. 19 preguntas fueron cerradas, con tres opciones de respuesta; en la pregunta restante, el niño respondía escribiendo un texto libremente (a esta pregunta la llamaremos “abierta”)<sup>20</sup>.

Con base en las respuestas de los niños a este instrumento, tenemos, de un lado, los niveles de desempeño estandarizados que el MEN estableció; y, de otro lado, los niveles por competencia que el equipo definió, tanto para la parte cerrada, como para la parte abierta. A continuación se explica cada uno de estos niveles de desempeño.

#### ***1.3.1 Categorización por estandarización***

Como los sistemas de calificación de las pruebas (ciencias, matemáticas y lenguaje) eran diferentes, sus puntajes se hacían inconmensurables. Por esto se llevó cada escala «a distribuciones homogéneas de media 50 y desviación estándar 10, lo cual demarca el intervalo 40-60 que define la categoría de puntajes *medios*; por debajo de 40 puntos se denominan *deficientes* [bajo]; y por encima de 60 puntos, *sobresalientes* [alto]» [MEN, 1997:24]. La “media” es el valor al rededor del cual se considera que está dispuesta la información; la “desviación estándar” (y su cuadrado, la “varianza”) mide el grado de dispersión de la información respecto a la media.

---

<sup>19</sup> En cada caso se aclarará cómo se hizo la redefinición.

<sup>20</sup> La prueba puede verse en Jurado et al., 1998, anexo 1.

Así explica el MEN [1997b:20] la estandarización: «Para describir el rendimiento de un estudiante, se acude a una escala de medición estandarizada construida a partir de las respuestas correctas de cada estudiante y sin tener en cuenta el número total de preguntas que contiene el examen; con esta escala, cada estudiante puede ser situado con referencia al número promedio de respuestas correctas de un grupo de alumnos que “estudian” el mismo contenido programático y abordan el mismo examen».

Con la estandarización hecha en la evaluación de impacto del PUEBP, resultan los desempeños de los niños clasificados en las categorías “bajo”, “medio” o “alto”, de acuerdo con la distancia del puntaje de cada uno respecto a la media. Es decir, esas categorías no están sustentadas desde una teoría del lenguaje como objeto a evaluar, sino que aparecen por rutina técnica.

Según el MEN [1997b:20], los resultados de este tipo de categorización se enjuician mediante la valoración cultural dada al sitio que se ocupa en la escala, con referencia al promedio del grupo (“bajo”, “medio” o “alto”), el cual se utiliza como estándar de rendimiento.

### ***1.3.2 Categorización por criterio***

Escogimos una categorización por criterio, tanto en la parte cerrada como en la abierta, pues nos interesaba establecer el desempeño de los niños en relación con una teoría genético-semiótica del desarrollo frente a los procesos de lectura y escritura [cf. Jurado et al., 1998]. No obstante, la decisión de cómo calificar a cada niño fue un acuerdo del grupo de investigación que, como toda decisión de esta naturaleza, es más o menos arbitraria (por eso, a continuación explicitamos los criterios).

#### **8 Parte cerrada (comprensión)**

La parte cerrada de la prueba tiene 19 preguntas que hemos convenido en llamar “de comprensión” (sin desconocer que en ello hay producción), pues se trata de indagar sobre niveles de aproximación del estudiante a los textos (entendidos en sentido semiótico) que la narrativa icónica materializaba y cuya comprensión las preguntas cerradas pretendían interrogar. Esta aproximación la teorizamos en tres niveles<sup>21</sup>. «Se entiende el “criterio” como de naturaleza interna, la prueba misma o lo que la prueba representa, para dar por sobreentendido que se trata de una prueba opuesta a aquellas que lo hacen con referencia a un estándar externo (externo a la prueba): el

---

<sup>21</sup> Sin embargo, admitimos que no todos los desempeños caben en nuestra teorización, razón por la cual postulamos, *a posteriori*, un nivel 0.



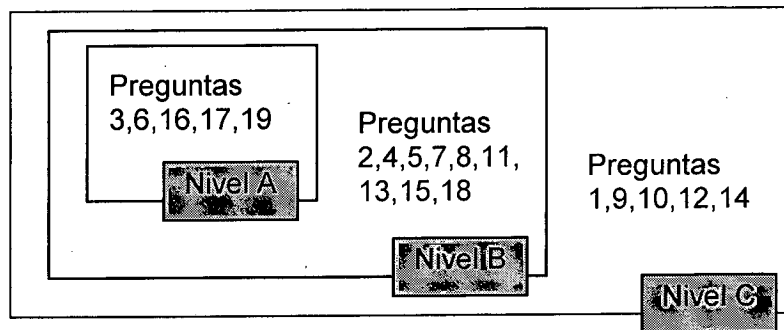
grupo o el estándar del grupo o el rendimiento promedio del grupo que se constituye en patrón de medida [...] La preocupación presente en una prueba de logro es qué tan válidamente representa los estados de acceso del sujeto al objeto de conocimiento» [MEN, 1997b:21].

A continuación se hace una síntesis de los niveles de comprensión desarrollados en Jurado et al. [1998: §3.2]:

- Nivel 0: Está configurado por los niños que nuestros criterios de clasificación no pueden “registrar”, pues no responden ninguna pregunta, o responden acertadamente máximo dos preguntas (en cambio, bajo la forma como se hace la estandarización, toda la muestra queda clasificada). En la investigación anterior [Jurado et al., 1998], definíamos este nivel como “no clasificados”. En realidad, los niños que quedan en esta categoría sí están clasificados, pues son aquellos cuyo desempeño no alcanza el nivel literal.
- Nivel literal: Son lecturas instauradas en el marco del “diccionario” o de los significados “estables” integrados a las estructuras superficiales de los textos. En este nivel hay dos variantes: a) Transcripción: el lector reconoce palabras y frases, con sus correspondientes significados de “diccionario” y las asociaciones automáticas con su uso. El lector reconoce y discrimina grafías y palabras. b) Paráfrasis: el lector hace una traducción semántica en donde palabras semejantes a las del texto leído ayudan a integrar el sentido. Se trata del reconocimiento del primer nivel de significado del mensaje, y se realiza cuando el lector parafrasea, glosa o resume lo que lee. Se activan aquí las macro-reglas, necesarias en toda comprensión de texto: generalizar, seleccionar, omitir e integrar información.
- Nivel inferencial: El lector establece relaciones y asociaciones entre los significados, lo cual conduce a relaciones de implicación, causa, tiempo, espacio, inclusión, exclusión, agrupación, etc., inherentes a la funcionalidad del pensamiento y constitutivas de todo texto. Se infiere lo no dicho en el acto de decir, pues el acto de leer conduce a permanentes deducciones y completaciones de los intersticios textuales. En la inferencia se hallan los procedimientos propios de la presuposición (o hipocodificación).
- Nivel crítico-intertextual: El lector hace conjeturas, determinadas ya no desde el *Diccionario* sino desde la *Enciclopedia*; es decir, la puesta en red de saberes de múltiples procedencias (esto sería lo intertextual). La explicación interpretativa se realiza por distintos senderos: va desde la reconstrucción de la macroestructura semántica (coherencia global del texto), pasa por la diferenciación genérico-discursiva (identificación de la superestructura) y desemboca en el reconoci-

miento de los puntos de vista tanto del enunciador textual y el enunciatario, como de las intencionalidades del autor empírico. También el lector se posiciona críticamente: emite juicios respecto a lo leído. Se trata de la abducción creativa, mediante la cual el lector conjetura y evalúa aquello que dice el texto e indaga por el modo como lo dice. Tales movimientos conducen a identificar intenciones de los textos y de los autores y, en consecuencia, a actualizar las representaciones de quien lee. En la lectura crítico-intertextual el lector pone en juego la capacidad para controlar la consistencia en las interpretaciones diversas y posibles que el texto puede soportar, en un proceso de semiosis que converge finalmente en el reconocimiento valorativo del mismo texto en relación con los otros textos de la cultura, y que se pueden manifestar, a manera de citación, de alusión o de imitación.

A continuación se ilustra la manera como se concibió previamente la clasificación de las preguntas, así como la concepción de niveles envolventes:



#### ● Parte abierta (producción)

La parte abierta de la prueba tiene una pregunta que hemos convenido en llamar “de producción” (sin desconocer que en ello hay comprensión), pues se trata de indagar sobre niveles de escritura de los niños, que teorizamos en tres; sin embargo, no todos los desempeños caben en dicha teorización, razón por la cual también en este caso postulamos un nivel 0. A continuación se hace una síntesis de los niveles de producción desarrollados en Jurado et al. [1998:§4.2]:

- Nivel 0: En este nivel estarían los niños cuya producción no alcanza a ser detectada por nuestro criterio de clasificación, o sea, los que no alcanzan el Nivel de coherencia local. En la investigación anterior [Jurado et al., 1998], también definíamos este nivel como “no clasificados”. En realidad, los niños que quedan en esta categoría sí están clasificados, pues son aquellos cuya producción no alcanza el primer nivel teorizado de producción.

- Nivel de Coherencia Local: Se refiere al nivel interno de la proposición (por lo tanto, se requiere la producción de al menos una) y es entendida como la realización adecuada de enunciados; constituye el nivel microestructural. Se tiene en cuenta la producción de proposiciones delimitadas semánticamente y su coherencia interna. En esta categoría se evidencia la competencia para establecer las concordancias de sujeto/verbo, género, número y la competencia del estudiante para delimitar proposiciones (segmentación).
- Nivel de Coherencia Lineal: Se refiere al seguimiento de un núcleo temático a lo largo de la producción. Constituye un nivel macroestructural: dar cuenta, en un sentido, de la globalidad del texto. Se considera que un texto responde a esta categoría cuando varias proposiciones coherentes siguen un hilo temático a lo largo del texto
- Nivel de Coherencia Global: Referida a la ilación de las proposiciones entre sí; es decir, al establecimiento de vínculos, relaciones y jerarquías entre las proposiciones para constituir unidades mayores de significado. Se logra mediante el empleo de recursos cohesivos como los conectores, señalizadores y los signos de puntuación, cumpliendo una función lógica y estructural; es decir, estableciendo relaciones de manera explícita entre las proposiciones, bien mediante conectores con función (cuando, además de producir más de una proposición coherente, establece algún tipo de relación estructural entre las proposiciones); bien mediante la explicitación de la/s relación/es interproposicional/es a través del uso de signos de puntuación con función lógica.

A continuación se ilustra la manera como se clasificaron las respuestas de los niños en la parte abierta, así como la concepción de niveles envolventes de sus desempeños:

Dimensión Textual					
Categorías	Coherencia local		Coherencia lineal	Coherencia global	
Subcategorías	1 Concordancia	2 Segmentación	3 Progresión Temática	4 Conectores con función	5 Signos puntuación con función
Condiciones	- Al menos una proposición - Concordancia sujeto/verbo	- Segmentarla debidamente - Mediante recurso explícito	Seguir un hilo temático a lo largo del texto	Establecer relación estructural explícita entre proposiciones	Evidenciar relaciones interproposicionales mediante recursos cohesivos con función lógica
Niveles	A		B	C	
	B				
	C				

Ambas escalas de niveles por competencias (de comprensión y de producción) expresan una dificultad creciente: quien está en el tercer nivel, también está en el segundo; y quien se encuentra en el cuarto, también se encuentra en el tercero y en el segundo<sup>22</sup>. «Los niveles de logro establecen un entramado jerárquico que va de los más bajos a los más altos o de los más simples a los más complejos, pero donde la naturaleza propia de los niveles superiores trasciende e incluye a los niveles inferiores y no viceversa» [MEN, 1997b:22]... algo que no se cumple en la estandarización.

No obstante, las cifras que en este trabajo se usarán de los desempeños por competencia corresponden a los estudiantes que *alcanzan máximo el correspondiente nivel* (por eso, los porcentajes sumarán 100%). Para facilitar y explicitar la lectura del tipo de asociación (en caso de que se evidencie), se observará principalmente la proporción de estudiantes que alcanzan las categorías extremas de los desempeños, en función de las modalidades de los respectivos factores asociables. Sin embargo, un énfasis sobre las categorías intermedias es igualmente plausible para otros propósitos.

---

<sup>22</sup> En esto, el Nivel 0 no hace serie, porque se refiere justamente a aquellos cuyo desempeño no ha podido quedar definido con la categorización usada para diseñar la prueba.

## 2. Plan de trabajo



Rius\*

### 2.1 La herramienta estadística usada aquí

En esta sección se explican las tablas de contingencia y algunos métodos para su análisis estadístico. Se incluyen algunos conceptos de clasificación, independencia y la estadística Ji-cuadrado para verificar la hipótesis de independencia entre las variables categóricas que conforman las tablas. Finalmente, se desarrolla un ejemplo tipo, siguiendo la metodología estadística del estudio.

#### 2.1.1 Clasificación

Se parte del supuesto de que alguien investido de autoridad teórica y/o empírica, clasifica los miembros de una población en categorías diferentes. Hay contextos en los que atribuimos esa autoridad con facilidad: por ejemplo, en relación con nuestro rH sanguíneo. Pero hay otros contextos (no por ello más arbitrarios) en los que resulta más claro que hay una decisión y no simplemente una comprobación: por ejemplo, el coeficiente intelectual “normal”. En la presente investigación tendremos ese tipo de clasificaciones: “alta” repitencia; lectura “literal”; coherencia “global”; haber recibido o no textos de español; ubicarse en una de las categorías de bajo, medio o alto en el desempeño de una prueba de lenguaje; tener muchos, intermedio, o pocos alumnos por docente, etc.

---

\* Rius, 1971:45.

De esta forma, se tienen variables cuyos valores son dos o más categorías (o modalidades). Éstas se asumen excluyentes y exhaustivas: por excluyentes se entiende que un individuo se ubica o asigna en tan sólo una de las modalidades de la variable, y por exhaustivas se hace referencia a que la variable brinda las categorías suficientes para ubicar allí a cualquier individuo de la población observada. Para garantizar el cumplimiento de esta última característica, en las variables de desempeño en lenguaje, el equipo decidió establecer una modalidad adicional llamada “nivel 0” [cf. § 1.3.2 y 1.3.3]<sup>23</sup>. Además, las categorías pueden ser transformadas mediante la unión o división de algunas de sus modalidades, con arreglo a concepto o por necesidades del análisis.

Cuando una población es clasificada en varias categorías, es posible contar el número de individuos que “caen” en cada modalidad. Estos conteos o frecuencias son los datos con los cuales se hace el análisis estadístico. La información se obtiene a través de un censo<sup>24</sup> o de una *muestra*<sup>25</sup> de la población de interés. El objetivo principal de la disciplina estadística es cuantificar el grado de incertidumbre de las inferencias hechas desde el examen de la información contenida en una muestra. Una característica esencial para este proceso es asegurar que la muestra tomada *represente* a la población. Una manera de garantizarlo es mediante una selección *aleatoria*, con la cual cada miembro de la población tiene igual oportunidad de ser incluido en la muestra. Además, se debe considerar el diseño de muestreo pertinente en cada caso o en cada población (muestreo estratificado, sistemático, por conglomerados, etc. [cf. Peña, 1997:203 a 208]).

### **2.1.2 Tablas de contingencia**

En el presente estudio se analizan datos que constituyen una muestra de una población clasificada respecto a dos o más variables cualitativas. La tabla siguiente muestra la clasificación de los estudiantes de la muestra nacional de acuerdo con las variables *desempeño estandarizado* y por *competencias* en la prueba de lenguaje realizada dentro de la evaluación del PUEBP [MEN, 1997].

---

<sup>23</sup> Podría pensarse que este nivel 0 —tanto en la parte cerrada como en la abierta— influiría negativamente en las asociaciones que se buscaron en este estudio; pero, como puede verse en el Anexo I, a) hay más asociaciones cuando la clasificación tiene nivel 0 que cuando no tiene (64 contra 60) y b) cuando hay asociación, en 7 casos la presencia del nivel 0 favorece que la asociación sea alta y no moderada; y en otros 7 casos es al contrario.

<sup>24</sup> Como en la actual evaluación de competencias del DC.

<sup>25</sup> Como en el caso de la Evaluación de la calidad en Colombia o de la evaluación del PUEBP.

		Competencias				Total
		Nivel 0	Literal	Inferencial	Crítico intertext.	
Estandarizado	Bajo	8	278	668	155	1109
	Medio	54	955	2394	556	3959
	Alto	11	340	841	187	1379
	Total	73	1573	3903	898	6447

Una tabla como la anterior se conoce con el nombre de *tabla de contingencia*; en este caso, la variable columna (desempeño por competencias) tiene 4 modalidades, mientras que la variable fila (desempeño estandarizado) tiene 3 modalidades. En general, una tabla de contingencia se asume como un arreglo bidimensional de  $f$ -filas por  $c$ -columnas ( $f \times c$ -celdas); en el ejemplo,  $f=3$  y  $c=4$ . Las entradas en las celdas de la tabla son las frecuencias, o sea, el número de individuos que se asignan a una modalidad de desempeño estandarizado (filas) y por competencias (columnas); así, por ejemplo, en la modalidad *medio-estandarizada e inferencial-de-competencias* se ubican 2394 alumnos. En general, se nota con " $n_{ij}$ " a la frecuencia de la  $i$ -ésima modalidad de la variable fila y la  $j$ -ésima de la variable columna; así, cuando  $i=2$  y  $j=3$  (fila 2 y columna 3), la frecuencia de esta celda es:  $n_{23}=2394$ .

El total por fila o por columna está formado por las frecuencias marginales, y se notan por  $n_{i.}$  (donde el punto señala que se suman columnas dentro de la misma fila, " $i$ ") y  $n_{.j}$  (donde el punto señala que se suman filas dentro de la misma columna, " $j$ "), respectivamente. En la tabla anterior, el número de niños asignados a la categoría alto-estandarizada es:  $n_{3.}=1379$  y es la frecuencia marginal para esta fila,

	Nivel 0	Literal	Inferencial	Crítico intertext.	Total
Alto	11	+ 340	+ 841	+ 187	= 1379

mientras que la frecuencia marginal para la columna de desempeño literal es  $n_{.2}=1573$ .

	Literal
Bajo	278
Medio	+ 955
Alto	340
Total	1573

La suma de las frecuencias por celda es igual a la suma de las marginales e igual al número total de individuos seleccionados y clasificados; se nota por “N”, en este caso  $N=6447$  y corresponde al número total de niños de los grados 3° y 4° de educación básica primaria evaluados en el marco de la evaluación de impacto del PUEBP.

La notación general, para una tabla de contingencia de  $f$ -filas y  $c$ -columnas, es la siguiente:

- La frecuencia o conteo de la  $i$ -ésima modalidad de la variable fila y la modalidad  $j$ -ésima de la variable columna, se escribe como  $n_{ij}$
- El total de observaciones en la  $i$ -ésima modalidad de la variable fila se nota por  $n_{i.}$ ; es decir:

$$[\text{Ec.1}] \quad n_{i.} = n_{i1} + n_{i2} + \dots + n_{ic}, \text{ donde las filas } i \text{ varían conforme a: } i=1, 2, \dots, f$$

- El total de observaciones en la  $j$ -ésima modalidad de la variable columna se nota por  $n_{.j}$ ; es decir:

$$[\text{Ec.2}] \quad n_{.j} = n_{1j} + n_{2j} + \dots + n_{fj}, \text{ donde las columnas } j \text{ varían conforme a: } j=1, 2, \dots, c$$

- El número total de observaciones (individuos) en la muestra se escribe con  $N$ , y es igual a la suma de los márgenes fila o columna (o a la suma de todas las celdas  $i$ - $j$ ); es decir:

$$[\text{Ec.3}] \quad N = n_{1.} + n_{2.} + \dots + n_{f.} = n_{.1} + n_{.2} + \dots + n_{.c}$$

Las frecuencias pueden ser transformadas en *proporciones* o *porcentajes*. Un primer porcentaje, para una tabla de contingencia, se obtiene dividiendo cada frecuencia  $n_{ij}$  por el número total de observaciones  $N$ ; este porcentaje se escribe como “ $f_{ij}$ ”; es decir:

$$[\text{Ec.4}] \quad f_{ij} = (n_{ij}/N) \times 100$$

$f_{ij}$  corresponde a la proporción o porcentaje de individuos que tiene los atributos “ $i$ ” y “ $j$ ”; o sea, los de las dos variables.

Un segundo porcentaje se obtiene al dividir cada frecuencia  $n_{ij}$  por la respectiva frecuencia marginal fila  $n_{i.}$ ; esto es:

$$[\text{Ec.5}] \quad f_{ji} = (n_{ij}/n_{i.}) \times 100$$



La cantidad  $f_{ji}$  es la proporción de individuos de cada celda, respecto al total de la fila “ $i$ ”. Donde “ $ji$ ” (se lee: “ $j$  dado  $i$ ”) significa “la columna  $j$ , a condición de estar en la fila  $i$ ”, es decir, se deja fija la fila  $i$  y se recorren sus columnas. Estas frecuencias corresponden al *perfil fila*. La tabla siguiente contiene el perfil de desempeño por competencias de los estudiantes cuyo desempeño estandarizado es *alto*:

		Nivel 0	Literal	Inferencial	Crítico intertextual	Total
Alto	$n_{3j}$	11	340	841	187	1379
	$f_{j 3}$	$f_{1 3} = 11/1379 \times 100 = 0.80$	$f_{2 3} = 340/1379 \times 100 = 24.66$	$f_{3 3} = 841/1379 \times 100 = 60.99$	$f_{4 3} = 187/1379 \times 100 = 13.56$	100%

Un tercer porcentaje se obtiene al dividir cada frecuencia  $n_{ij}$  por la respectiva frecuencia marginal columna  $n_{.j}$ ; esto es:

$$[Ec.6] \quad f_{ij} = n_{ij} / n_{.j} \times 100$$

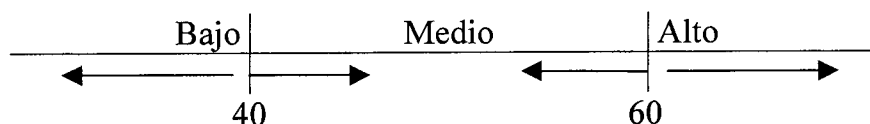
La cantidad  $f_{ij}$  es la proporción de individuos de cada celda, respecto al total de la columna “ $j$ ”. Donde “ $ij$ ” (se lee: “ $i$  dado  $j$ ”) significa “la fila  $i$ , a condición de estar en la columna  $j$ ”, es decir, se deja fija la columna  $j$  y se recorren sus filas. Estas frecuencias corresponden al *perfil columna*. La tabla siguiente contiene el perfil de desempeño estandarizado de los estudiantes cuyo desempeño por competencias es *literal*:

		Literal	
		$n_{i2}$	$f_{i 2}$
Bajo	278	$f_{1 2} = 278 / 1573 \times 100 = 17.67$	
Medio	955	$f_{2 2} = 955 / 1573 \times 100 = 60.71$	
Alto	340	$f_{3 2} = 340 / 1573 \times 100 = 21.61$	
Total	1573	= 100%	

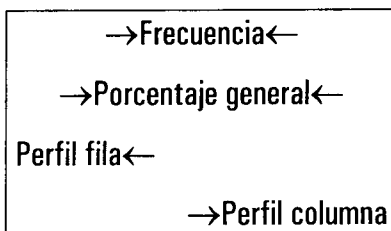
En consecuencia, se pueden obtener tres tipos de tablas adicionales: la primera hace referencia al porcentaje de cada celda con relación al número total de individuos  $N$ ; la segunda, al porcentaje de cada celda respecto al total de la respectiva fila (éste se conoce como el perfil fila); y el tercero, al porcentaje de cada celda con relación al total de la respectiva columna (se conoce como el perfil columna).

Es importante resaltar que los datos que se presentan en cualquiera de las últimas formas, originalmente proceden de conteos o frecuencias de variables categóricas más que de mediciones continuas. Es más, variables continuas pueden transformarse

a variables categóricas mediante la asignación de intervalos sobre su escala; así, por ejemplo, el puntaje estandarizado de la prueba cerrada, que es una variable continua, se transformó en una variable categorizada al definir como “bajo” al puntaje menor que la media menos una desviación estándar (40), “medio” al puntaje ubicado entre la media y más o menos una desviación estándar (entre 40 y 60) y “alto” al que está por encima de la media más una desviación estándar (60):



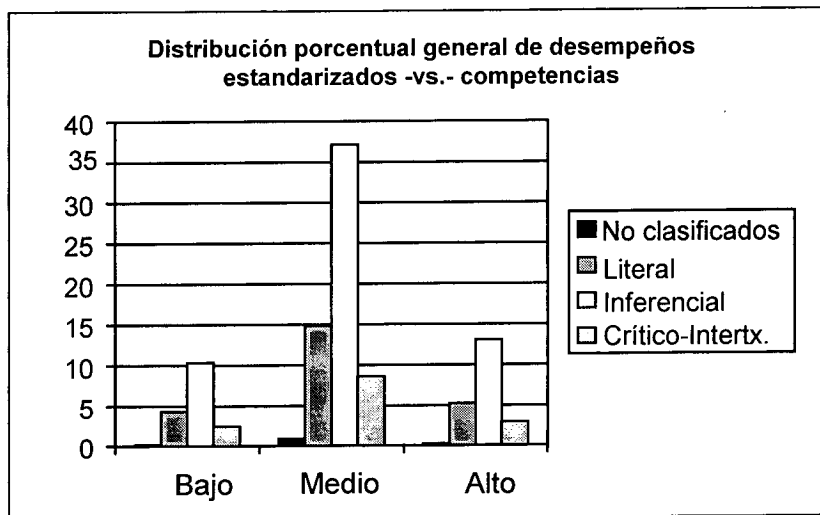
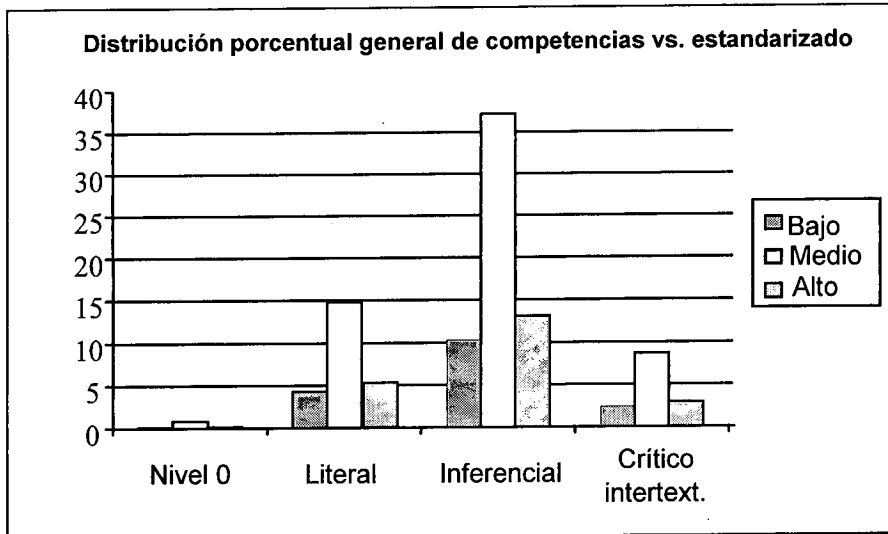
La tabla siguiente condensa los cuatro tipos de tablas (frecuencia, porcentaje general, perfil-fila y perfil-columna). En cada celda, los datos están organizados de la siguiente manera:



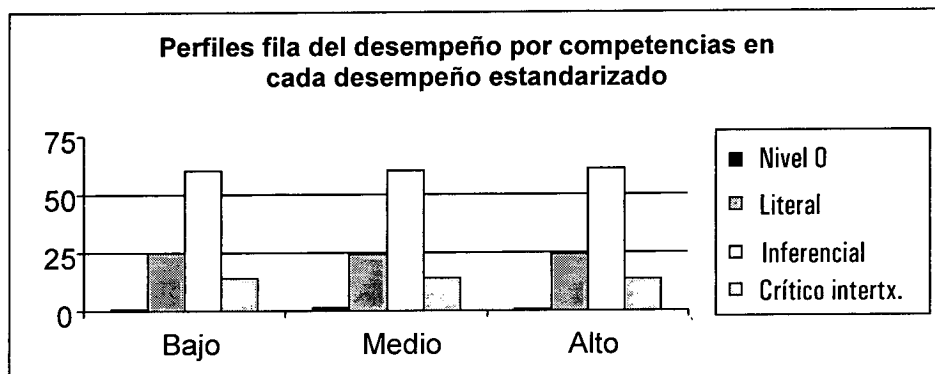
		Desempeño por Competencias				Total Fila
		Nivel 0	Literal	Inferencial	Crítico inter- tertext.	
Desempeño Estandarizado	Bajo	8 0.12 0.72 10.96	278 4.31 25.07 17.67	668 10.36 60.23 17.12	155 2.40 13.98 17.26	1109 17.20
	Medio	54 0.84 1.36 73.97	955 14.81 24.12 60.71	2394 37.13 60.47 61.34	556 8.62 14.04 61.92	3959 61.41
	Alto	11 0.17 0.80 15.07	340 5.27 24.66 21.61	841 13.04 60.99 21.55	187 2.90 13.56 20.82	1379 21.39
Total Columna		73 1.13	1573 24.40	3903 60.54	898 13.93	6447 100.00

Cada una de estas tablas se puede expresar gráficamente mediante un histograma. A continuación se presentan los histogramas para el porcentaje general: el primero

muestra el desempeño por competencias en relación con el desempeño estandarizado; el segundo muestra el desempeño estandarizado en relación con el desempeño por competencias.



La gráfica siguiente muestra los perfiles condicionados a los desempeños estandarizados (perfiles fila):



Se advierte una alta similitud entre los perfiles fila, son isomorfos; esto permite aventurar la hipótesis de no asociación entre las variables fila y columna, porque al cambiar de una modalidad a otra del desempeño estandarizado, no provoca una alteración considerable en los porcentajes de las modalidades ligadas al desempeño por competencia; si hubiera asociación, al desplazarse de una modalidad a otra de una variable, se observaría algún patrón de cambio en los porcentajes relacionados con las otras modalidades del otro desempeño. Por ejemplo, al cambiar de bajo a alto, se esperarían un cambio progresivo ascendente en el desempeño Crítico intertextual, pues se supone que es el desempeño más complejo. En otras palabras, si hubiera asociación positiva, se esperaría que la mayoría de los desempeños altos quedara altamente concentrada en los desempeños crítico-intertextuales, que los desempeños bajos quedaran altamente concentrados en los desempeños literales, etc.

De esta manera, se puede afirmar que los desempeños en cada una de la variables son independientes; así, el desempeño clasificado como bajo, medio o alto no da cuenta del desempeño literal, inferencial, crítico intertextual (incluso del nivel 0). Esta última afirmación se sustentará, más formalmente, en § 2.1.4 y 2.1.5.

### 2.1.3 Independencia de la clasificación

Una vez que se han dispuesto los datos en una tabla de contingencia, la pregunta más importante, en general, es si las variables cualitativas que conforman la tabla de contingencia *son independientes o no*. Veamos qué encierra esta independencia: en el ejemplo anterior se ve que si la forma como se clasifican los alumnos mediante el criterio de la estandarización es independiente de su clasificación mediante las competencias, entonces la proporción (o porcentaje) de los que se clasifican en las categorías “Nivel 0”, “literal”, “inferencial” y “crítico intertextual” es la misma para cada uno de los niveles “bajo”, “medio” y “alto”; es decir, que los respectivos histogramas deben ser geoméricamente congruentes, tal como es el caso de las variables estandarizada y de competencias. Si estas proporciones difieren, esto puede ser *aso-*

ciado más con algunas de las modalidades de la otra variable; por supuesto que puede tenerse algún grado de diferencia en las proporciones debido a la aleatoriedad de la muestra y a otras causas no identificables. Lo que se necesita es saber si se puede o no asegurar si las diferencias observadas entre las dos proporciones es demasiado grande como para que sea atribuida a tales causas. Por esto, en §2.1.4 se mostrará un contraste estadístico para esta afirmación.

Supóngase que en la población de la que se extrae la muestra, la probabilidad<sup>26</sup> (“p”) de que una observación pertenezca a la  $i$ -ésima modalidad de la variable fila y a la  $j$ -ésima modalidad de la variable columna es  $p_{ij}$ ; en consecuencia, el número de observaciones que se espera “caigan” en la celda  $i-j$  (“ $F_{ij}$ ”) de la tabla de acuerdo con las  $N$  observaciones muestreadas es dada por:

$$[\text{Ec.7}] \quad F_{ij} = Np_{ij}$$

Ahora,  $p_i$  representa la probabilidad, en la población, de que una observación pertenezca a la  $i$ -ésima categoría de la variable fila; de manera semejante,  $p_j$  representa la probabilidad correspondiente para la  $j$ -ésima categoría de la variable columna. Entonces, la independencia entre las dos variables, en la población, se garantiza por el cumplimiento de la siguiente igualdad [Ato y López, 1996:§2.1.4]:

$$[\text{Ec.8}] \quad p_{ij} = p_i p_j$$

Es decir, que las dos variables son independientes, si y solo si, la probabilidad de que una observación “caiga” en la celda  $i-j$  es igual al producto de las respectivas probabilidades marginales. Desde el punto de vista de la frecuencia esperada en la tabla de contingencia [Ec.7], por la independencia [Ec.8], al remplazar  $p_{ij}$  se tiene:

$$[\text{Ec.9}] \quad F_{ij} = Np_i p_j$$

Como los valores de  $p_{ij}$ ,  $p_i$  y  $p_j$  son valores desconocidos en la población sobre la cual precisamente se indaga en este estudio, la verificación de la igualdad anterior no es fácil de realizar, razón por la que “debemos” conformarnos con el contraste a través de los datos muestrales. De esta manera, los valores de  $p_{ij}$ ,  $p_i$  y  $p_j$  deben estimarse (calcularse) a través de la información contenida en la muestra. Desde los datos muestrales disponibles, los estimadores para cada una de estas cantidades desconocidas son:

$$[\text{Ec.10}] \quad \hat{p}_{ij} = n_{ij}/N, \hat{p}_i = n_i /N \text{ y } \hat{p}_j = n_j /N$$

---

<sup>26</sup> Número entre 0.0 y 1.0, que indica la proporción de veces que un evento, bajo las mismas circunstancias, se puede presentar.

donde las “ $\hat{p}$ ” representan la probabilidad estimada mediante los datos muestrales. Así, la frecuencia esperada estimada, notada por “ $E_{ij}$ ” (de acuerdo con [Ec.9] y [Ec.10]), es igual a

$$\begin{aligned} \text{[Ec.11]} \quad E_{ij} &= N \hat{p}_i \hat{p}_j \\ &= N(n_{i.}/N)(n_{.j}/N) = (n_{i.} n_{.j}) / N \end{aligned}$$

Es decir, bajo independencia, la frecuencia esperada en cada celda es igual al producto de la frecuencia de las márgenes dividido por el número total de individuos de la tabla. Cuando las dos variables son independientes, la frecuencia estimada usando [Ec.11] diferiría de las frecuencias observadas  $n_{ij}$  en cantidades muy pequeñas y atribuibles a cambios en las variables únicamente. Sin embargo, si las dos variables no son independientes se esperaría que se presentasen diferencias grandes entre las frecuencias observadas y las esperadas. Esta es la base para una medida que registre la independencia o no entre dos variables de tipo categórico, la cual se describe a continuación.

#### 2.1.4 Prueba Ji-cuadrado

Para verificar la independencia entre dos variables se requiere averiguar por la veracidad de la hipótesis que de este supuesto se deriva, es decir por:

$$\text{[Ec.12]} \quad p_{ij} = p_i p_j$$

En general, esta hipótesis se referirá como la *hipótesis nula* (independencia). Como se anotó, la verificación —que no la prueba— de esta hipótesis debería basarse en la diferencia entre los valores estimados de las frecuencias que se esperan bajo la hipótesis nula  $E_{ij}$  y las frecuencias observadas en la muestra  $n_{ij}$ . Tal prueba, sugerida por Pearson (1904)<sup>27</sup>, emplea la estadística  $\chi^2$  (Ji-cuadrado), dada por

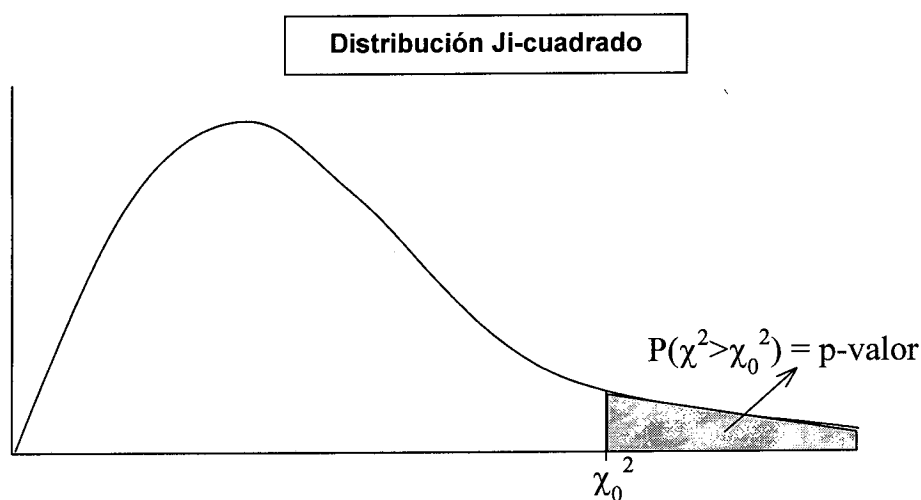
$$\begin{aligned} \text{[Ec.13]} \quad \chi_0^2 &= \sum_{i=1}^f \sum_{j=1}^c (\text{frec. observada} - \text{frec. esperada})^2 / \text{frec. esperada} \\ \chi_0^2 &= \sum_{i=1}^f \sum_{j=1}^c (n_{ij} - E_{ij})^2 / E_{ij} \end{aligned}$$

La notación  $\chi_0^2$  resalta que la estadística se calcula con los datos observados; la doble suma indica que se deben sumar todas las celdas, lo cual se consigue “desplazándose” por filas o por columnas. Se observa que la magnitud de esta estadística

---

<sup>27</sup> Citado por Everitt [1994:7].

depende de los valores que asuman las diferencias  $(n_{ij}-E_{ij})^2$ ; nótese que la fórmula tiene implicada la distancia al cuadrado (euclidiana) entre las frecuencias observadas y las esperadas; de tal manera, la fórmula registra qué tan “distante” está lo observado de lo esperado bajo la hipótesis de independencia. Si las dos variables son independientes, estas diferencias deberían ser menores que en el caso contrario (dependencia); en consecuencia,  $\chi_0^2$  será más pequeña cuando los datos “apoyan” la hipótesis nula que cuando no. Por tanto, se necesita de un método mediante el cual se pueda o no rechazar la hipótesis nula. Tal procedimiento se basa en la *distribución de probabilidades* de la estadística  $\chi^2$  bajo el supuesto de que la hipótesis de independencia es cierta. Para cada valor que tome la estadística  $\chi_0^2$  se tiene una probabilidad de que haya valores mayores o iguales (o menores) que tal valor, de manera que la distribución de probabilidades de  $\chi^2$  corresponde a las probabilidades asociadas con los diferentes valores de  $\chi^2$ . Por construcción, la estadística Ji-cuadrado toma valores positivos, pues es una suma de cantidades elevadas al cuadrado. La siguiente gráfica ilustra la distribución de probabilidades asociada con la estadística Ji-cuadrado:



La decisión de no rechazar<sup>28</sup> o de rechazar la hipótesis nula de independencia entre las variables, se basa en la probabilidad de los valores obtenidos para  $\chi^2$ ; pues valores con baja probabilidad (de ser mayores o iguales que éste) permiten el rechazo de la hipótesis nula y, en caso contrario, su no rechazo.

Como se nota en la gráfica, valores sobre el eje horizontal ( $\chi^2$ ) con baja probabilidad a la izquierda tienen probabilidades altas a la derecha; y, recíprocamente, valores

<sup>28</sup> Se habla de *no rechazar*, pues aquí lo consideraremos distinto de *aceptar*.

con alta probabilidad a su izquierda, tienen baja probabilidad a la derecha (como el área sombreada<sup>29</sup> en la gráfica). Como se explicó, valores de  $\chi^2$  con baja probabilidad (a la derecha), provocan el rechazo de la hipótesis nula; pues esto dice que la probabilidad de que *no existan discrepancias* entre los valores observados y los esperados, bajo la hipótesis de independencia, es baja, pues como se señaló, la magnitud de la estadística  $\chi^2$  depende de los valores que asuman las diferencias  $(n_{ij}-E_{ij})^2$ : si las dos variables son independientes, estas diferencias deberían ser menores que en el caso contrario (dependencia).

Este es el procedimiento usual para derivar la llamada *significancia de la prueba*. Se entiende por “significancia” la probabilidad de rechazar la hipótesis nula (no asociación) cuando ésta es verdadera (asociación). En general, se consideran como valores de significancia “bajos”, a valores de probabilidad menores que 0.05 (5%); valores por debajo de 0.01 (1%) se consideran como extremadamente bajos, éstos son referidos como *niveles de significancia* de la prueba estadística. Una de las inquietudes con las que se encuentran los usuarios de estos métodos estadísticos es respecto al umbral a partir del cual las conclusiones derivadas adquieren un carácter “objetivo” y aceptable. Es decir, ¿qué tan grande debe ser el nivel de significancia para ser considerado como bajo, aceptable o alto? Esta respuesta *no* la tiene la estadística, pues depende del ámbito de investigación, de los intereses que sobre el objeto en cuestión se planteen, e incluso, del grado de ignorancia admitido sobre el objeto. En la investigación social usualmente se consideran significativos estadísticas con un p-valor hasta el 10% (0.100).

Respecto a los valores obtenidos para las distintas asociaciones, se debe tener la prevención de su valor, pues el alto tamaño de muestra (6447, en el caso del presente estudio) coadyuva en mostrar asociaciones altamente significativas, situación que valida el ejercicio de tipo exploratorio que se hace en este texto con los datos. Por esto, en el presente estudio se asumirán como altamente significativos valores de probabilidad inferiores a 1% (<0.010); y como moderadamente significativos, valores entre 1% y 10% (de 0.010 a 0.100).

### **2.1.5 Distribución Ji-Cuadrado**

La distribución Ji-cuadrado se define como la suma de los cuadrados de distribuciones normales estándar independientes; es decir, distribuciones con media cero y varianza uno. La forma de la distribución Ji-cuadrado depende del número de su-

---

<sup>29</sup> Esta área, que corresponde a la probabilidad de obtener un valor por encima de  $\chi^2_0$ , se conoce como p-valor.



mandos libres involucrados en la suma. Por ejemplo, una Ji-cuadrado conformada con  $k$ -variables normales estándar independientes, escrita como

$$\chi^2 = Z_1^2 + Z_2^2 + \dots + Z_k^2,$$

tiene una distribución Ji-cuadrado que depende sólo de  $k$ . En general, *el número de sumandos libres* asociados con una distribución Ji-cuadrado se conoce como los *grados de libertad* de la distribución. En la siguiente condición: “tres números cuya suma sea 20”, una vez que se han establecido dos sumandos (por ejemplo, 7 y 5), no hay libertad alguna para elegir el tercero (8); entonces, hay dos grados de libertad en dicha suma. Para el caso de tablas de contingencia de  $f$  filas y  $c$  columnas, el número de grados de libertad de la distribución es  $(f-1)(c-1)$ ; así, por ejemplo, para la tabla anterior [§2.1.2] los grados de libertad son  $(3-1)(4-1)=6$ . La mayoría de los textos de estadística básica contienen tablas para los valores de probabilidad asociados con un distribución Ji-cuadrado, la cual queda determinada por el número de grados de libertad.

Para algunos grados de libertad y varios niveles de significancia se han elaborado tablas estadísticas en donde se encuentran los respectivos valores de  $\chi^2$ ; de esta manera, se obtiene el valor de la respectiva distribución  $\chi^2$  (determinada por los grados de libertad) que deja un área a la derecha igual a la significancia asumida. La prueba estadística de la hipótesis de independencia se puede desarrollar también mediante la comparación de la estadística  $\chi_0^2$  (calculada con los datos muestrales) con los valores tabulados para la distribución Ji-cuadrado. De manera que si el valor  $\chi_0^2$  calculado mediante los datos es mayor que el dispuesto en una tabla, para un número de grados de libertad y nivel de significancia establecidos, entonces se rechaza la hipótesis de independencia; en caso contrario, no se rechaza. Dada la actual disponibilidad de programas en estadística para computador (paquetes), el uso de las tablas Ji-cuadrado ha sido desplazado, pues la decisión de rechazar o no la hipótesis de independencia se hace de acuerdo con el p-valor suministrado por la salida del paquete estadístico.

En el presente texto, las tablas de contingencia contendrán, en su mayoría los perfiles ligados a los factores asociables, y el respectivo p-valor, de manera que no se tendrán tablas de contingencia en el sentido descrito, sino tablas de perfiles<sup>30</sup>. Además, como dijimos atrás [cf. §1.3], se observará principalmente la proporción de estudiantes que alcanzan las categorías extremas de los desempeños, en función de las modalidades de los respectivos factores asociables.

---

<sup>30</sup> Los interesados en obtener las tablas de contingencia completas, pueden consultar el anexo [en CD] a esta investigación.

• *Un ejemplo*

Para el caso que se ha venido tratando se desarrollan los cálculos conducentes a obtener el valor de la estadística  $\chi^2$ , con el cual se verifica la hipótesis nula: “la clasificación de los estudiantes a través de la estandarización es independiente de la clasificación por competencias”.

Los valores para las frecuencias esperadas  $E_{ij}$  se calculan conforme a la fórmula [Ec.11], así para la celda de la fila 1 y columna 1 (bajo y Nivel 0), el valor de la frecuencia esperada  $E_{11}$  es dado por:

$$E_{11} = (n_{1.} \cdot n_{.1}) / N$$

$$E_{11} = (1109 \times 73) / 6447$$

$$E_{11} = 12.55731$$

Los demás valores  $E_{ij}$  ( $E_{12}, E_{13}, E_{14}, \dots, E_{34}$ ) se calculan en forma semejante. La siguiente tabla contiene los resultados.

		Desempeño por competencias-C			
		Nivel 0	Literal	Inferencial	Crítico intertextual
Desempeño estandarizado	Bajo	$(E_{11})$ 12.5573	$(E_{12})$ 270.5843	$(E_{13})$ 671.3862	$(E_{14})$ 154.4722
	Medio	$(E_{21})$ 44.8281	$(E_{22})$ 965.9542	$(E_{23})$ 2396.77	$(E_{24})$ 551.4475
	Alto	$(E_{31})$ 15.6145	$(E_{32})$ 336.4615	$(E_{33})$ 834.8436	$(E_{34})$ 192.0803

El valor de la estadística  $\chi^2$  se obtiene de calcular cada uno de los 12 valores de  $(n_{ij} - E_{ij})^2 / E_{ij}$  y luego sumarlos, estos cálculos se muestran en forma condensada a continuación:

$$\chi^2 = (8-12.5573)^2/12.5573 + (278-270.5843)^2/270.5843 + \dots + (187-192.0803)^2/192.0803$$

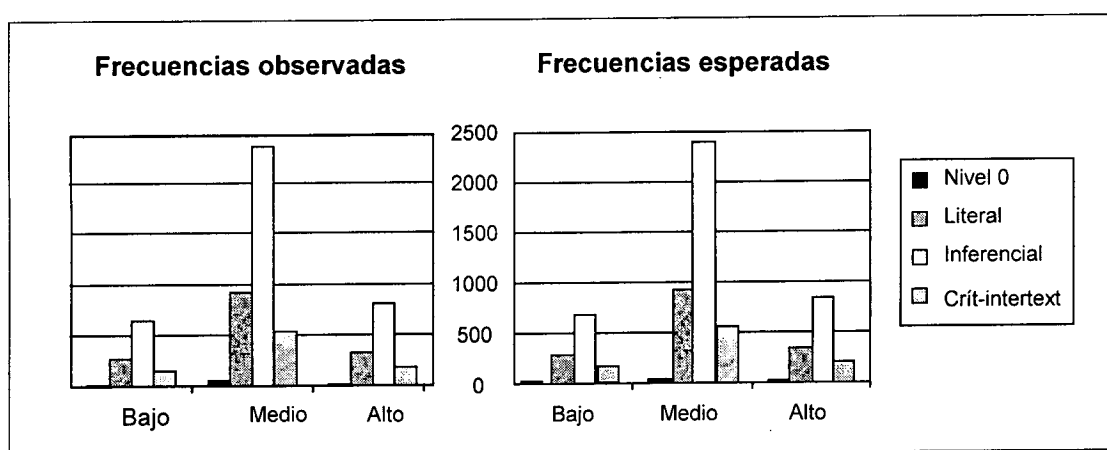
$$\chi^2 = 1.65394502 + 0.203236 + \dots + 0.1343$$

$$\chi^2 = 5.49835828$$

El valor de la variable  $\chi^2$ , que se obtiene desde una tabla Ji-cuadrado con 6 grados de libertad y un nivel de significación de 0.05 (5%), es 12.592; en consecuencia, como

el valor calculado para  $\chi^2 = 5.49835828$  es menor que el valor de la tabla, no se rechaza la hipótesis de independencia entre las dos variables; es decir se puede afirmar, con un nivel de incertidumbre del 5%, que las variables son independientes (no están asociadas).

Un procedimiento alternativo —el cual se seguirá a lo largo del presente estudio— para apoyar la decisión estadística es el de la probabilidad de encontrar un valor mayor o igual que  $\chi^2 = 5.49835828$ , en símbolos,  $P(\chi^2 \geq 5.49835828) = p$ , éste es el p-valor, que, de acuerdo con el procedimiento FREQ del paquete SAS<sup>31</sup>, es = 0.482, de manera que el nivel de incertidumbre para rechazar la hipótesis de independencia es alto, de donde se llega a la misma conclusión anterior.



Se ve que las diferencias entre las frecuencias observadas y las frecuencias esperadas son bajas, lo cual se registró en la estadística Ji-cuadrado. En resumen, la clasificación hecha respecto a los desempeños en la prueba de lenguaje a través de la estandarización es independiente de la clasificación hecha por competencias.

## 2.2 Metodologías estadísticas alternas

### 2.2.1 Modelos jerárquicos

De acuerdo con la estructura del actual sistema escolar, en un sentido conjuntista se tiene un arreglo jerárquico de las unidades de observación: los alumnos —que se pueden considerar como la unidad de observación primaria— conforman grados (y comparten profesores, aulas, saberes, entre otros); los grados, a su vez, están conte-

<sup>31</sup> Statistical analysis system.

nidos en una institución educativa; la cual a su vez es parte de una localidad, una zona un municipio, etc. Como puede advertirse, se dispone de unidades de observación en extensión y complejidad distintas.

Una técnica estadística usada frecuentemente en los estudios de factores asociables a los desempeños de los estudiantes, ha sido la *regresión clásica*. Con esta técnica se construye un modelo donde los resultados de la variable logro se muestran en función de otras variables que ponen en juego a un mismo nivel características de los estudiantes, de los docentes, del plantel, de la localidad, etc. Otra alternativa, menos arrasadora, construye modelos de regresión para cada uno de los niveles de las unidades de observación, de manera que, por ejemplo, se propone un modelo de regresión para planteles mediante la agregación de las variables individuales de los estudiantes en los respectivos colegios o planteles.

Uno de los inconvenientes del primer caso es el incumplimiento del supuesto de independencia entre las observaciones, pues los alumnos de la misma aula, grado o plantel, por ejemplo, tendrán los mismos valores en cada una de las variables que representan a estas unidades. El problema con la segunda alternativa es que se pierde información sobre la variación dentro de las unidades (aula, grado, plantel, etc.), de manera que la asociación (colinealidad) entre las variables agregadas es más fuerte y diferente que entre las variables sin agrupar. De todas formas, la regresión clásica, así considerada, no permite explicar el logro en términos de las características del niño, del grado o aula, y del plantel, por ejemplo.

En educación, la estructura de la información corresponde a una conformación jerárquica de las unidades de observación. Así, los niños están contenidos en cursos o aulas las cuales a la vez están contenidas en las instituciones. Esto hace que las variables observadas expliquen de manera conjunta y distinta los resultados con relación a una variable de interés.

Se trata de modelar algunas variables en función de un conjunto de variables explicativas, las cuales están ligadas a las respectivas unidades de cada nivel. Este tipo de modelamiento corresponde a los llamados *modelos jerárquicos o multinivel*, con los cuales se podrá dar cuenta de los resultados, atribuibles a variables de los niños, a variables de los grados, de las instituciones, de las localidades<sup>32</sup>, etc. [cf. Goldstein, 1987].

El siguiente esquema, muestra cómo los desempeños de los estudiantes tienen una variabilidad (información) atribuible a la variabilidad de los mismos estudiantes, a la

---

<sup>32</sup> Esta fue la metodología aplicada en el estudio de factores asociables en el DC [cf. SED-Corpoeducación-DNP, 1999].

variabilidad de los grados, a la variabilidad de los planteles<sup>33</sup> y a lo que quede por fuera de los anteriores (ruido):

$$\text{Var. (desempeños)} = \text{Var. (estudiante)} + \text{Var. (Grado)} + \text{Var. (Plantel)} + \text{Var. (no explicada)}$$

La descomposición anterior permite calcular la “importancia” de cada efecto, el cual expresa la proporción de la variabilidad de cada nivel con relación a la variabilidad total. Como se anotó, el problema es determinar el tipo de asociación existente entre las diferentes unidades de observación que hace que éstas tengan una jerarquía específica. No obstante, es una herramienta útil para apoyar la asociabilidad de un factor con los desempeños (resultados, logros, competencias...) de los estudiantes.

### *2.2.2 Análisis de correspondencias*

Es una técnica para explorar la posible relación entre las variables que se dispongan en una tabla de contingencia (bidimensional o múltiple). En este contexto, una tabla de contingencia es una tabla cuyas celdas contienen alguna medida de correspondencia (asociación, similitud, afinidad o interacción) entre las variables. En §2.1 se muestra una forma de detectar y medir la fuerza de esta correspondencia a través de la estadística ji-cuadrado.

Con esta técnica se procede a encontrar un número pequeño de dimensiones (a la manera de factores, componentes o ejes principales) que expliquen las asociaciones presentes en las variables dispuestas en la tabla, de tal modo que la primera dimensión explique la mayor parte de la asociación total entre las filas y las columnas, la segunda dimensión explique la mayor parte de la asociación no explicada por la primera, y así sucesivamente con el resto de las dimensiones.

Cada una de las dimensiones (eje principal) es una suma ponderada de las categorías de las variables puestas para el análisis de correspondencias. Las puntuaciones (ponderaciones) se consideran como el grado de importancia de cada una de las categorías de una variable en la definición de la dimensión; de manera que mediante la magnitud de tales ponderaciones se puede “acuñar” un nombre genérico para la respectiva dimensión. Este ejercicio coadyuva a la identificación de factores posiblemente influyentes en el desempeño de los estudiantes.

---

<sup>33</sup> Se conoce como modelo a tres niveles, éstos son: estudiantes, grado y plantel.

Una vez que se han encontrado los “nuevos” ejes o dimensiones, se pueden ubicar, por ejemplo en un plano determinado por los dos primeros ejes, cada una de las categorías y/o los individuos. Este gráfico o mapa posibilita la exploración de la asociación entre las diferentes variables (por ejemplo, factores y desempeños [cf. Díaz, 1999a]).

### ***2.2.3 Modelos log-lineales***

La mayoría de las investigaciones en las ciencias humanas involucra grandes tablas multidimensionales. En estas tablas se mide un número tal de variables, que la interrelación entre ellas resulta compleja. Frecuentemente, cuando el investigador se enfrenta a tablas de dimensión alta, el tratamiento ordinario que le hace a los datos contenidos en éstas es seleccionar de a dos variables, colapsando las demás. Aunque estos procedimientos pueden ser útiles, tienen al menos los siguientes inconvenientes:

- Se pierde habilidad para estudiar las interacciones entre más de dos variables;
- Las asociaciones marginales entre dos variables es muy distinta de la asociación que podría presentarse entre las mismas variables pero en presencia de otras;
- No se pueden estudiar simultáneamente todos los pares de variables.
- La asociación marginal entre variables puede ser distinta de la asociación entre las mismas variables bajo un factor fijo de una tercera variable<sup>34</sup>.

Cuando se estudian varias mediciones simultáneamente, resulta útil conseguir una descripción parsimoniosa (con el menor número de variables) de los datos a través de un *modelo matemático* que explique, de alguna forma, las observaciones; a esto se le denomina *modelamiento*. El modelamiento consiste en la aplicación de una serie de procesos (estimación, contrastes de simplificación, diagnosis y replanteamiento) con el objeto de conseguir una explicación apropiada del comportamiento de una variable respuesta (datos) a partir de una función ponderada de una o más variables explicativas (modelo). La explicación en general, por fortuna, no suele ser perfecta, y por tanto, es pertinente observar la discrepancia entre los datos y el modelo. A esta diferencia se le denomina error o residual. La siguiente igualdad ilustra el concepto de modelo:

---

<sup>34</sup> Este resultado se conoce como *Paradoja de Simpson* [cf. Díaz, 1999b]. Una ilustración aparece en este estudio en § 3.2.1.

$$\text{Datos} = \text{Modelo} + \text{Error}$$

La igualdad anterior señala que los datos observados, objeto del modelado, son función de un componente estructural (sistemático), representado por algún modelo teórico apropiado y un componente aleatorio que representa la discrepancia o error entre los datos empíricos y el modelo teórico propuesto. Uno de los ejemplos clásicos de modelos estadísticos es el de regresión lineal múltiple, en el cual una variable respuesta (Y) es “explicada” a través de unas variables  $X_1, X_2, \dots, X_p$  mediante el modelo estadístico:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$$

Los  $\beta$  que resulten ser significativamente distintos de cero dan cuenta de la asociabilidad de la respectiva variable X con la variable respuesta Y. En un caso en el que todos los  $\beta$  no son significativamente distintos de cero, lleva a concluir que los factores que las variables representan, no influyen, desde el punto de vista estadístico, sobre la variable respuesta en cuestión.

El propósito del modelamiento estadístico es la búsqueda del modelo más simple (parsimonioso o “simple”, según dice la lingüística) que sea capaz de explicar los datos con el mínimo error posible [Ato y López, 1996:79]. Esto equivale a buscar un modelo que se ajuste adecuadamente a los datos. Para el investigador, modelar un conjunto de datos le implica:

- seleccionar un modelo teórico simple que, en el conocimiento sustantivo del objeto de su investigación, contiene una función ponderada de las variables explicativas;
- estimar los parámetros a través de la información disponible;
- calcular una medida de discrepancia para evaluar el ajuste entre los valores observados y los valores pronosticados por el modelo estimado;
- optar, de acuerdo con la evaluación observada en el punto anterior, por alguna de las siguientes alternativas:

- si el ajuste es adecuado, pero el modelo es complejo, acepta el modelo temporalmente y retorna al comienzo del proceso en busca de un modelo más simple;
- si el ajuste es adecuado, acepta el modelo de manera tentativa, lo interpreta e integra al marco conceptual donde están inscritos los datos;
- si el ajuste no es adecuado, lo rechaza y retorna al comienzo del proceso en busca de un modelo más apropiado (partiendo de la veracidad del marco conceptual asumido por investigador).

Los modelos para los datos de una tablas de contingencia son similares a los usados para datos cuantitativos, por ejemplo, los modelos empleados en el análisis de varianza. Una consecuencia de esta similaridad es que el término “interacción” se usa como una alternativa para describir y cuantificar la asociación entre las variables cualitativas que forman una tabla de contingencia. Cuando se quiera hacer referencia a la interacción entre un par de variables se hablará de “interacción de primer orden”, para la interacción entre tres variables se le dirá “interacción de segundo orden”, y así sucesivamente.

Los principales atractivos del modelamiento de tablas de contingencia son, de un lado, que los modelos suministran una aproximación sistemática al análisis de tablas complejas y, de otro, que proveen una estimación de la magnitud de los efectos de interés, sean ellos principales o de interacción (de asociación) y, en consecuencia, suministran la importancia de diferentes efectos a juzgar [cf. Díaz, 1999b].

#### **2.2.4 Análisis bayesiano**

La mirada desde la estadística clásica se ampara sobre un hecho coyuntural de la realidad (muestra), que sin embargo intenta dar cuenta de una situación más global y cambiante en el tiempo y en el espacio (es como un fotograma contra una película de la que hace parte). Entonces, como una muestra, en general, se corresponde con una mirada sobre una realidad inscrita en un tiempo y un espacio determinados, las conclusiones que desde esta información se puedan obtener sólo apuntan hacia tal incrustación de los datos.

En cambio, el enfoque *bayesiano* permite incluir información inicial sobre los parámetros (análisis secuencial); esta información surge del conocimiento que el investigador tiene a priori, lo cual expresa el estado de conocimiento o ignorancia sobre la realidad que se indaga. Además, cada muestra que se haga sobre la población per-

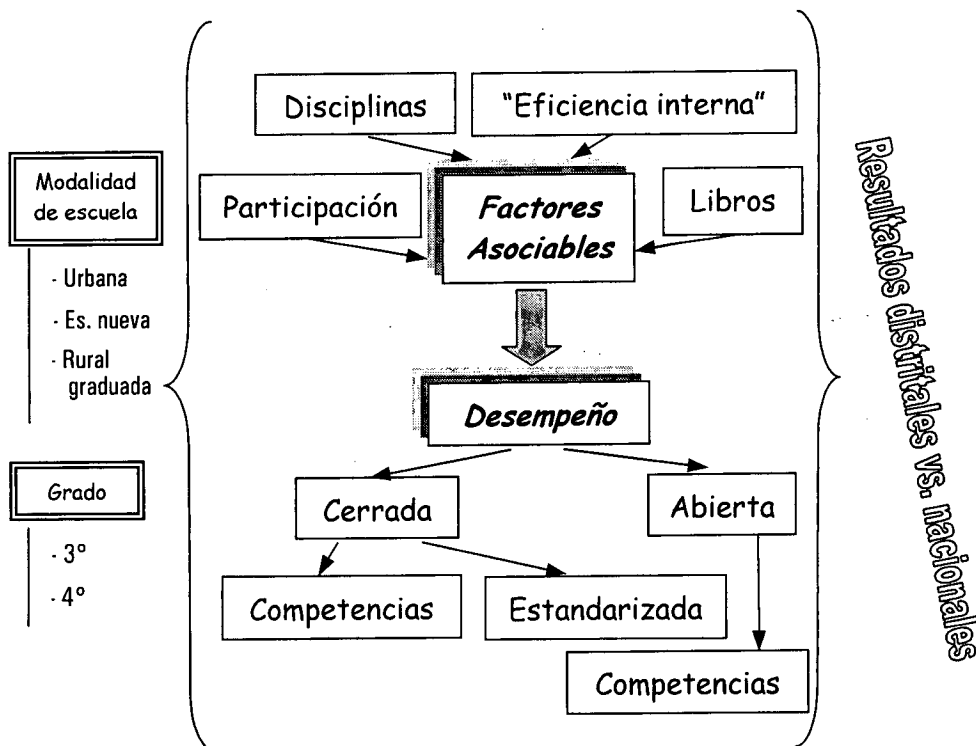


mite actualizar la información disponible; de manera que esta incorporación de conocimiento (a través de los parámetros en juego) produce una aproximación cada vez mayor sobre el objeto construido.

Los procedimientos bayesianos son una alternativa para no hacer “borrón y cuenta nueva” cada que se evalúa la calidad de la educación, mediante la indagación de los desempeños estudiantiles.

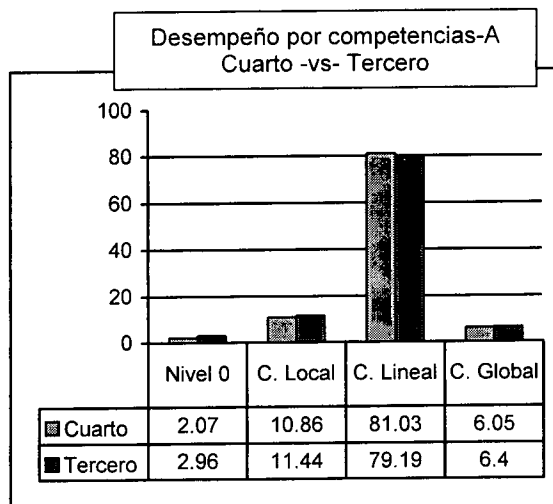
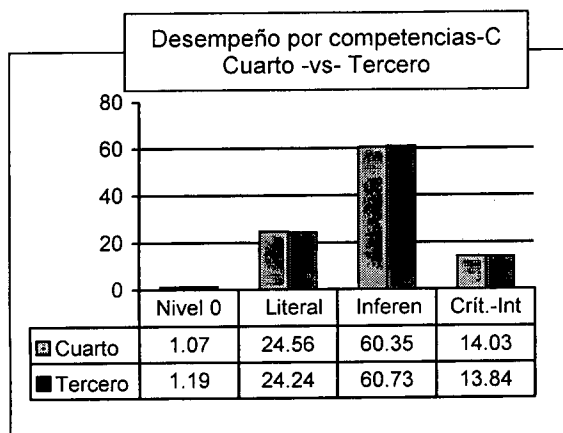
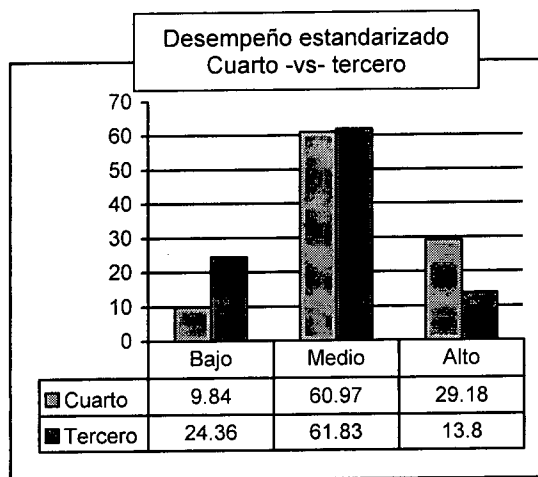
## 2.3 Escenarios de trabajo

Se analizarán los resultados de la prueba de desempeño en el área de lenguaje, realizada en el marco de la evaluación de impacto del PUEBP, desde las tres perspectivas descritas: estandarizada, por competencias-C y por competencias-A; este análisis se confrontará con factores tales como participación de los niños en los diversos comités que la actividad escolar genera, y en la gestión de eventos escolares, libros al alcance de los niños, repitencia, deserción, alumnos por docente, y desempeños en ciencias y matemáticas (“disciplinas”); para cada caso, explicaremos por qué aparece como factor asociable y si, finalmente, lo es —desde nuestro punto de vista—. Además, se analizarán los datos por modalidad de escuela y por grado. Todo esto a nivel nacional. Luego se hará la comparación con las cifras del DC. El siguiente diagrama resume el recorrido del trabajo:



Se comparan los grados (tercero y cuarto) y las modalidades de escuela (urbana, nueva y rural graduada) en el entendido de que en cada caso hay diferencias significativas entre estos escenarios. Esto, que parecería evidente, no lo es necesariamente, de modo que a continuación haremos la comparación de los desempeños en cada caso, con arreglo a cada una de las maneras de interpretarlo. Lo que ahí encontramos amerita el desglose anunciado.

En primer lugar, la diferencia de grado:



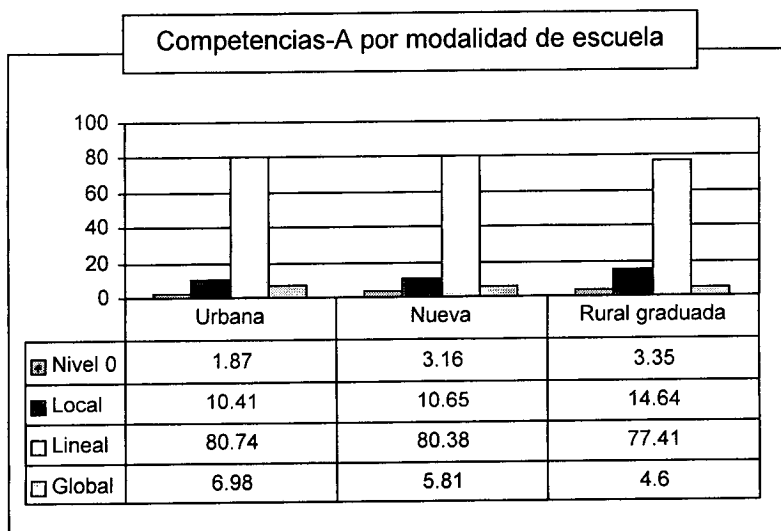
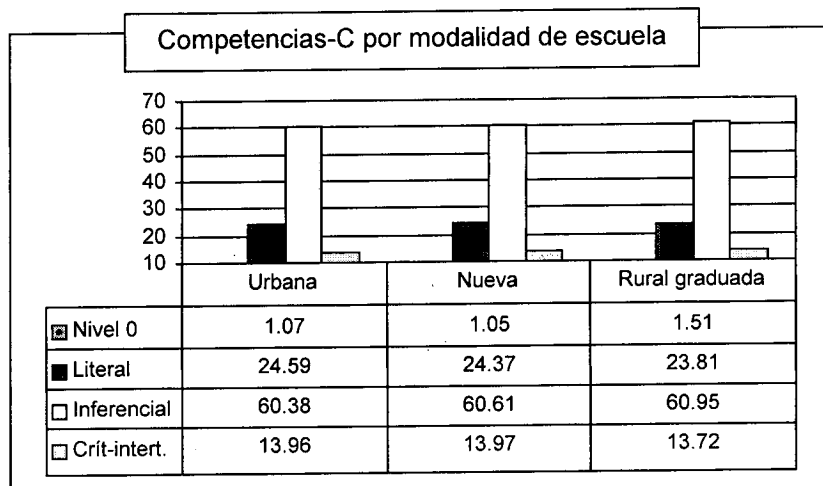
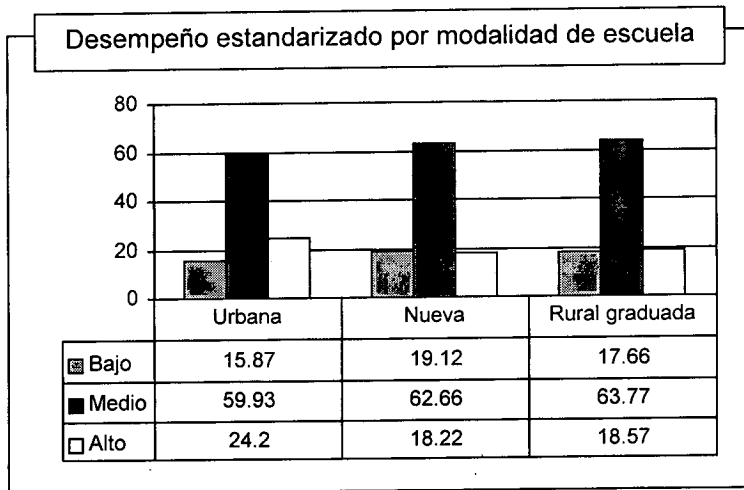
Como puede apreciarse, de cara al desempeño en el área de lenguaje (indagado con la prueba que se aplicó en la evaluación del PUEBP), solamente cuando los desempe-

ños se analizan con las categorías estandarizadas habría clara diferencia entre pertenecer al grado cuarto y pertenecer al grado tercero: los perfiles en los desempeños bajos y altos son muy diferentes, mientras que en el desempeño medio son casi iguales; en otras palabras, los niños de grado cuarto tienden más a avanzar a desempeños altos que los niños de grado tercero, y los niños de grado tercero tienden a quedarse más en desempeños bajos que los de cuarto.

En cambio, con las categorías por competencia, los perfiles —tanto en prueba cerrada como en prueba abierta— son casi idénticos; es decir, frente al desempeño en la prueba de lenguaje, desde las interpretaciones por competencias no hay mayor diferencia entre ser de grado tercero y ser de grado cuarto:

- en el análisis por competencias-C, un 0.32% más de niños de cuarto que de tercero se quedan en nivel literal; un 0.38% más de niños de tercero que de cuarto pasa a nivel inferencial; y un 0.19% más de niños de cuarto que de tercero alcanza el nivel crítico-intertextual. Sin ser significativas estas cifras, su tendencia no es clara, pues los niños de cuarto tenderían tanto a quedarse en el nivel elemental de la comprensión, como a pasar al nivel más complejo;
- en el análisis por competencias-A, un 0.58% más de los niños de tercero que de cuarto se queda en coherencia local; un 1.84% más de niños de cuarto que de tercero pasan a coherencia lineal; y un 0.35% más de niños de tercero que de cuarto avanza hasta coherencia global. Al contrario del caso anterior, aquí los niños de tercero son los que tienden a quedarse en el primer nivel, pero también a alcanzar el máximo nivel (¡más niños de tercero que de cuarto alcanzan el más alto nivel de producción!). No obstante, aquí las diferencias tampoco son significativas.

En segundo lugar, con respecto a la comparación entre modalidades de escuela (urbana, nueva y rural graduada), vale la pena hacer la comparación de sus desempeños, con arreglo a cada una de las maneras de interpretarlo:



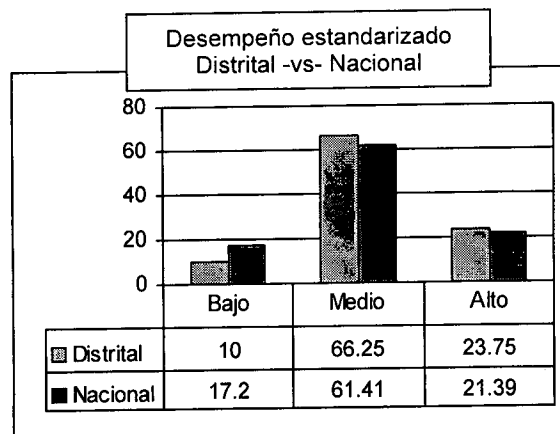
Desde el análisis estandarizado del desempeño, se observa que la escuela urbana es la de mejores desempeños: tiene el menor porcentaje de desempeños bajos y el mayor de desempeños altos. De las escuelas rurales, la escuela graduada tiene mejores desempeños que escuela nueva: tiene menor porcentaje de desempeños bajos y ligeramente superior los desempeños altos.

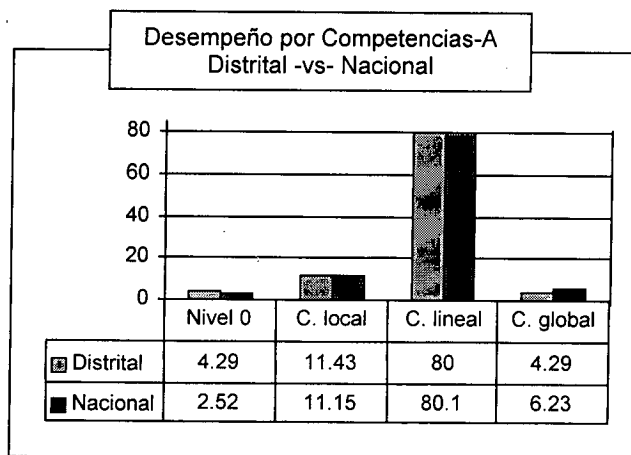
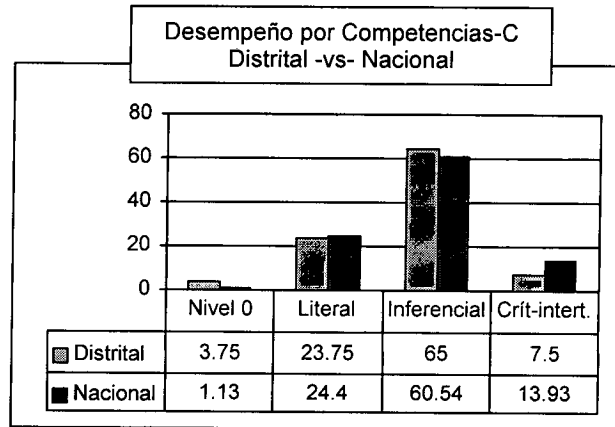
Desde la perspectiva del análisis por competencias-C, los perfiles son casi coincidentes: en ninguna de las categorías del desempeño, la diferencia entre las modalidades de escuela alcanza la unidad porcentual. Es decir, los perfiles por escuela son casi congruentes, que también es equivalente a la no asociación entre desempeño por competencias-C y modalidad de escuela.

Y desde la perspectiva del análisis por competencias-A, en cambio, los perfiles son diferentes: la escuela urbana es en la que mejor se discrimina la prueba, pues es la que menos alumnos sin clasificar deja (Nivel 0); a este respecto, las modalidades rurales son casi idénticas. La escuela urbana es la que más tiende a que sus estudiantes pasen de la coherencia local hacia mejores desempeños, aunque es muy cercano a escuela nueva. Y, de nuevo, la escuela urbana es aquella en la que más quedan estudiantes del nivel más alto. Entre las escuelas rurales, a escuela nueva le va mejor que a la graduada: tiene menor porcentaje de estudiantes en desempeño de coherencia local y mayor porcentaje de estudiantes que alcanzan la coherencia global.

Las interpretaciones del desempeño para las que la diferencia de modalidad es pertinente son la estandarizada y la de competencias-A. En ambos casos, los mejores desempeños están en la escuela urbana. No hay acuerdo, en cambio, en relación con los desempeños más bajos: la primera interpretación ubica allí a escuela nueva, mientras la segunda ubica a la escuela rural graduada.

Y, en tercer lugar, la comparación de los desempeños entre escalas (distrital y nacional), arroja los siguientes resultados:





Desde el análisis estandarizado, el DC tiene mejor desempeño que el promedio nacional: a escala nacional, hay una mayor tendencia a quedarse en desempeño bajo, mientras a escala distrital hay una mayor tendencia a pasar a mejores desempeños; a desempeño alto alcanzan a llegar más niños del distrito que del promedio nacional. Los análisis por competencias también establecen una diferencia entre las escalas en juego, pero en sentido inverso al anterior: tanto en comprensión (parte cerrada) como en producción (parte abierta), más niños del promedio nacional tienden a llegar al nivel alto de cada competencia. En el nivel inferencial, en el caso de las competencias-C, las diferencias están a favor de los niños del DC. En el caso de las competencias-A, las diferencias se establecen en los extremos; el desempeño en coherencia local y lineal es prácticamente idéntico en ambas escalas.

De todas maneras —como se verá más adelante—, estas diferencias de grado, modalidad de escuela y escala se muestran pertinentes cuando se discriminan los estudiantes frente a cada uno de los factores asociables.

**CAPÍTULO II:**  
**ALGUNOS FACTORES ASOCIABLES AL DESEMPEÑO EN LENGUAJE**

### 3. Participación del alumno

El objeto de evaluación de la prueba de lenguaje, en la evaluación de impacto del PUEBP, es la *competencia comunicativa*. Desde esa categoría —como lo veremos—, la participación es fundamental. En este aparte se explorará lo relativo a la participación en comités y gestión de asuntos escolares, tal como fue indagado en la mencionada evaluación (se transcriben las preguntas respectivas en §3.2.1 y 3.2.2).

#### 3.1 La participación como factor asociable al desempeño

Vamos a situar el asunto de la gestión y de la participación del alumno en dos momentos: escuela nueva y legislación reciente. Se hace esta división en tanto en escuela nueva hay un modelo pedagógico explícito<sup>35</sup> en el que la gestión y la participación *aparecen* como importantes. De otro lado, cerca de la tercera parte de la población encuestada en la evaluación de impacto del PUEBP pertenece a esa modalidad de escuela (mientras que la mitad pertenece a escuela urbana y el resto a escuela rural graduada).

##### 3.1.1 Escuela Nueva

En la reunión de ministros de educación en Ginebra, en 1961, se dispuso crear las “Escuelas unitarias”: un maestro para atender los cinco grados de la educación primaria. Se trataba de ampliar la cobertura de la educación básica primaria en las zonas rurales, *sin aumentar el número de docentes*; se dice [Siabato, 1998:105] que había un problema de carencia de docentes en las zona rural. Manteniendo el criterio de ampliar la cobertura, estas escuelas se reemplazaron por escuela nueva<sup>36</sup>, dado que ni la formación de los maestros, ni la educación que los niños recibían fueron consideradas las mejores. La nueva propuesta agregó entonces materiales y formación del docente para que manejara los cinco grados de la escuela. Antes de estable-

---

<sup>35</sup> Mientras que la escuela graduada —urbana o rural—no está definida por un modelo pedagógico explícito.

<sup>36</sup> Este programa no se corresponde con los postulados filosóficos y metodológicos de la “Escuela activa” implementada por Agustín Nieto Caballero y que fue puesta en práctica en el Gimnasio Moderno en Colombia, hacia principios del siglo XX.



cerse como modalidad específica de escuela, fue durante diez años un programa de capacitación de maestros.

En 1976 aparece el primer manual de escuela nueva, escrito por Colbert y Mogollón [1980]. Desde sus inicios, los ejes sobre los cuales se estructura la propuesta son: administración, comunidad, capacitación-seguimiento, y currículo; componentes que debían someterse a un proceso regular de evaluación, para determinar «el mejoramiento de la calidad del programa» [3]. A continuación se desarrollarán uno por uno.

En primer lugar, lo administrativo se enuncia como una intención de descentralizar el mannejo de las escuelas y responder a los requerimientos y necesidades de maestros y escuelas, a través de los supervisores (que, a la vez, son capacitadores) y de un comité departamental.

En segundo lugar, lo comunitario aparece como la búsqueda de un fortalecimiento de las relaciones entre escuela y comunidad, la cual brindaría apoyo a los niños, participando de las actividades curriculares de la escuela. Esto debido a la incredulidad de los habitantes del campo frente a las supuestas ventajas que ofrecía la escolarización de los niños, pues no encontraban que los aprendizajes escolares tuvieran aplicación en sus necesidades cotidianas.

En tercer lugar, la capacitación y el seguimiento se concebía en forma de talleres que abarcaran la metodología del programa y los conocimientos, habilidades y destrezas buscados [Colbert y Mogollón, 1980:2]. Los maestros debían reconocer las formas de trabajo del programa, mediante una capacitación de “metodología vivencial” en la que se reflexionara sobre la práctica, se interactuara con quien imparte la capacitación, y se experimentara lo que posteriormente se va a hacer en la escuela: manejo de las guías y conocimiento del manual<sup>37</sup> (la estructura de este manual es la misma que la de las guías de los alumnos). La capacitación iba dirigida al supervisor departamental y a los maestros rurales.

Y, en cuarto lugar, el eje curricular consistía en materiales educativos, guías de trabajo<sup>38</sup>, para profesores, supervisores y alumnos, en las que se desarrollan los pro-

---

<sup>37</sup> En la oficina de educación rural del MEN está la última edición del manual *Hacia la escuela nueva*, publicado en 1992, que se usa actualmente para la capacitación y que tiene muy pocas modificaciones frente a la tercera edición de 1982.

<sup>38</sup> Como las guías presuponen que el usuario sabe leer y escribir, el grado primero de primaria careció de ellas durante largo tiempo. Una primera elaboración no se implementó, pues no tenía una buena fundamentación. Otras guías tampoco fueron utilizadas y reposan en las escuelas, debido, según los docentes, a la falta de capacitación.

gramas de naturales, sociales, matemáticas, y español<sup>39</sup>. Posteriormente, se agregan las de estética, educación física y religión (con arreglo al currículo reglamentado por el Decreto 1419 de 1978 [cf. MEN, 1981]). Las guías buscaban: dar contenidos para trabajar en el medio y seguir estudiando; desarrollar destrezas y habilidades para aprender a aprender; adquirir hábitos de investigación (interesarse por conocer el medio); y desarrollar actitudes de cooperación y participación en la vida comunitaria. En su inicio, las guías contenían: un objetivo que determina el cambio de conducta que se pretende en el niño; un “cuento pedagógico” (para aclarar, afianzar y ampliar el conocimiento adquirido); y una ficha para que el maestro consignara cambios que creyera necesarios, para lo cual recibía capacitación y criterios. Luego, las guías tuvieron unas actividades (marcadas con letras) correspondientes a: “A”, el reconocimiento de lo que se sabe acerca de un determinado tema; “B”, la exposición de lo que académicamente se sabe del mismo, de acuerdo con el nivel de desarrollo de los niños; “C”, la interiorización-mecanización; y “D”, la confrontación con el contexto<sup>40</sup>. Incluían órdenes para realizar actividades individuales, grupales y con el profesor; indicadores de evaluación del número de actividades libres (“excelente”, “bien” y “aceptable”); y un tiempo estimado para el desarrollo de la guía.

Este material remite a unos “rincones de trabajo” para que el niño opere, experimente, investigue, practique y manipule objetos: «El niño no puede aprender qué es un imán, si no manipula un imán» [Colbert y Mogollón, 1980:168]. Estos objetos, que deben ser conseguidos en colaboración con la comunidad, son clasificados como: material real, de experimentación, impreso, para la creación estética, y producido por los niños. Las escuelas debían organizar los rincones de naturales, sociales, español y literatura<sup>41</sup>, matemática y estética. Se verán aparecer los rincones y las guías en la pregunta sobre la pertenencia a comités [cf. §3.2.1], hecha con el objeto de buscar factores asociados al desempeño.

Como puede verse, en esta modalidad de escuela permanentemente la participación está implicada o supuesta:

- una interacción con la comunidad para convencerla de las “bondades” de tener los hijos en la escuela;

---

<sup>39</sup> En un comienzo el área sólo contemplaba la enseñanza del español. La implementación de la enseñanza de la literatura se da posteriormente y es utilizada para la enseñanza de las estructuras gramaticales. Por ejemplo, si se leía una poesía, la actividad consistía en marcar los sustantivos y los verbos. Otra actividad que era considerada como literatura era realizar rimas y trabalenguas.

<sup>40</sup> La actividad “D” tiene relación con uno de los objetivos que el programa tenía en sus inicios: convencer a los padres y a la comunidad de la importancia de que el niño asistiera a la escuela, partiendo de demostrarle que lo que allí aprendía era “útil” para su vida diaria.

<sup>41</sup> «Haga una lista de 10 leyendas, poesías y dichos propios de la comunidad para el rincón de español y literatura» [Colbert y Mogollón, 1980:174].

- una capacitación basada en la participación y la interacción, pues se trataba de capacitar al maestro de la misma manera en que él iba a participar en la formación de los niños;
- materiales de trabajo que implican todo el tiempo la colaboración y la autogestión (pues no había un docente para cada curso que estuviera pendiente del desarrollo de las actividades);
- actividades sugeridas por las guías en las que los niños interactúan entre ellos para llevar a cabo las actividades conducentes al aprendizaje;
- actividades sugeridas por las guías en las que los niños interactúan con los miembros de su comunidad, para introducir la escuela en la cotidianidad; e
- interacción con la comunidad para la consecución de materiales destinados a los rincones.

### ***3.1.2 Reciente legislación***

Por fuera de la modalidad de escuela nueva, aparecen las siguientes menciones a la participación del alumno:

- Durante los períodos de la Misión alemana y de la *Renovación curricular*, la preocupación fundamental en los documentos era el diseño instruccional que se realizaba en el MEN. Los programas por curso contenían las actividades, los ejemplos, los indicadores de evaluación, etc.; el niño, en consecuencia, era el objeto de la educación, la cual tenía claro el perfil con el que debería modelárselo.
- Luego, en el marco de la nueva *Constitución Nacional*, supuestamente con el fin de adaptar el contexto jurídico del país a los procesos político-económicos internacionales, se introdujo un cambio nominal: de una democracia representativa a una democracia participativa. Se trataba de una de las condiciones para implementar cierta política, que, al descomprometer al Estado con los servicios sociales, apela a la “sociedad civil” para que se haga cargo de los costos y las responsabilidades (o parte de ellos) que abandona<sup>42</sup>. En este panorama, el texto de la Constitución “reparte” participación en muchos ámbitos; la educación es uno de

---

<sup>42</sup> Por ejemplo, el Plan de desarrollo del Gobierno de Pastrana dice que «Los *nuevos colegios* tendrán manejos sobre todos los recursos financieros y humanos, de modo que puedan hacerse responsables de los resultados de su servicio educativo» [DNP, 1999:41-42].

ellos: «La comunidad educativa participará en la dirección de las instituciones de educación» [art.68].

- ✿ Ya durante la evaluación de la calidad de la educación, el MEN plantea una posición frente a los “factores asociados” que podría ser una sustentación de la idea de que participar es algo asociado a los resultados educativos: habla de intersubjetividad, en el sentido en que «la interacción social es un proceso general de creación e intercambio de significados» [MEN, 1992:48]; no obstante, las preguntas que materializarían esa conceptualización en los instrumentos no tocan el tema. Según el documento, los “factores asociados” son características ambientales o personales que pueden incidir en el aprendizaje. Inexplicablemente (porque no parece provenir de “ambiente” ni de “persona”), de esto se desprende un intento de aproximación a las modalidades de interacción social del alumno en sus contextos socializadores; ámbito en el que podría perfectamente inscribirse una serie de interrogaciones sobre la participación del alumno en las actividades; sin embargo, las variables del estudiante [MEN, 1992:50] son: sexo, edad, residencia, experiencia educativa (preescolar, repitencia, movilidad, transporte, calificaciones), trabajo infantil, jornada cotidiana, actividad escolar (frecuencia de tareas, tiempo diario dedicado), otras actividades (TV), actitudes (hacia escuela, áreas y maestro), percepción de su entorno (interés familiar, de los maestros) y expectativas (educativas, laborales).
- ✿ Luego, la Ley 115 ordena que el alumno sea el centro del proceso educativo<sup>43</sup>. Entonces, su formación —que se pretende “integral”— incluye su participación, en la medida en que se busca favorecer «el desarrollo equilibrado y armónico de las habilidades de los educandos, en especial las capacidades para la toma de decisiones, la adquisición de criterios, el trabajo en equipo, la administración eficiente del tiempo, la asunción de responsabilidades, la solución de conflictos y problemas y las habilidades para la comunicación, la negociación y la *participación*» [art.92] (salvo que, al final, el beneficiado es el desarrollo socioeconómico del país).
- ✿ También para la Ley 115, los estudiantes son parte de la “comunidad educativa”<sup>44</sup>. Esto impone lógicamente alguna representación ante los órganos de dirección institucional así como alguna participación: efectivamente, en los Consejos Directivos habrá un representante estudiantil, elegido por ellos mismos, de

---

<sup>43</sup> «El alumno o educando es el centro del proceso educativo y debe participar activamente en su propia formación integral. El Proyecto Educativo Institucional reconocerá este carácter» [art.91].

<sup>44</sup> «La comunidad educativa está conformada por estudiantes o educandos, educadores, padres de familia o acudientes de los estudiantes, egresados, directivos docentes y administradores escolares» [Ley 115, art.6].

acuerdo con la reglamentación de la institución [cf. Ley 115, arts.93 y 143]<sup>45</sup>; así mismo, en el Gobierno escolar «serán consideradas las iniciativas de los estudiantes [...]» [art. 142] y participarán, según su competencia, en el diseño y ejecución y evaluación del PEI [art.6]; y en cada establecimiento, los estudiantes elegirán un personero para que promueva sus derechos y deberes [art.94].

- ❁ En este contexto, el *Salto educativo* —Plan de desarrollo sectorial de Samper—, para mejorar la calidad de la educación, se proponía, entre otros, la “promoción de la educación activa”: «Con el fin de superar el carácter pasivo de la enseñanza tradicional, se promoverá la transformación de los sistemas pedagógicos rutinarios en sistemas activos, de tal manera que el alumno *participe y pueda organizar efectivamente su propio aprendizaje*» [DNP, 1994:21].
- ❁ Por su parte, el Decreto 1860/94 ratifica, en el artículo 15, la participación de todos los estamentos de la comunidad educativa en la formulación del PEI, y, en el artículo 29, establece el Consejo de estudiantes, «máximo órgano colegiado que asegura y garantiza el continuo ejercicio de la participación por parte de los educandos», y que estará integrado por los voceros de cada grado (preescolar y los tres primeros grados sólo tienen un vocero), elegidos en asambleas por grado que convoca el Consejo Directivo. Sus funciones son:
  - a. *Darse su propia organización interna;*
  - b. *Elegir el representante de los estudiantes ante el Consejo Directivo del establecimiento y asesorarlo en el cumplimiento de su representación;*
  - c. *Invitar a sus deliberaciones a aquellos estudiantes que presenten iniciativas sobre el desarrollo de la vida estudiantil; y*
  - d. *Las demás actividades afines o complementarias con las anteriores que le atribuyan el manual de convivencia.*
- ❁ Ortiz y Burgos [1995:465] dicen que uno de los pilares de los postulados educativos «lo constituye el papel que la escuela debe jugar como espacio para el aprendizaje y no como tradicionalmente se le ha tenido, un lugar para la enseñanza»; citan en su apoyo a Sánchez y Ochoa (*Métodos activos para enseñanzas media y secundaria*), quienes dicen que lo importante no es que los profesores enseñen, sino que los alumnos aprendan, que lo esencial es el aprendizaje y no la enseñanza. Criterio que introduce la participación, pues para ellos se trata de la construcción de democracia en Colombia, *vía la escuela*.
- ❁ Todas las condiciones mencionadas, de realizarse, evidentemente comprometen la gestión y la participación de los alumnos (lo que no ocurría cuando se les exi-

---

<sup>45</sup> En donde no participan es en el Consejo Académico [cf. Ley 115, art.145].

gía solamente en el ámbito del aula, y para repetir información). En tal dirección, Ortiz y Burgos [1995:459-460] hablan de las implicaciones que para los estudiantes tienen estas medidas: i) organizar la elección del representante (planeación, cronograma, reglas), con lo que se desarrolla la capacidad para hablar, para hacer propuestas orales y escritas, y para diseñar propagandas, actividades que prefiguran la participación en diversos niveles de la futura vida social; ii) convertir la institución en objeto de reflexión e investigación, para poder llevar inquietudes al Consejo, con lo que aparece la necesidad de formular proyectos tendientes a solucionar los problemas; iii) pasar de ser receptores de normas y exigencias de los adultos, a participar activamente del quehacer escolar a través del personero.

Participar en el diseño, ejecución y evaluación del PEI, conformar el Consejo de estudiantes, darle su organización, convocar las iniciativas para la elección de los voceros de cada grado —en cumplimiento de la norma—, también implica otro tipo de acciones, interacciones y evaluaciones que articularían diversas formas de gestión y participación del alumno (como negociaciones comunicativas, trabajo en equipo, organización de acciones extracurriculares, etc.).

☛ Samper —a quien le tocaron los primeros 4 años de aplicación de la Ley 115 y algo menos del decreto reglamentario— hace un “balance” parcial de las mencionadas medidas:

- Dice [Samper, 1998:253] que el PEI posibilitó efectivamente la participación.
- Menciona «el despertar progresivo de una conciencia ciudadana en los estudiantes y docentes en torno a la paz, la democracia y los derechos humanos como resultado de la definitiva incorporación de estos temas a lo largo de todo el currículo; de la promoción de una cultura democrática en las escuelas y colegios con el establecimiento del gobierno escolar y sus instituciones representativas de padres de familia, educadores y estudiantes» [:254];
- habla de una democratización de las estructuras de poder en las instituciones educativas «a través de la formación y conformación de gobiernos escolares, de personeros estudiantiles y manuales de convivencia» [:254].
- Y dice haber atendido las consultas de los personeros, trazado lineamientos para ellos y organizado eventos específicos [:255].

☛ Por su parte, la *Misión ciencia, educación y desarrollo* [1997:194] recomienda «incorporar en las organizaciones educativas lo más avanzado de las técnicas de gestión». Para ello propone, de un lado, utilizar la disposición sobre PEI de la Ley

115 «para que el gobierno escolar defina el sistema de gestión que debe adoptar» [:194]; y, de otro, estimular la experimentación y la innovación en temas relacionados con la organización y el gobierno de los planteles, tales como metodologías de participación, proyectos de vinculación al entorno y a la comunidad, desarrollo de habilidades para el diálogo y la negociación, manejo de servicios internos por parte de los educandos, evaluación conjunta del desempeño de las entidades y otros que contribuyan a provocar el interés por los temas de la gestión de organizaciones educativas» [:195]<sup>46</sup>. Incluso dedica un aparte completo al tema: «El papel de la gestión» [:129].

- ❁ En el Plan de desarrollo en educación del gobierno de Pastrana [cf. DNP, 1999], hay pocas menciones a la participación; allí se reduce a la función de garantizar un terreno de aplicabilidad de las políticas (recordemos que se llama “plan educativo *para la paz*”): «La escuela es el primer ámbito de lo público para el individuo y, por lo tanto, escenario privilegiado para la formación de los valores y prácticas de solidaridad social, resolución creativa y productiva de conflictos, participación social y convivencia ciudadana. Esto hace que sea primordial fortalecer los gobiernos escolares como ámbito de formación y participación social y de fundamentación en los valores y prácticas para la democracia» [:34]. Es decir, no se trata de la participación como ligada a cierta concepción de sujeto y de sociedad, sino como mecanismo para implementar una política coyuntural... pese a que en el *Proyecto de presupuesto nacional de la nación 2000* [MINHACIENDA-DNP, 1999:661] se dice que el MEN busca «el logro permanente de la siguiente misión: [...] 4. Promover la participación activa de los jóvenes en los organismos que tengan a su cargo la protección, la educación y el progreso de la juventud».
- ❁ La tendencia a convertir la “gestión” en algo importante para los logros educativos la vemos en la siguiente afirmación del informe de la investigación sobre factores asociados, hecha con la evaluación censal del DC: «de una visión centrada en el proceso de enseñanza-aprendizaje, como núcleo explicativo de la calidad de la educación, se ha pasado a una visión más sistémica, que enfatiza el contexto y la gestión institucional» [SED-Corpoeducación-DNP, 1999:i].
- ❁ No obstante, según Posada [1999], en el propio contexto escolar la participación se entiende más en la dirección de una formación en valores, democratización de la escuela, búsqueda de la paz... tal vez como efecto de la situación política del país. En el marco de una interacción con maestros del programa RED de la Universidad Nacional, Posada encuentra que los maestros entienden el concepto de “comunidad educativa” en relación con: el contexto, la convivencia y la partici-

---

<sup>46</sup> Y, sin embargo, Heyneman [1991] dice que las innovaciones resultan asociadas negativamente al desempeño...

pación (en la organización escolar, el gobierno escolar y el aula), y no, por ejemplo, tal como la norma lo define.

- Por su parte, el equipo de investigación del presente estudio considera que la interacción —tanto escolar como extraescolar— es una condición fundamental de la cualificación de la competencia comunicativa (que es el objeto definido para la evaluación de impacto del PUEBP en el área de lenguaje). Valga esta idea para señalar las limitaciones que tienen los instrumentos de factores asociables cuando se elaboran independientemente de los instrumentos para evaluar el desempeño de los estudiantes; sobre todo teniendo en cuenta que los resultados de ambos instrumentos se van a cruzar después, en busca de correlaciones.

El concepto de *competencia comunicativa* proviene de una discusión sobre conceptos previos: competencia lingüística y competencia ideológica. Ulloa [1984] recuerda que el concepto de competencia lingüística, planteado por Chomsky, fue objeto de crítica en la medida en que postula un hablante-oyente ideal, ubicado en una comunidad lingüística homogénea. De ahí que Eliseo Verón hablara de competencia ideológica, en el sentido de reglas semánticas que condicionan la producción de mensajes, entendidos éstos como manifestaciones ideológicas. El planteamiento de Chomsky también se ve objetado por Dell Hymes, quien asegura que el uso de la lengua no es arbitrario —como creyera de Saussure— sino que está perfectamente estructurado, como si a su vez fuera una lengua. Para Dell Hymes, entonces, la competencia es la capacidad de un hablante históricamente determinado para producir enunciados con ayuda de todos los sistemas semióticos, en contextos específicos que están codificados y constituidos socialmente por la interacción.

Así, la escuela es un espacio privilegiado de interacción, pues los niños interactúan para llegar a ella, para entrar al salón, para llevar a cabo las actividades propias de instrucción que el maestro imparte, para jugar —tanto en los momentos y espacios formalmente dedicados a ello, como en los informales—, para comunicarse, para responder a las preguntas y requerimientos del maestro y de los directivos, para hacer solicitudes, para organizar sus propias iniciativas individuales o grupales, etc. En la educación se encuentran todas las interacciones humanas (pragmática, matética, sociocultural y estética); sin embargo, «se hace necesario señalar que las interacciones establecidas entre los participantes por el uso del lenguaje como instrumento de significación se realizan, fundamentalmente, mediante una transacción semántica» [Baena y Bustamante, 1992:37-38].

Pues bien, en el formulario de factores asociables de la evaluación de impacto del PUEBP se recogen algunas actividades de los estudiantes, observadas solamente en la



escuela; esto se hace mediante preguntas que se refieren, de un lado, a la participación de los niños en los diversos comités que la actividad escolar genera y, de otro, a la participación de los niños en la gestión de ciertos eventos escolares.

### **3.2 Algunos factores-participación asociables al desempeño en lenguaje**

Nos referiremos a los factores que tienen que ver con la participación en los diversos comités y a la gestión que realizan los estudiantes en una serie de tareas; por supuesto, esto se hará con base en la manera como la encuesta del MEN lo indagó, pues, por ejemplo, la única pregunta sobre el gobierno escolar —cuya importancia para la participación se ha señalado— es si votó o no votó...

#### **3.2.1 Comités**

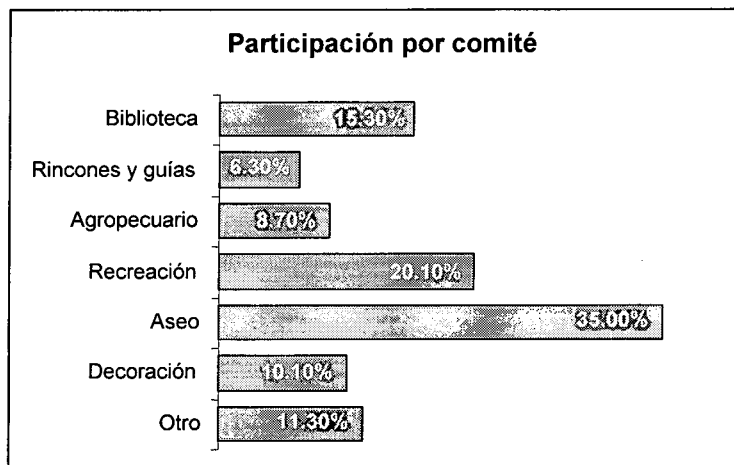
En relación con los comités se hace una única pregunta; el formulario de ciencias y matemáticas, que incluye las preguntas sobre factores asociables que se le hacen al alumno, interroga:

Aquí, en la escuela, ¿a qué comité o comités perteneces?	
<input type="checkbox"/> A	Biblioteca
<input type="checkbox"/> B	Al de aseo
<input type="checkbox"/> C	Al de decoración
<input type="checkbox"/> D	Al de rincones y guías*
<input type="checkbox"/> E	Al de recreación
<input type="checkbox"/> F	Al de actividades agropecuarias
<input type="checkbox"/> G	A otro diferente a los anteriores
<input type="checkbox"/> H	A ninguno

Los estudiantes participan de la siguiente manera en los diversos comités:

---

\* Se refiere a escuela nueva [cf. §3.1], en la que se dispone, de un lado, de "rincones", para concentrar el material de cada área; y, de otro, de guías de autoformación.



Se nota una mayor participación en los comités de aseo y recreación, mientras que la suma de la participación en los comités que podríamos llamar “académicos” [cf. §3.3.1] (biblioteca, rincones y guías, y agropecuario) no supera a la de aseo. El hecho de que el objeto de la participación sea cualitativamente diferente, permitiría esperar —sin detrimento de las actividades propias de la escuela— mayor participación en los comités académicos. Ahora veamos los desempeños en el área de lenguaje, vistos desde la estandarización, contra la participación en los diversos comités:

Comité	Participa	Desempeño estandarizado			p-valor
		Bajo	Medio	Alto	
Biblioteca	Sí	25.43	60.18	14.39	0.001
	No	15.71	61.63	22.66	
Rincones y Guías	Sí	23.96	62.84	13.20	0.001
	No	16.74	61.31	21.94	
Agropecuarias	Sí	20.00	61.43	18.57	0.079
	No	16.94	61.41	21.66	
Recreación	Sí	18.63	60.66	20.71	0.297
	No	16.84	61.60	21.56	
Aseo	Sí	17.88	63.21	18.90	0.002
	No	16.83	60.43	22.73	
Decoración	Sí	22.50	60.09	17.41	0.001
	No	16.61	61.56	21.84	
Otro	Sí	17.15	56.24	26.61	0.001+
	No	17.21	62.07	20.72	

Los datos revelan que la pertenencia a estos comités (excepto recreación) muestran asociación con el desempeño; salvo agropecuarias, cuya asociación es moderada, las

asociaciones son altamente significativas. No obstante, pertenecer a estos comités no favorece el paso de desempeños bajos a desempeños altos, sino que lo desfavorece; es decir, hay una asociación *negativa*<sup>47</sup>.

El único comité que, además de mostrar correlación, favorece los desempeños altos y desfavorece los bajos (es decir, con una asociación *positiva*) es el rotulado como “otro”; es decir, el espectro de comités indagado (A. a F. en la pregunta) no cubre los que responden a la asociación esperada.

Ahora veamos los desempeños por competencia-C, contra pertenencia a comités:

Comité	Participa	Desempeño por competencias-C				p-valor
		Nivel 0	Lectura			
			Literal	Inferencial	Crit.-intert.	
Biblioteca	Sí	1.01	21.68	62.51	14.79	0.169
	No	1.15	24.89	60.18	13.77	
Rincones y Guías	Sí	0.49	19.07	62.59	17.85	0.008
	No	1.18	24.76	60.40	13.66	
Agropecuarias	Sí	1.43	23.39	59.29	15.89	0.460
	No	1.10	24.49	60.66	13.74	
Recreación	Sí	1.23	24.40	59.66	14.70	0.789
	No	1.11	24.40	60.76	13.73	
Aseo	Sí	0.97	24.61	59.85	14.56	0.549
	No	1.22	24.28	60.91	13.59	
Decoración	Sí	1.69	25.89	57.78	14.64	0.279
	No	1.07	24.23	60.85	13.85	
Otro	Sí	0.96	25.51	58.98	14.54	0.757
	No	1.15	24.26	60.74	13.85	

Visto desde el análisis de los desempeños por competencias-C, solamente la participación en el comité de rincones y guías muestra una asociación (tiene un p-valor de 0.008). Esta asociación resultó como la esperada: altamente significativa y positiva, es decir, opuesta a la asociación vista desde el análisis estandarizado. Los p-valores restantes no apoyan la idea de correlaciones. En otras palabras, desde este análisis, la pertenencia a esos comités es indiferente para el desempeño en el área de lenguaje: los perfiles de quienes pertenecen y de quienes no pertenecen son muy similares. En otras palabras, participar o no en comités no permite predecir el desempeño de un

<sup>47</sup> La calificación de “negativa” o “positiva” no es valorativa ni se refiere a los significados matemáticos de dichas palabras. Llamaremos “negativa” (y lo notaremos mediante el signo [-] al lado del p-valor) a una asociación cuyos valores se relacionan de manera inversa; “positiva” (y lo notaremos mediante el signo [+] al lado del p-valor), a una asociación cuyos valores se relacionan de manera directa; y “no lineal” (y lo notaremos mediante el signo [±] al lado del p-valor) cuando no hay relación directa ni inversa definidas. Cuando esta característica no coincida con la valoración esperada, haremos la aclaración pertinente.

estudiante; éste parece depender de otras cosas. La escuela no le hace un favor al niño “permitiéndole” participar.

Es de destacar que, aunque la información es la misma (o sea, está expuesta al mismo ruido), en esta manera de considerar los desempeños se muestra más consistencia con la idea que de competencia comunicativa asumió el equipo, aunque muestra más certidumbre de no asociación; es decir, más necesidad de investigar de otra manera, si se quiere indagar por la asociación buscada. Todo esto —ante el hecho de que la misma información dé resultados opuestos— abre interrogaciones también sobre la forma de mirar: el instrumento del área de lenguaje, el concepto mismo de competencia comunicativa y la manera de “leer” los datos desde la estadística. Que la única correlación que la mirada sobre el desempeño desde las competencias-C arroja sea justamente en “rincones y guías”, apunta a algo que tiene consistencia con la modalidad de escuela nueva. Otros factores que se indagan en la pregunta tal vez sean dispersos de un tipo de escuela a otro; si esto es así, son impertinentes para una indagación masiva y homogénea: no se le puede preguntar a todos por aspectos que son propios de modalidades que el mismo sistema discrimina.

Ahora haremos el análisis para los datos desde el desempeño en la producción escrita, o sea, la parte abierta de la prueba:

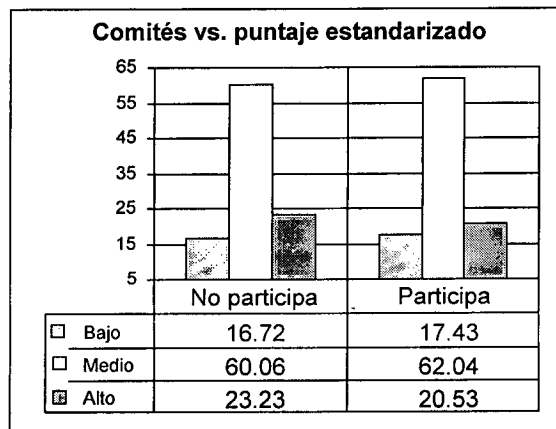
Comité	Participa	Desempeño por competencias-A				p-valor
		Nivel 0	Coherencia			
			Local	Lineal	Global	
Biblioteca	Sí	2.26	12.78	78.63	6.34	0.356
	No	2.57	10.86	80.37	6.21	
Rincones y Guías	Sí	1.28	14.03	79.59	5.10	0.088±
	No	2.61	10.95	80.13	6.31	
Agropecuarias	Sí	1.69	10.90	81.20	6.20	0.634
	No	2.60	11.18	79.99	6.23	
Recreación	Sí	2.54	11.07	79.08	7.30	0.387
	No	2.51	11.17	80.35	5.96	
Aseo	Sí	2.80	11.85	79.00	6.35	0.384
	No	2.37	10.77	80.70	6.16	
Decoración	Sí	2.61	14.36	76.67	6.36	0.063±
	No	2.51	10.79	80.48	6.21	
Otro	Sí	1.93	12.59	80.30	5.19	0.283
	No	2.59	10.97	80.07	6.36	

Desde el desempeño de los niños en la parte abierta, en algunos casos las pequeñas diferencias muestran perfiles paralelos entre los que participan y los que no en cada comité; es decir, la competencia para producir textos por parte de los niños no tiene, en esos casos, relación con su participación en comités. Dicha capacidad dependería

de otra cosa. En otros casos, muestran asociaciones estadísticas moderadas (rincones y guías, con un p-valor de 0.088, y decoración con un p-valor de 0.063) en sentidos contrarios, o sea, que no se puede decir si es positiva o negativa (este es el tipo de asociación que identificaremos como “no lineal”, sin la connotación matemática estricta). Hay correlación pero no se puede identificar en qué sentido una cosa favorece a la otra; es decir, hay que investigarlo de otra manera. Esto es diferente a los casos anteriores.

Para continuar indagando la participación en los comités (y hacerla comparable con el DC), resolvimos generar dos grupos: i) los niños que participan en alguno de estos comités (literales A. a G. en la pregunta), y ii) los que no (literal H). De tal manera, sólo se destaca la participación, sin diferenciar entre sí los diversos comités, con lo cual se “suaviza” un poco la heterogeneidad que habría entre modalidades de escuela.

Pues bien, de acuerdo con los resultados de la encuesta, se encuentra que la mayoría de los niños participa en algún comité; a saber, el 68.1%. Era de esperarse (por las consideraciones expuestas en §3.1) que el desempeño en la prueba de lenguaje de los niños que participan se viera afectado fuerte y positivamente; sin embargo, la asociación es moderadamente significativa: el p-valor es de 0.048:



En otras palabras, los desempeños parecen depender en alguna medida de este tipo de participación, aunque puede verse una tendencia contraria a la hipótesis, en tanto tienden más a pasar a desempeños altos quienes no participan, y en tanto tienden a quedarse en desempeños bajos quienes participan; es decir, una asociación negativa.

El análisis desde los niveles de logro del mismo desempeño, pero obtenidos mediante las categorías teóricas del equipo, refleja perfiles muy parecidos (p-valor: 0.880). Es decir, que visto desde esa perspectiva, el desempeño no parece depender de la participación. De acuerdo con los datos (destacamos los extremos), para lograr

un nivel de desempeño crítico-intertextual, sería indiferente participar o no, pues del total de niños que presentó la prueba, los que participan son un 13.9% y los que no participan son un 13.99%; así mismo, para quedarse en un nivel de desempeño literal, también sería indiferente participar (24.17%) o no participar (24.88%).

Visto desde los desempeños de los niños en la prueba abierta, tenemos igualmente un p-valor no significativo: 0.661, es decir, que no muestra correlación significativa entre participación y desempeño. Para los desempeños de coherencia local, la participación en comités tiene un 10.93% y la no participación 11.63%; para la coherencia lineal, participar, con 80.17%, y no, con 79.94%; y para la coherencia global, participar, con 6.24%, y no, con 6.20%; perfiles tan semejantes indicarían que el factor investigado es indiferente ante los desempeños en la prueba abierta. O sea, los niños producen textos independientemente de las maneras de participación *preguntadas*.

Veamos los desempeños (desde las tres perspectivas) cruzados con este factor, pero ahora desagregado por grado:

			Grado			
			Cuarto		Tercero	
			Participa		Participa	
			Sí	No	Sí	No
<b>Desempeño</b>	Estandarizado	Bajo	10.51	8.47	24.02	25.12
		Medio	61.51	59.87	62.54	60.26
		Alto	27.98	31.67	13.43	14.62
		p-valor	0.038		0.436	
	Competencias cerrada	Nivel 0	0.93	1.35	1.25	1.08
		Lectura Literal	24.05	25.60	24.29	24.14
		L. Inferencial	60.95	59.10	60.72	60.75
		L. Crít.-intertex.	14.06	13.96	13.75	14.03
		p-valor	0.531		0.976	
	Competencias Abierta	Nivel 0	2.18	1.85	3.12	2.61
		Coherencia Local	10.65	11.28	11.20	11.99
		C. Lineal	81.28	80.51	79.12	79.35
		C. Global	5.89	6.36	6.57	6.05
		p-valor	0.831		0.746	

Cuando se trata de cuarto de primaria, los desempeños en la parte cerrada de la prueba son más indicativos de una asociación con la participación: un p-valor de 0.038 (moderadamente significativo, pero negativo), contra un 0.436 en tercero, que no permite hablar de correlación; esto para el análisis estandarizado. En el caso del análisis por competencias-C, no puede afirmarse que exista asociación (como cuan-

do no estaba desagregado por grado), los perfiles son casi idénticos; en otras palabras, que participar o no participar en los comités, es indiferente para el desempeño. El desempeño en la parte abierta de la prueba, en ninguno de los dos grados permite hablar de correlación con la participación (como cuando no estaba desagregado por grado). En definitiva, el desglose por curso, más puntual, no permite encontrar más asociaciones. Para la parte cerrada, desde los análisis estandarizado y por competencias, las cifras desglosadas, aunque con distinta magnitud, dicen lo mismo.

Ahora veamos los desempeños (desde las tres perspectivas) cruzados con este factor, pero ahora desagregado por modalidad de escuela:

		Modalidad de escuela						
		Urbana		Nueva		Rural graduada		
		Participa		Participa		Participa		
		Sí	No	Sí	No	Sí	No	
<b>Desempeño</b>	<b>Estandarizado</b>	Bajo	16.94	14.07	18.40	21.29	16.76	19.93
		Medio	60.29	59.31	62.89	61.98	65.35	59.79
		Alto	22.77	26.62	18.71	16.73	17.89	20.28
		p-valor	0.011		0.268		0.253	
	<b>Competencias cerrada</b>	Nivel 0	1.00	1.20	0.95	1.33	1.69	1.07
		L. Literal	24.72	24.38	24.19	24.90	22.54	27.05
		L. Inferencial	60.25	60.59	60.73	60.27	62.82	56.23
		L. Crít.-intertext.	14.04	13.83	14.13	13.50	12.96	15.66
		p-valor	0.945		0.864		0.192	
	<b>Abierta</b>	Nivel 0	2.18	1.36	2.98	3.70	3.36	3.32
		C. Local	9.77	11.47	10.98	9.67	14.16	15.87
		C. Lineal	81.06	80.20	80.35	80.45	77.23	77.86
		C. Global	6.99	6.97	5.69	6.17	5.26	2.95
		p-valor	0.191		0.721		0.451	

Vistos los desempeños con relación a la participación en comités, dentro de cada una de las modalidades de escuela, de las nueve posibles, solamente hay una asociación; ésta es moderadamente significativa y va en sentido inverso al esperado; se presenta en la escuela urbana, visto el desempeño desde el análisis estandarizado. En este tipo de análisis, desagregar la información por modalidad de escuela arrojaría la idea de que la tendencia moderada del consolidado nacional tendría que ver con una tendencia fuerte de asociación en la escuela urbana: sólo allí no es indiferente la participación.

En el Distrito Capital, los resultados son como sigue:

			Distrito	
			Comités	
			No	Sí
<b>Desempeño</b>	<b>Estandarizado</b>	Bajo	4.35	12.28
		Medio	73.91	63.16
		Alto	21.74	24.56
		p-valor	0.503	
	<b>Competencias cerrada</b>	Nivel 0	4.35	3.51
		L. Literal	26.09	22.81
		L. Inferencial	65.22	64.91
		L. Crít.-intertext.	4.35	8.77
		p-valor	0.912	
	<b>Abierta</b>	Nivel 0	0.00	6.00
		C. Local	10.00	12.00
		C. Lineal	85.00	78.00
		C. Global	5.00	4.00
		p-valor	0.712	

Tanto a escala nacional como distrital, la interpretación del desempeño hecha desde las competencias no muestra asociación con la participación en comités; es decir, que en ninguno de los dos casos puede afirmarse que dicha participación coadyuve a desarrollar la competencia comunicativa de los niños. Se participe o no, el desempeño puede o no tender a los niveles más altos. En esto, la mirada sobre lo distrital no suministra más información.

En cambio, la interpretación del desempeño desde la estandarización sí muestra diferencias: mientras a escala nacional hay una asociación moderadamente significativa y negativa (p-valor: 0.048), en el DC continúa evidenciándose una no asociación, como si a esta escala la participación en comités —tal como se indagó— no tuviera incidencia en el desempeño, desde las tres categorizaciones puestas en juego en el presente estudio para dar cuenta del desempeño en el área de lenguaje.

### 3.2.2 Gestión

La pregunta que se hizo para obtener la información de lo que hemos denominado “gestión” de los niños en la escuela fue<sup>48</sup>:

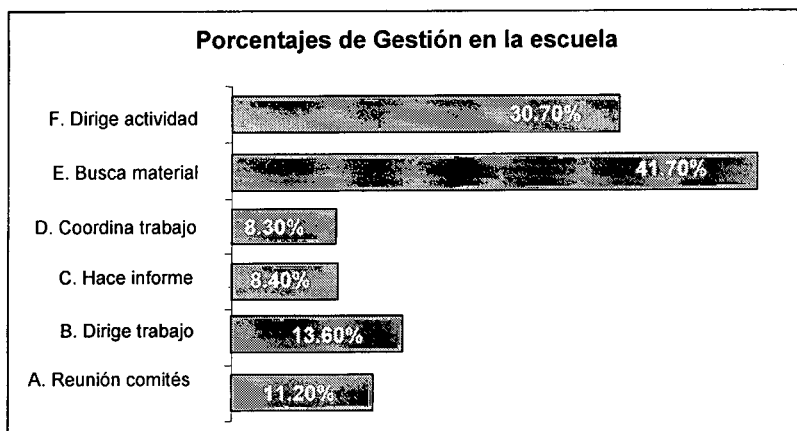
<sup>48</sup> Es la novena pregunta del formulario de desempeño para niños en ciencias y matemáticas.



Aquí, en la escuela, ¿participas de algunas de las siguientes tareas?

- A Dirigir las reuniones de comités
- B Dirigir el trabajo de los ayudantes
- C Hacer informes sobre el trabajo de comités
- D Coordinar el trabajo de los ayudante
- E Buscar el material (guías, libros) para las clases
- F Dirigir algunas actividades que se realizan en la escuela

Las respuestas a estas preguntas, distribuyen la población encuestada de la siguiente manera:



Como los niños podían responder de manera no excluyente a las alternativas ofrecidas, esto da lugar a que los mismos niños puedan estar en más de una categoría (de ahí que la suma de los porcentajes sea mayor de 100%). Por esta razón, se buscarán asociaciones para cada modalidad de gestión, ya que el análisis estadístico que se está manejando supone categorías mutuamente excluyentes y exhaustivas [cf. §2.1.1].

Aunque las actividades de gestión “Buscar el material (guías, libros) para las clases” y “Dirigir algunas actividades que se realizan en la escuela” presentan la mayor participación de los estudiantes, no se encuentran asociadas con el desempeño en el área de lenguaje, de acuerdo con el análisis estandarizado. Las restantes, que sí se encuentran asociadas con el desempeño, están asociadas en dirección opuesta a la hipótesis del equipo; es decir, cuando se espera que la gestión favorezca el paso a las modalidades de desempeño alto, y desfavorezca quedarse en las de bajo, los datos evidencian lo contrario: no gestionar favorece el tránsito a los mejores desempeños

en el área de lenguaje, vistos desde el análisis estandarizado. Estos resultados no deberían sorprender,

- si, de un lado, la competencia comunicativa no tuviera mucha relación con la *manera* como se gestiona;
- y, de otro lado, si los subgrupos generados por cada una de las modalidades, por sus tamaños distintos, no fueran equiponderables; estos tamaños parecen desproporcionados cuando se desagregan en aquellos que participan y aquellos que no. De tal manera, cada individuo del subgrupo tiene una ponderación diferente a la que tiene en el grupo; esto puede conducir a la aparición de asociaciones en los grupos pequeños y de no asociaciones en los grupos grandes, como en este caso, pero podría ser la inversa. Los datos son los siguientes:

Gestión	Participa	Desempeño estandarizado			p-valor
		Bajo	Medio	Alto	
Dirige reuniones de comités	Sí	27.26	62.73	10.01	0.001
	No	15.94	61.24	22.82	
Dirige trabajo de ayudantes	Sí	24.14	62.93	12.93	0.001
	No	16.11	61.17	22.72	
Hace informes trabajo comités	Sí	23.70	62.22	14.07	0.001
	No	16.61	61.33	22.06	
Coordina trabajo ayudantes	Sí	23.09	59.40	17.50	0.001
	No	16.67	61.59	21.74	
Busca material para clases	Sí	16.36	61.25	22.39	0.132
	No	17.80	61.52	20.68	
Dirige actividades en la escuela	Sí	16.81	60.52	22.67	0.247
	No	17.38	61.80	20.82	

Analizando el desempeño con arreglo a las categorías por competencia (cf. tabla siguiente), con excepción de “Dirigir algunas actividades que se realizan en la escuela” —donde aparece una asociación positiva, moderadamente significativa (p-valor: 0.013)—, no encontramos p-valores que den cuenta de la asociación entre gestión y desempeño; de ahí que los perfiles entre los que gestionan y los que no, sean casi idénticos. O sea, no es que no haya asociación entre gestión y desempeño, sino que la gestión así asumida no es “detectable” con los datos disponibles. Los datos son como sigue:

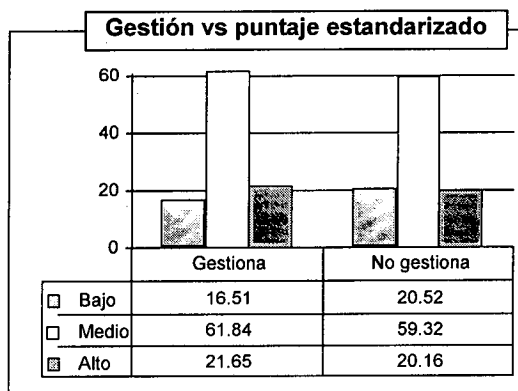
Gestión	Participa	Desempeño por competencias-C				p-valor
		Nivel 0	Lectura			
			Literal	Inferencial	Crit.-intertext	
Dirige reuniones de comités	Sí	1.25	23.92	61.06	13.77	0.973
	No	1.12	24.46	60.47	13.95	
Dirige trabajo de ayudantes	Sí	1.14	24.60	58.70	15.56	0.472
	No	1.13	24.37	60.83	13.67	
Hace informes trabajo comités	Sí	1.30	24.26	58.52	15.93	0.526
	No	1.12	24.41	60.72	13.75	
Coordina trabajo ayudantes	Sí	1.30	24.39	57.73	16.57	0.277
	No	1.12	24.40	60.80	13.69	
Busca material para clases	Sí	1.15	24.73	59.61	14.50	0.578
	No	1.12	24.16	61.20	13.52	
Dirige actividades en la escuela	Sí	1.26	23.32	59.46	15.95	0.013
	No	1.07	24.88	61.02	13.03	

Y, en relación con el desempeño en producción, prueba abierta (cf. tabla siguiente), hay asociaciones moderadas y negativas en “dirección de reuniones de comités” y en “dirección de trabajo de ayudantes”; el hecho de que sean negativas implica que la gestión, al contrario de lo esperado, coadyuvaría a que los desempeños de los estudiantes tiendan a quedarse en desempeños locales, y que la no gestión coadyuvaría a que la producción de los estudiantes tenga coherencia global. Las demás formas de gestión tienen p-valores superiores al 10%. Veamos los datos:

Gestión	Participa	Desempeño por competencias-A				p-valor
		Nivel 0	Coherencia			
			Local	Lineal	Global	
Dirige reuniones de comités	Sí	4.12	12.50	77.94	5.44	0.017
	No	2.32	10.98	80.37	6.33	
Dirige trabajo de ayudantes	Sí	3.40	13.49	76.91	6.20	0.032
	No	2.38	10.79	80.60	6.23	
Hace informes trabajo comités	Sí	2.75	14.12	76.67	6.47	0.147
	No	2.50	10.88	80.41	6.21	
Coordina trabajo ayudantes	Sí	2.57	12.65	77.67	7.11	0.529
	No	2.52	11.02	80.32	6.15	
Busca material para clases	Sí	2.52	10.87	80.04	6.57	0.760
	No	2.52	11.36	80.14	5.98	
Dirige actividades en la escuela	Sí	2.41	11.81	79.00	6.79	0.396
	No	2.57	10.86	80.59	5.98	

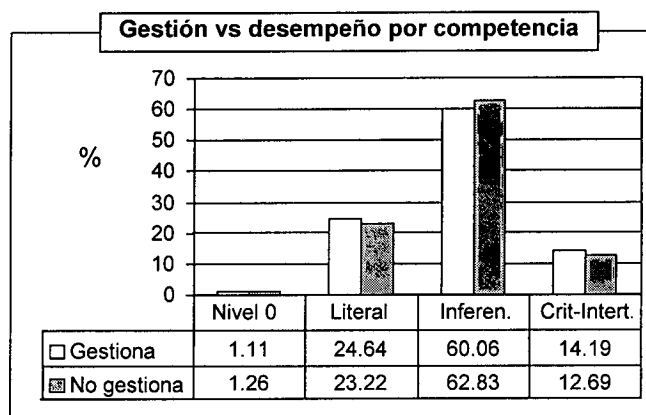
Para tener una panorámica (y hacerlo comparable con el DC), haremos una recategorización con dos modalidades de gestión: gestiona y no-gestiona. De acuerdo con los resultados de la encuesta, se encuentra que la mayoría de los niños (el 82.8%) que

presentaron la prueba, lleva a cabo labores de gestión (superan en 14.7 puntos a los que participan en comités). Dada la relación que hipotetizamos entre desempeño y gestión, podría presuponerse que la minoría que no gestiona (17.2%) tendría menores niveles de logro. Los niveles de desempeño obtenidos mediante la estandarización, cuando enfrentamos niños que gestionan y niños que no gestionan, arroja lo siguiente:



Se aprecia una asociación: en el nivel bajo hay 4% más de niños que no gestionan; y en el nivel alto hay 1.5% más de niños que gestionan. En otras palabras, no gestionar desfavorece el desempeño en lenguaje y gestionar lo favorece. Esto lo corrobora el coeficiente de p-valor para la contingencia entre niveles y gestión, que es de 0.005, altamente significativo y positivo. Pero la asociación, que resulta positiva cuando los factores de gestión están aglutinados, recordemos que era negativa cuando estaban desglosados (esto es un ejemplo de lo que conocemos como *paradoja de Simpson*).

Cuando observamos los desempeños desde las categorías por competencia, los resultados varían:



Las pequeñas diferencias no permiten afirmar nada sobre las ventajas o desventajas que la gestión preguntada tendría. En este caso, el p-valor es de 0.309. Dicho de otra

manera, el desempeño no depende de la gestión: a más gestión de aquella por la que responden los estudiantes en el cuestionario, no se puede decir que haya mayores probabilidades de pasar a desempeños de lectura crítico-intertextual.

Tampoco los desempeños en la parte abierta de la prueba permiten hablar de correlación (p-valor de 0.490); los perfiles entre quienes gestionan y no gestionan son casi idénticos: entre los que se quedan en el nivel de desempeño local, la diferencia de los que gestionan y no gestionan es de 0.63%; entre los que se quedan en el nivel de desempeño lineal, la diferencia de los que gestionan y no gestionan es de 0.81%; y entre los que alcanzan el nivel de desempeño global, la diferencia de los que gestionan y no gestionan es de 1.05%.

Desglosándolo por grado, tenemos las siguientes cifras:

				Grado			
				Cuarto		Tercero	
				Gestiona		Gestiona	
				Sí	No	Sí	No
<b>Desempeño</b>	Estandarizado	Bajo	10.19	7.92	22.99	30.11	
		Medio	60.89	61.46	62.82	57.69	
		Alto	28.93	30.63	14.19	12.20	
		p-valor	0.279		0.001+		
	Competencias cerrada	Nivel 0	1.04	1.25	1.18	1.27	
		L. Literal	24.37	25.63	24.92	21.39	
		L. Inferencial	60.11	61.67	60.02	63.71	
		L. Crítico-intertex.	14.48	11.46	13.88	13.63	
		p-valor	0.354		0.282		
	Abierta	Nivel 0	2.08	2.01	3.11	2.34	
		C. Local	10.48	12.98	11.62	10.70	
		C. Lineal	80.96	81.43	78.93	80.27	
		C. Global	6.48	3.58	6.34	6.69	
		p-valor	0.058+		0.679		

En lo que a prueba cerrada corresponde, hay una diferencia entre tercero y cuarto, para el análisis estandarizado. Sólo hay asociación, altamente significativa, en el análisis estandarizado para 3°: la gestión favorecería el paso a desempeños altos y desfavorecería quedarse en los bajos (correlación positiva). En 4°, en cambio, ambos aspectos son independientes. Para el análisis por Competencias-C, el desglose por curso no aporta información nueva: como en el consolidado nacional, ninguno de los dos grados presenta asociación. En cuanto a los desempeños en la prueba abierta, sí hay un cambio: se ve asociación moderada para el desempeño de los niños de cuarto; allí la gestión favorece el paso a desempeños altos y la no gestión favorece la

detención en desempeños bajos (correlación positiva). ¿Qué podría decirse? Hay muchas opciones: por ejemplo, que los instrumentos diseñados sólo permiten referirse a los desempeños de los niños de tercero, cuando se analizan desde la estandarización; y que sólo permiten referirse a los desempeños de los niños de cuarto, cuando se analizan desde la categorización de la prueba abierta. En cualquier caso, no puede leerse sencillamente que la gestión se asocie con el desempeño en un grado y, no obstante, en el grado siguiente, la asociación desaparezca. Asuntos como este, antes que autorizarnos a hacer afirmaciones sobre la realidad educativa, nos deben ayudar a interrogar nuestros medios de conocimiento.

La desagregación por modalidad de escuela de los desempeños vistos en función de la gestión, muestra los siguientes datos:

		Modalidad de escuela						
		Urbana		Nueva		Rural graduada		
		Gestiona		Gestiona		Gestiona		
		Sí	No	Sí	No	Sí	No	
<b>Desempeño</b>	<b>Estandarizado</b>	Bajo	15.36	18.43	17.82	25.20	17.70	17.49
		Medio	60.11	59.03	62.79	62.06	65.84	54.64
		Alto	24.54	22.54	19.39	12.74	16.46	27.87
		p-valor	0.164		0.001+		0.001-	
	<b>Competencias cerrada</b>	Nivel 0	1.11	0.89	0.93	1.63	1.49	1.64
		L. Literal	24.86	23.26	24.71	22.76	23.76	24.04
		L. Inferencial	59.86	62.97	60.19	62.60	60.52	62.84
		L. Crítico-intertext.	14.18	12.88	14.18	13.01	14.23	11.48
		p-valor	0.574		0.478		0.804	
	<b>Abierta</b>	Nivel 0	1.90	1.73	3.28	2.60	3.47	2.81
		C. Local	10.41	10.36	11.20	8.09	12.85	22.47
		C. Lineal	80.50	81.96	79.39	84.97	79.31	69.10
		C. Global	7.18	5.95	6.13	4.34	4.37	5.62
		p-valor	0.769		0.128		0.008±	

Para la modalidad de escuela urbana ninguno de los análisis del desempeño evidencia dependencia entre las variables. Es como si la asociación a escala nacional no permitiera ver que lo buscado es más pertinente en la educación del campo que en la de la ciudad. No obstante, mientras el consolidado tenía una asociación positiva, el desglose por modalidad de escuela muestra en el campo asociaciones contrarias: en escuela nueva, asociación positiva, y en escuela rural graduada, asociación negativa. Es también un caso de Paradoja de Simpson: una tendencia alta positiva, desglosada en tres, da no asociación, asociación positiva y asociación negativa.

La gestión, más propia de la modalidad de escuela nueva, resulta mostrando más asociaciones en la escuela rural graduada (una negativa y otra no lineal). La única manera de interpretar el desempeño que no muestra asociaciones por modalidad de escuela (como en lo nacional) es la de competencias-C. Y la que más muestra correlaciones es la estandarizada.

Para el DC, las cifras son como sigue:

			Distrito	
			Gestión	
			Sí	No
<b>Desempeño</b>	<b>Estandarizado</b>	Bajo	10.67	0.00
		Medio	69.33	20.00
		Alto	20.00	80.00
		p-valor	0.009	
	<b>Competencias cerrada</b>	Nivel 0	4.00	0.00
		L. Literal	22.67	40.00
		L. Inferencial	65.33	60.00
		L. Crít.-intertext.	8.00	0.00
		p-valor	0.750	
	<b>Abierta</b>	Nivel 0	3.08	20.00
		C. Local	12.31	0.00
		C. Lineal	81.54	60.00
		C. Global	3.08	20.00
		p-valor	0.069	

Cuando en lo nacional las modalidades de gestión estaban separadas, las asociaciones para el análisis estandarizado se mostraban negativas. En tal sentido, el resultado distrital es similar, pues la asociación es negativa y altamente significativa (p-valor:0.009), como puede verse en las cifras: los estudiantes tienden más a ser ubicados en desempeños bajos cuando gestionan que cuando no lo hacen; en cambio, tienden más a ser ubicados en desempeños altos cuando no gestionan que cuando gestionan.

En cambio, en el resultado nacional agregado, la asociación es positiva (paradoja de Simpson), es decir, que los estudiantes que gestionan —en la manera como ello es evaluado— tienden a avanzar hacia mejores desempeños. En esta nueva situación, el DC se comporta de manera opuesta a la nación (de manera que lo que podría parecer una sugerencia válida a escala nacional [impulsar la gestión de los estudiantes], puede no serlo a escala distrital).

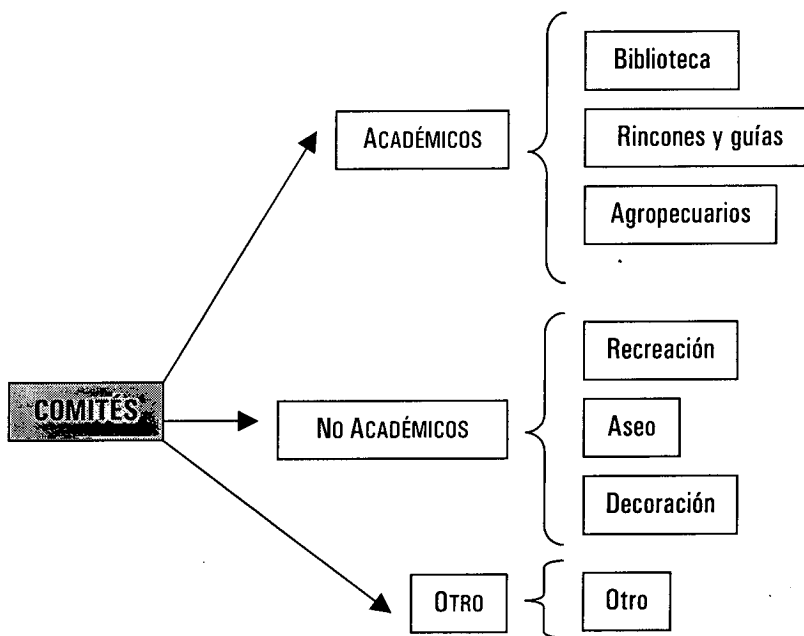
Cuando las modalidades de gestión están desagregadas, los análisis por competencias muestran muy pocas asociaciones a escala nacional; y cuando se colapsan, no

aparecen asociaciones a esa escala, pero a escala distrital sí: se muestra asociación moderadamente significativa en el análisis por competencias-A; la tendencia es negativa, tal como en las dos asociaciones que se presentan a nivel nacional cuando los factores están desagregados: como puede verse en la tabla anterior, hay más clasificados en los desempeños de Coherencia local entre quienes gestionan, y hay más clasificados en los desempeños de Coherencia global entre quienes no gestionan.

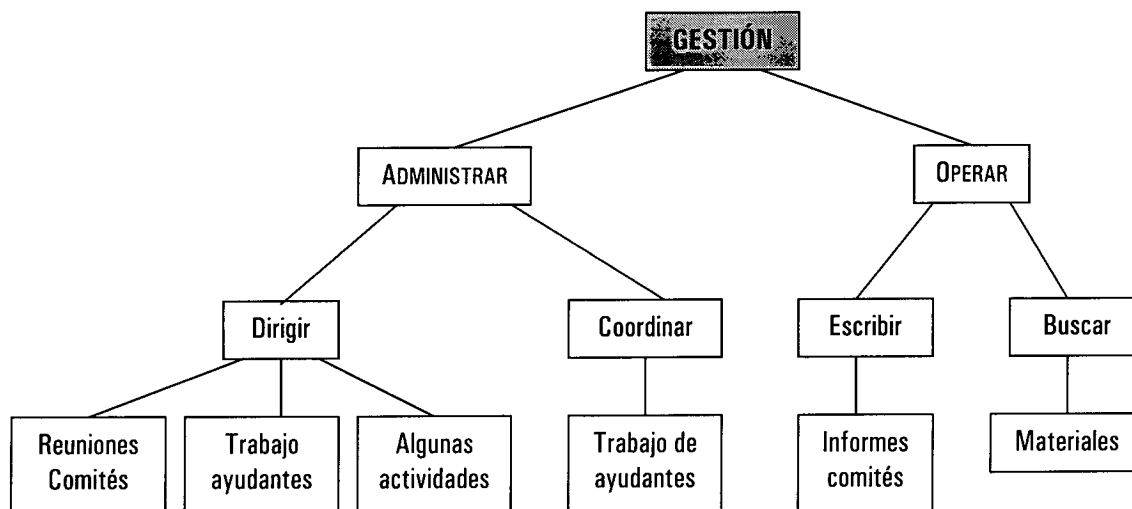
En resumen, el análisis para el DC tiene más coherencia con el nivel nacional visto desde los factores desgregados; con los factores colapsados, la tendencia del DC es contraria a la nacional para el análisis estandarizado, ninguno presenta asociación para el análisis por competencias-C y sólo el distrital muestra asociación para el análisis por competencias-A.

### 3.3 Resumen parcial

3.3.1 A continuación hacemos un esquema de la concepción que aparece de participación en comités y de gestión, en las preguntas respectivas hechas en los instrumentos:

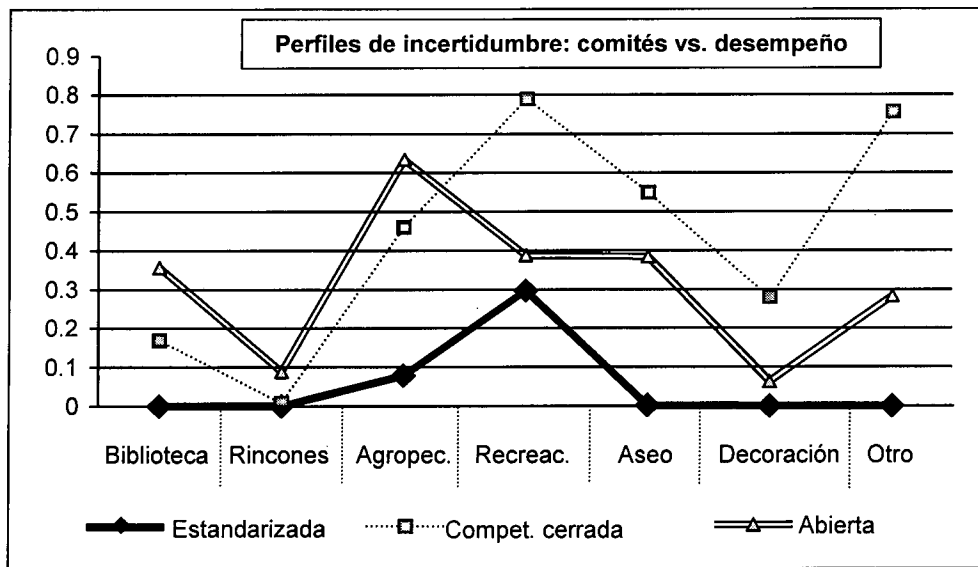






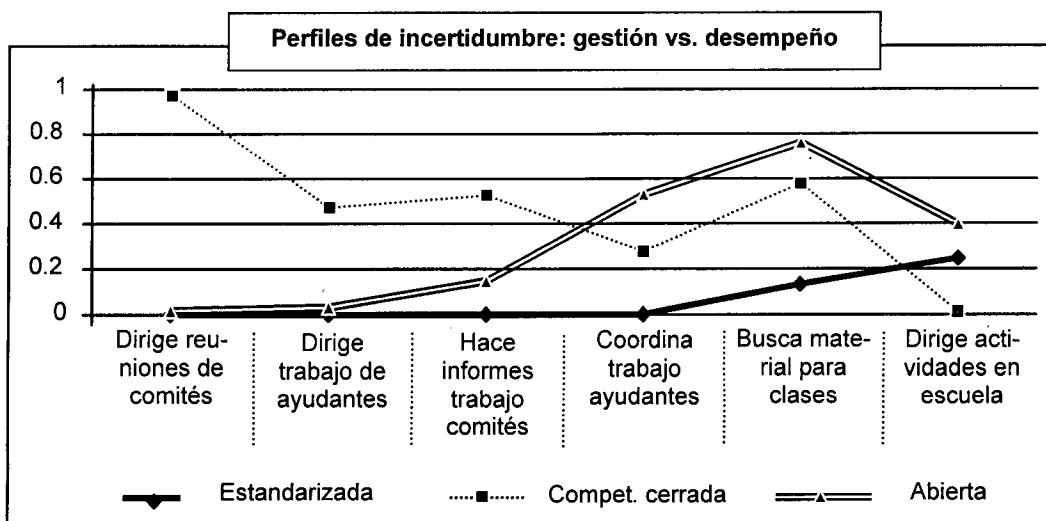
3.3.2 A continuación, aparecen en una tabla los p-valores de las asociaciones entre las tres maneras de analizar el desempeño, y los factores relativos a participación en los comités encuestados durante la evaluación de impacto del PUEBP. A la derecha de cada p-valor [p-v] escribimos qué tipo de asociación puede sugerirse con base en la herramienta estadística adoptada, en los instrumentos construidos y en los datos recogidos. Luego se grafican estos valores para visualizar el comportamiento de las distintas modalidades de la participación en comités, así como el comportamiento de las distintas maneras de analizar el desempeño.

Factores -comité	Desempeño					
	Estandarizado		Competencias-C		Competencias-A	
	p-v	Asociación	p-v	Asociación	p-v	Asociación
Biblioteca	0.001	Alta-Negativa	0.169	No	0.356	No
Rincones y Guías	0.001	Alta-Negativa	0.008	Alta-Positiva	0.088	Moderada-Negativa
Agropecuarias	0.079	Moderada-Negativa	0.460	No	0.634	No
Recreación	0.297	No	0.789	No	0.387	No
Aseo	0.002	Alta-Negativa	0.549	No	0.384	No
Decoración	0.001	Alta-Negativa	0.279	No	0.063	Moderada-No lineal
Otro	0.001	Alta-Positiva	0.757	No	0.283	No



A continuación se muestran los niveles de incertidumbre (p-valores) de la asociación entre los desempeños y los factores considerados de gestión; de esta tabla se elabora también la gráfica correspondiente:

Factores-gestión	Desempeño					
	Estandarizado		Competencias-C		Competencias-A	
	p-v	Asociación	p-v	Asociación	p-v	Asociación
Dirige reuniones de comités	0.001	Alta-Negativa	0.973	No	0.017	Moderada-Negativa
Dirige trabajo de ayudantes	0.001	Alta-Negativa	0.472	No	0.032	Moderada-Negativa
Hace informes trabajo ctés.	0.001	Alta-Negativa	0.526	No	0.147	No
Coordina trabajo ayudantes	0.001	Alta-Negativa	0.277	No	0.529	No
Busca material para clases	0.132	No	0.578	No	0.760	No
Dirige actividades escuela	0.247	No	0.013	Moderada-posit.	0.396	No



3.3.3 A continuación se comparan las cifras del análisis nacional con las de grado, modalidad de escuela y DC, en relación con la participación.

#### Comités

	Nacional	4°	3°	Urbana	Nueva	Rural grad	Distrital
Estandarizada	0.048	0.038	0.436	0.011	0.268	0.253	0.503
Competencias-C	0.880	0.531	0.976	0.945	0.864	0.192	0.912
Competencias-A	0.661	0.831	0.746	0.191	0.721	0.451	0.712

En los comités, la desagregación por grado y por modalidad de escuela no mostró cambios en el tipo de asociación para los desempeños vistos desde los análisis por competencias. En cambio, para el análisis estandarizado, sólo se mantiene la asociación vista desde lo nacional, en grado 4° y en escuela urbana.

#### Gestión

	Nacional	4°	3°	Urbana	Nueva	Rural grad	Distrital
Estandarizada	0.005	0.279	0.001	0.164	0.001	0.001	0.009
Competencias-C	0.309	0.354	0.282	0.574	0.478	0.804	0.750
Competencias-A	0.490	0.058	0.679	0.769	0.128	0.008	0.069

En la gestión, la desagregación mencionada tampoco mostró cambios en el tipo de asociación para el desempeño visto desde el análisis por competencias-C. En cambio, para el análisis estandarizado, se mantiene la asociación vista desde lo nacional, en grado 3° y en escuela nueva; sin embargo aparecen asociaciones en sentido contrario en escuela rural graduada y en el DC. De otro lado, el desglose permite visualizar asociaciones para el análisis por competencias-A en grado 4°, en escuela rural graduada y en el DC, todas en sentidos distintos.

### 3.4 Un análisis parcial

☛ Cinco perspectivas parecen vislumbrarse [cf. §3.1] frente a la participación:

- Con la nueva *Constitución*, se promueve la idea de una educación gratuita<sup>49</sup> en cuya gestión participará la comunidad. Pero faltaba reglamentarlo.
- Durante el comienzo de la evaluación de la calidad de la educación no son relevantes aspectos relativos a la participación [cf. MEN, 1992]. En ese enton-

<sup>49</sup> Claro que, simultáneamente, se dice: «la educación será gratuita en las instituciones del Estado, sin perjuicio del cobro de derechos académicos a quienes puedan sufragarlos» [art.67].

ces, la educación colombiana todavía estaba bajo el régimen de la *Renovación curricular*, para cuyos documentos el asunto de la participación no era tan importante.

- Con la Ley 115 —que reglamentaba la *Constitución*—, se promueve la idea de que el niño está en el centro del dispositivo educativo y que la comunidad debe participar. Luego, el Decreto reglamentario 1860 precisa una serie de mecanismos para esa participación.
  - Recientemente, la mayoría de instancias entiende la participación como actividades de los miembros de la comunidad educativa, dirigidas a una gestión entendida como administración (incluso exportable de instituciones “exitosas” a otras instituciones [DNP, 1999:42]). En el *Plan sectorial de educación* del DC [cf. SED, 1998], esta tendencia se hace más clara, hasta el punto de considerar la gestión-administración como uno de los factores fundamentales que se asocian con los resultados, y de invertir \$30 mil millones en “eficiencia”. Es exactamente la misma postura del DNP, cuando plantea que entre los cinco factores que se han mostrado ligados al mejoramiento de la calidad educativa están la cualificación de los métodos de gestión, la oportunidad de la información y la asignación eficiente de los recursos [Ávila: 1999:37]; en ninguno de los dos casos queda claro con base en qué estudios se hacen tales afirmaciones [sobre esto, cf. §7.1].
  - Y en el actual contexto escolar [cf. Posada, 1999], la idea de participación se entiende más en la dirección de una formación en valores, democratización de la escuela, búsqueda de la paz... los maestros entienden el concepto de “comunidad educativa” en relación con: el contexto, la convivencia y la participación, todo ello como forma de afrontar las condiciones *específicas* de las escuelas.
- ❁ Pese a todo esto, en la evaluación de impacto del PUEBP la participación se reduce a la pertenencia en comités y al involucramiento en ciertas actividades escolares, preguntado de manera que, a) la gestión mezcla actividades propiamente operativas con actividades en las que habría que administrar algo (no son del mismo nivel llevar los materiales al salón de clase que dirigir reuniones de comité); incluye, en las actividades “administrativas”, la “dirección” y la “coordinación”, sin que sean claras las diferencias (o, mejor, las relaciones en el marco de las cuales se permite diferenciar entre dirigir y coordinar); introduce, entre las opciones que al niño se le presentan, una actividad que incluye a las otras (o sea, las opciones constituyen una clasificación caótica). Y b) en los comités se incluye unos que serían propiamente “académicos” y otros que no lo son. Es decir, los

instrumentos parecen recoger muy poco la expectativa, tanto oficial como de la comunidad educativa; y parecen no tener una concepción clara sobre la naturaleza de los efectos que la participación podría tener en el desempeño de los niños (es tanta la separación, que la actividad de *escribir informes de comité* no se encuentra asociada con la parte de la prueba de lenguaje que indaga justamente por la escritura de los niños, a escala nacional, mientras que a escala distrital se encuentra asociada *negativamente*).

- ❁ La lectura de los resultados frente a la participación está limitada a las condiciones descritas y a la manera como se las interroga. No hay ninguna sensibilidad ante el hecho que los diversos contextos de la interacción están codificados y que, por ende, no es lo mismo “participar” en unos que en otros; no obstante, las preguntas —como se vio— se limitan a verificar la participación o no en comités y en actividades, sin explorar sus respectivos condicionamientos comunicativos<sup>50</sup>.
- ❁ Los comités que muestran asociaciones, muestran asociaciones *negativas*, es decir, algo contrario a lo esperado. Entonces, ante el “hecho” (numérico, claro está) de que a los que no participan les va mejor en la prueba de desempeño en lenguaje, ¿habría que trazar una política educativa de desestimulación de la participación en los comités? No; lo que llamamos una “asociación” (al menos moderada) es la disminución de una incertidumbre; o sea, una indicación de que podría haber una relación a investigar entre factores. Que aumente o disminuya la incertidumbre no es automáticamente una verdad sobre los aspectos cuya relación se indaga; con mayor razón cuando se sabe, de un lado, las limitaciones de las preguntas frente al aspecto de “participación” y, de otro, que la muestra con la que estamos trabajando es una de entre muchas posibles. Por su parte, en gestión, más de la mitad de los casos no muestra correlación estadística. Esto no quiere decir necesariamente que la dependencia entre las variables no exista; también puede afirmarse que la manera de indagarla (el instrumento de factores) y de analizarla (la metodología estadística seguida aquí) no es la mejor; o una combinación de ambas... En el peor de los casos, se trataría de que las cifras van en el sentido de que *esa* participación indagada es contraproducente... pero no es la única.
- ❁ Así, la hipótesis de que la participación influye positivamente en el desarrollo de la competencia comunicativa no puede convertirse en su contraria (“hay que acabar con los comités”), ni simplemente desestimarse por estos resultados, pues podemos mantenerla a la espera de que indagaciones distintas sobre este aspecto

---

<sup>50</sup> En la evaluación censal del DC no se indaga en este sentido, pues no hay preguntas sobre la participación de los estudiantes en la vida escolar.

den otras luces. En todo caso, no se trataría solamente de declarar si se participa o no (independientemente de que se disponga de una extensa lista de comités), sino fundamentalmente del *tipo* de participación; es decir, se hace necesario investigar *cómo* participan los niños, *cómo* solicita y orienta la escuela la participación, qué relaciones se generan entre niños, entre los niños y los maestros, y entre los niños y la comunidad, sobre la base de la participación<sup>51</sup>. Podría “nombrarse” a los niños en las distintas funciones, sin que esto transforme las condiciones de su inserción en la escuela. Además, este aspecto se articula con otros, de manera que la falta de participación podría suplirse con otro tipo de actividades; sería muy interesante, por ejemplo, investigar cómo se realizan efectivamente las actividades que en las cartillas de escuela nueva se denominan “D”, que deben llevarse a cabo con miembros de la comunidad.

- ❁ Investigar el *cómo* de la participación no autoriza a hacer promedios, pues dos participaciones “en el comité de biblioteca”, por ejemplo, no significan lo mismo en dos contextos distintos, de modo que no son homogéneas, no se pueden sumar, no se pueden promediar.
- ❁ Como la participación en comités está agregada en una serie grande de factores, tal vez el evaluador se ve obligado a hacer sobre ella una pregunta muy general, con la cual es complejo determinar las maneras específicas de llevarse a cabo. Sin embargo, la concepción expresada en las preguntas se satisface solamente con la idea del “sí” o el “no”: es el caso de las dos preguntas que la evaluación de impacto del PUEBP hizo sobre gobierno escolar: «¿Votaste en las pasadas elecciones para presidente?» [pregunta N° 10] y «¿Votaste en las pasadas elecciones para gobierno escolar, es decir para elegir personero y representante al Consejo Directivo?» [pregunta N° 11].
- ❁ Lo curioso es que para esta pregunta sobre el voto en las elecciones de gobierno escolar, con semejante generalidad, se muestre asociación alta para el análisis estandarizado (p-valor de 0.001), con tendencia positiva: los que votaron avanzan a niveles más altos, los que no votaron tienden a quedarse en niveles bajos<sup>52</sup>.
- ❁ Es esperable que el sentido de “gobierno” cambie según la modalidad de escuela (escuela nueva, por ejemplo, tiene gobierno escolar antes de que la reciente le-

---

<sup>51</sup> En este sentido, se esperaría que las evaluaciones en el DC consulten los resultados de otras evaluaciones, con el fin de mejorar los instrumentos. Por ejemplo, de precisar más la naturaleza de la participación —si se la considera relacionada con la competencia comunicativa— y preguntar en consecuencia.

<sup>52</sup> Para ambos análisis por competencias, no hay asociación (Competencias-C, p-valor de 0.888 y Competencias-A, p-valor de 0.254).

gislación lo generalizara), según la edad (los investigadores postulan niveles de comprensión de lo social [cf. Carretero, 1993:§4.2]), según la región (el país vive fenómenos políticos que se presentan en distintas intensidades de acuerdo con la región), etc. Estas diferencias se pierden cuando todos los estudiantes deben dar un “sí” o un “no”; no habría, en consecuencia, manera de investigarlas con instrumentos que se construyen sin tenerlas en cuenta.

Podemos concluir que, como cita San Fabián [1996:200], «una cosa es hablar de participación y otra muy distinta aunar metas y proponer soluciones que integren la acción educativa de los ambientes escolares y familiares». Posada [1999:20] muestra que la participación en la escuela no queda garantizada ni porque esté de moda hablar de ella, ni porque esté en la norma; no se trataría de un agregado curricular (como se ufanaba Samper [cf. §3.1.2]) ni de una forma específica de la “moderna” gestión escolar (como en los planes de desarrollo para educación de Pastrana y del DC), sino «de un elemento inherente de la formación democrática»; de ahí que Posada traiga a cuento investigaciones que muestran que los consejos de gobierno escolar son formalismos que no inciden realmente en las decisiones de la escuela y que funcionan de manera paralela a la cotidianidad. «Lo que se ha dado es un modelo de participación esporádico y puntual, que la mayoría de las veces se manifiesta, y agota, en la votación y en la delegación de responsabilidades» [Posada, 1999:22]. Así entendida, no podría hablarse de participación en abstracto (ni pretender que su investigación masiva es suficiente, por fuera de perspectivas más “etnográficas”); ésta tendría diversas facetas, no necesariamente jerarquizadas; pensamos en al menos las siguientes<sup>53</sup>:

- No participación; debida, ente otras cosas, al tipo de convocatoria que la escuela le hace a sus agentes.
- Recepción; la escuela convoca para tratar temas disciplinarios (para regañar), dar información, entregar resultados.
- Legitimación; convoca para avalar decisiones ya tomadas.
- Ejecución; llamado a asumir responsabilidades económicas, académicas, administrativas y de gestión, decididas por otros.
- Inter-gestión e inter-decisión; participación comunicativa, no estratégica, como en los otros casos.

---

<sup>53</sup> Basados en Posada [1999:21], quien a su vez se basa en Quintina y Antúñez.

- ❁ Curiosamente, el único comité que muestra asociación *positiva* en el análisis estandarizado es el rotulado como “otro”; es decir, el espectro de comités *indagado* no cubre los que responden a la asociación esperada. Esto mostraría:
- o bien que el instrumento con el que se indagó sobre los comités no permite identificar el tipo de participación que influiría positivamente en el desempeño. Así, posiblemente, no sea con arreglo a las nociones explícitas de la escuela (“rincones y guías”, por ejemplo, pertenece a la jerga de escuela nueva) como se ha de indagar la participación, sino con categorías que teoricen sobre la idea de “participar” en relación con la idea de “desempeño”, particularmente en el área de lenguaje;
  - o bien que —pese a las limitaciones instrumentales— la escuela no permite la participación y, en consecuencia, hay una asociación negativa entre lo que la escuela llama “participar” y el desempeño. Faltaría saber, entonces, a qué se refieren los niños cuando llenan en el cuestionario la opción G.: pertenecer a un comité diferente de los que han sido listados de A. a F.;
  - por supuesto que habría una tercera opción: la articulación entre las limitaciones instrumentales y la dificultad para identificar mediante las nociones escolares aquello que incide en la competencia comunicativa; aun siendo nociones formales, algo ocurre en relación con la participación, pero, ¿qué?
- ❁ Podría creerse que mientras más puntual sea el análisis (por ejemplo, cuando se hacen desgloses por curso, por modalidad de escuela, o para el DC), es posible encontrar más asociaciones, eliminar el ruido a favor de una variable; es el caso en el DC, donde se muestran más asociaciones que en el análisis nacional, para el caso de los comités. Pero esto no se cumple para todos los casos: a veces el análisis globalizado muestra más certidumbre (de que la hipótesis de no asociación sea falsa) que el análisis del mismo desempeño desglosado por curso o por modalidad de escuela (cf. el caso de participación en comités contra desempeño en la parte abierta, globalizado y por grado); a veces cambia la dirección de las asociaciones (por ejemplo, mientras lo nacional muestra una asociación positiva, el DC muestra dos asociaciones negativas)... Estos casos se conocen en estadística como Paradoja de Simpson. A veces el análisis globalizado muestra la misma incertidumbre que el análisis del mismo desempeño desglosado por curso o por modalidad de escuela (cf. el caso de la participación en comités contra desempeño en la parte cerrada, desde los análisis estandarizado y por competencias). Pero, entonces, ¿cuál resultado escoger cuando, por ejemplo, una asociación (gestión vs. estandarizado), que resulta positiva cuando los factores de gestión están aglutinados, se muestra negativa cuando están desglosados? Pensamos que no se



trataría de escoger, pues podría producirse una manera no sistemática de selección de resultados: cuando conviene la correlación estadística positiva, se desechan las negativas, se ignoran las ausencias de correlación, etc. No se trataría de una cacería ciega de factores hasta obtener los que obedezcan a un interés preestablecido, sino de poner en forma explícita la conceptualización (en este caso: ¿qué se entiende por participación y por gestión?) y desde allí confrontar nuestras aproximaciones a aquello que queremos entender.

- Veamos el siguiente caso: en la gestión, la desagregación no mostró cambios para la asociación con el desempeño por competencias-C; para el análisis estandarizado, se mantiene la asociación vista desde lo nacional, en grado 3° y en escuela nueva; sin embargo aparece una asociación en sentido contrario en escuela rural graduada. Si esto se leyera de manera parcial, se diría que hay que impulsar la modalidad de escuela nueva en toda la zona rural (para que cambie de dirección la asociación: de negativa a positiva). Pero recordemos que los mismos datos autorizarían a decir que no vale la pena, porque cuando los niños pasan a grado 4° la dependencia entre variables desaparece. En el análisis por competencias-A mientras en lo nacional no se mostraba asociación, aparecen asociaciones en grado 4° (positiva) y en escuela rural graduada (no lineal).
- Hay indicios para profundizar en la hipótesis de que la gestión por parte de los niños es un factor que incide en el desarrollo de su competencia comunicativa; de todas maneras, sería necesario hacer investigaciones más puntuales sobre la manera como se decide la gestión, las relaciones que ello promueve a la existencia, la incidencia real que la gestión del niño tiene. No es desestimable la posibilidad de que los niños gestionen asuntos de poca monta y en el contexto de jerarquías y relaciones determinadas de antemano por la escuela y el docente; un tipo de participación como ésta podría no coadyuvar al desarrollo de la competencia comunicativa (como parece verse en el DC), o a hacerlo en una magnitud inferior a la esperada.
- Esta “prudencia” es absolutamente necesaria, pues con los datos disponibles nos veríamos obligados a decir que la gestión por parte de los niños poco contribuye a mejorar sus desempeños en lenguaje y que, en consecuencia, los esfuerzos conducentes a establecerla son pérdida de tiempo o, acaso, de dinero. No; los datos de los que disponemos nos autorizan a decir que esta idea de “gestión” no basta para ser considerada como factor asociable al logro cognitivo de los estudiantes; en cambio posiblemente sí un *tipo* de gestión que involucre realmente al niño, que permita tomar decisiones importantes, que tienda a generar nuevas relaciones entre los agentes educativos, etc. Por supuesto, eso lo demostraría una investigación de otro orden, la cual no podría “recomendar” la gestión ideal para

que los desempeños mejoren, sino que establecería las condiciones bajo las cuales se dan efectivamente las diversas formas de lo que las comunidades educativas llaman de esa manera. Así podría saberse las condiciones de posibilidad de cualquier recomendación de política educativa. A escala distrital, por ejemplo, se habla de gestión de una manera que excluye prácticamente a los estudiantes [cf. SED, CORPOEDUCACIÓN Y Misión Social-DNP: §1.1].

- ✿ Hay muy pocas coincidencias entre los tres tipos de análisis del desempeño. El análisis estandarizado es el que más se presta para hablar de asociación estadística entre participación en comités y desempeño en el área de lenguaje; no obstante, las asociaciones que evidencia son negativas (salvo “otro”, pero justamente es aquello sobre lo que nada puede saberse). Pese a que los análisis por competencia son más inciertos (en comités y desempeño), sus perfiles van casi de manera paralela; podría interpretarse que indagan de manera similar.
- ✿ Sólo para el comité “rincones y guías” los tres tipos de clasificación del desempeño en lenguaje muestran asociaciones, pero sólo el categorizado por competencias-C muestra asociaciones positivas. Hay aquí, no obstante, algo muy particular: rincones y guías es propio de la metodología de escuela nueva [cf. §3.1.1]; ¿por qué preguntar, entonces, a niños (que constituyen las dos terceras partes de la población encuestada) sobre una estrategia que no pertenece a su modalidad de escuela? Este sesgo tal vez concuerda con el interés oficial por destacar los resultados de escuela nueva; por ejemplo, el Banco Mundial et al. [1996] dicen que diferentes evaluaciones han encontrado que el modelo de escuela nueva ha logrado, entre otros, un rendimiento promedio superior al de las escuelas tradicionales; así mismo, el *Plan educativo para la paz* [DNP, 1999:30] dice que «los resultados de la Escuela Nueva son mejores que los de la escuela tradicional rural». No obstante, en §2.3 mostramos que bajo *ninguno* de los análisis del desempeño escuela nueva es mejor: desde el análisis estandarizado, ocupa el último lugar de las tres modalidades de escuela; desde el análisis por competencias-C, no hay diferencia en el desempeño de las modalidades; y desde el análisis por competencias-A, escuela nueva ocupa el segundo lugar.
- ✿ El sesgo anterior es un error metodológico, pues “rincones y guías” no es algo en relación con lo cual la población encuestada esté expuesta de manera homogénea; en otras palabras, con este tipo de mecanismo se favorecería —si las cosas funcionaran como dicen los objetivos y si las pruebas fueran objetivas, sin más— a un sector de la población encuestada que de antemano se sabe que tiene ventaja en relación con el aspecto indagado. Dadas estas condiciones, se esperaría que en las modalidades urbana y rural graduada —en las que la mencionada estrategia no se utiliza— no hubiera correlaciones con el desempeño en el área de lenguaje,

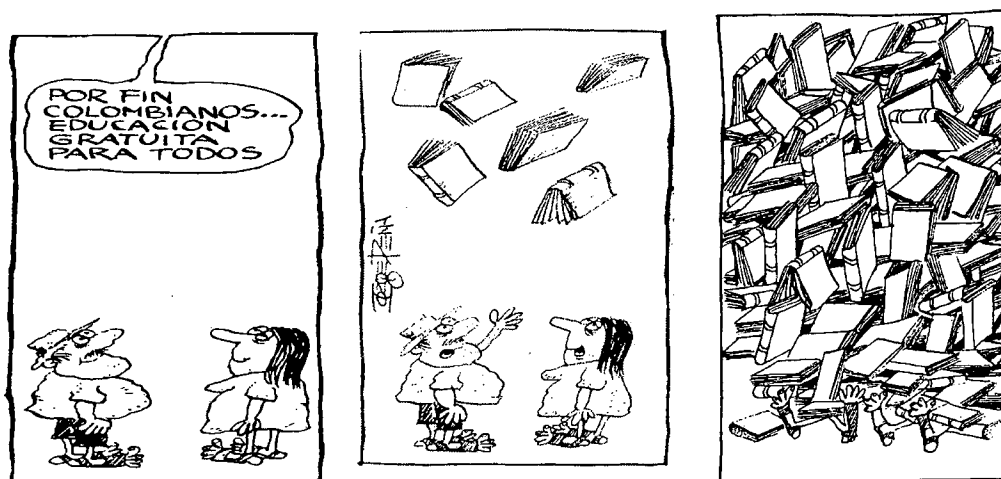
mientras que en escuela nueva sí. Sin embargo, el cruce con el análisis estandarizado del desempeño en lenguaje muestra correlación, y negativa, pero en las modalidades donde se supone que no hay rincones y guías; el cruce con el análisis por competencias-C muestra dos asociaciones moderadas positivas en ambas escuelas rurales (una de las cuales se supone que no tiene rincones y guías); y el cruce con el análisis por competencias-A muestra una sola correlación, justamente en escuela nueva, pero negativa. A continuación se transcriben los p-valores para cada caso:

		Desempeño	Asociación con "rincones y guías" (p-valor)
Modalidad	Urbana	Estandarizado	0.001
		Competencias-C	0.928
		Competencias-A	0.711
	Nueva	Estandarizado	0.187
		Competencias-C	0.021
		Competencias-A	0.095
	Rural graduada	Estandarizado	0.093
		Competencias-C	0.065
		Competencias-A	0.638

- ☛ De igual manera, la gestión, que es más propia de la modalidad de escuela nueva, muestra más asociaciones en la escuela rural graduada.
- ☛ En el análisis de la gestión, los perfiles de estandarizada y abierta, de un lado, están más cercanos entre sí; parecerían percibir de manera semejante la correlación explorada. Y, de otro lado, muestran más asociación con las actividades de gestión relacionadas con "dirección de...". La interpretación del desempeño desde la perspectiva de la competencia-C sólo muestra una asociación, en relación con "dirigir actividades en la escuela"; curiosamente, las otras actividades de dirección estarían incluidas en ésta, en tanto son de la escuela y, no obstante, no muestran asociación en esta interpretación del desempeño. Vale la pena destacar que es ésta la única asociación positiva. O sea que, salvo este caso, podría decirse que *para las cifras disponibles el desempeño no depende de esta manera de preguntar lo que hemos agrupado con el nombre de "gestión"*.
- ☛ En relación con la asociación entre el desempeño y la participación en comités, mientras desde las competencias la mirada sobre lo distrital no suministra más información que la nacional, desde la estandarización sí hay diferencias, pues en el DC no se muestra asociación. Y en relación con la asociación entre el desempeño y gestión, el análisis para el DC tiene más coherencia con el nivel nacional

cuando los factores están desagregados; con los factores colapsados, la tendencia del DC es contraria a la nacional para el análisis estandarizado, ninguno presenta asociación para el análisis por competencias-C y sólo el distrital muestra asociación para el análisis por competencias-A.

## 4. Libros



Jorge Peña\*

### 4.1 Los libros como factor asociable al desempeño

La bibliografía, la investigación y la política educativa (sin ser excluyentes entre sí) destacan el valor de los libros en la educación; Piñeros y Rodríguez [1999:41] encontraron que «todos los estudios sobre este insumo demuestran su importancia en el proceso de aprendizaje y, por tanto, su significancia estadística sobre el rendimiento académico».

Sin embargo, en lugar de aceptar esto como un hecho, nos parece que sería útil saber de qué depende tal acuerdo, pues las distintas fuentes que hacen dicho planteamiento no lo hacen necesariamente con los mismos presupuestos conceptuales, y las investigaciones no lo hacen con las mismas metodologías ni con los mismos instrumentos; unos y otros no lo hacen con idénticos intereses. Veamos algunas de estas fuentes:

---

\* Profesor Universidad Nacional de Colombia, Facultad de Artes, sede Bogotá. La tira cómica fue publicada en el periódico UN N°1 de la Universidad Nacional de Colombia, Agosto 15 de 1999.

- ☞ Desde que apareció el programa de escuela nueva, la entrega de materiales ha sido una condición fundamental cuyo fundamento no se ha explicitado; sencillamente se da por hecho que *debe haber* material impreso para los capacitadores (supervisores), para el maestro y para los estudiantes, con el fin de dar contenidos, desarrollar destrezas, investigar y participar. La estructura de la guía era la que garantizaba estos propósitos, pues para cada tema siempre empezaba con el reconocimiento de lo que el estudiante sabía al respecto, seguía con la información académica respectiva, luego con una interiorización y, por último, con una confrontación con su medio.
  
- ☞ El Banco Mundial ha llegado a afirmar que inciden más en el mejoramiento de la calidad de la educación los libros para los estudiantes, que la capacitación de docentes: «Las pruebas disponibles indican que las inversiones en libros de texto producirán ganancias para el aprendizaje, y que es más probable que esto ocurra como resultado de una inversión en libros de texto que como resultado de otras intervenciones educacionales, como la capacitación docente» [Heyneman et al., 1991:134].
  
- ☞ La reunión de expertos latinoamericanos en el marco del Proyecto de renovación de las prácticas pedagógicas en la formación de lectores y escritores, realizada en Bogotá entre los días 5 y 9 de Octubre de 1993, hace a los gobiernos una serie de recomendaciones; entre éstas está la de garantizar a corto plazo «la entrega de acervos bibliográficos de calidad a todas las escuelas primarias públicas de la región, así como la entrega gratuita de textos escolares a todos los niños que asisten a dichas escuelas» [CERLALC, 1994:35].
  
- ☞ El proyecto «Podemos leer y escribir», del Centro Regional para el Fomento del Libro en América Latina y el Caribe [CERLALC], apoya a los países de la región que se proponen mejorar la calidad mediante un énfasis en lenguaje y matemáticas, para lo cual dedican esfuerzos a la dotación de libros. De ahí los *acervos bibliográficos* que ha regalado el gobierno mexicano a los países participantes: «Hasta la fecha se han entregado 1035 acervos bibliográficos latinoamericanos en 9 países: Colombia, Costa Rica, Chile, Cuba, El Salvador, Honduras, Nicaragua, Uruguay y Venezuela» [Hederich y Charria, 1999:66].
  
- ☞ En la investigación del MEN y CERLALC sobre el uso del texto escolar en la educación básica primaria se plantea [Venegas, 1993:23] que el Plan Nacional de dotación de Textos Escolares (componente del PUEBP) fue un medio para asegurar el mejoramiento cualitativo de la educación; que la dotación de textos tiene un potencial educativo del que los docentes deben ser conscientes, pues puede constituir un enorme paso en la solución definitiva del problema educativo; y que

el programa desplegó enormes esfuerzos de todo tipo para asegurar que en la primaria oficial los docentes dispusieran de textos escolares suficientes. Además, la autora usa como referencia la idea de Daza y Jiménez [1977] de que «la carencia de textos escolares en grandes sectores de la población estudiantil del país es un problema que tiene incidencia en la calidad de la enseñanza impartida en tales sectores» [Venegas, 1993:23], y anota que «en Colombia, más de la mitad de los niños en edad escolar carecen de texto propio» y que algunos docentes subutilizan y subvaloran los recursos bibliográficos por pensar que son innecesarios, inadecuados o reemplazables por los cuadernos.

- ☞ El plan de desarrollo del gobierno de Gaviria en educación —*Plan de Apertura Educativa*—, entre los objetivos específicos de la política educativa para el período 1991-1994 tiene el de “elevar la calidad de la educación”, lo cual significa una serie de acciones entre las cuales está «Proveer gratuitamente a todos los niños de la primaria pública textos y material pedagógico, y a los docentes guías didácticas y material de apoyo. Dotar a los establecimientos educativos oficiales con laboratorios básicos, bibliotecas y material pedagógico» [DNP, 1991:26]. Y, más adelante: «Se harán las inversiones necesarias para la provisión de juegos didácticos, textos, manuales, laboratorios y demás recursos necesarios para el logro de los objetivos» [:35].
- ☞ En Colombia, desde el comienzo de la evaluación de la calidad educativa, se ha buscado una asociación entre textos y logro cognitivo. En el primer informe del SINECE se dice que en las escuelas públicas urbanas se destaca, entre otras, la asociación entre logro y la relación número de textos de área/alumno [MEN, 1992:88], que en el área de lenguaje se destacan variables del docente entre las que se cita la disponibilidad de textos [:94], que entre las variables asociadas que deben ser investigadas en profundidad está la disponibilidad de textos [:98], y que se recomienda el impulso a la dotación de textos, en tanto es un factor decisivo de la calidad [:100]. Y en uno de los últimos informes [MEN, 1997b:45], se plantea que entre las variables más importantes asociadas al “logro promedio” en el sector oficial urbano, está «el factor dotación de la biblioteca como complemento de la dotación de material de apoyo pedagógico, [que] se asocia positivamente con un mayor logro promedio. Este factor se encuentra representado por tres indicadores: la biblioteca tiene enciclopedias, diccionarios y cuentos».
- ☞ Por su parte, la Ley 115, en su artículo 102, dice que, a partir de 1995, el Gobierno Nacional «destinará anualmente para textos y materiales o equipos educativos para uso de los estudiantes de las instituciones educativas del Estado o contratadas por éste un monto no menor a la cantidad resultante de multiplicar el equivalente a un salario mínimo legal mensual, por el número total de los educadores

oficiales». Y en su artículo 141, dice que los establecimientos educativos deberán contar, entre otros elementos, con una biblioteca o —en caso de poblaciones de menos de 20.000 habitantes— con convenios con bibliotecas públicas o instituciones cercanas que permitan cumplir tal función.

☞ El Decreto 1860 de 1994, reglamentario de la Ley 115, en su artículo 42, dice que los textos escolares deben seleccionarse y adquirirse «para ofrecer al alumno soporte pedagógico e información relevante sobre una asignatura o proyecto pedagógico y guiar o encausar al estudiante en la práctica de la experimentación y de la observación, apartándolo de la simple repetición memorística». Y frente a los textos escolares que el plan de estudios establezca, dice que se hará mediante el sistema de bibliobanco: «el establecimiento educativo pone a disposición del alumno, en el aula de clase o en un lugar adecuado, un número de textos suficientes, especialmente seleccionados y periódicamente renovados». La biblioteca se conformaría así: «los bibliobancos de textos escolares y los libros de consulta, tales como diccionarios, enciclopedias temáticas, publicaciones periódicas, libros y otros materiales».

☞ Después de promulgada esta Ley 115, se convocó la *Misión ciencia, educación y desarrollo* (en Julio de 1994, los comisionados reportaron su documento), que también se inscribe en la idea de la afectación positiva que la entrega de textos tiene en la educación; a propósito de esto, en el acápite de «Recomendaciones acerca de la educación», punto II, aparte “reformar la educación básica”, dice:

*Garantizar la disponibilidad de distintos tipos de textos. Los colegios de la educación oficial dispondrán de un número y diversidad suficiente de textos de calidad para prestárselos a cada uno de los niños durante todo el año escolar. No se trata únicamente de los libros de texto usuales en nuestras escuelas y colegios, sino también de cualquier tipo de impresos que estimulen la curiosidad, propicien el diálogo y la investigación y hagan sentir el placer de aprender [Misión..., 1997:200].*

☞ El plan de desarrollo del gobierno de Samper (*El salto educativo*), plantea 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» [DNP, 1994:11]. Entre los objetivos y metas de este Plan, está mejorar la calidad de la educación básica, para lo cual, entre otros, «se dotará con textos escolares a todos los estudiantes de establecimientos oficiales y se pondrán en funcionamiento aproximadamente 2100 bibliotecas y centros de recursos educativos» [:18]. Y entre las estrategias y pro-



gramas para universalizar la educación, está la «dotación de textos, bibliotecas y material educativo» [:23].

- ☞ De esta manera, la administración Samper distribuyó, en 1996, «casi 3 millones 500 mil cartillas para lenguaje y matemáticas, y luego en 1997 otros 15 millones 200 mil juegos de textos que comprendían también ciencias naturales y ciencias sociales, por un valor superior a los \$15.000 millones» [Samper, 1998:264]. Dice el documento que más de 7 millones de niños de la educación pública serán beneficiados. «Su impacto en la calidad de la educación se empieza a evaluar, no obstante, experiencias internacionales indican que su adquisición se justifica plenamente porque los estudios muestran que los más importantes predictores de éxito académico son los textos, especialmente en las poblaciones más pobres» [:264].
- ☞ Así mismo, durante 1994, destinó a textos la tercera parte de un préstamo de 90 millones de dólares para el programa PACES, que tenía como propósito contribuir al mejoramiento de la calidad de la educación secundaria; de fondo había una presuposición sobre cómo incidir positivamente en dicho mejoramiento, pues el préstamo se invertía en a) becas para los estudiantes, b) inversión en infraestructura de las instituciones escolares y c) textos. Para llevar a cabo la inversión en textos, el MEN expidió la Circular Ministerial N° 28 [MEN, 1994] con el fin de que las instituciones educativas tuvieran unos criterios para la selección de textos mediante formato del Fondo de Cofinanciación para la Inversión Social [FIS, 1994].
- ☞ El Programa Nacional de Incentivos (1996 a 1998), entre sus justificaciones argumentaba: «Nuestros esfuerzos deben centrarse en asegurarle a los estudiantes los materiales más elementales para el funcionamiento adecuado de las clases. Los alumnos deben tener dónde sentarse cómodamente, acceso a textos o guías de aprendizaje, con qué y dónde escribir».[MEN, 1996a:5].
- ☞ El *Plan decenal de educación*, en su segunda estrategia, literal h. (sobre producción y distribución de textos y otros materiales), para elevar la calidad de la educación en los próximos 10 años, dice que «Corresponde a todas las instituciones que conforman el Sistema Educativo Nacional, a las entidades territoriales y a la empresa privada, en proyectos concertados, la producción y distribución de textos, libros y materiales de trabajo pedagógico, buscando garantizar que los niños y niñas de menores recursos puedan disponer gratuitamente de ellos para su adecuada formación» [MEN, 1996b:25].

- ☞ Así mismo, la Corporación para el Desarrollo de la Educación Básica (CORPOEDUCACIÓN), interesada en informar sobre la investigación en los factores que inciden sobre el desempeño de los estudiantes, plantea: «Hasta ahora, los estudios han contribuido a reorientar las políticas educativas en el nivel central. En efecto, con los resultados de las primeras evaluaciones el MEN se propuso, en sus planes de desarrollo, impulsar acciones como las siguientes: a) inversión en textos escolares, procurando que éstos lleguen a todas las escuelas y a cada uno de los alumnos [...]»<sup>54</sup>[Ávila, 1999:35].
- ☞ El Diagnóstico del Plan de desarrollo del actual gobierno (*Un plan educativo para la paz*) plantea: «Otros factores decisivos para la calidad de la educación que presentan deficiencias son, en primer lugar, la insuficiente dotación de textos, bibliotecas y medios pedagógicos [...]» [DNP, 1999:11]. En las Políticas del Plan dice: «El gobierno impulsará un programa que resalte la importancia del libro como instrumento de aprendizaje, de difusión del conocimiento, como medio indispensable para que los colombianos tengan mayor capacidad de lecto-escritura y, en general, para la transmisión de los valores de la cultura. Con este propósito se apoyarán y desarrollarán acciones que conduzcan a que cada día, cada colombiano tenga mayores posibilidades de tener y leer libros» [:39]. Y en los Programas motores, el de “Educación es Calidad” tiene entre sus proyectos el de Política del libro [:65].
- ☞ Finalmente, en el *Plan sectorial de educación, 1998-2001*, la Secretaría de Educación del DC —en el capítulo llamado “Estrategias y programas”, en el aparte de “Focalización de los recursos del Estado hacia la población de menores ingresos”, en el literal a., “Redistribución de recursos”— plantea: «se asignarán por institución las plantas de personal, así como los recursos necesarios para la dotación de materiales, textos y espacios físicos» [SED, 1998:31]. Para la dotación oficial, entonces, la SED diseñó un instrumento para la selección de libros de texto [SED, 1999a]. Y en el aparte “Calidad”, del mismo capítulo, literal e., “Nuevos espacios para la educación”, plantea que se construirán cuatro bibliotecas en zonas estratégicas de la capital, teniendo en cuenta distancia, acceso y seguridad; cada una «contará con una colección inicial de 30.000 libros, además de videos, discos, revistas y periódicos nacionales e internacionales, computadores y conexión a Internet. Para su administración, se convocará al sector privado y a las Cajas de Compensación» [SED, 1998:40].
- ☞ En cumplimiento de la obligación de dotar material escrito a las escuelas, la SED emitió el documento *Pretextos. Reflexiones para elegir libros de texto* [SED,

---

<sup>54</sup> Las otras acciones son las relativas a repitencia, tiempo de estudio y planteles con infraestructura mínima que ofrezcan toda la educación básica.



1999a], en el que los libros de texto son como “una prenda necesaria” [:1]; un objeto cuya adquisición nos obliga a “vitriniar” en determinada época del año [:1]. Se trata de una evidencia que ni siquiera se toma el trabajo de enunciarse; es un pre supuesto del documento, que es escrito por María Clemencia Venegas (“textos e investigación bibliográfica”), de quien ya reseñamos un documento [1993].


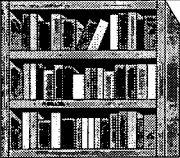

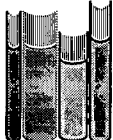
Todo este panorama muestra una convicción (proveniente o no de la investigación) en el sentido de la incidencia positiva que tendría el material escrito en el desempeño de los estudiantes; se dice que: informa; permite la participación; produce ganancias para el aprendizaje; sirve para una formación adecuada; cualifica la educación; estimula la curiosidad, el diálogo, la investigación; está asociado positivamente con el logro estudiantil; permite el funcionamiento adecuado de la clase; es instrumento de difusión y transmisión de los valores culturales; es predictor de éxito académico... En principio, también nosotros somos de la idea de que si un niño dispone de material escrito y otro no, el primero tendría ventajas para desempeñarse en las áreas de conocimiento en la escuela (aquí nos restringiremos al área de lenguaje, que es en la que disponemos de información). Es una hipótesis ligada a la idea de que la tradición escrita es el mecanismo para trascender el contexto particular hacia la cultura universal y que es ella y su “código” la que predomina en la escuela [cf. Mockus et al., 1994].

Entonces, la evaluación de impacto del PUEBP nos brindó información de primera mano para investigar al respecto. Veremos si las cifras obtenidas permiten validar la hipótesis.

#### **4.2 Algunos factores-libro asociables al desempeño en lenguaje**

Nos referiremos a los factores que tienen que ver con los textos y los libros, en la manera como la encuesta del MEN los indagó:

	Libros en la casa del estudiante (se refiere a libros y textos escolares).
	Libros recibidos por la escuela (se refiere a textos escolares, guías de escuela nueva y libros varios para la biblioteca).

	<p>Guías de escuela nueva (se refiere al material impreso que el MEN distribuye para las escuelas de la zona rural que ejercen la modalidad de escuela nueva).</p>
	<p>Libros para la biblioteca de la escuela (por la pregunta del cuestionario [transcrita en §4.2.2], suponemos que se trata de indagar por los libros distintos a los textos escolares).</p>
	<p>Textos escolares en general (se refiere a los textos escolares en las áreas de ciencias sociales, naturales, castellano, matemáticas y Guías de escuela nueva).</p>
	<p>Textos escolares en el área de lenguaje (de 1° a 9° grados).</p>

Para la comparación de las escalas nacional y distrital, nos referiremos solamente a los dos primeros factores-libro de la tabla anterior; de modo que, con esta decisión estamos enfrentando los posibles inconvenientes derivados de una muestra pequeña. No obstante, los factores del tercero al sexto de la tabla anterior son, como se verá, una desagregación del segundo.

#### 4.2.1 Libros en la casa

En la encuesta de factores asociables aplicada a los niños se les preguntó:

¿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

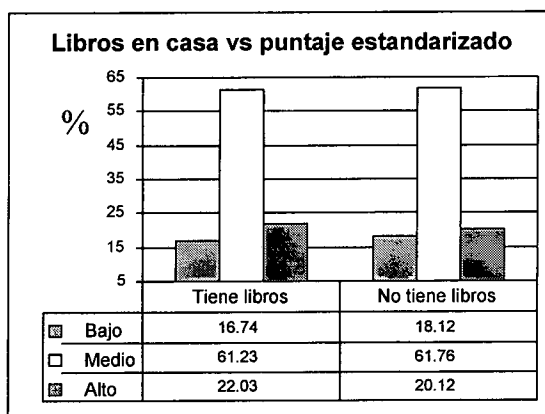
D Los textos escolares y más de 50 libros

Como en esta pregunta se hace una diferencia entre los “textos escolares” y los “libros”<sup>55</sup>, una indagación sobre éstos habría de considerar a los niños que responden la opción A como que no poseen libros en la casa (el material escrito disponible sólo

<sup>55</sup> En la pregunta que se reproduce en §4.2.2 también puede verse que se conciben los textos escolares y los libros como conjuntos excluyentes.

habría ingresado a casa por la obligación de dotar al niño de texto escolar). Pues bien, el 33.38% respondió de esa manera. Y como establecer los rangos de los literales B a D de la pregunta nos ha parecido complejo para niños con las edades de los que fueron encuestados (además de la omisión del literal C)<sup>56</sup>, los que respondieron alguna de esas opciones hemos considerado solamente que sí poseen libros en la casa: el 66.62%.

Con esta readecuación, la siguiente gráfica muestra el desempeño de los niños en el área de lenguaje, visto desde la estandarización, frente a la posesión o no de libros:



Como se observa, los perfiles son muy similares; por eso el p-valor es de 0.130.

De otro lado, analizar los desempeños desde las categorías por competencia, tampoco permite concluir sobre una incidencia positiva de la presencia de libros en casa; de igual manera, los perfiles son casi idénticos (en ningún caso, la diferencia alcanza a ser de un punto):

		Desempeño por competencias-C			
		Nivel 0	Lectura		
			Literal	Inferencial	Crítico inter-textual
	No tiene	0.93	24.72	60.18	14.17
	Tiene	1.23	24.24	60.72	13.81
	p-valor		0.677		

Casi lo mismo ocurre cuando miramos los desempeños en la parte abierta de la prueba, frente a la presencia de libros en la casa: para la coherencia local, la diferencia es de 0.18 (sí tiene libros en casa, 11.21%, y no tiene libros en casa, 11.03%); para la coherencia lineal, la diferencia es de 0.5 (sí, con 79.93%, y no, con 80.43%);

<sup>56</sup> En ese literal falta la palabra "libro", pero lo transcribimos tal como aparece en el formulario.

y para la coherencia global, la diferencia es de 0.29 (sí, con 6.13%, y no, con 6.42%); ante diferencias tan pequeñas y perfiles tan parecidos, el p-valor es de 0.505.

Ninguna de las tres perspectivas permite afirmar que el desempeño del niño en el área de lenguaje se vea afectado (ni positiva ni negativamente) por los libros que hay en su casa, pese a que la tercera parte de la población respondió que no tenía. Hay que tener “prudencia” para no hacer afirmaciones apresuradas, tales como que si no se ha probado que tener libros en casa favorece el logro de los niños en lenguaje (que se viera una tendencia a tener desempeños más altos entre los niños que tuvieran libros en casa, y más bajos entre quienes no tuvieran), entonces habría que de-stimular tal perspectiva.

Con los datos disponibles, no puede apuntarse una política educativa sobre el tema de los libros en la casa. No obstante, mantenemos la hipótesis de que este factor puede hacer un aporte al desempeño de los estudiantes en el área de lenguaje, lo que sería “detectable” con otro tipo de mirada, investigaciones más puntuales, que exploren un poco más allá de *cuántos* libros se poseen, que tengan en cuenta categorías más acordes con la relación entre escritura y competencia comunicativa (que es el objeto de la evaluación de impacto para el desempeño en lenguaje). En tal sentido, no podrían excluirse las revistas ni los periódicos (como sí hace explícitamente la pregunta): es previsible que, en ciertas circunstancias, impresos como los periódicos sean los *únicos* materiales escritos de un niño en su casa.

No se puede descalificar ningún tipo de texto para efectos de coadyuvar a desarrollar la competencia comunicativa, porque el libro no tiene una “esencia” que lo haría “bueno” o “malo”, sino que el sentido se produce en la interacción [Jurado y Bustamante, 1994:41-42]; no hay textos fáciles o difíciles, malos o buenos, sino lecturas diferentes.

En países como el nuestro, la cotidianidad familiar de la mayor parte de la población atendida por las escuelas públicas está alejada de la escritura [Hederich y Charria, 1999:67]; podemos inferirlo de las cifras manejadas en este estudio: la tercera parte de la población encuestada *dice* tener sólo textos escolares en casa (sin el requerimiento escolar, no habría libros). Pero la cifra podría ser mayor si algunos niños no escogieron la opción A, pensando en algún estereotipo como, por ejemplo, que se vería con mejores ojos a una familia que tenga libros.

Entonces, sería pertinente averiguar por el *sentido* de la lengua escrita para la población atendida por la escuela pública; por los espacios sociales generados de cierta manera al margen de la lengua escrita; no va a ser posible investigar este sentido y

esta condición de posibilidad solamente preguntando “¿cuántos?”, ni va a ser posible cambiarlos con una “dotación”. Y, para el caso de quienes dicen tener libros en sus casas, sería fundamental saber cómo son las prácticas domésticas de lectura y qué relación tienen con las lecturas escolares. Es necesario hacer distinciones de los tipos de libros que hay en las casas, saber sobre la utilización que tienen; por ejemplo: ¿los padres leen?, ¿para qué?, ¿se producen interacciones a propósito de los hechos de lectura?, etc. De esta manera, en el caso de que las cifras hubieran marcado alguna asociación positiva, tampoco podría implicar que se haga una entrega indiscriminada de libros para las casas de los estudiantes; ello puede que no sea útil por sí mismo, sin un conocimiento del sentido cultural del manejo de la lengua escrita en las distintas regiones y estratos de nuestro país.

En la búsqueda de otros factores que den cuenta del impacto de la tenencia de libros en la casa, separamos los desempeños de los niños por grado:

			Grado			
			Cuarto		Tercero	
			Hay	No hay	Hay	No hay
<b>Desempeño</b>	<b>Estandarizado</b>	Bajo	9.68	10.18	23.68	25.70
		Medio	60.59	61.75	61.87	61.76
		Alto	29.73	28.07	14.45	12.53
		p-valor	0.605		0.205	
	<b>Competencias cerrada</b>	Nivel 0	1.13	0.95	1.34	0.91
		L. Literal	24.80	24.07	23.68	25.34
		L. Inferencial	59.61	61.85	61.82	58.58
		L. Crítico-intertext.	14.47	13.13	13.16	15.17
		p-valor	0.607		0.148	
	<b>Abierta</b>	Nivel 0	2.20	1.81	3.25	2.39
		C. Local	10.81	10.95	11.61	11.11
		C. Lineal	80.74	81.61	79.13	79.31
		C. Global	6.25	5.63	6.00	7.18
		p-valor	0.802		0.340	

Como puede verse, en ningún caso se presenta asociación; es decir, que el desempeño no dependería de los libros en la casa en ninguno de los dos grados. Esto puede querer decir muchas cosas, pero no podría cederse a la tentación de decir que es más incierta la asociación entre libros en la casa y desempeño cuando se trata de niños de cuarto que cuando se trata de niños de tercero, pues las diferencias entre los p-valores quedan en el rango de lo no significativo. Como se aplicó la misma prueba de lenguaje en ambos grados, podría decirse que dicha prueba ayuda más al establecimiento de la relación buscada cuando se trata de los niños de tercero; en otras palabras, esta diferencia entre los grados implica al menos un par de preguntas sobre la

prueba de lenguaje: ¿es más apta para niños de tercero? (esto cuando se cree que hay una asociación entre libros en la casa y desempeño en el área de lenguaje y, además, se confía en la manera como el instrumento de factores indagó este factor-libro, pues para tercero es menos incierta la posibilidad de asociación); o ¿es más apta para niños de cuarto? (esto cuando se confía en la prueba de lenguaje, y en atención a que los desempeños de los niños de cuarto son mejores en esa prueba, y, entonces, se afirmaría que es la búsqueda del factor-libro la que debe ser afinada).

También daría la posibilidad de interrogar el papel jugado por los adultos: los padres y los maestros podrían estar ayudando a los niños de manera diferencial en relación con el uso de los libros disponibles en la casa.

La descripción del desempeño de los niños de cuarto es idéntica en las dos categorizaciones de los resultados de la parte cerrada de la prueba de lenguaje (p-valores de 0.605 y 0.607); en este grado, es mayor la certeza de no asociación en el caso de la prueba abierta. Aunque los p-valores no hablan de asociación en el caso del grado cuarto, las cifras irían en el sentido de la asociación positiva esperada entre la tenencia de libros y los desempeños (para las tres miradas).

Para buscar otros factores que den cuenta del impacto de la tenencia de libros en la casa, separamos los desempeños de los niños por modalidad de escuela:

			Modalidad de escuela					
			Urbana		Nueva		Rural graduada	
			Hay	No hay	Hay	No hay	Hay	No hay
<b>Desempeño</b>	<b>Estandarizado</b>	Bajo	15.04	17.97	19.25	18.93	18.21	16.80
		Medio	59.76	60.36	62.60	62.76	64.24	63.05
		Alto	25.20	21.67	18.15	18.32	17.55	20.16
		p-valor	0.027		0.982		0.554	
	<b>Competencias cerrada</b>	Nivel 0	1.08	1.06	1.25	0.73	1.82	1.03
		L. Literal	24.00	26.11	23.87	25.15	25.99	20.41
		L. Inferencial	61.33	57.93	60.56	60.68	58.61	64.60
		L. Crítico-intertext.	13.59	14.90	14.32	13.43	13.58	13.95
		p-valor	0.336		0.592		0.141	
	<b>Abierta</b>	Nivel 0	1.87	1.89	3.50	2.63	4.50	1.59
		C. Local	9.99	11.44	10.99	10.12	16.44	11.90
		C. Lineal	81.13	79.78	80.27	80.55	74.57	81.75
		C. Global	7.02	6.89	5.25	6.70	4.50	4.76
		p-valor	0.690		0.369		0.014	



En una lectura vertical de la tabla anterior, se ve que escuela nueva es la modalidad que menos permite hablar de dependencia entre libros en la casa y desempeño en el área de lenguaje.

En esta misma dirección de lectura, la modalidad Rural graduada es la más cercana a la hipótesis. Particularmente, hay allí una asociación, moderadamente significativa entre el desempeño en la parte abierta de la prueba y la tenencia de libros en la casa; esta asociación es negativa: se llega a mejores desempeños y hay una mayor capacidad de discriminar desempeños (un valor bajo del Nivel 0) cuando los niños dicen no tener libros en la casa; y hay detención en los desempeños más elementales y menor capacidad para discriminar desempeños (un valor alto del Nivel 0) cuando los niños dicen tener libros en casa. También hay una correlación moderadamente significativa en la modalidad urbana, entre el desempeño estandarizado y la tenencia de libros en la casa; es una asociación positiva: mejores logros donde se dice tener libros.

Y, en una lectura horizontal, el único análisis del desempeño que no muestra asociaciones es el categorizado por competencias-C de la prueba.

Para la escala distrital, las cifras son las siguientes:

		<b>Distrito</b>		
		<b>Libros en la casa</b>		
		<b>Hay</b>	<b>No hay</b>	
<b>Desempeño</b>	<b>Estandarizado</b>	Bajo	9.38	12.50
		Medio	64.06	75.00
		Alto	26.56	12.50
		p-valor	0.491	
	<b>Competencias cerrada</b>	Nivel 0	3.13	6.25
		L. Literal	25.00	18.75
		L. Inferencial	67.19	56.25
		L. Crít.-intertext.	4.69	18.75
		p-valor	0.245	
	<b>Abierta</b>	Nivel 0	5.45	0.00
		C. Local	12.73	6.67
		C. Lineal	78.18	86.67
		C. Global	3.64	6.67
		p-valor	0.670	

En relación con los libros en la casa, la asociación con el desempeño en el área de lenguaje —evidenciada con estos datos— es similar en ambas escalas (nacional y distrital). Para el análisis estandarizado los p-valores son, respectivamente 0.130 y

0.491; para el análisis por competencias-C, son 0.677 y 0.245; y para el análisis por competencias-A, son 0.505 y 0.670.

En otras palabras, en el DC se replican los resultados nacionales sobre libros en la casa, que afirmarían que éstos no inciden en los desempeños en el área de lenguaje de los niños. Sacar la submuestra (más técnicamente, “conglomerado”) distrital de la muestra maestra (o sea, la nacional), no suministra información adicional; en este caso, nada dice el subconjunto que no lo diga el conjunto. Si estas cifras se leyeran literalmente podría pensarse que toda campaña a favor de la adquisición de libros por parte de las familias es un desperdicio, que todo esfuerzo de las familias por adquirir libros no coadyuvan a impulsar la competencia comunicativa de los niños, etc. No obstante, las cifras disponibles han sido obtenidas de manera que no puede precisarse la incidencia que tiene la manera de preguntar en relación con esto; es decir, queda abierta la posibilidad de hacer análisis cuyo objetivo sea indagar por la relación en mención, para lo cual habría que teorizarla primero, con lo que la indagación no estaría dando tantos “palos de ciego”, sino que iluminada por, y en, el problema pueda buscar minuciosamente con referencia a esa teorización. Al respecto, confrontar la diferencia entre la pregunta sobre libros en la casa que hace la evaluación de impacto del PUEBP y la que hace la evaluación censal del DC [ver § 7.1.2].

#### **4.2.2 Libros recibidos por la escuela**


Los formularios de factores asociables aplicados a la escuela, hacen la siguiente pregunta:

¿Después de 1989, cuántos textos y/o libros de la lista siguiente se recibieron para uso en la escuela?		
		<i>Escriba un número</i>
<input type="checkbox"/> A	Ciencias sociales y naturales .....	___
<input type="checkbox"/> B	Castellano de 1° a 3r grados .....	___
<input type="checkbox"/> C	Matemáticas de 1° a 3r grados .....	___
<input type="checkbox"/> D	Castellano para grados del 4° al 9° .....	___
<input type="checkbox"/> E	Ciencias naturales integradas para grados del 4° al 9° ..	___
<input type="checkbox"/> F	Ciencias sociales integradas para grados del 4° al 9° ...	___
<input type="checkbox"/> G	Guías de Escuela Nueva para alumnos .....	___
<input type="checkbox"/> H	Libros varios para la biblioteca .....	___

Para el aspecto de si se recibieron textos y/o libros para la escuela (de A a H), después de varios intentos de categorización sobre esta variable<sup>57</sup>, optamos por la siguiente: si la escuela recibió 30 o menos libros, se considera que “recibió bajo”; y si recibió más de 30 libros, se considera que “recibió alto”. De esta manera, se estableció que el 40.44% recibió bajo; y que el 59.56% recibió alto. Este umbral se seleccionó en atención al único criterio explícito que habría sobre la distribución de libros: en escuela nueva se entrega un texto para cada tres niños; y si admitimos que hay un promedio de 23 alumnos por docente en la escuela rural oficial [MEN, 1997b:34], en una escuela nueva que tenga 4 grados de educación básica primaria (la evaluación es hasta grado 4°), se espera que se hayan distribuido como mínimo cerca de 30 textos por escuela.

Hipotetizamos que el hecho de que la escuela reciba libros es algo que coadyuva al desempeño de los niños.

Al observar el desempeño desde el análisis estandarizado, las cifras quedan como sigue:


		Desempeño estandarizado		
		Bajo	Medio	Alto
	Recibió bajo	15.77	62.98	21.25
	Recibió alto	18.18	60.34	21.48
	p-valor	0.029 ±		

Los que se desempeñaron en nivel alto, recibir bajo o recibir alto fuera indiferente para alcanzar este nivel de desempeño, pues los porcentajes son casi idénticos: respectivamente, 21.25% y 21.48%. En cambio, la asociación moderadamente significativa que se presenta (p-valor: 0.029), parece deberse al porcentaje de los desempeños que se retiene en las categorías bajo y medio. En conjunto, la asociación se la puede calificar de no-lineal. En el nivel bajo es donde más se ven las diferencias: recibir alto favorece que los estudiantes se queden en el desempeño bajo; y recibir bajo favorece que los desempeños de los estudiantes pasen a la categoría medio.

Con la hipótesis de que tener libros en la escuela es ventajoso, era de esperarse que los niños que estudian en escuelas donde se había recibido alto, a diferencia de los que recibieron bajo, estuvieran cargados porcentualmente hacia los mejores desempeños; no obstante, como se ve en la tabla, los perfiles no son muy disímiles para nuestra mirada.

<sup>57</sup> Después de haber hecho el análisis nacional con una categorización de tres niveles, su aplicación a escala distrital no presenta frecuencias en algunas celdas.

Los perfiles del análisis con las categorías por competencias-C son muy parecidos y, por lo tanto, no permiten hacer afirmaciones concluyentes. La estadística (p-valor de 0.662) insinúa no asociación entre recepción de libros en la escuela y el desempeño, visto desde las categorías por competencia-C.


		Desempeño por competencias-C			
		Nivel 0	Lectura		
			Literal	Inferencial	Crítico inter-textual
	Recibió bajo	0.96	24.51	60.91	13.62
	Recibió alto	1.25	24.32	60.29	14.14
	p-valor	0.662			

En conclusión, desde estos dos ángulos del desempeño en la parte cerrada de la prueba no podemos afirmar que recibir libros en la escuela, por sí mismo, sea un factor que coadyuve a tener buenos desempeños en el área de lenguaje. Por supuesto que el hecho de no poder afirmarlo, con la información disponible, no quiere tampoco decir lo contrario. Ello porque puede resultar fundamental conocer el destino de los libros recibidos. Si «el modelo de escuela que ha prevalecido en la región [Centroamérica y el Caribe] está lejos de la cultura escrita, tanto desde el punto de vista administrativo, como del punto de vista pedagógico» [Hederich y Charria, 1999:67], puede ocurrir, por ejemplo, que se reciban los libros y se almacenen —so pretexto de cuidarlos y tenerlos en buen estado para cuando haya que dar cuenta del inventario de la escuela—, o se mantienen a la vista en una biblioteca que permanece con llave para que no los dañen... con lo que tal vez no tengan incidencia alguna en el desempeño de los niños en el área de lenguaje.

Esto llama la atención en el sentido de investigar *cómo* la institución, los directivos, los profesores y los niños se articulan para utilizar los libros en la escuela, qué vida académica generan, para qué son utilizados, quién y con qué criterios los entrega a los niños, si generan lecturas más allá de los compromisos del salón de clase, etc. Tal vez una indagación masiva no permita ver esos detalles y se preste para hacer preguntas generales que —¿por limitaciones técnicas?, ¿por intereses no explícitos?— terminan ignorando matices que podrían aportar mucho más a una mirada sobre la dependencia buscada entre las variables. Las estadísticas, per se, hacen que los datos pierdan su individualidad; este es el precio que se paga cuando la atención se centra en las cifras. Que los libros y los textos lleguen a la escuela (lo que ya es bastante, pues a veces se quedan en el camino) no necesariamente quiere decir que lleguen a cada niño; es más: que lleguen a cada niño no es garantía de que éstos los usen... y todavía está por discutir en qué consistiría ese uso de acuerdo con las diversas perspectivas de lo que es una educación de “calidad”.

En este sentido, el proyecto «Podemos leer y escribir»<sup>58</sup> propone unas categorías de investigación mucho más acordes con el objeto en mención: en primer lugar, se plantea un objetivo de naturaleza cultural: «la apropiación de la cultura escrita por parte de cada uno de los participantes de la experiencia»; en segundo lugar, trabaja con instituciones que hayan hecho un proyecto pedagógico sobre lectura y escritura para poder recibir los acervos bibliográficos; y, en tercer lugar, se propone indagar por la *disponibilidad* de los textos (ubicación, clasificación y préstamo en la escuela), el número de textos *cuyos títulos se recuerdan*, razones sobre *gusto o disgusto* por los libros, *estrategias* pedagógicas y los *lugares* donde se lee; esto para escuela, maestros, alumnos y padres [cf. Hederich y Charria, 1999].

El análisis de las categorías por competencias-A en relación con la recepción de textos en la escuela no muestra asociación (p-valor de 0.190). Veamos los datos:

		Desempeño por competencias-A			
		Nivel 0	Coherencia		
			Local	Lineal	Global
	Recibió bajo	2.23	10.85	80.01	6.91
	Recibió alto	2.72	11.36	80.16	5.76
	p-valor	0.190			

De acuerdo con la tabla, el porcentaje de niños cuyo desempeño pasa hasta los desempeños de coherencia global (parte abierta de la prueba) es levemente mayor entre aquellos cuyas escuelas reciben menos libros. Esto parece no respaldar el axioma oficial de que “a más libros, mejores desempeños”. Pero como el análisis estadístico no establece de manera directa relaciones de causalidad, sino que alerta sobre un comportamiento de los datos, sería irresponsable argumentar que los libros no contribuyen al mejoramiento de la competencia comunicativa. Además, la recepción de libros no la podemos concebir de manera aislada de otros factores.

La recepción baja, a nuestro juicio, sería más sensible a que algo se detecta con los instrumentos; tal vez las modalidades de recepción postulan unos rangos distantes de los asuntos más importantes por averiguar, y entonces se convierten en maneras de preguntar que no resultan útiles (tal vez estereotipadas).

Al desagregar los desempeños en la prueba de lenguaje por curso, de acuerdo con la cantidad de libros recibidos por la escuela, obtenemos la siguiente tabla:

<sup>58</sup> Este proyecto hizo una evaluación de impacto de la entrega de acervos bibliográficos. Pese a que entre las variables que cruzó no estaba el desempeño de los niños, aquí la consideramos porque utilizó categorías más cercanas al objeto, sin detrimento de un análisis estadístico.

			Grado			
			Cuarto		Tercero	
			Recepción		Recepción	
			Baja	Alta	Baja	Alta
<b>Desempeño</b>	Estandarizado	Bajo	9.03	10.40	22.41	25.68
		Medio	61.85	60.37	64.10	60.31
		Alto	29.11	29.23	13.49	14.02
		p-valor	0.421		0.066±	
	Competencias cerrada	Nivel 0	1.00	1.11	0.91	1.38
		L. Literal	24.56	24.56	24.47	24.09
		L. Inferencial	59.85	60.69	61.97	59.90
		L. Crít.-intertext.	14.59	13.63	12.65	14.63
		p-valor	0.877		0.236	
	Abierta	Nivel 0	1.55	2.43	2.91	2.99
		C. Local	12.52	9.71	9.20	12.96
		C. Lineal	79.13	82.34	80.87	78.06
		C. Global	6.79	5.53	7.02	5.99
		p-valor	0.013±		0.010±	

Se observa que se tiene más certeza para hacer alguna afirmación sobre tercero que sobre cuarto. La asociación con más certidumbre es la observada en tercero, bajo la óptica de la clasificación por Competencias-A; a pesar de eso, la dependencia entre variables no apoya la expectativa: en dos casos, es no lineal y en el otro la baja recepción favorece el tránsito a desempeños altos (la asociación sería negativa).

En general, el desempeño en Competencias-A sugiere la existencia de asociaciones entre recepción de libros por la escuela y desempeño de los estudiantes en lenguaje, pero no en el sentido de la hipótesis (la asociación que se presenta en cuarto es no lineal y la que se presenta en tercero es negativa). El análisis por Competencias-C, en ninguno de los dos grados muestra asociación; mientras que el análisis estandarizado muestra asociación solamente en un grado (tercero), con tendencia no lineal.

El análisis por grado permite aflorar la sensibilidad de la prueba abierta a detectar asociación con los libros recibidos por la escuela, pues a escala nacional no se mostraba asociación. Para Competencias-C, es invariante, se confirma la no asociación. Y para el análisis estandarizado, se conserva la tendencia no lineal del análisis nacional, pero sólo para grado tercero (téngase en cuenta la paradoja de Simpson).

Y en relación con modalidad de escuela, las cifras se distribuyen de la siguiente manera:

			Modalidad de escuela					
			Urbana		Nueva		Rural graduada	
			Recibe		Recibe		Recibe	
			Bajo	Alto	Bajo	Alto	Bajo	Alto
Desempeño	Estandarizado	Bajo	15.00	17.00	16.83	19.51	18.56	17.04
		Medio	61.95	57.30	64.36	62.37	66.83	61.67
		Alto	23.05	25.70	18.81	18.12	14.60	21.29
		p-valor	0.024 <sup>+</sup>		0.548		0.029 <sup>+</sup>	
	Competencias cerrada	Nivel 0	1.00	1.17	0.66	1.11	0.99	1.87
		L. Literal	25.63	23.24	22.11	24.75	21.04	25.72
		L. Inferencial	60.47	60.25	63.04	60.20	61.39	60.65
		L. Crítico-intertext.	12.89	15.35	14.19	13.94	16.58	11.75
		p-valor	0.126		0.658		0.057 <sup>+</sup>	
	Abierta	Nivel 0	1.95	1.77	3.28	3.14	2.82	3.71
		C. Local	9.63	11.44	9.49	10.84	17.44	12.72
		C. Lineal	80.80	80.66	81.02	80.27	75.64	78.62
		C. Global	7.62	6.13	6.20	5.75	4.10	4.95
		p-valor	0.168		0.914		0.197	

No aparecen asociaciones, ni en el análisis por competencias-A (para todas las modalidades de escuela), ni en escuela nueva (para los tres análisis del desempeño). De los casos restantes, en tres hay asociaciones moderadamente significativas: una — entre desempeño estandarizado y escuela urbana— no lineal; una negativa, entre desempeño por competencias-C y escuela rural graduada; y una positiva, entre desempeño estandarizado y escuela rural graduada. La escuela rural graduada sería el caso donde hay más asociaciones, pero un análisis del desempeño muestra una asociación negativa, mientras otro análisis muestra una asociación positiva.

Antes de desglosar por modalidad de escuela, el análisis estandarizado del desempeño mostró asociación no lineal (p-valor de 0.029). Desglosado, no muestra asociación en escuela nueva; para las otras dos modalidades de escuela, en una conserva la tendencia no lineal y en la otra la tendencia se vuelve positiva. Ni antes ni después de desglosar por modalidad de escuela, el análisis del desempeño por competencia-A muestra asociación. Antes de desglosar por modalidad de escuela, el análisis del desempeño por competencias-C no mostró asociación; desglosado en las tres modalidades de escuela, muestra una asociación negativa.

Para la comparación con la escala distrital, las cifras son las siguientes:

			Distrito	
			Libros recibidos	
			Bajo	Alto
<b>Desempeño</b>	Estandarizado	Bajo	5.00	11.67
		Medio	60.00	68.33
		Alto	35.00	20.00
		p-valor	0.325	
	Competencias cerrada	Nivel 0	0.00	5.00
		L. Literal	5.00	30.00
		L. Inferencial	80.00	60.00
		L. Crít.-intertext.	15.00	5.00
		p-valor	0.049	
	Abierta	Nivel 0	5.56	3.85
		C. Local	16.67	9.62
		C. Lineal	77.78	80.77
		C. Global	0.00	5.77
		p-valor	0.631	

En la parte cerrada de la prueba, se muestra una inversión entre la escala nacional y la distrital: mientras el análisis estandarizado está asociado de manera moderada a escala nacional, en la distrital no permite asociar el desempeño de los estudiantes con los libros recibidos por la escuela; y mientras el análisis por Competencias-C no permite ver asociaciones en la escala nacional, en la distrital sí, de manera moderada y con tendencia negativa. Es decir, para el DC el desempeño de los estudiantes en lenguaje o bien es independiente de los textos recibidos por la escuela, o bien se ve afectado negativamente por éstos: allí donde se reciben libros hay menos tránsito de los desempeños de los estudiantes desde los niveles elementales hacia los niveles más complejos; y allí donde no se reciben libros, los desempeños tienden hacia esos niveles, o sea que su influencia es contraria a la esperada.

En lo que respecta a la la parte abierta de la prueba, tanto en lo nacional como en lo distrital la producción escrita del niño no permite hablar de asociación entre el desempeño y los libros recibidos por la escuela; es como decir que los libros entregados han sido indiferentes para los desempeños: los niños avanzan a mejores desempeños o se quedan en los niveles básicos, hayan o no recibido libros sus escuelas. La desagregación a escala distrital no da información adicional a la nacional.

Esto, por supuesto, no puede ser un argumento a favor de considerar que los libros entregados no hayan servido; advierte tanto sobre la necesidad de indagar por la calidad de dichos materiales, como por su papel en las prácticas escolares. En otras palabras, no podemos concluir de manera causal. Pero como la información acerca del factor-libros se obtuvo en la escuela como unidad de observación, mientras el



desempeño se observó en el niño, se trataría de indagar las condiciones de la dotación y, sobre todo, del uso en la práctica docente y, particularmente, del uso que hace el niño de los libros.

### ***4.2.3 Guías de escuela nueva***

De la pregunta transcrita en §4.2.2, aquí se analizará solamente el literal G, que es el que se refiere específicamente a esta modalidad de escuela. Las escuelas encuestadas en todo el país se seleccionaron de manera tal que el conjunto constituyera un corpus representativo de la educación básica pública; esto incluía la composición de escuela urbana y rural. En la zona rural, la mayor parte de las escuelas siguen la metodología denominada “escuela nueva” [cf. §3.1], modalidad a la que, como hemos dicho, el Estado entrega de manera gratuita unas Guías de trabajo que sirven para que los niños regulen sus propias actividades, en interacción con sus compañeros, con el profesor y con la comunidad.

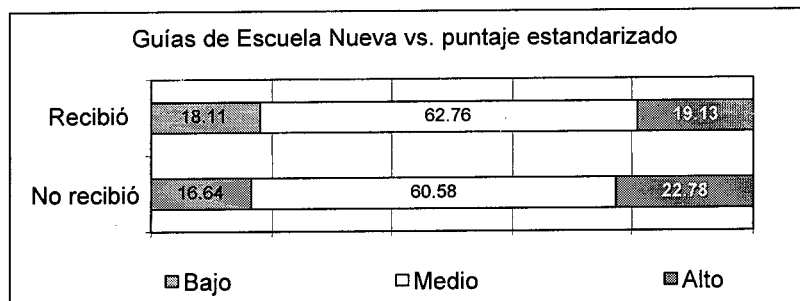
Los desempeños de lenguaje, vistos desde la categorización inducida por la respuesta a la pregunta de si la escuela recibió las Guías de escuela nueva, implican confrontar los resultados de dicha modalidad de educación en la zona rural, frente a la llamada “escuela graduada” (de la que también hay una pequeña proporción, la sexta parte, en el campo<sup>59</sup>). De otro lado, los niños que no reciben las Guías no es que estén desprovistos de libros, propios o de la escuela. De tal manera, esta mirada podría decir algo sobre el manejo de Guías enfrentado a textos que no es posible determinar o, incluso, a carencia de textos; con los datos disponibles, no podemos saber más.

Según los datos recogidos, el 38.1% de las escuelas recibió Guías de escuela nueva (o sea, prácticamente las escuelas de esa modalidad, pues en la muestra el 32.5% de los niños son de escuela nueva; la diferencia posiblemente se deba a que los niños contestaron con base en las opciones del formulario: podrían pertenecer a escuelas urbanas marginales que trabajan con la misma metodología). En atención a lo dicho, entonces, no sabemos qué recibió el 61.9%. De otro lado, esta proporción daría cuenta en forma aproximada de la composición de la muestra en relación con la modalidad: nueva vs. graduada.

Veamos las cifras para el desempeño, visto desde la categorización estandarizada, contra la recepción de Guías:

---

<sup>59</sup> El 52.1% de los niños fueron de escuela urbana; el 32.5% de escuela nueva; y el 15.4% de escuela rural graduada.



De acuerdo con los datos, no recibir las Guías favorece el tránsito a puntajes altos (en casi 4 puntos) y desfavorece los puntajes bajo y medio (en alrededor de 2 puntos). Este tipo de correlación resulta más claro: cuando muestra un comportamiento que se puede leer en todos los niveles de logro; si favorece quedarse en los niveles bajo y medio, se puede considerar un factor asociado a esos niveles de logro; pero si, además, no favorece el paso a logros altos, se puede considerar un factor asociado negativamente al logro. Por cualquier desempeño que se lo tome, resulta consistente. Es decir, en este caso, los datos hablarían a favor de una asociación *negativa* entre el desempeño en el área de lenguaje y la recepción de Guías de escuela nueva. Tiene un p-valor de 0.002, altamente significativo.


De otro lado, los datos hablan a favor de una asociación *positiva*, altamente significativa, entre el desempeño en el área de lenguaje y *no* recibir Guías de escuela nueva. Podríamos decir que los libros o el material escrito de los que disponen las escuelas que tienen la modalidad graduada (o sea, el material que haría las veces de las Guías en las escuelas con la modalidad de escuela nueva) favorece más el desempeño de los estudiantes encuestados.

El análisis mediante los desempeños por competencia-C no permite hablar de asociación, pues los perfiles de los que recibieron Guías y de los que no, son muy parecidos en cada nivel de logro:

		<b>Desempeño por competencias-C</b>			
		Nivel 0	Lectura		
			Literal	Inferencial	Crítico intertextual
	No recibió	1.06	23.85	61.09	14.00
	Recibió	1.18	24.74	60.20	13.88
	p-valor		0.826		

Es decir, desde las categorías que definió el grupo para la parte cerrada de la prueba de lenguaje, trabajar con las Guías de escuela nueva o con otros libros, prepara a los estudiantes de igual forma para desempeñarse en el área de lenguaje (p-valor: 0.826). En otras palabras, no se puede afirmar nada en el sentido de que las Guías de escuela nueva favorezcan o no ese desempeño, o que lo hagan mejor que otro tipo de material.

Con base en la información contenida en la siguiente tabla, analicemos la influencia de las Guías sobre el desempeño, visto desde la producción escrita:

		Desempeño por competencias-A			
		Nivel 0	Coherencia		
			Local	Lineal	Global
	No recibió	1.94	11.51	80.07	6.47
	Recibió	3.45	10.58	80.14	5.83
	p-valor	0.002			

El p-valor (0.002) sugiere una asociación altamente significativa. La diferencia parece provocarse fundamentalmente por el “Nivel 0” y la “Coherencia local”. Según los datos, quienes no reciben Guías de escuela nueva tienden más a tener una producción clasificable como con algún grado de coherencia (superar el nivel 0). Así mismo, alcanzar solamente desempeños de coherencia local, o llegar hasta desempeños de coherencia global, también parece estar “favorecido” cuando no se reciben Guías de escuela nueva.

Para que todo esto no se convierta en una sugerencia implícita de recoger las cartillas de escuela nueva, este resultado llama la atención en varios sentidos: ¿qué tipo de material se usa en las escuelas graduadas?; ¿qué uso se le da a las Guías? y ¿qué uso se le da a los textos en la escuela graduada?; ¿cada cuánto y con qué criterios se revisan las Guías?, pues, en atención a la competencia que tiene lugar entre los textos escolares comerciales, posiblemente éstos se revisen con mayor frecuencia; ¿qué incidencia tiene el hecho de que la Guía sea de la escuela y que se calcule la distribución pensando en que no va a corresponder una por cada niño, mientras que en la escuela graduada cada niño puede que tenga su texto?, etc.

Para responder preguntas como éstas se podría, por ejemplo, investigar grupos de desempeño parecido, unos con Guías y otros con otro tipo de libro (ojalá con diversidad de libros); además, sería indispensable “controlar” las utilizaciones, los contextos y las interacciones que se generen a propósito del uso de los libros; no se trataría de perpetuar unas prácticas basadas en los textos escolares, sino de pensar cómo sería una escuela para la cultura escrita en la que la diversidad es una condición de posibilidad y, en consecuencia, habría que reaprender a “enseñar” bajo esas nue-

vas condiciones. Investigaciones como éstas harían falta para la evaluación que periódicamente el MEN hace al material impreso que reparte con objetivos de enseñanza y aprendizaje (no se trataría sencillamente de “actualizarlas”). Es decir, se sugiere diversificar las investigaciones, pues una investigación como la de la evaluación de impacto del PUEBP constituye una panorámica general que es necesario desagregar.

Desglosados por curso los desempeños, se observa:

			Grado			
			Cuarto		Tercero	
			Recibir		Recibir	
			No	Sí	No	Sí
<b>Desempeño</b>	Estandarizado	Bajo	8.95	11.29	24.12	24.76
		Medio	59.99	62.57	61.15	62.94
		Alto	31.06	26.13	14.73	12.30
		p-valor	0.004		0.147	
	Competencias cerrada	Nivel 0	1.22	0.82	1.14	1.29
		L. Literal	25.27	23.41	24.22	24.28
		L. Inferencial	59.23	62.16	61.15	60.05
		L. Crítico-intertex.	14.29	13.60	13.49	14.39
		p-valor	0.325		0.863	
	Abierta	Nivel 0	1.56	2.90	2.32	3.99
		C. Local	11.48	9.84	11.54	11.29
		C. Lineal	80.71	81.55	79.45	78.78
		C. Global	6.25	5.71	6.69	5.94
		p-valor	0.042+		0.055	

Para referirse a la dependencia entre la recepción de Guías de escuela nueva y el desempeño, hay más certidumbre cuando se trata de grado cuarto. En grado cuarto, hay dos asociaciones: entre recepción de Guías y clasificación estandarizada, en el sentido de que el material que hace las veces de las Guías favorece más el desempeño alto (altamente significativa, negativa); y clasificación por competencias-A: moderadamente significativa y no lineal. Y hay asociación moderadamente significativa y negativa entre la clasificación por competencias-A y Guías, en el grado tercero. El desempeño en la parte abierta siempre presenta asociación, o sea que es una manera de indagar el desempeño que se presta más para la búsqueda en mención.

Y desagregado por modalidad de escuela, tenemos para las guías de escuela nueva y el desempeño:


			Modalidad de escuela					
			Urbana		Nueva		Rural graduada	
			Recibe		Recibe		Recibe	
			No	Sí	No	Sí	No	Sí
<b>Desempeño</b>	Estandarizado	Bajo	16.15	13.00	20.61	18.88	17.17	18.54
		Medio	59.76	61.67	64.86	62.30	62.52	66.01
		Alto	24.09	25.33	14.53	18.82	20.31	15.45
		p-valor	0.359		0.198		0.166	
	Competencias cerrada	Nivel 0	1.14	0.33	0.68	1.11	1.57	1.40
		L. Literal	24.88	21.67	23.31	24.54	24.72	22.19
		L. Inferencial	60.05	63.67	64.53	59.97	58.90	64.61
		L. Crítico-intertext.	13.93	14.33	11.49	14.38	14.80	11.80
		p-valor	0.328		0.390		0.337	
	Abierta	Nivel 0	1.78	2.83	3.97	3.03	1.81	6.03
		C. Local	10.84	6.01	7.58	11.16	16.45	11.49
		C. Lineal	80.48	83.39	80.87	80.30	77.80	76.72
		C. Global	6.90	7.77	7.58	5.52	3.95	5.75
		p-valor	0.049+		0.153		0.001+	

El análisis estandarizado, que mostraba en lo nacional una asociación alta y negativa, en ninguna de las tres modalidades de escuela muestra asociación. El análisis por competencia-C, revela no asociación tanto en el análisis sin desglosar por modalidad de escuela, como en cada modalidad. Y, por su parte, el análisis por competencias-A no hereda la alta asociación negativa de lo nacional en escuela nueva, y se invierte el sentido de las asociaciones para las otras dos modalidades (se vuelve positivo lo que en lo nacional era negativo).

En este caso, el desglose permite pensar que mientras más puntual sea el análisis, es posible encontrar más cosas. Pero esto no se cumple para todos los casos; en casos anteriores vemos que el análisis globalizado dice más, o el análisis desglosado es más disperso o dice lo mismo. Esta es la paradoja de Simpson.


#### 4.2.4 Libros varios para biblioteca

De la pregunta transcrita en §4.2.2, aquí se analizará solamente el literal H. El 72.2% de las escuelas recibió libros para su biblioteca (los restantes, ¿para dónde son?). En consecuencia con la hipótesis, se esperaría que sus estudiantes se vean beneficiados en el desempeño en el área de lenguaje. No obstante, el p-valor para la asociación con el puntaje estandarizado es de 0.187, o sea, no significativa; los perfiles de ambas poblaciones son muy parecidos, como se ve a continuación:

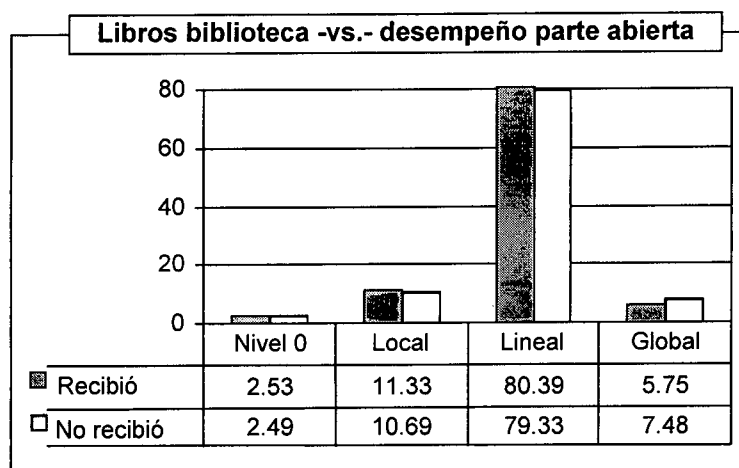
		Desempeño estandarizado		
		Bajo	Medio	Alto
	No recibió	18.48	61.03	20.49
	Recibió	16.71	61.55	21.74
	p-valor	0.187		

Es decir, el desempeño en el área de lenguaje parece no depender de la recepción de libros para la biblioteca de la escuela.

En la asociación entre la recepción de libros varios para la biblioteca y el desempeño interpretado con los niveles por competencia-C, los perfiles también son muy parecidos (p-valor: 0.339). Una mínima diferencia en el nivel inferencial a favor de recibir los libros (1.88%), aparece en el nivel crítico-intertextual a favor de no recibirlos (1.66%). La cercana diferencia con el análisis por puntaje estandarizado es inversa. Veamos los datos:

		Desempeño por competencias-C			
		Nivel 0	Lectura		
	Literal		Inferencial	Crítico intertextual	
	No recibió	1.12	24.57	59.18	15.13
	Recibió	1.14	24.33	61.06	13.47
	p-valor	0.339			

La siguiente gráfica exhibe el comportamiento de los datos de recepción de libros en biblioteca, en relación con el desempeño en producción (parte abierta):



Con estas cifras, aparece una correlación moderadamente significativa (p-valor: 0.089) entre los dos aspectos. Recibir libros para la biblioteca favorece el avance hasta niveles de coherencia local y lineal, mientras que la no recepción favorece el paso a mayores logros en la producción, es decir, los de coherencia global (nótese la disparidad entre las barras). O sea, la asociación sería *negativa*, inversa a la esperada en la hipótesis con la que abrimos este punto.

No obstante, estos análisis no pueden utilizarse para no proveer de libros a las bibliotecas escolares. La base para una política educativa al respecto no podría reducirse a investigaciones generales; por ejemplo, de acuerdo con la investigación de Otálora et al. [1993] (hecha en doce colegios de Bogotá, públicos y privados), *los textos no tienen sentido por sí mismos, sino que lo ganan si están inscritos en una propuesta de trabajo que compromete a sus agentes*<sup>60</sup>. Es decir, cuando una institución educativa posee muchos libros, pero carece de una propuesta así, éstos no necesariamente incidirán de manera positiva en el desempeño de los estudiantes; por el contrario, una institución educativa con pocos libros sí verá afectado positivamente el desempeño de sus alumnos por este factor, gracias a la existencia de una iniciativa de trabajo más allá de las meras condiciones normativas del ejercicio de la educación; de ahí el acierto de la medida mencionada en §4.2.2 de entregar los acervos bibliográficos solamente a aquellas instituciones que construyan un proyecto relativo a la lectura y la escritura (esto, por supuesto, no es garantía, pero es al menos un comienzo que muestra el tipo de participación que se espera y se requiere para producir una relación con la cultura escrita. De donde lo buscado y lo producido se confunden<sup>61</sup>).

Si bien la mayoría de las escuelas encuestadas durante la evaluación de impacto del PUEBP manifestaron haber recibido libros para la biblioteca, tal situación no es diferenciadora del desempeño de sus estudiantes. De acuerdo con la información disponible, el resultado no es sorprendente ni debe descalificar el abastecimiento de libros para las bibliotecas de las escuelas, sino unirlo a los factores que lo desencadenarían como factor asociado. En una escuela en la que el centro de la acción es el aula, no la cultura, es esperable que los libros de la biblioteca todavía tengan —en el mejor de los casos— el papel de “auxiliares” de las tareas, incluso de los textos escolares; ya anotábamos otros destinos posibles de las bibliotecas escolares: prisiones para los libros. La cultura, en cambio, está expresada de manera que pueda fijarse y perdurar; eso incluye, por supuesto, libros, pero también mapas, pinturas, diagramas, escultu-

---

<sup>60</sup> La propuesta de la que se habla no se refiere al PEI, pues la investigación se hizo antes de la Ley 115, que fue la que lo estableció.

<sup>61</sup> Es el caso de la “vitrina escolar” en el DC: las instituciones educativas tenían que listar necesariamente libros de texto; muchas iniciativas en otro sentido tuvieron que deponerse ante tal exigencia.

ras y, hoy en día, medios magnéticos de archivo de datos (discos duros y flexibles, discos compactos, DVD, etc.).

Desagregada por cursos la dotación de libros para la biblioteca en relación con el desempeño en el área de lenguaje, muestra las siguientes cifras:

			Grado			
			Cuarto		Tercero	
			No	Sí	No	Sí
<b>Desempeño</b>	Estandarizado	Bajo	11.98	9.02	24.83	24.18
		Medio	60.79	61.05	61.26	62.05
		Alto	27.23	29.93	13.91	13.77
		p-valor	0.0267		0.910	
	Competencias cerrada	Nivel 0	1.02	1.09	1.21	1.19
		L. Literal	24.18	24.71	24.94	23.97
		L. Inferencial	58.98	60.87	59.38	61.25
		L. Crítico-intertext.	15.82	13.33	14.46	13.60
		p-valor	0.348		0.805	
	Abierta	Nivel 0	1.93	2.12	3.04	2.93
		C. Local	11.23	10.71	10.16	11.94
		C. Lineal	78.74	81.90	79.91	78.92
		C. Global	8.09	5.26	6.89	6.22
		p-valor	0.0304		0.524	

La tabla permite pensar que el curso tercero queda por fuera de toda posibilidad de hacer afirmaciones sobre la asociación entre libros para la biblioteca y desempeño. Como si el hecho de que la escuela recibiera libros para la biblioteca no influyera en el desempeño de estos niños. ¿Será que por ser “más pequeños” no los llevan a la biblioteca?, ¿o que por mostrar todavía “problemas” de escritura, se piensa que no pueden utilizar libros de la biblioteca (y, por ejemplo, sí los “textos escolares”)? Con los cursos en cifras unificadas, el análisis estandarizado no mostraba asociación; desglosados los cursos, cuarto muestra asociación positiva, moderada; el análisis por competencias-C no muestra asociación ni antes ni después del análisis por curso; y el análisis por competencias-A, continúa mostrando asociación moderada negativa (para estos datos).

Y desagregado por modalidad de escuela, tenemos:



			Modalidad de escuela					
			Urbana		Nueva		Rural graduada	
			Recibe		Recibe		Recibe	
			No	Sí	No	Sí	No	Sí
Desempeño	Estandarizado	Bajo	18.20	14.89	18.91	19.18	18.73	17.04
		Medio	60.26	59.79	59.91	63.39	64.46	63.38
		Alto	21.54	25.32	21.18	17.43	16.80	19.59
		p-valor	0.011+		0.185		0.503	
	Competencias cerrada	Nivel 0	1.01	1.10	1.37	0.97	1.10	1.75
		L. Literal	26.09	23.97	24.37	24.37	20.66	25.64
		L. Inferencial	57.94	61.39	60.14	60.74	61.43	60.67
		L. Crítico-intertext.	14.96	13.54	14.12	13.93	16.80	11.94
		p-valor	0.294		0.905		0.071-	
	Abierta	Nivel 0	2.48	1.62	3.66	3.03	1.16	4.59
		C. Local	9.59	10.75	10.98	10.57	13.29	15.41
		C. Lineal	80.17	80.98	76.59	81.38	80.35	75.74
		C. Global	7.76	6.65	8.78	5.03	5.20	4.26
		p-valor	0.202		0.026-		0.024-	

El análisis estandarizado del desempeño, que en lo nacional no mostraba asociación, continúa no mostrando asociación en las dos modalidades de escuela rural; la asociación moderada que muestra en escuela urbana es positiva. El análisis por competencias-C continúa mostrando no asociación en la escuela urbana y en la escuela nueva; la asociación moderada que muestra en la escuela rural graduada es negativa. Y el análisis por competencias-A, que mostraba asociación moderada negativa, sigue mostrando este tipo de asociación para las dos modalidades de escuela rural, mientras que para la escuela urbana no muestra asociación. Las asociaciones negativas, así como el p-valor más alto, están en lo rural. La única asociación positiva está en lo urbano (contra el análisis estandarizado); ¿qué relación podría tener esto con el hecho de que en lo rural haya menos bibliotecas?


#### 4.2.5 Textos escolares en general

De la pregunta transcrita en §4.2.2, aquí se analizarán, colapsados, los literales de A a G (incluimos las *Guías* de escuela nueva en la clase “textos escolares” y sólo excluimos los libros varios para la biblioteca). De esta manera, obtenemos la información de que el 87% de las escuelas ecuestadas recibió textos escolares, y que el 13% no recibió. Una débil prueba de lo que decíamos en el punto anterior —una escuela volcada sobre los textos escolares y no sobre la cultura— parece estar en la compa-

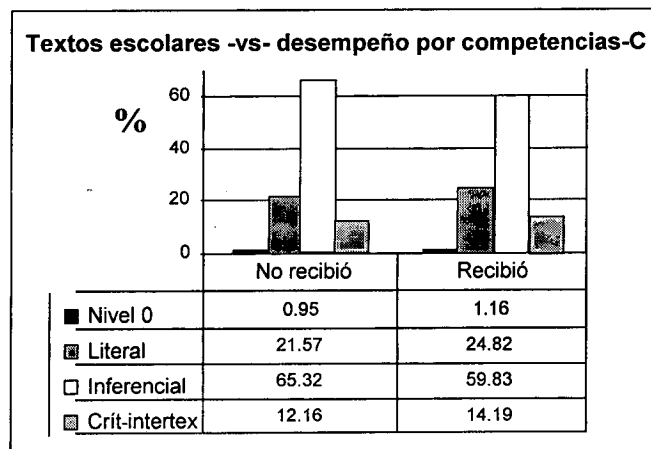
ración de las siguientes cifras: mientras el 72.2% de las escuelas recibió libros para la biblioteca, el 87% recibió textos escolares, o sea, un 14.8% más.

Esta diferencia explícita entre “libros varios” y “textos escolares” es fundamental; sin embargo, para efectos de buscar las asociaciones en este punto, continuamos con la hipótesis de que el material impreso (así sea texto escolar) favorece el desempeño de los niños en lenguaje. Es decir, que si un niño tiene disponibilidad de al menos texto escolar, estaría en ventaja sobre aquel que ni siquiera tiene texto escolar.

La clasificación estandarizada de los desempeños muestra una asociación moderadamente significativa con la recepción de textos escolares: tiene un p-valor de 0.042. Para quienes se quedan en los logros bajos, es indiferente haber recibido textos; para quienes pasan a desempeños intermedios, resulta ventajoso con un 3.46%; pero para los que alcanzan desempeños altos, los estudiantes que han recibido textos son un 3.77% menos que los que no han recibido, lo cual resulta en contravía de lo esperado; es decir, la correlación es negativa:


		Desempeño estandarizado		
		Bajo	Medio	Alto
	No recibió	16.92	58.4	24.67
	Recibió	17.24	61.86	20.90
	p-valor	0.042		

El análisis de la relación entre el desempeño en la parte cerrada —visto desde las categorías del equipo— y la recepción de textos escolares en general, concluye a favor de una asociación moderadamente significativa entre ambos (p-valor: 0.026). Para el desempeño crítico-intertextual, hay una diferencia de 2.03% a favor de recibir textos escolares; sin embargo, para los desempeños literales, también la recepción de textos parece ventajosa (en 3.25%). Esto se desprende de la información contenida en la siguiente gráfica:



En otras palabras, recibir textos escolares favorecería mínimamente el hecho de que los estudiantes avancen a buenos desempeños... pero también el que permanezcan en desempeños bajos; tenemos, entonces, una asociación no lineal; las ventajas se neutralizan con las desventajas. Por eso es fundamental investigar *cómo* resultan favorecidos los desempeños altos, si este favorecimiento depende de la estructura de los textos escolares, del hecho de ser material escrito, o de las prácticas pedagógicas; e investigar *cómo* resultan favorecidos los desempeños bajos y de qué depende esto. La relación entre textos escolares y desempeños en la parte cerrada, en ambos tipos de análisis, no se muestra sistemática.

Por su parte, el análisis de los desempeños por competencias-A también muestra una asociación moderadamente significativa (p-valor: 0.033) con los textos escolares en general. En este caso, como en el estandarizado, la asociación es negativa: permanecer en los desempeños más elementales (producir al menos una frase con coherencia sujeto/verbo) se ve favorecido por la recepción de textos escolares, mientras que evolucionar a logros más complejos (producir un texto con coherencia y cohesión) se ve favorecido por no recibir textos escolares:

Desempeño por competencias-A					
	Nivel 0	Coherencia			
		Local	Lineal	Global	
	No recibió	3.64	9.79	79.05	7.53
	Recibió	2.35	11.36	80.26	6.03
	p-valor	0.033			

De tal manera, con los datos disponibles y con las perspectivas de análisis asumidas, no podemos hacer afirmaciones contundentes sobre la dependencia entre recepción de textos escolares y desempeño. Es el único caso en que las tres perspectivas de análisis coinciden en la significatividad de sus p-valores (moderados), pero dos muestran una asociación negativa y uno muestra una asociación no lineal. Por idénticas razones, no se pueden sustentar ni una disminución del abastecimiento de textos, ni una campaña a su favor, sino estudios que digan más de lo que las cifras han permitido; si algunos estudios dicen haber mostrado asociación positiva entre textos escolares y desempeño, no siempre aclaran la incidencia que en esto tiene el hecho de que algunos maestros usen dichos textos *en ausencia de otro material escrito* e, incluso, que los usen para *develar sus contradicciones* [cf. Jurado y Bustamante, 1994:48-49].

No obstante, en la política oficial no parece importar este tipo de análisis, pues el abastecimiento de libros de texto es parte de los programas: por ejemplo, las Guías de escuela nueva forman parte de esa modalidad de trabajo que el MEN no está inte-

resado en cambiar, pese a que la misma Ley 115 habla de una selección del material impreso por parte de cada PEI.

Desagregado por cursos, los textos escolares en general contra el desempeño arrojan las siguientes cifras:

			Grado			
			Cuarto		Tercero	
			Recibir		Recibir	
			No	Sí	No	Sí
<b>Desempeño</b>	Estandarizado	Bajo	9.88	9.84	23.82	24.45
		Medio	55.90	61.74	60.85	61.98
		Alto	34.22	28.43	15.33	13.58
		p-valor	0.046		0.620	
	Competencias cerrada	Nivel 0	1.20	1.05	0.71	1.27
		L. Literal	22.17	24.92	20.99	24.73
		L. Inferencial	64.10	59.78	66.51	59.87
		L. Crítico-intertext.	12.53	14.25	11.79	14.14
		p-valor	0.382		0.065±	
	Abierta	Nivel 0	1.04	2.22	6.08	2.48
		C. Local	10.88	10.85	8.76	11.86
		C. Lineal	80.05	81.17	78.10	79.36
		C. Global	8.03	5.75	7.06	6.30
		p-valor	0.156		0.001	

Antes de desglosarse por curso, los tres análisis del desempeño mostraban asociación moderada; en dos casos, negativa y no lineal en el otro. En este caso, son los desempeños de los niños del curso tercero los que más tienden a estar asociados a la presencia de textos escolares en la escuela (recordemos que era inverso en relación con biblioteca): en dos análisis del desempeño se encontró asociación, en el otro no; dos análisis del desempeño de cuarto no están asociados, uno sí. La asociación del análisis estandarizado en cuarto sigue siendo moderada y negativa. La asociación del análisis por competencias-C en tercero sigue siendo no lineal, como antes del desglose. Y la asociación del análisis por competencias-A en tercero ahora es altamente significativa, y sigue siendo negativa.

Desagregados por modalidad de escuela, los datos se organizan así:

		Modalidad de escuela						
		Urbana		Nueva		Rural graduada		
		Recibe		Recibe		Recibe		
		No	Sí	No	Sí	No	Sí	
Desempeño	Estandarizado	Bajo	13.21	16.11	17.44	19.48	21.35	16.77
		Medio	55.00	60.38	56.40	63.99	67.19	62.95
		Alto	31.79	23.51	26.16	16.53	11.46	20.28
		p-valor	0.007*		0.001*		0.013*	
	Competencias cerrada	Nivel 0	0.71	1.10	0.54	1.16	2.08	1.38
		L. Literal	25.36	24.52	21.25	25.03	16.67	25.53
		L. Inferencial	62.86	60.15	65.12	59.65	69.27	58.95
		L. Crítico-intertext.	11.07	14.23	13.08	14.16	11.98	14.14
		p-valor	0.461		0.204		0.031*	
	Abierta	Nivel 0	2.96	1.77	3.48	3.09	4.95	2.97
		C. Local	5.93	10.83	10.43	10.70	14.29	14.73
		C. Lineal	84.81	80.36	78.26	80.83	71.98	78.68
		C. Global	6.30	7.04	7.83	5.38	8.79	3.62
p-valor		0.039*		0.346		0.010*		

Por modalidad de escuela, sólo la escuela rural graduada muestra asociaciones para los tres análisis del desempeño. El análisis estandarizado, que mostraba asociación moderada negativa, muestra asociaciones en todos los casos: altas para la escuela urbana y para escuela nueva, ambas negativas; y para la escuela rural graduada la asociación moderada es positiva. El análisis por competencias-C, que mostraba una asociación no lineal, muestra una sola asociación moderada, que sigue siendo no lineal. Y el análisis por competencias-A, que mostraba asociación moderada negativa, muestra asociaciones moderadas en escuela urbana y en escuela rural graduada, una no lineal y la otra negativa. Escuela nueva es la que menos asociaciones muestra, cuando allí el texto escolar es una condición de la modalidad...

#### 4.2.6 Textos escolares de lenguaje

De la pregunta transcrita en §4.2.2, aquí se analizan, colapsados, los literales B (textos escolares de lenguaje para 1°, 2° y 3°) y D (de 4° a 9°). De las escuelas encuestadas, el 86.4% declara haber recibido textos escolares de lenguaje (o sea, 0.6% por debajo del promedio de los textos escolares en general) y el 13.6% declara no haber recibido.


En relación con el desempeño de los niños en lenguaje, los libros de texto del área no tienen, para nosotros, un valor especial por encima de los otros, pues la evalua-

ción de desempeños que diseñamos no se hizo con arreglo a *contenidos* del área respectiva —caso en el que sí los pondría claramente en ventaja frente a los otros en lo que a desempeño de lenguaje se trata—, sino con arreglo a *competencia comunicativa*, que es una capacidad producida (de manera heterogénea) en todos los ámbitos de la vida y, particularmente, en todas las asignaturas y actividades escolares.

No obstante, como uno de los objetivos explícitos del área de lenguaje es desarrollar las “habilidades lingüísticas”, podría hipotetizarse que si los profesores de las demás áreas evaden su implicación en el asunto de dichas “habilidades” en los estudiantes —por ejemplo, por no parecerles específica—, entonces los textos escolares del área de lenguaje tendrán una incidencia más significativa en el desempeño, que los textos escolares en general, en función del trabajo que con ellos harían los profesores cuya meta específica es desarrollar las “habilidades lingüísticas” (entonces ya no sería propiamente asunto de textos... además, en cada disciplina hay una “gramática” específica).

La asociación, vista desde el análisis del desempeño estandarizado también es moderadamente significativa, como con los textos escolares en general; vista desde el desempeño por competencias-C no es detectable, a diferencia de los textos escolares en general, donde era moderada; y sólo desde el análisis del desempeño en la parte abierta la asociación con los textos de lenguaje se encontró que es más significativa que para los textos en general. Veámoslo, caso por caso.

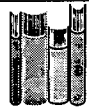
Desde el análisis estandarizado del desempeño, se revelan las siguientes cifras:

		Desempeño estandarizado		
		Bajo	Medio	Alto
	No recibió	17.16	58.35	24.49
	Recibió	17.21	61.89	20.90
	p-valor	0.048		

Según vemos, tener o no tener texto de lenguaje no influye en la no superación de logros bajos, el perfil es casi idéntico; en cambio, para pasar a logros intermedios, es favorable tener texto de lenguaje; mientras que para llegar a logros altos es desfavorable (con una diferencia de 3.59%). El coeficiente p-valor es de 0.048, o sea, asociación moderadamente significativa.

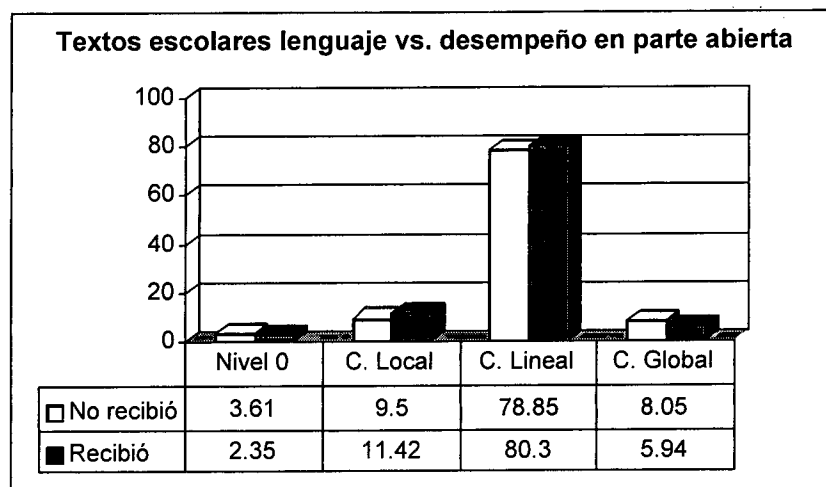
Por su parte, la asociación frente al desempeño en la parte cerrada, establecido con las categorías del grupo de investigación, no es detectable (p-valor: 0.154). En este análisis, el logro literal tiene un puntaje más alto en el caso de haber recibido texto

de lenguaje, con una diferencia de casi tres puntos (2.81); también el logro crítico-intertextual, pero con una diferencia de menos de un punto (0.89).

		Desempeño por competencias-C			
		Nivel 0	Lectura		
			Literal	Inferencial	Crítico-intertext.
	No recibió	0.92	21.97	63.96	13.16
	Recibió	1.17	24.78	60.00	14.05
	p-valor	0.154			

Entonces, nada podemos decir sobre la influencia que tiene el hecho de poseer libros de texto en lenguaje, sobre el desempeño en la misma área. Como en los otros casos de recepción de textos, es primordial investigar cómo se articulan en las diversas prácticas (particularmente en las pedagógicas, que podrían estar condicionando la detención en desempeños básicos, gracias a un énfasis en la lingüística de la frase).

En el análisis del desempeño en la parte abierta, las cifras presentan un p-valor de 0.007, es decir, una asociación altamente significativa. Quienes reciben textos de lenguaje tienden a ubicarse en logros elementales (escribir con coherencia local), con una ventaja de casi dos puntos, frente a los que no reciben; mientras que quienes no reciben textos ven favorecido su tránsito a logros más complejos (escribir con coherencia global, que presupone la lineal y la local), en un poco más de dos puntos. La asociación, entonces, es negativa (en contra de la hipótesis), como se visualiza en la siguiente gráfica:



Desagregado el desempeño por grado, frente a los textos escolares de lenguaje, obtenemos lo siguiente:

		Grado				
		Cuarto		Tercero		
		No	Sí	No	Sí	
<b>Desempeño</b>	<b>Estandarizado</b>	Bajo	9.91	9.83	24.32	24.37
		Medio	55.76	61.80	60.91	61.97
		Alto	34.33	28.37	14.77	13.65
		p-valor	0.033		0.812	
	<b>Competencias cerrada</b>	Nivel 0	1.15	1.06	0.68	1.27
		L. Literal	23.27	24.76	20.68	24.80
		L. Inferencial	62.21	60.05	65.68	59.96
		L. Crítico-intertext.	13.36	14.13	12.95	13.97
		p-valor	0.845		0.106	
	<b>Abierta</b>	Nivel 0	1.23	2.20	5.85	2.49
		C. Local	10.37	10.93	8.67	11.89
		C. Lineal	79.75	81.23	77.99	79.39
		C. Global	8.64	5.64	7.49	6.23
		p-valor	0.072		0.001	

El análisis estandarizado mostraba en lo nacional una asociación moderada negativa; ahora no muestra asociación en tercero y continúa mostrando el mismo tipo de asociación en cuarto. El análisis por competencias-C continúa no mostrando asociación. Y el análisis por competencias-A, que mostraba asociación alta negativa, muestra asociación moderada en cuarto y alta en tercero, ambas negativas. Las tendencias de lo nacional continuaron. Dicho de otro modo, en este caso nada ganamos con el desglose.

Y desagregado por modalidad de escuela, obtenemos:



			Modalidad de escuela					
			Urbana		Nueva		Rural graduada	
			Recibe		Recibe		Recibe	
			No	Sí	No	Sí	No	Sí
Desempeño	Estandarizado	Bajo	13.00	16.15	18.32	19.30	21.35	16.77
		Medio	55.33	60.38	56.28	64.08	67.19	62.95
		Alto	31.67	23.47	25.39	16.62	11.46	20.28
		p-valor	0.005		0.001		0.013	
	Competencias cerrada	Nivel 0	0.67	1.11	0.52	1.17	2.08	1.38
		L. Literal	24.67	24.58	22.51	24.78	16.67	25.53
		L. Inferencial	60.67	60.35	63.87	59.88	69.27	58.95
		L. Crítico-intertext.	14.00	13.96	13.09	14.17	11.98	14.14
		p-valor	0.917		0.396		0.031	
	Abierta	Nivel 0	2.76	1.78	3.61	3.06	4.95	2.97
		C. Local	5.52	10.90	10.28	10.74	14.29	14.73
		C. Lineal	83.45	80.47	78.61	80.77	71.98	78.68
		C. Global	8.28	6.85	7.50	5.43	8.79	3.62
		p-valor	0.021		0.444		0.011	

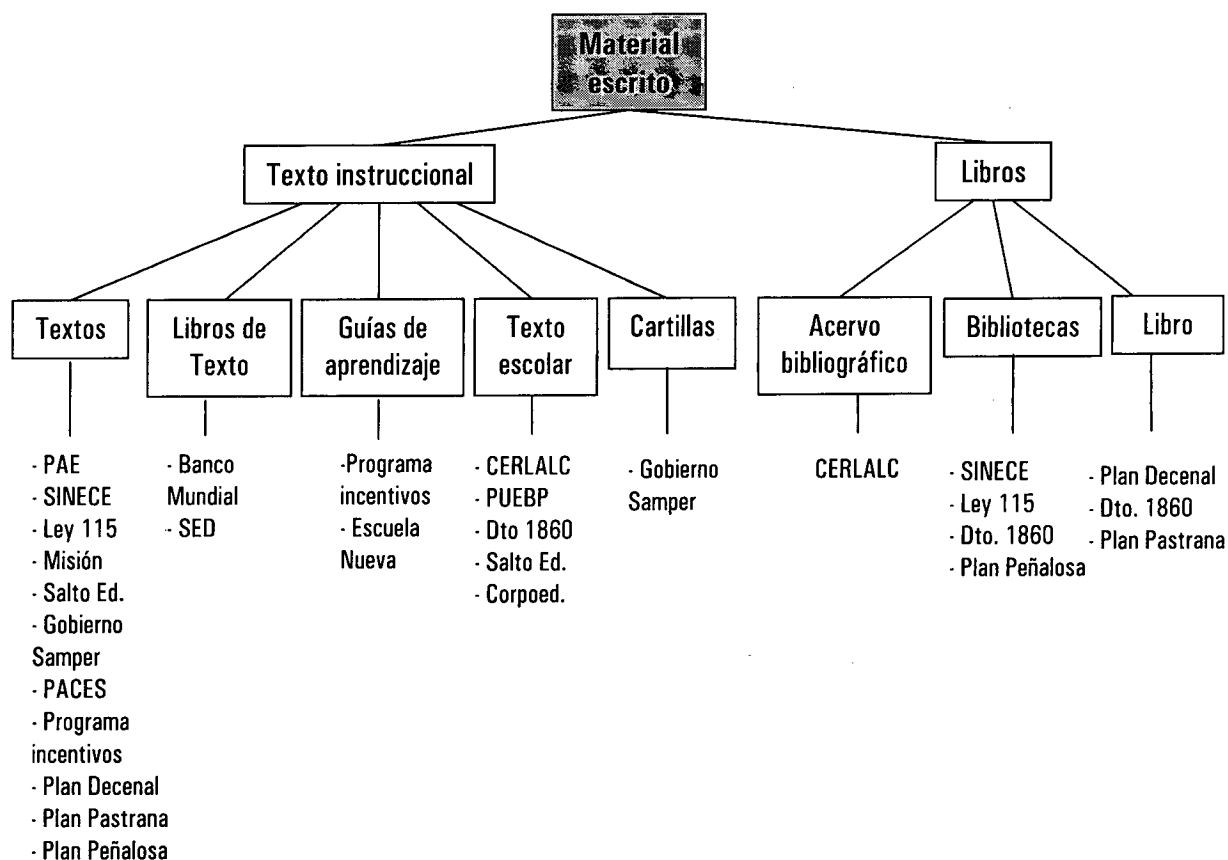
Escuela rural graduada es la única que muestra asociación para los tres análisis del desempeño en el área de lenguaje; escuela urbana muestra asociación en dos casos; y escuela nueva solamente en uno. El análisis estandarizado, que mostraba asociación moderada negativa, muestra dos asociaciones altas negativas (escuelas urbana y nueva) y una moderada positiva (escuela rural graduada); esto quiere decir que los textos no tendrían el mismo sentido en cada modalidad, de donde la política educativa al respecto no puede ser unánime. El análisis del desempeño por competencias-C, que no mostraba asociación, continúa no mostrándola en escuelas urbana y nueva, pero marca una asociación no lineal en escuela rural graduada. Y el análisis por competencias-A, que marcaba asociación alta negativa, marca —en las modalidades diferentes a escuela nueva— dos asociaciones moderadas, ambas negativas.

### 4.3 Resumen parcial

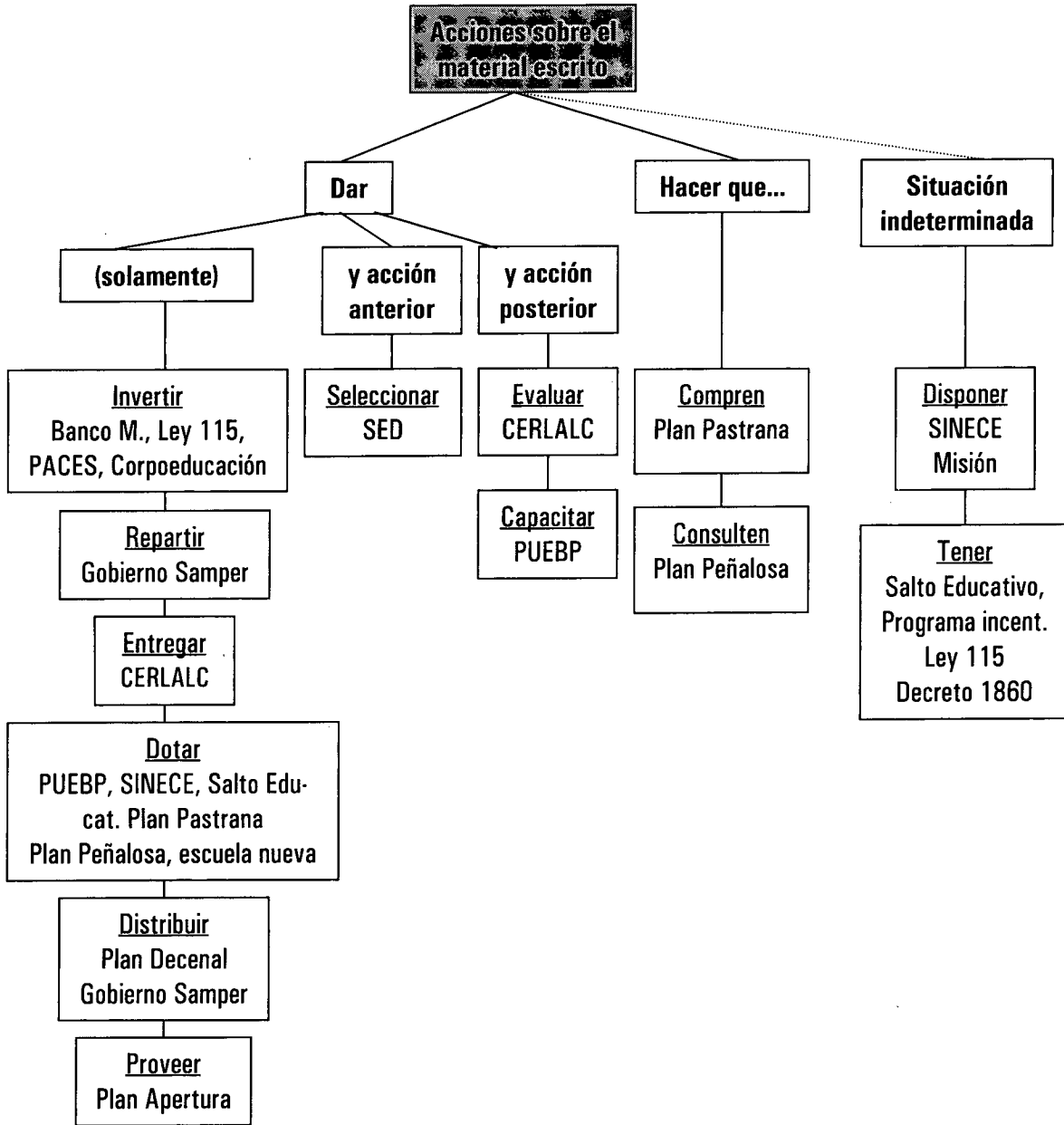
4.3.1 De los estudios, investigaciones y políticas educativas que revisamos [cf. §4.1], esquematizamos las posiciones frente al *tipo de material* al que se refieren y a las *acciones* que dicen ejercer con él.

En cuanto al tipo de material, encontramos dos grupos: de un lado, *textos escolares*, denominados “textos”, “libros de texto”, “guías de aprendizaje” y “texto escolar”; y, de otro, *libros*, denominados “acervos bibliográficos”, “bibliotecas” y “libros” (de-

nominaciones que no excluyen los “textos”). Como se ve, la mayoría del material referenciado se encuentra en el primer grupo. El siguiente gráfico lo muestra:



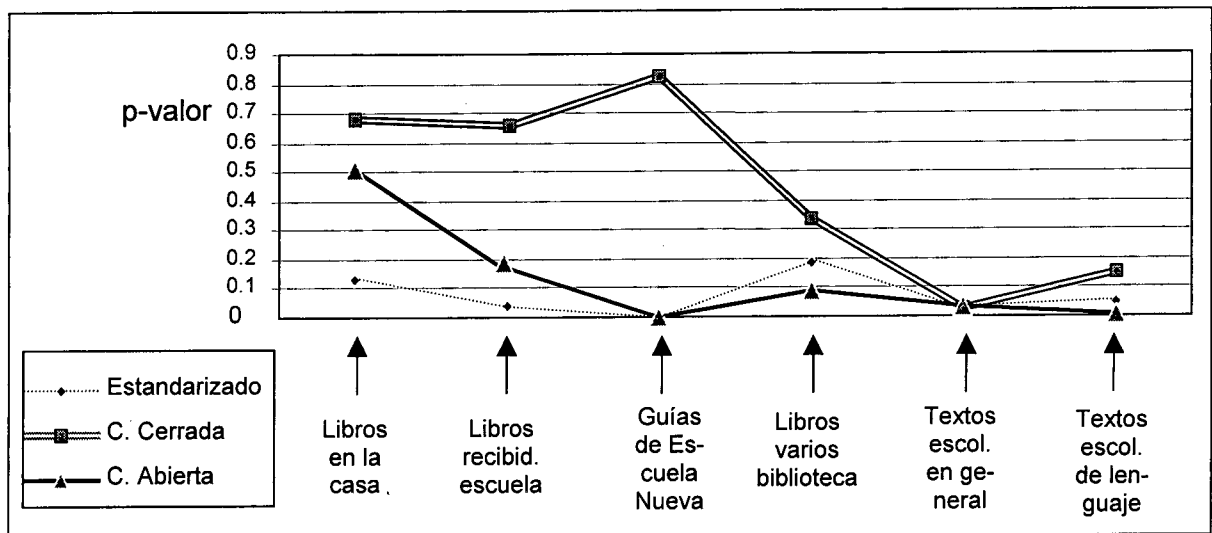
Y, en cuanto a las acciones a realizar con el material escrito, se mencionan fundamentalmente dos tipos: en primera instancia, *dar*, manifiesta en expresiones como “invertir”, “repartir”, “entregar”, “dotar”, “distribuir” y “proveer”, en las que se enfatiza solamente el hecho de dar, y *dar y otra acción* que se agrega al hecho de dar. En segunda instancia, *hacer que* compren, que consulten. Quedan unas situaciones indeterminadas del tipo, “tener” y “disponer”, no se sabe si como efecto de un otorgamiento o de una compra. (Comentarios al respecto se incluyen en §4.4). Veámoslo en el siguiente gráfico:



4.3.2 A continuación, escribimos en una tabla los p-valores de las asociaciones entre las tres maneras de analizar el desempeño (estandarizado, competencias-C y competencias-A) y los factores relativos al libro (factores-libro) encuestados durante la evaluación de impacto del PUEBP. A la derecha de cada p-valor [p-v] escribimos qué tipo de asociación puede sugerirse con base en la herramienta estadística adoptada, en los instrumentos contruídos y en los datos recogidos.

Desempeño	Estandarizado		Competencias-C		Competencias-A	
	p-v	Asociación	p-v	Asociación	p-v	Asociación
Factores-libro						
Libros en la casa	0.130	No	0.677	No	0.505	No
Libros recibid. x escuela	0.029	Moderada-NL	0.662	No	0.190	No
Guías de escuela nueva	0.002	Alta-Negativa	0.826	No	0.002	Alta-Negativa
Libros para biblioteca	0.187	No	0.339	No	0.089	Moderada-Negativa
Textos escolares en gral.	0.042	Moderada-Negativa	0.026	Moderada-NL	0.033	Moderada-Negativa
Textos de lenguaje	0.048	Moderada-Negativa	0.154	No	0.007	Alta-Negativa

Una manera de visualizar la asociabilidad entre el factor libro y el desempeño es mediante la graficación de los perfiles de los p-valores para cada modalidad de análisis del desempeño, lo que permite compararlos entre sí. La gráfica queda como sigue:



4.3.3

Libros en la casa

	Nacional	4°	3°	Urbana	Nueva	Rural grad	Distrital
Estandarizada	0.130	0.605	0.205	0.027+	0.982	0.554	0.491
Competencias-C	0.677	0.607	0.148	0.336	0.592	0.141	0.245
Competencias-A	0.505	0.802	0.340	0.690	0.369	0.014-	0.670

Libros recibidos por la escuela

	Nacional	4°	3°	Urbana	Nueva	Rural grad	Distrital
Estandarizada	0.029+	0.429	0.066±	0.024±	0.648	0.029+	0.325
Competencias-C	0.662	0.877	0.236	0.126	0.658	0.057-	0.049-
Competencias-A	0.190	0.013±	0.011-	0.168	0.914	0.197	0.631

Guías de Escuela Nueva

	Nacional	4°	3°	Urbana	Nueva	Rural grad
Estandarizada	0.002-	0.004-	0.147	0.359	0.198	0.166
Competencias-C	0.826	0.325	0.863	0.328	0.390	0.337
Competencias-A	0.002-	0.042±	0.055-	0.049+	0.153	0.001+

Libros varios para biblioteca

	Nacional	4°	3°	Urbana	Nueva	Rural grad
Estandarizada	0.187	0.026+	0.910	0.011+	0.185	0.503
Competencias-C	0.339	0.348	0.805	0.294	0.905	0.071-
Competencias-A	0.089-	0.030-	0.524	0.202	0.026-	0.024-

Textos escolares en general

	Nacional	4°	3°	Urbana	Nueva	Rural grad
Estandarizada	0.042-	0.046-	0.620	0.007-	0.001-	0.013+
Competencias-C	0.026±	0.382	0.065±	0.461	0.204	0.031±
Competencias-A	0.033-	0.156	0.001-	0.039±	0.346	0.011-

Textos escolares de lenguaje

	Nacional	4°	3°	Urbana	Nueva	Rural grad
Estandarizada	0.048-	0.033-	0.812	0.005-	0.001-	0.013+
Competencias-C	0.154	0.845	0.106	0.917	0.396	0.031±
Competencias-A	0.007-	0.072-	0.001-	0.021-	0.444	0.011-

#### 4.4 Un análisis parcial

☞ Como se ve en la tabla de §4.3.2, en 9 casos (el 50%), no hay asociaciones entre libros y desempeño en el área de lenguaje a escala nacional (bajo las tres modalidades: estandarizado, por competencias-C y por competencias-A); es decir, que hay bastante incertidumbre para relacionar ambos aspectos o, en otras palabras, que el desempeño dependería de otros factores. En el resto de casos, hay seis asociaciones moderadas, cuatro de ellas negativas, y tres asociaciones altamente significativas, todas negativas. De dos asociaciones moderadas (entre textos escolares en general y desempeño por competencias-C, y entre libros recibidos por la escuela y desempeño estandarizado), no se puede decir si es positiva o negativa (son las asociaciones que hemos llamado no lineales). En resumen, el estudio no respalda la idea de que los libros y el desempeño están asociados. Solamente un estudio, entre todos los consultados, coincide en señalar una asociación negativa entre libros y desempeño: se trata de la investigación de Piñeros y Rodríguez [1999]. Para el resto aparece como evidente o, supuestamente, demostrado.

☞ Así, la hipótesis de las ventajas que para el desempeño presenta la disposición de libros, si bien no se ha visto respaldada por la información recogida en este estudio, podría mantenerse, pero sólo a condición de que sea de manera más compleja: el material escrito coadyuvaría al desarrollo de la competencia comunicativa, pues la escuela es un espacio social que pretende aproximarse al conocimiento y, en consecuencia, despliega una serie de interacciones basadas —directa o indirectamente— en el acervo cultural *escrito* [cf. Mockus et al., 1994]. Esto implica preguntarse por las maneras de investigar cómo dicho acervo se relaciona con la escuela y cómo ésta a su vez se relaciona con su contexto específico. Y, de otro lado, el material escrito coadyuvaría al desarrollo de la competencia comunicativa si se dan ciertas condiciones que los datos disponibles, por no haberlas buscado, no permiten precisar; sin embargo, éstas estarían en relación, con el hecho de que el “uso” del material escrito no es independiente de las maneras como los involucrados en la educación entienden y hacen efectiva la formación del sujeto, la formación de lo social, la génesis del conocimiento, la función de la interacción en la escuela... Habíamos dicho que el material impreso y su código dominan en la escuela, de modo que indagar sobre los libros presupone indagar fundamentalmente sobre el acceso a ese código.

☞ En este sentido, para investigar sobre la asociación entre material escrito y desempeño ha de tenerse en cuenta categorías más acordes con la relación entre escritura y competencia comunicativa; en tal dirección, no se puede excluir ningún tipo de material escrito cuando se trata de desarrollar dicha competencia (en ciertas circunstancias, el *único* material escrito de un niño en su casa podría ser,

por ejemplo, el periódico); ni se pueden privilegiar unos sobre otros (por ejemplo, el llamado “texto escolar” tendría que ser considerado como uno de los múltiples libros y no como el único o el más importante, tal como parece ser para la SED [1999a]). No hay una “esencia” del material escrito que lo haría “bueno” o “malo”: la interacción del lector con el texto y con otros lectores es fundamental en la producción del sentido; en el caso de la escuela, es fundamental la interacción entre niños, maestros y escritos.

- ☞ El análisis de la información no muestra que sea innecesario “dar” libros, sino que no es suficiente. Esto no es evidente para la mayoría de los estudios, investigaciones y políticas educativas que revisamos, que hablan de “invertir”, “repartir”, “entregar”, “dotar”, “distribuir” y “proveer” [cf. §4.3.1]; la distribución indiscriminada que caracteriza a esta manera de entender la relación con el material escrito pudo verse en la pasada “vitrina escolar” del DC. Por su parte, en escuela nueva el criterio para la dotación de los textos es por el número de alumnos de la escuela, dando dos o tres ejemplares de cada área por grupo de alumnos (no hay ejemplar por alumno, para “obligar” al trabajo de grupo).
- ☞ Con base en los resultados, podemos afirmar que el factor *libros* no es asociable a cualquier situación. Cierta tipo de actividades con los libros pueden hacer que éstos *comiencen* a tener sentido (no lo tienen por sí mismos); y entonces son tales actividades las que vale la pena impulsar, investigar y evaluar. Actividades que prescindan del material escrito o que lo usen de manera estereotipada, podrían privar a dicho material de sus posibilidades en relación con el desempeño de los estudiantes.
- ☞ En este sentido nos parece importante la idea que hay en una experiencia como la de CERLALC cuando restringe la entrega de libros a establecimientos que construyan un proyecto pedagógico en el campo de la lectura y la escritura [Hederich y Charria, 1999:66]. O sea, tal vez los resultados advierten la necesidad de inscribir la distribución de los libros en otro tipo de acciones que les den sentido. Incluso los libros existentes (entregados/recibidos) deberían inscribirse en proyectos que indaguen sobre su uso, su calidad y su efecto sobre el desempeño, pero visto desde teorías de la lectura, la escritura y la producción de conocimiento. Curiosamente, sólo dos estudios mencionan algo más allá de entregar: evaluar el destino de los materiales<sup>62</sup> y capacitar a los docentes (la SED habla pero de una acción previa: seleccionar, sólo que se aplica únicamente a libros de texto). Nos parece que efectivamente no basta con la entrega, que la evaluación es necesaria,

---

<sup>62</sup> En su informe al Congreso, Samper [1998:264] dice que las cartilla entregadas “comienzan a evaluarse”; sin embargo, a continuación sentencia esta acción con la idea de que los estudios internacionales indican que la *adquisición* de libros es un predictor de éxito académico.

pero siempre y cuando se la entienda como investigación y que también se aplique al material escrito que se distribuye para la escuela; y que la “capacitación” es necesaria siempre y cuando no se la entienda como la información que los que “ya saben” van a dar a los docentes que “no saben” cómo usar los libros (esto se evidencia en el documento que la SED [1999a] emitió para que las instituciones seleccionaran sus textos en el marco de la “vitrina escolar”: se asume que el documento tiene unas verdades sobre el asunto que los maestros no tienen y que adquirirán gracias a su lectura). También la capacitación ha de pensarse como investigación (autoformación de los docentes) sobre los quehaceres educativos —tanto en el aula como fuera de ella— en los que los libros intervienen.

☞ Hay dos grupos de factores relativos a los libros claramente diferenciados [cf. §4.3.2]: de un lado, i) libros en la casa y ii) libros recibidos en la escuela, para los que los tres tipos de análisis del desempeño (estandarizado, por competencias-C y por competencias-A) son muy diferentes; y, de otro lado, i) Guías de escuela nueva, ii) libros para la biblioteca, iii) textos escolares en general y iv) textos escolares de lenguaje, para los que los tres tipos de análisis son más cercanos. En el primer grupo, las tres perspectivas de análisis no se apoyan; cualquier afirmación sobre la asociación entre textos y desempeño en lenguaje correspondería al tipo de interpretación que se esté haciendo y no se puede universalizar, aunque se marque o no asociación. En el segundo, pueden remitirse unos análisis a los otros en busca de confrontaciones para la investigación.

☞ La gráfica de perfiles [cf. §4.3.2] permitiría pensar que los encuestados están haciendo una especie de clasificación del material escrito: de un lado, aquel que está formalmente ligado al trabajo en el aula (“texto escolar”, Guías de escuela nueva); y, de otro, aquel que no parecería estarlo de manera directa (libros en la casa, en la biblioteca). Esto revela que posiblemente nuestra educación está habituada a sacar provecho del primero y concibe el segundo como accesorio o complementario, cuando el desarrollo de la competencia comunicativa exige una interacción con los libros *en general*<sup>63</sup>, no solamente los instruccionales (por simple contención de conjuntos, los textos instruccionales son un subconjunto de los libros). Esto se relaciona con la manera como los estudios, investigaciones y políticas educativas que revisamos mencionan el material impreso: en su mayoría lo traen como “textos escolares”, en sus diferentes acepciones (“textos”, “libros de texto”, “guías de aprendizaje” y “texto escolar”). Muy pocos hablan de “libros”. El hecho de que, no obstante, las cifras para los primeros muestren asociaciones *negativas*, y que para los segundos casi no muestre correlaciones, podría ser un llamado de atención en dos sentidos: primero, comprender tanto el ti-

---

<sup>63</sup> Pero, curiosamente, la SED, que hace una evaluación de competencias básicas, emite una cartilla para selección únicamente de libros de texto [cf. SED, 1999a].



po de prácticas que se llevan a cabo con los “textos escolares”, como su estructura (contenidos, diseño, tipo de tareas, tratamiento de los discursos que se recontextualizan, etc.); y, segundo, valorar de otra manera el material escrito que parece no relacionarse de manera directa con el aula (para ciertas prácticas pedagógicas), haciendo indagaciones más apropiadas para este tipo de material.

- ☞ La tenencia de libros en la casa es una variable que también se ve ligada a aspectos socioeconómicos; no se trata de personas frente a un factor neutral, como si todo el mundo estuviera en igualdad de condiciones para tener libros en la casa; al decir “aspectos socioeconómicos”, no sólo nos referimos a la posibilidad económica de tener libros, sino también a la posibilidad cultural [cf. Bodnar, 1992].
- ☞ El modelo que promueve la dotación de textos (no sólo en Colombia) en este momento muestra una preocupación de “ajuste fiscal”; en consecuencia, pone la mirada inquisidora sobre los libros, no sobre los textos escolares: en los estudios, investigaciones y políticas educativas que revisamos, uno solo, el Plan de desarrollo del gobierno Pastrana, se refiere —además de la dotación de textos— a la necesidad de resaltar la importancia de los libros para que cada colombiano tenga mayores posibilidades de tener (léase: de comprar).
- ☞ Una mirada a los perfiles [cf. §4.3.2] muestra que el análisis con categorías estandarizadas y el análisis con categorías por competencias-A están más próximos entre sí, pese a que en el primer caso se trata de una distribución aproximada a la “normal”, mientras que en el otro se trata de una distribución por la producción de los niños encuestados en relación con unos niveles de desempeño teórica y previamente definidos (o sea, con arreglo a criterio). Hay más proximidad teórica entre los dos enfoques por competencia y, sin embargo, en las cifras, el análisis de la comprensión (parte cerrada) por competencias es el menos asociable a los libros (no así en el caso del DC, donde sólo el análisis por competencias-C se asocia).
- ☞ El análisis que más muestra alta asociación es el del desempeño en producción, o sea el análisis que da cuenta de la parte abierta de la prueba. Podría afirmarse, con arreglo a tal hallazgo que, para los datos trabajados, la prueba masiva con preguntas cerradas, de respuesta única entre varias opciones, pese a tener mayor ductibilidad y prestarse para el manejo estadístico, es menos propicia para buscar una relación entre libros y desempeño de los niños en el área de lenguaje. Pese a que la calificación de preguntas con opción de respuesta abierta —en la que el niño escribe de manera más o menos libre—, presenta una dificultad a nivel de la confiabilidad (cómo asegurar que los diversos calificadores coincidan en la apli-

cación de los criterios establecidos para analizar las producciones de los niños), con los datos que se manejaron en este estudio se puede afirmar que es apta para buscar la asociación entre libros y desempeño en el área de lenguaje. No obstante, la calificación aparentemente más confiable (la estandarizada) dice menos de la asociación libros/desempeño-lenguaje (sólo muestra una asociación alta). En otras palabras, mientras se pretende eliminar el “ruido” del calificador mediante preguntas cerradas, se observa que es más el “ruido” que aparece en las respuestas de los estudiantes; y mientras más “ruido” se esperaría en el calificador de preguntas abiertas, menos “ruido” aparece en las respuestas de los estudiantes.

☞ “Textos escolares en general” es el único factor relativo a los libros en el que las tres formas de análisis muestran asociación. Este hecho, sin embargo, no quiere decir que efectivamente haya asociación en los procesos educativos (aquí el acierto no se asigna por mayoría). Quiere decir, si somos consecuentes con los principios estadísticos en juego, que la incertidumbre de que haya asociación ha disminuido ostensiblemente, para estas cifras. Tanta sorpresa debe producirnos el que no coincidan como el que coincidan dos o las tres formas de análisis del desempeño; y, en consecuencia, cualquier caso debe ser objeto de investigación. De tal manera, creemos, se justifica investigar acerca de la incidencia del material escrito en prácticas pedagógicas, acciones focalizadas (como la que actualmente adelanta la SED).

☞ La asociación que aparece entre “textos escolares en general” y las tres formas de analizar el desempeño es, no obstante, *negativa*: en las escuelas donde no se recibieron textos escolares en general, hay cifras de mejores desempeños; y en las escuelas donde sí se recibieron textos escolares en general, resultan favorecidas las cifras de los desempeños más elementales (es decir, los que se busca trascender). Ahora, como las modalidades hacen referencia, salvo en un caso, a la *recepción* o no de textos, una conclusión posible pero inmediatista —lo cual no quiere decir, de entrada, que sea falsa— sería que los textos entregados han sido contraproducentes. No obstante, dadas las limitaciones de los instrumentos, también cabría la pregunta de qué material impreso utilizaron alternativamente aquellos que no recibieron el material que el PUEBP distribuía y/o qué prácticas tuvieron que desplegar. De esta manera no se cierra la investigación con la sencilla decisión de acabar el “dar” textos escolares; el asunto como se ve es complejo: la investigación se abre hacia la indagación de los posibles materiales y/o prácticas alternativas a los textos escolares que han dado —según el estudio— mejores resultados, hacia la utilidad de la diversidad de material escrito en la escuela (no solamente ni principalmente textos escolares), así como hacia la precisión de la diferencia entre libros y textos escolares (o sea que los textos no son un factor aislado: cuando no están, no se nota mucho su ausencia).

- ☞ El que no se muestre asociación en unos casos y el que las asociaciones sean negativas en los otros, merece una aclaración: esto ocurre, pero *para los datos recogidos, obtenidos con los instrumentos tal y como éstos indagaron sobre los libros, para las perspectivas manejadas y para la población encuestada*. Si cualquiera de estos elementos se cuestiona, habrá que cuestionar todo el estudio. Esta precisión es fundamental, pues de tomar como “objetivos” los instrumentos, tendríamos que concluir como si éstos permitieran “ver” la realidad tal como es.
  
- ☞ A manera de ilustración sobre la incidencia de la postura del investigador sobre sus hallazgos, veamos cómo distintas elecciones frente a los mismos datos podrían dar resultados distintos. En la tabla de textos de lenguaje contra desempeño en grado 3° en la prueba abierta [cf. §4.2.6], al observar los desempeños de coherencia local y de coherencia global (los extremos), hemos señalado una asociación negativa. Pero al observador podrían interesarle otras lecturas; por ejemplo, atendiendo a que los extremos no suman el 20% de la población, podría interesarse por la mayoría restante (el medio), en cuyo caso observaría una asociación positiva; o, considerando que el nivel más alto es el esperado en condiciones óptimas (que no son las de nuestra escuela), podría centrar su atención más bien en los primeros niveles, en cuyo caso también vería una evolución positiva; aquel que rigiera su preocupación en relación con los mejores desempeños, en tanto indicadores de avanzados horizontes pedagógicos (para no regirse mediocremente), observaría sobre todo el nivel más alto, con lo cual encontraría una asociación negativa...
  
- ☞ La investigación puede volcarse sobre sí misma para ver cómo busca. Aparecen, así, todas las decisiones y limitaciones instrumentales con las que se indagó tanto los factores-libro asociables como el desempeño. Dos ejemplos: el primero es el de los rangos de significatividad: las diferencias entre lo “altamente significativo”, lo “moderadamente significativo” y lo “no significativo” fueron decisiones del equipo; si éstas hubieran sido más drásticas (como efectivamente lo son en otros estudios), estaríamos hablando de menos asociaciones estadísticas: si hubiéramos contemplado la significatividad a partir de un p-valor de 0.010, estaríamos hablando solamente de un 16.6% de correlaciones, contra el 50% que tenemos, gracias a unos márgenes mayores para la significatividad. El segundo ejemplo es el de la discriminación del material impreso [cf. §4.2.1 y 4.2.2], la cual proviene de una forma de concebir la educación en la que se privilegian los llamados “textos escolares”.
  
- ☞ De los estudios, investigaciones y políticas educativas revisadas [cf. §4.1] sólo unos pocos se refieren a las características del material escrito. A los demás, pa-

rece bastarles el hecho de entregarlo. Curiosamente, las características mencionadas —que tienen que ver con los libros y no con los textos escolares— se refieren a la calidad y a la variedad (CERLALC y La Misión). Acorde con las categorías usadas para el diseño de la prueba de lenguaje, valdría la pena investigar las características del material escrito (libros y textos escolares), en relación con propiedades como: *Potencialidad discursiva para generar interlocución* en los contextos escolares y extraescolares; *Universalidad* con el objetivo de cualificar abstracciones y acciones para trascender el contexto; *Estructuración no dogmática*, en función de la historia que hace posible un conocimiento; *Capacidad para sugerir redes intertextuales e intercognitivas* para propiciar la recuperación de los saberes y la exploración de las fuentes; *Seducción hacia la lectura y la escritura*; y *Aportes para la investigación de los profesores* [cf. Jurado y Bustamante, 1994:50-51]. Todo esto, por supuesto, requiere repensar la escuela y la pedagogía, lo que no quiere decir que no existan hoy en día experiencias en esta dirección.

☞ Para los niños encuestados, cuando a la diferencia de grado agregamos los factores-libro, comienza a aparecer una pequeña diferencia, gracias a la cual podría decirse que para buscar asociaciones entre libros y desempeño no sería exactamente lo mismo estar en tercero que estar en cuarto: en grado cuarto se presentan más asociaciones (8) que en grado tercero (6). De nuevo, como en el análisis del desempeño<sup>64</sup>, esto parece ser detectable desde la perspectiva estandarizada: cuatro asociaciones en grado cuarto, contra una en tercero; mientras que en el análisis por competencias, el asunto queda con una mínima diferencia a favor de tercero: mientras en el análisis por competencias-C sólo aparece una asociación (en grado tercero), en el análisis por competencias-A aparecen tres factores-libro en los que ambos cursos presentan asociación, otro factor en el que la asociación se presenta para tercero y un último factor para el que la asociación se presenta en cuarto. En todos los casos esta tendencia nacional no explica el desempeño de los niños vistos contra cada factor...

☞ En cuanto al desglose por grado, frente al resultado nacional,

- para el análisis estandarizado, en un caso se conserva la ausencia de asociación tanto a nivel nacional como en ambos cursos. En el resto de casos, uno de los cursos conserva el tipo de asociación y la tendencia;

---

<sup>64</sup> Recuérdese que en la introducción al capítulo II se menciona que, en cuanto al desempeño, el análisis estandarizado es el único desde el que se puede percibir una diferencia entre pertenecer a un grado o a otro.

- para el análisis por competencias-C, en todos los casos —salvo uno— a nivel nacional no se presentaba asociación, tendencia que se conserva en los dos grados. Para la excepción, el tipo de asociación y la tendencia se conserva para un grado; y
- para el análisis por competencias-A, en la mitad de los casos se conserva la asociación y la tendencia del nivel nacional. En el resto de casos, hay dos en los que la asociación se conserva para uno de los grados (una vez para cada uno). Y solamente en un caso, el desglose es completamente distinto al resultado nacional (libros recibidos por la escuela): a nivel nacional no presentaba asociación, mientras que en ambos cursos sí presenta asociaciones moderadas.

En resumen, en la mitad de los casos el desglose por grado no aporta información adicional. En los restantes —salvo uno—, uno de los dos grados tal vez introducía “ruido” cuando los datos se tomaban conjuntamente en el conglomerado nacional (efectivamente, el p-valor para la asociación en el grado en que se presenta es en casi todos los casos menor al p-valor nacional respectivo). Y en un caso, donde la paradoja de Simpson es flagrante, podríamos decir que la asociación no puede buscarse a nivel de la escuela o de otra instancia que contenga a los estudiantes (por ejemplo, tampoco en el DC se muestra asociación), sino a nivel del grado. Lo que el desglose sí deja claro es que la investigación no puede ir haciendo diferenciaciones cada vez mayores con la idea de ir corroborándose. Las tendencias de la asociación sólo se cambian en dos casos.

📖 En cuanto al desglose por modalidad de escuela, frente al resultado nacional, cinco veces (de 18) se conserva para las tres modalidades de escuela el tipo de asociación registrado a nivel nacional; 11 veces se conserva el tipo de asociación en dos de las modalidades; una vez se conserva el tipo de asociación en sólo una de las modalidades; y una vez en ninguna de las modalidades se conserva. Tal como hemos venido diciendo, podría decirse que, en el caso del desglose por modalidad de escuela sirve, en la mayoría de los 11 casos, para hipotetizar sobre la modalidad para la cual el factor que se indaga es indiferente para el desempeño (6 casos) o es justamente aquel para el que hay dependencia entre el factor y el desempeño (5 casos). De 17 casos en los que una asociación se presenta en lo nacional y se repite en la modalidad, 6 veces la tendencia de la asociación se cambia (también paradoja de Simpson).

📖 En este contexto, es sorprendente que escuela nueva sea aquella modalidad para la que son indiferentes los libros recibidos por la escuela, las guías de escuela nueva, los textos escolares en general, y los textos escolares de lenguaje, casi

siempre de cara a la producción (parte abierta). Es escandaloso que la única modalidad donde las guías de escuela nueva no se asocian al desempeño sea justamente en aquella modalidad para la que fueron hechas. El material que hace las veces de las guías favorece más el tránsito a desempeños altos. Todas las asociaciones que presenta escuela nueva coinciden con las asociaciones nacionales, en cambio, en el contexto de las otras modalidades, no presenta varias de las asociaciones de nivel nacional. ¿Esto quiere decir que los instrumentos no son sensibles a esa modalidad de escuela? En primera instancia, no sería el caso de la prueba de lenguaje, hecha con arreglo a la teoría que ha iluminado el área desde la Renovación Curricular y, de manera tal vez más específica en escuela nueva; pero, ¿acaso basta esto para garantizar que entre la teoría y el instrumento hay concordancia?

- ☞ En cambio, de un lado, escuela urbana aparece asociada casi tanto como el nivel nacional, y, de otro, escuela rural graduada aparece asociada más veces que el nivel nacional (y más que las otras dos modalidades juntas). Ésta es la modalidad que parece ser más sensible a los instrumentos; además, tiene la mitad de las asociaciones positivas —en el sentido esperado— para todos los análisis del desempeño en ambas escalas, en ambos grados y en las tres modalidades de escuela.
- ☞ La escala distrital casi no se distingue de la nacional: en el caso de los libros en la casa, en ninguna de las dos escalas se muestran asociaciones; o sea, en el DC se replican los resultados nacionales: los libros en la casa no incidirían en los desempeños de los niños en el área de lenguaje. Obtener la submuestra distrital de la muestra maestra, no suministra información adicional; es decir, nada dice el subconjunto que no diga el conjunto. Pero, como hemos dicho de la escala nacional, esto no se puede leer literalmente para hablar en contra, por ejemplo, de las campañas para que las familias adquieran libros, o de sus esfuerzos en tal sentido en tanto los libros en la casa no desarrollarían la competencia comunicativa de los niños. Queda abierta la posibilidad, señalada también para la escala nacional, de hacer otros análisis para indagar sobre la relación en mención. Y, en el caso de los libros recibidos por la escuela, cada escala muestra asociación con el desempeño en lenguaje sólo una vez: en lo nacional, con el análisis estandarizado; y a escala distrital, con el análisis por competencias-C. Es decir, para el DC el desempeño de los estudiantes en lenguaje o bien es independiente de los textos recibidos por la escuela, o bien se ve afectado negativamente por éstos: allí donde se reciben libros hay menos evolución de los desempeños hacia los niveles altos, y allí donde no se reciben libros, los desempeños tienden hacia esos niveles (su influencia es contraria a la esperada).

## 5. “Eficiencia interna”

No importa de lo que estén hablando, hablan de dinero\*

Como se anotó en §1.1, los referentes de la política educativa son: la cobertura, la tasa de escolarización, el número de docentes y la llamada “eficiencia interna”, la cual se entiende en relación con índices que miden la retención, la escolaridad, el tiempo promedio de duración de los estudios, el número de alumnos por profesor y los costos [cf. DNP, 1981]. Tanto así que aunque hace una década se estén introduciendo otros —mediante las indagaciones de factores asociables—, en el *Proyecto de presupuesto general de la nación 2000* (también conocido como “presupuesto de la verdad”), los indicadores que se tienen en cuenta para educación no han cambiado mucho: matrícula, cobertura (bruta y neta), número de establecimientos educativos, número de docentes en servicio, tasa de alumnos por docente y tasa de analfabetismo [Minhacienda-DNP, 1999:685].

En la evaluación de impacto del PUEBP se indagó por algunos de estos factores. Retomaremos tres entre ellos: la deserción, la repitencia (negativos de la retención y de la promoción [cf. §5.3.1]) y el número de alumnos por docente. No es que, de entrada, demos fe de que los factores de “eficiencia interna” sean los indicios para comprender y valorar nuestros procesos educativos (por ejemplo, que tengan una incidencia directa en el desempeño de los estudiantes), sino que quienes trazan la política educativa llegan a poner la calidad en función de factores como esos; haremos, entonces, el análisis para mostrar, desde su propia lógica, algunos de los problemas que comporta o a los que conduce.

### 5.1 La “eficiencia interna” como factor asociable al desempeño

El *Salto educativo* de Samper planteaba que la deficiente calidad de la educación explicaba los problemas de eficiencia (cobertura y retención) [DNP, 1994:11]. Por su parte, el *Plan educativo para la paz* de Pastrana define así la eficiencia del gasto: «la

---

\* Bloch, 1981:57.

calidad alcanzada (medida por las pruebas de Saber) en relación con los recursos financieros invertidos (gasto por alumno) [...]» [DNP, 1999:30].

A continuación aparecerán discursos oficiales y no oficiales que sustentan la llamada “eficiencia interna” (en lo relativo a deserción, repitencia y número de alumnos por docente) como factor asociable al desempeño de los estudiantes. Lo dividiremos en dos partes, tomando como eje las medidas sobre evaluación del rendimiento, en tanto ellas son puestas como foco de la política destinada a mejorar la “eficiencia interna”.

### ***5.1.1 Sustentaciones en general***

- § Durante Octubre de 1991 y Junio de 1992, el equipo de Harvard Institute for International Development, el CIUP (Centro de investigaciones de la Universidad Pedagógica Nacional) y el MEN realizaron en Colombia un estudio sobre repitencia en el que se presupone que la repitencia i) baja la autoestima y las expectativas de los niños; ii) baja las expectativas de los padres; iii) causa deserción; y iv) no sirve para mejorar los aprendizajes [MEN. 1993c.:6].
- § En los instrumentos de factores asociables suele preguntarse al estudiante por su permanencia en la institución educativa. En la evaluación que dio lugar al informe del MEN de 1997b se pregunta incluso de manera redundante: «¿Ha estudiado en otro colegio?», «Grados que ha realizado en este colegio» y «¿Cuántos años lleva estudiando en este colegio?». Al saber el grado del estudiante, con la respuesta sobre los grados que ha cursado en ese colegio podría inferirse si ha estudiado en otro colegio; ¿por qué tanta insistencia? Sin esperar los resultados de la aplicación y análisis del instrumento, ya se está haciendo un manejo de lo educativo en el que eficiencia y eficacia —entendidas como altos grados de retención y eliminación de la deserción— se consideran fundamentales. Esto puede comprobarse en el hecho de que el *Plan de Apertura Educativa*, que implementó el SINECE para investigar si asuntos como retención y deserción estaban asociados a la calidad, ya hablaba en estos términos; es decir, no se trata de un hallazgo de la “investigación” en factores asociables, sino de una política preestablecida.
- § Tales términos están también en los subsiguientes planes de gobierno en educación, *El salto educativo* y *Plan educativo para la paz*, así como en el *Plan decenal de educación*, casos en los que el asunto se menciona pero ya como resultado de la investigación.



- § El plan de desarrollo en educación del gobierno de César Gaviria [DNP, 1991, tabla 5] dice: «la mayor cobertura lograda por contratación de nuevo personal es el aumento por utilización de más recursos, mientras la ampliación debida al mejoramiento de la relación alumno-docente y a la disminución de la repitencia y deserción es aumento por mayor eficiencia». En esta dirección, se habla de ineficiencia porque «más de una quinta parte de los cupos disponibles son utilizados para atender a jóvenes por encima de la edad correspondiente, por repetición o por haber ingresado tarde» [:12]; se menciona que la promoción automática «no ha logrado reducir apreciablemente las tasas de repetición» [:9], que el PUEBP «no ha sido muy eficaz hasta ahora para aumentar la cobertura y mejorar la calidad» [:9].
- § El plan de desarrollo en educación del gobierno de Ernesto Samper habla del gasto por estudiante y lo compara con otros países y entre niveles educativos [DNP, 1994:9,10]; menciona que la mayor dificultad en educación es «la inadecuada retención del estudiante en el sistema» [:10,11]; que «los logros recientes en el aumento de cobertura se han visto contrarrestados por la baja eficiencia interna del sistema», es decir, poca retención y mucha repitencia [:11]. Dedicó todo un cuadro a la “trayectoria y eficiencia de una cohorte” [:12], en el que se habla de “eficiencia bruta” (en la que la deserción es considerada ineficiencia) y de “eficiencia neta” (en la que sólo la repitencia se considera ineficiencia). Otro cuadro [:40] está dedicado a las “tasas de promoción, deserción y repitencia”.
- § En el *Plan decenal* de educación, uno de los objetivos, relativo a cobertura, dice: «Disminuir el abandono y la repitencia para conseguir que el 90% de los niños que ingresen al grado primero, concluyan el 9° en no más de 9 años de escolaridad» [MEN, 1996a:22]. En el mismo documento se dice: «El sistema está afectado por serios problemas de ineficiencia: de 100 niños que inician el primer grado de primaria, 60 terminan el quinto y sólo 40 de ellos lo hacen en cinco años; de cada 100 niños y niñas que entran al primer grado, únicamente 30 terminan la educación básica y 7 lo hacen sin repetir año. Las tasas de repitencia y deserción son particularmente elevadas en los grados primero (25.5%), sexto y séptimo» [MEN, 1996a.]. Más adelante: «Solamente reconstruyendo la institución escolar, fortaleciéndola, ligándola a la comunidad, dándole recursos, capacidad decisoria y claras responsabilidades, podrán ser exitosas las acciones para transformar la gestión del sistema, hacerla eficiente, ampliar su cobertura, mejorar su calidad y hacerla mucho más competitiva» [:16].
- § La *Misión ciencia, educación y desarrollo* [1997:77] dice: «Aun cuando las estadísticas actuales indican que Colombia está en un nivel superior al de otros países en vías de desarrollo, el sistema educativo acusa serios problemas que se re-

flejan en las altas tasas de repitencia, deserción, deficiencia docente y pedagógica, inadecuados materiales e infraestructura, indisciplina y falta de educación para la democracia y la competencia». Agrega que una buena gestión (o sea, “eficiente”), «disminuye el costo social del desperdicio, no sólo para sus usuarios directos, sino para las comunidades alejadas y las generaciones futuras» [:130].

- § El Plan de desarrollo en educación del gobierno de Pastrana dice: «Si se emplearan los recursos con la máxima eficiencia, el sistema podría atender el 90% de los niños entre 7 y 17 años, y no sólo el 75% que actualmente atiende» [DNP, 1999:29]. Más adelante [:31] agrega: «El tiempo que tomaba hacer la primaria para quienes la terminaron en 1994 era 5.6 años, un mes menos que para quienes la hicieron 10 años antes<sup>65</sup>, lo cual representa una reducción de 1.6%, pero el costo real para promover un alumno con primaria completa aumentó en 19%. Es decir, el sistema en primaria se ha hecho más caro y sólo un poco más eficiente». La búsqueda de la eficiencia se pone incluso como condición para asegurar un derecho constitucional: «es importante mejorar la eficiencia del gasto social para desarrollar el principio constitucional que hace de la educación un derecho fundamental» [:33].
- § El exministro de educación Jaime Niño Díez, en la «Carta a los maestros» que sirve de presentación al documento *Factores asociados al logro cognitivo de los estudiantes. Grados 3º, 5º. 1993-1995* [MEN, 1997b], plantea que frente a la educación la empresa moderna, la más decisiva, es aquella que aboga por la calidad de la educación: «se trata de que nuestros niños *permanezcan* en la escuela, encuentren en el saber sentido y placer y se preparen en las aulas para conquistar un futuro mejor» [:6]. No es casual que la primera idea asociada a “calidad” sea la de “permanencia”, es decir, retención y promoción: no deserción y no repitencia.
- § El *Plan de desarrollo* de Pastrana plantea que «La política de equidad del Plan tiene una doble dimensión: la financiera y la de ampliación del acceso y la permanencia» [DNP, 1999:34]. Curiosamente, la inequidad estructural no va a ser atacada; solamente se pretende ampliar la cobertura *para todos*, siendo que el mismo documento habla de las dificultades para acceder a la educación para aquellos que no tienen cómo pagarla.

---

<sup>65</sup> Esto nos recuerda el cuento de Woody Allen en el que «Los astrónomos hablan de un planeta habitado llamado Quelm, tan alejado de La Tierra que un hombre que viajase a la velocidad de la luz tardaría seis millones de años en llegar a él, aunque se estudia una nueva ruta, más rápida, que reduciría en dos horas el viaje» [Allen, 1980:178-179].

- ❧ Igual situación es interpretada por la SED de la siguiente manera: en el capítulo que el plan sectorial dedica al fomento a la retención, dice que la institución escolar no logra desarrollar estrategias que motiven suficientemente a los alumnos para que permanezcan en ella; en respuesta a esto, «trabaja directamente con las instituciones educativas donde se identifiquen niños y jóvenes en “alto riesgo” de abandono escolar debido a sus condiciones afectivas y familiares. El objetivo es promover en ellos nuevas conductas de socialización que les permitan disfrutar del espacio escolar, sentir la aceptación social y considerar lo pertinente para su vida» [SED, 1998:32-33; cf. también :29]. En tal sentido, la Subdirección de Apoyo a la Comunidad Educativa de la SED, en 1998, estudió las situaciones que inciden en la deserción escolar y elaboró el proyecto “Fomento a la retención escolar”, que se aplica desde enero de 2000; se busca que los maestros cambien sus relaciones con los alumnos y que los alumnos adopten nuevas conductas, que se desempeñen autónomamente como ciudadanos [Carrillo, 2000:12]. No es claro en el documento, pero estos objetivos pretenden lograrse también con el programa “Opciones”, desarrollado por un consultor internacional y un equipo que tienen como propósito diseñar clases «que los alumnos esperen ansiosamente» [:13], con lo que, supuestamente, se disminuirá el conflicto escolar, se mejorará el rendimiento, se aumentará el nivel de compromiso y se disminuirá la deserción escolar [:13].
- ❧ De ahí que el ministro Germán Bula diga: «Existe consenso en los diferentes actores de la comunidad educativa sobre la necesidad de evitar el fracaso escolar que se manifiesta principalmente en la deserción y en la repitencia. Por esta razón, en los últimos años se han tomado medidas en las que se le reconoce un papel determinante a la evaluación y a la promoción en el aula» [Bula, 1999:3]. No se cita la fuente del “consenso”.
- ❧ El plan de desarrollo educativo del actual gobierno reduce el asunto educativo a un problema económico. Por ejemplo, para asegurar a los niños de 5 a 15 años la oportunidad de acceder a la educación básica, se ampliará la cobertura «con base en mejoras en la eficiencia, mediante la reforma del actual esquema de asignación de recursos, aumentando progresivamente la relación de alumnos por docente [...] hasta un promedio del orden de 30» [DNP, 1999:35; cf. también :37, 56]. Se dice que el actual esquema de financiación se cambiará por uno basado en resultados de cobertura, calidad y logros de eficiencia [:37].
- ❧ En el último informe de la evaluación de la calidad [MEN, 1997b:47] se reproduce una tabla en la que “no repetir” aparece influyendo positivamente en el desempeño de lenguaje de los niños de grado 3º, pertenecientes al sector oficial rural, que estudian en calendario B (o sea, no incide positivamente en matemáticas,

ni en 5°, ni en los demás sectores, ni en el otro calendario). Esto no quiere decir que la recíproca sea válida: que la repitencia incida negativamente en ese caso (con seguridad, no salió asociada). Paradójicamente, repetir “en el mismo plantel” ¡está asociado positivamente! al desempeño de lenguaje de los niños de grado 5°, pertenecientes al sector urbano oficial, que estudian en calendario A (o sea, no incide positivamente en matemáticas, ni en 3°, ni en los demás sectores, ni en el otro calendario). En calendario A, “Nunca haber repetido cursos” influye positivamente sólo en el logro individual de 5° [:64], en lenguaje, en el sector privado; y, en 3° [:67], sólo influye en lenguaje en el sector oficial rural y en el privado. O sea, nunca haber repetido cursos no influye ni negativa ni positivamente en la mayoría de los casos (por ejemplo, nunca en matemáticas).

- § En la misma tabla, no aparece la deserción, pero sí aparece el “ausentismo” debido a la necesidad de cuidar a los hermanos o de trabajar, o por falta de motivación, casi siempre asociado negativamente.
- § La Fundación social [1998:17] se une al coro: «La cobertura del sistema escolar ha aumentado, pero los excesivos niveles de repitencia y deserción son síntomas de la incapacidad de retención del sistema escolar y la baja calidad de la educación que recibe la mayoría de los niños y jóvenes colombianos».
- § La ley 508 de 1999 [Plan Nacional de Desarrollo, art. 17] dice: «el Plan tendrá por finalidad cumplir la tasa de asignación de personal docente por alumno, definida periódicamente por el MEN y el DNP, de acuerdo con la densidad de la población estudiantil y con las necesidades de cada entidad territorial para lograr la distribución equitativa de los docentes [...]».
- § Por su parte, la SED planea ampliar la cobertura de la siguiente manera: «con base en el análisis de la capacidad instalada de los colegios estatales del Distrito se determinará el número de niños que, de acuerdo con la infraestructura física y una adecuada relación alumno-docente deberían estar matriculados en cada uno de los centros educativos. Esta capacidad será contrastada con la matrícula actual, con el fin de establecer el número adicional de niños que podrían ingresar al sistema utilizando la infraestructura existente» [SED, 1998:29-30].
- § En relación con la tasa de alumnos por docente, los docentes siempre han sostenido que mientras más alumnos haya en un salón de clase, menos opción hay de llevar a cabo una labor personalizada: menos posibilidades de dedicarse a cada uno, a cada escrito, etc. Es decir, en términos del modelo de la política educativa en boga, “a más estudiantes, menos calidad”.

- ❧ Contra esta intuición —no demostrada, por supuesto— se levanta el pensamiento administrativo de la educación que sostiene haber hecho reiteradamente el hallazgo empírico de que el número de alumnos por docente no influye en los logros de los estudiantes. De ahí que se niegue a disminuir la proporción existente, arguyendo 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]<sup>66</sup>; e, incluso, que se proponga aumentar el número de alumnos por docente (tanto en lo nacional como en lo distrital, según vimos más atrás). Si ciertos instrumentos no encuentran correlación significativa entre cantidad de alumnos por curso y logro, la política de eficiencia, según la cual hay que aumentar la cobertura sin aumentar el número de profesores, quedaría plenamente justificada<sup>67</sup>.
- ❧ El último informe de la evaluación de la calidad reproduce una tabla [MEN, 1997b:47] en la que el promedio de alumnos por curso aparece influyendo negativamente en el desempeño individual de lenguaje de los niños de grado 3°, pertenecientes al sector urbano oficial, que estudian en calendario B (o sea, no incide negativamente en matemáticas, ni en 5°, ni en los demás sectores, ni en el otro calendario). En relación con el “logro promedio” [:61], el número de alumnos por docente influye sólo para matemáticas de grado 5° en el sector urbano; mientras que para grado 3° [:65] no influye en lenguaje para el sector rural ni en matemáticas para el sector privado urbano, y en el resto de casos influye en Calendario B (salvo matemáticas rural, sólo A). El logro promedio, cuando se mira contra el número de alumnos por curso, aparece asociado, pero «con bajo aporte a la explicación» [:66] y sólo en el sector privado.

### 5.1.2 Sustentaciones desde la evaluación del rendimiento

Para atacar la repitencia y la deserción se ha usado la evaluación escolar. Por la importancia que representa para el asunto de la “eficiencia interna”, a continuación ampliamos lo relativo a la evaluación, sobre todo a la promoción automática:

- para el sector rural se adoptó desde 1975 la modalidad de escuela nueva, en la que se habla de una *Promoción flexible*; se tenía en cuenta un factor decisivo para la zona rural: los ciclos de cosecha y siembra vinculan a los niños a ritmos laborales que no siempre dejan un tiempo constante para la escuela; de tal manera,

---

<sup>66</sup> La cita habla de alumnos por curso. Aunque no sea exactamente igual, la hacemos equivaler a alumnos por docente.

<sup>67</sup> No se piensa que al variar un factor (por ejemplo, aumentar el promedio de alumnos por docente) pueda comenzar a hacerse significativo si no lo era (era significativo pero no en la proporción medida).

se implementó dicha modalidad para que los niños regularan su propio aprendizaje, fundamentalmente mediante las Guías [cf. §3.1.1]: un niño podía adelantar trabajo por su propia cuenta y, si le tocaba ausentarse durante un tiempo, volvía al punto donde había quedado y continuaba. Así, el año lectivo podía durar más o menos que un año calendario (de todas maneras, el niño no puede pasar de una unidad a otra sin que las mismas estén totalmente terminadas). En esta modalidad de escuela, la evaluación se concibe como “formativa” y “sumativa”: de un lado, resultado de un seguimiento al trabajo adelantado por el estudiante, quien podrá pasar al nivel siguiente luego del visto bueno del profesor; y, de otro, labor del maestro ayudado por una prueba que recibe ya diseñada en un folleto anexo a las guías, las cuales contienen las respuestas “correctas” para que el niño las compare con las suyas. Los indicadores de evaluación se dan en términos de “excelente”, “muy bueno” y “satisfactorio”, y los aspectos socio-afectivos del niño se juzgan a través de su participación en el gobierno escolar;

- el decreto 1860 [art.7] plantea que el proceso pedagógico de la educación básica debe permitir la formación integral, la evaluación por logros y la *permanencia* del educando dentro del servicio educativo; así mismo, establece [art.52] condiciones para la promoción de los estudiantes no al finalizar cada grado, sino al finalizar sexto y noveno grados, de manera que se prolongue el tiempo de la decisión de una repitencia, con lo que el estudiante tendría más tiempo para recuperarse y no repetir e incluso no desertar:

*La promoción en la educación básica se fundamenta en el reconocimiento de la existencia de diferencias en el ritmo del aprendizaje de los alumnos. Por tanto, los educandos deben tener oportunidades de avanzar en el proceso educativo, según sus capacidades y aptitudes personales.*

*El Consejo Académico conformará comisiones de promoción integradas por docentes, para definir la promoción de los alumnos que al finalizar los grados sexto o noveno presenten deficiencias en la obtención de los logros definidos en el presente decreto.*

*Para tal efecto, las comisiones revisarán las evaluaciones practicadas en los grados precedentes, con el fin de determinar las actividades complementarias especiales que requieran cumplir para satisfacer debidamente los logros.*

*Las comisiones también podrán decidir la promoción anticipada de los alumnos que demuestren persistentemente la superación de los logros previstos para un determinado grado.*

*Parágrafo: Los establecimientos educativos que no ofrezcan el sexto o el noveno grado, transitoriamente definirán la promoción de los alumnos al finalizar el quinto grado y el último que ofrezcan.*

- desde 1987 rige la *Promoción automática*, que consiste —grosso modo— en que los docentes deben concentrar su trabajo en la idea de que ningún niño repita el año; en manos del docente estarían todas las posibilidades de realizar actividades que conlleven al aprendizaje de los niños; y si algunos temas no pueden ser comprendidos cabalmente, se puede flexibilizar el tiempo de manera que el niño continúe por los grados establecidos y comprenda más adelante aquello en lo que tiene dificultades. Se trataba de no echar por la borda un año de trabajo, a nombre de unos contenidos cuya “medición” era muy incierta, y de no volcar la responsabilidad fundamentalmente en el niño. Sabemos que esta política tiene que ver con medidas internacionales que mencionan la repitencia como un síntoma de mala calidad educativa y como un factor que encarece la educación de un país<sup>68</sup>;

La promoción automática fue impulsada durante el gobierno de Virgilio Barco, como parte del programa denominado «Acción nacional educativo-cultural». Fue reglamentada por el Decreto 1469 de agosto 3 de 1987 que, a su vez, reglamentaba el artículo 8° del decreto 0088 de 1976 sobre la ampliación de la cobertura de la educación básica primaria. Con ella se buscaba superar las altas tasas de deserción y repitencia que se presentaban en la básica primaria, registradas en las siguientes estadísticas [MEN, 1987:3-4]:

	Zona		Grado	
	urbana	rural	1°	5°
Cobertura en básica primaria	90%	70%		
Niños de 7 a 11 años sin acceso a educación básica	30%			
Retención (estudiantes que terminan 5° de primaria)	60%	20%		
Promedio de años requeridos para concluir primaria	7	20		
Grados promedio de escolaridad alcanzada	3.7	1.7		
Repitencia			20%	70%

Analfabetismo	Absoluto	2.25 millones de personas
	Funcional	5.0 millones de personas
	Entre 12 y 44 años	54.1%
	Mayores de 44 años	45.9%

Con ayuda de estos datos, se asociaban los altos niveles de deserción y repitencia, con la ineficiencia del sistema escolar, y con los conceptos de valores y calidad. En

<sup>68</sup> Más allá de esta lectura, se puede pensar que dicha medida abría la posibilidad de tener en cuenta los diversos ritmos de aprendizaje.

esta lógica, aumentar el cubrimiento de la básica primaria a las zonas marginales rurales y urbanas, y evitar la repitencia, debía corresponder con el mejoramiento de la calidad de la educación. Buena parte de las causas de la “mala calidad de la educación” era atribuida a las formas de evaluación numérica, que supuestamente creaba situaciones de “frustración” y “tristeza” en los niños. En este sentido, el PUEBP implicaba, no sólo que los niños tuvieran acceso a la educación, sino que permanecieran y avanzaran. Aumentar el cubrimiento en la básica primaria y mejorar su calidad, debía contribuir a erradicar la pobreza absoluta, bandera del gobierno de Virgilio Barco.

Identificada la evaluación como la causa de la baja calidad de la educación, era necesario impulsar un *Plan nacional de evaluación* que tuviera como fundamento los procesos de desarrollo de los niños, que reconociera sus capacidades y que permitiera afianzar los valores democráticos. En este orden de ideas, el Decreto 1469 [Artículo 2º] hace una serie de definiciones:

- Entiende por Evaluación escolar el «proceso de seguimiento y valoración permanente del estado en que se encuentra la institución educativa en sus aspectos organizacionales y administrativos, pedagógicos y de desarrollo del alumno, frente a los fines y objetivos del sistema educativo colombiano. Este proceso tiene como propósito la identificación de los logros y dificultades, y la aplicación de correctivos que hagan posible una alta calidad educativa».
- Entiende por Promoción automática «el proceso por medio del cual después de un seguimiento permanente de evaluación escolar, todo niño que cursa un grado en el nivel de Educación Básica Primaria, es promovido al grado siguiente al finalizar el año lectivo o antes, si sus capacidades y logros se lo permiten».
- Entiende por Calificación el «juicio de valor sobre el desarrollo y progreso del alumno expresado en una escala convencional».
- Y entiende por Actividades de recuperación «el conjunto de acciones planeadas y desarrolladas a lo largo del año escolar, con el propósito de que el alumno logre los objetivos de grado no alcanzados en las diferentes áreas de formación. Dichas actividades de recuperación se desarrollarán permanentemente en el tiempo escolar de que trata el Decreto 1002 de 1984 y como parte del proceso regular de enseñanza-aprendizaje».

Definidos los conceptos sobre los cuales se habría de trabajar, el Decreto establece los ámbitos de la vida escolar sobre los cuales debe recaer el Sistema: procesos pedagógicos, organizacionales y administrativos.



En cuanto a los procesos pedagógicos (entendidos éstos como «el conjunto de acciones orientadas por los educadores» [MEN, 1987:11]), el Decreto 1469 define los fines de la Educación Básica Primaria como «el desarrollo del conocimiento, actitudes, valores, habilidades, destrezas psicomotrices y del proceso de socialización». Dichos procesos implican una mirada sobre el currículo<sup>69</sup>, el cual debe ser de carácter regional para permitir el reconocimiento de las diferencias culturales y los niveles de aprendizaje de cada niño. Estos dos últimos aspectos definen la idea de “ritmos de aprendizaje” la cual a su vez define y diferencia entre promoción “automática” y “flexible”: gracias a la primera, el niño debe ser promovido al grado siguiente al terminar el año lectivo; gracias a la segunda, el niño puede ser promovido al grado siguiente, independientemente de la terminación del año escolar, siempre y cuando haya logrado los objetivos.

En cuanto a los procesos organizacionales, se crean los “Comités de evaluación” (resolución 13676 reglamentaria del párrafo del artículo 7° del Decreto 1469), buscando ampliar la participación de los agentes educativos para consolidar la propuesta de evaluación integral: según el artículo 3, formarán parte del comité de evaluación: el director o quien haga sus veces; el coordinador de primaria (si hubiere); el orientador o quien haga sus veces (si hubiere); un docente, elegido cada año por el consejo de maestros; un representante de los padres de familia, designado por la junta directiva de la asociación (si hubiere), o por los asistentes a asamblea general. La resolución, además de reglamentar y definir los participantes del comité, define sus funciones, entre las que se cuentan: orientar y controlar el proceso evaluativo; promover la actualización permanente de docentes; asesorar a docentes, alumnos y padres en la programación y ejecución de las actividades evaluativas; diseñar y/o modificar formatos para libros de calificaciones y boletines informativos de que tratan los artículos 3 y 4 del Decreto 1469 de 1987<sup>70</sup>; conceptuar y decidir sobre los casos excepcionales a que dé lugar la aplicación del Decreto 1469 de 1987 (particularmente, los artículos séptimo, octavo y noveno<sup>71</sup>); dar informes periódicos a la comunidad educativa y a los interesados sobre las acciones del comité; conceptuar, en última instancia, sobre conflictos presentados durante el proceso evaluativo.

---

<sup>69</sup> Las decisiones para la promoción automática, en los aspectos curriculares, recaen sobre «el dominio adquirido por el niño en el área de la comunicación y en el área lógico-matemática» [MEN, 1987:16].

<sup>70</sup> Estos artículos hacen referencia a la responsabilidad de los docentes en el proceso evaluativo.

<sup>71</sup> Que establecen, respectivamente: a) la asistencia a clase como criterio de evaluación: no se considerará cursado un grado si hay 20% o más de fallas no justificadas con respecto al número de horas clase realizadas; b) Según el cual, el comité de evaluación puede promover al grado siguiente a quien haya logrado todos los objetivos al finalizar el segundo periodo académico del año lectivo. Medida aplicable también para transferir alumnos de un calendario a otro; y c), El cual establece que el comité de evaluación puede promover alumnos con dificultades excepcionales que no logren los objetivos.

Y en cuanto a los procesos administrativos y organizacionales, el artículo 5 del Decreto faculta a los rectores de las instituciones para establecer un “Plan de evaluación coherente con la filosofía y políticas del Decreto”, lo cual compromete: las relaciones entre los sujetos educativos (docentes, alumnos, padres, administrativos) y de éstos con la comunidad; la comunicación y la información dentro y hacia afuera de la institución escolar; la participación y los procedimientos para la toma de decisiones; la elaboración, organización, ejecución y supervisión del plan institucional, incluida la administración de los recursos financieros y de planta física; y la normatividad institucional [MEN, 1987:13].

Se observa que el Decreto atribuye una serie de responsabilidades a los agentes educativos:

- Los docentes —responsables del proceso pedagógico que se adelante con los niños— deben reconocer los niveles de desarrollo cognitivo en cada niño<sup>72</sup> y, de acuerdo con ellos, determinar el lugar que les corresponde en el sistema de evaluación. Ante los estudiantes, el maestro sería el responsable —por su acción frente al sistema de evaluación— de su “felicidad” o de su “frustración” (con la promoción automática, los alumnos siempre tendrán una nueva oportunidad<sup>73</sup>). Ante el sistema educativo, sería el responsable de la retención de los niños, al igual que de la “buena calidad” de la educación. Es responsabilidad de los maestros la aplicación de la promoción automática, al igual que definir en qué situación aplica promoción Flexible o Automática. Esto implica para el maestro adelantar un proceso de formación permanente, lo cual lo colocaría en el camino de la “innovación”.
- Los estudiantes no tendrían excusa para retirarse del sistema escolar; gracias a la promoción automática, pueden lograr que el aprendizaje y el desarrollo adquieran “valor por sí mismos” [MEN, 1987:17] a diferencia de lo que ocurría con la evaluación numérica, en donde lo aprendido era proporcional al miedo que se tuviera a la evaluación. A partir de la aplicación de la promoción automática será “insuficiente” no el alumno, sino sus logros en el aprendizaje y su desarrollo: en el sistema de evaluación numérico, la pérdida era absoluta, lo cual significaba la expulsión; en cambio, con la promoción automática, aparece la deuda permanente.

---

<sup>72</sup> Para lo cual estarían las teorías cognitivas y constructivistas.

<sup>73</sup> A este aspecto corresponden las actividades de recuperación. Artículo 2º, Decreto 1469 de 1987.

- Y para las instituciones educativas y las directivas, su papel está centrado en el control permanente del desarrollo del sistema de evaluación, así como en su administración.

## **5.2 Algunos factores-“eficiencia interna” asociables al desempeño en lenguaje**

Analizaremos las posibles asociaciones entre los factores *repitencia*, *deserción* y *alumnos por docente* —ligados a la llamada “eficiencia interna”— y el desempeño en la prueba de lenguaje aplicada en la evaluación de impacto del PUEBP.

### **5.2.1 Repitencia**

Cuando en la evaluación de impacto del PUEBP se habla de “repitencia”, se hace referencia al número de estudiantes que cada escuela reporta que repiten, información que en el archivo de datos ha sido “colgada” a cada niño. De otro lado, la repitencia no procede de procesos homogéneos; es decir, que un grupo de niños repita no implica que todos hayan sido sancionados de esa manera bajo los mismos conceptos: no es posible unificar los criterios de los profesores (o de las distintas “comisiones de promoción” [cf. Decreto 1860/94]) al respecto. Dos repitencias no significan lo mismo.

La indagación oficial sobre la relación entre repitencia y desempeño parece ser indiferente ante discordancias como las señaladas. Parece más importante buscar a toda costa una dependencia entre la llamada “eficiencia interna” de la institución y el desempeño (pero así, se encuentra lo que se busca).

La evaluación de impacto se hizo, como hemos dicho, a niños que empezaban o que habían acabado de terminar los grados 3° y 4° de primaria, pues se trataba de evaluar una política enfocada a la educación básica. Es de suponerse [cf. §5.1] que estos niños han estado influidos por la promoción automática, la promoción flexible y las últimas medidas sobre evaluación en la reciente legislación educativa. Con todos estos antecedentes, era esperable *no encontrar* repitencia entre la población evaluada. No obstante, se encontró, y alta (hay alguna repitencia en el 79.2% de las escuelas), tal como se registra en la siguiente tabla, correspondiente a las escuelas:

	Repitencia		
	Baja	Media	Alta
Casos de repitencia	De 0 a 2	Entre 3 y 7	8 o más
Porcentaje	37.0	36.2	26.8

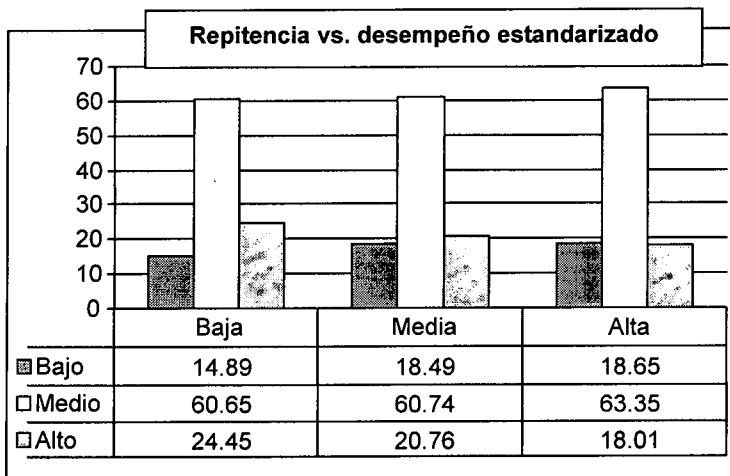
Se han escogido estos límites, a partir del comportamiento observado en los datos del rango de repitencia (que va entre 0 y 70, de manera no consecutiva). La repitencia baja se ha definido de 0 a 2 porque no hay distinción relevante entre escuelas en las que ha repetido 0, 1 o 2 estudiantes. Y a partir de 8 casos de repitencia, las cifras disminuyen considerablemente (obsérvese que, no obstante ir de 8 a 70, es el rango de menor porcentaje, con 10 puntos menos que los otros dos rangos). Pero se advierte que considerar estas categorías de repitencia como ordinales (en forma creciente o decreciente) puede conducir a error o a arbitrariedad en la interpretación, pues como no todas las escuelas tienen el mismo número de estudiantes, los porcentajes no necesariamente engloban números de repitencia de escuelas del mismo tamaño o por lo menos similar. Así, un número cualquiera de repitencia correspondería a porcentajes distintos según el tamaño de la escuela.

De todas maneras, haremos la lectura tal como hacen quienes determinan la política educativa, pues las distintas maneras de interpretar el desempeño en el área de lenguaje (contra las que se cruzará la repitencia) nos darán lugar a ver las dificultades que hay para justificar la posición que ya se tiene construída de antemano: evitar la repitencia como indicador de “eficiencia interna”, lo que, a su vez, sería indicador de calidad.

La razón por la que podría influir en el desempeño de los estudiantes —según el discurso oficial— es porque una escuela con mayor repitencia sería una escuela de menor calidad. Se quiere mostrar la calidad en relación directa con el gasto, pues calidad —en el discurso actual sobre educación— pasa por la eficiencia y la eficacia.

Con el presente trabajo, intentamos mostrar que sacar ese tipo de conclusiones es muy apresurado y que, ajustándonos a los datos, se pueden sacar otro tipo de conclusiones y hacer nuevas preguntas para investigar.

Veamos, entonces, la información disponible en la evaluación de impacto del PUEBP sobre repitencia:



Como puede observarse, el mejor desempeño alto estaría en las escuelas donde hay una repitencia baja. ¿Habría que recomendar, entonces, que se haga perder el año a no más de dos niños para que la calidad de la escuela mejore? (eso sería consistente con la manera como interpretan quienes pasan de los datos a la política educativa, sin mediación). El tercer puntaje alto está entre quienes más repiten. De donde, podríamos concluir que la baja repitencia parece favorecer el paso a desempeños altos. Mirando desde quienes permanecen en desempeños bajos, se nota que el mayor porcentaje está entre quienes estudian en escuelas con repitencia alta; y el menor desempeño bajo está entre quienes estudian en escuelas con repitencia baja; lo que resulta consistente con el dato sobre el paso a desempeños altos. En otras palabras, desde este análisis puede afirmarse que la repitencia baja tiende a favorecer el paso a desempeños altos y a desfavorecer la permanencia en desempeños bajos. Tiene un p-valor de 0.001; es decir, correlación altamente significativa, negativa (que es la esperada, en tanto la repitencia se considera negativa): según las cifras, a menos repitencia, mejores desempeños.

Desde los desempeños por competencias-C, aparecen las siguientes cifras:

		Repitencia		
		Baja	Media	Alta
Competencias-C	Nivel 0	0.96	1.46	0.93
	L. Literal	25.96	21.83	25.71
	L. Inferencial	58.35	62.89	60.39
	L. Crít-intertext.	14.72	13.83	12.97
p-valor		0.003(-)		

Vemos que los desempeños literales se mantienen más o menos constantes; es decir, que para quedarse en lecturas literales, es indiferente estar en una escuela con baja o con alta repitencia (incluso podría preguntarse si dicha competencia requiere de una enseñanza formal...). En cambio, para hacer lecturas crítico intertextuales, se ve que en las escuelas donde hay más repitencia, hay menos niños con este tipo de desempeño. Hay una asociación altamente significativa, con una tendencia negativa que no es consistente en todos los desempeños. Los desempeños de los niños ubicados en escuelas con repitencia media muestran la tendencia más consistente con la concepción de lectura en la que se inspiró la prueba de lenguaje; leído taxativamente con la intención de relacionar “eficiencia interna” y desempeño, habría que concluir que se necesita cierto grado de repitencia en las escuelas para que los niños evolucionen en su proceso de lectura; cosa que se contradice con el propósito mismo que ilumina la búsqueda. Tal vez no habría que tratar de probar algo que puede sencillamente ponerse como un propósito que nadie discutiría: es mejor que los niños no repitan el año; el problema es que ese propósito estaría justificado desde intereses muy distintos: desde el ahorro presupuestal hasta el sentido de la educación.

Desde los desempeños por competencia-A, aparecen las siguientes cifras:

		Repitencia		
		Baja	Media	Alta
<b>Competencias-A</b>	Nivel 0	2.31	1.89	3.65
	C. Local	9.56	11.40	13.01
	C. Lineal	82.04	79.13	78.72
	C. Global	6.09	7.59	4.62
p-valor		0,001+		

La correlación es altamente significativa, con un sentido no lineal; si bien los desempeños locales se ven favorecidos por la repitencia alta, los desempeños globales muestran la tendencia contraria hasta la repitencia media, es decir, hasta ahí tendríamos una asociación no lineal; sin embargo, en la repitencia alta el desempeño global decrece por debajo del correspondiente a la repitencia baja. Visto por los extremos, tendríamos una clara correlación negativa.

Vistos los desempeños por grado, se obtienen las siguientes cifras para la correlación con el factor repitencia:

			Grado					
			Cuarto			Tercero		
			Repitencia			Repitencia		
			Baja	Media	Alta	Baja	Media	Alta
<b>Desempeño</b>	Estandarizado	Bajo	7.05	10.63	12.76	22.73	26.21	24.11
		Medio	60.32	59.29	64.26	60.99	62.17	62.50
		Alto	32.63	30.08	22.98	16.28	11.62	13.39
		p-valor	0.001(-)			0.013(-)		
	Competencias cerrada	Nivel 0	0.92	1.56	0.60	1.01	1.36	1.23
		L. Literal	26.17	23.16	24.19	25.76	20.53	27.12
		L. Inferencial	58.22	61.71	61.49	58.47	64.04	59.38
		L. Crítico-intert.	14.68	13.57	13.72	14.77	14.08	12.28
		p-valor	0.188			0.008(-)		
	Abierta	Nivel 0	2.57	1.21	2.52	2.05	2.54	4.69
		C. Local	9.41	10.43	13.49	9.71	12.35	12.56
		C. Lineal	82.17	80.26	80.45	81.92	78.02	77.11
		C. Global	5.86	8.10	3.53	6.32	7.08	5.63
		p-valor	0.001(-)			0.002(-)		

El grado tercero muestra asociaciones contra todos los análisis del desempeño: el análisis estandarizado muestra una correlación moderadamente significativa, en cuyos extremos habría una tendencia negativa; el análisis por competencias-C muestra una asociación altamente significativa, con tendencia negativa (salvo la repitencia media en el desempeño literal); y el análisis por competencias-A también muestra una correlación altamente significativa, con tendencia negativa (salvo la repitencia media en el desempeño global).

En grado cuarto, se muestran dos correlaciones (el desempeño en la parte cerrada, analizado por competencias, no muestra correlación), ambas altamente significativas y de tendencia negativa. Es notorio que los mayores porcentajes de desempeños de coherencia global en la parte abierta, en ambos grados, está entre los niños cuyas escuelas tienen un nivel de repitencia media, lo que altera la tendencia negativa general, como en el caso de tercero.

Por modalidad de escuela, se obtienen las siguientes cifras para la dependencia entre los desempeños y el factor repitencia:

		Escuela urbana			
		Repitencia			
		Baja	Media	Alta	
Desempeño	Estandarizado	Bajo	13.52	18.51	15.91
		Medio	60.47	58.33	62.80
		Alto	26.01	23.16	21.29
		p-valor	0.002±		
	Competencias cerrada	Nivel 0	0.91	1.33	0.86
		L. Literal	25.42	21.61	30.54
		L. Inferencial	58.91	62.32	59.57
		L. Crítico-intert.	14.76	14.75	9.03
		p-valor	0.001±		
	Abierta	Nivel 0	1.64	1.68	3.18
		C. Local	9.17	11.03	12.73
		C. Lineal	82.82	79.62	77.05
		C. Global	6.37	7.67	7.05
		p-valor	0.044±		

En escuela urbana, cuando la repitencia por escuela es más alta, los desempeños bajos de sus estudiantes, en el análisis estandarizado, son más altos, lo cual querría decir que el nivel de repitencia de las escuelas estaría directamente relacionado con la detención de los niños en rendimientos bajos (escuelas de más repitencia, tendrían más proporción de estudiantes con rendimientos bajos); sin embargo, no hay linealidad: las escuelas con repitencia media tienen el mayor porcentaje de niños que se quedan en desempeños bajos. Por su parte, los desempeños altos decrecen conforme aumenta la repitencia. Es decir, se describe una tendencia negativa entre los desempeños y los niveles de repitencia, en el marco de una correlación altamente significativa.

Para los desempeños vistos desde las competencias-C, también tenemos una asociación altamente significativa. La tendencia, entre los extremos de repitencia es negativa; pero el menor logro literal no está en repitencia baja, sino en media, y para los desempeños crítico intertextuales, parece indiferente la diferencia entre deserción baja y media. En la parte abierta, la correlación es moderadamente significativa; la repitencia favorece tanto los desempeños bajos (el mayor puntaje en coherencia local está en deserción alta) como los altos (es más alto el puntaje de desempeños globales en repitencia alta que en baja; en consecuencia, la asociación es no lineal. El mejor puntaje en desempeño global está en deserción media... como si poca repitencia no fuera lo mejor.



		Escuela nueva			
		Repitencia			
		Baja	Media	Alta	
Desempeño	Estandarizado	Bajo	16.70	20.72	19.08
		Medio	64.64	62.03	62.16
		Alto	18.66	17.25	18.76
		p-valor	0.507		
	Competencias cerrada	Nivel 0	1.30	0.97	0.98
		L. Literal	26.46	21.28	25.74
		L. Inferencial	57.48	64.12	59.43
		L. Crítico-intert.	14.75	13.63	13.85
		p-valor	0.286		
	Abierta	Nivel 0	3.09	2.52	3.70
		C. Local	7.60	10.22	12.47
		C. Lineal	81.24	80.15	80.14
		C. Global	8.08	7.11	3.70
		p-valor	0.002		

En escuela nueva, la única correlación que se muestra es en la prueba abierta (aunque en las otras dos las tendencias de las cifras sean parecidas a los casos anteriores, los perfiles de los desempeños son muy similares). La asociación es altamente significativa y de tendencia claramente negativa: en coherencia local, a medida que aumenta el nivel de repitencia en las escuelas, se registra un mayor porcentaje de estudiantes; y el paso a desempeños de coherencia global va en sentido inverso al aumento de la repitencia en la escuela. Esta modalidad de escuela es la menos relacionada; las otras dos se correlacionan en las tres formas de interpretar el desempeño. Posiblemente esto tenga que ver con el hecho de que, de las tres, es la modalidad de escuela que nació con una forma evaluativa dirigida a evitar la repitencia (medidas como la promoción automática se dan *a posteriori* del funcionamiento de las modalidades graduadas, sea en la ciudad o en el campo). Esto nos lleva a buscar las tasas de repitencia por modalidad de escuela. Curiosamente, escuela nueva tiene el menor porcentaje de repitencia baja (21.98%, contra 45,79% y 38.85% de escuela urbana y rural graduada, respectivamente) y ¡el mayor porcentaje de repitencia alta! (43.73%, contra 13.84% y 34.81% de escuela urbana y rural graduada, respectivamente).

		Escuela rural graduada			
		Repitencia			
		Baja	Media	Alta	
<b>Desempeño</b>	Estandarizado	Bajo	18.18	12.26	21.16
		Medio	56.62	69.73	67.25
		Alto	25.19	18.01	11.59
		p-valor	0.001(-)		
	Competencias cerrada	Nivel 0	0.78	3.45	0.87
		L. Literal	27.53	24.52	19.13
		L. Inferencial	57.14	62.45	64.06
		L. Crítico-intert.	14.55	9.58	15.94
		p-valor	0.002(+)		
	Abierta	Nivel 0	4.08	1.20	4.13
		C. Local	13.32	16.47	14.75
		C. Lineal	79.89	73.90	77.29
		C. Global	2.72	8.43	3.83
		p-valor	0.008±		

En la escuela rural graduada, para todos los análisis del desempeño, éste y la repitencia presentan correlaciones altamente significativas. En el análisis estandarizado, los porcentajes de desempeños altos decrecen a medida que aumenta la repitencia; esto es solidario con los extremos del desempeño bajo: aumento a medida que aumenta la repitencia; es decir, salvo la disminución del desempeño bajo en la repitencia media, la asociación sería negativa. En el análisis por competencias-C, pasa exactamente lo contrario: quedarse en desempeños literales ocurre en escuelas con poca repitencia, mientras los extremos del desempeño crítico intertextual van en sentido inverso a la repitencia; salvo por la disminución del desempeño crítico intertextual en repitencia media, la correlación sería positiva. El análisis por competencias-A es un poco menos claro: entre los extremos de la producción local y global, los porcentajes aumentan a medida que aumenta la repitencia (pero en ambos casos el de repitencia media es el más alto); tendríamos, entonces, una tendencia no lineal.

En conclusión, donde más se presentan asociaciones altas es en la escuela rural graduada. Parece que los desempeños en esta modalidad fueran “más sensibles” (tiene tres asociaciones altamente significativas) a la relación con repitencia. En la modalidad de escuela nueva, donde no se debería presentar repitencia dada su estrategia de promoción, se presenta una correlación alta en producción. En la modalidad urbana, se presentan dos asociaciones altas y una correlación moderada. La prueba de desempeño que siempre está asociada es la de producción (también fue así por grado y en lo nacional); en términos psicométricos, se podría calificar de “consistente”. La

prueba de comprensión aparece altamente asociada en escuela urbana y en rural graduada. La detección de la correlación entre desempeño y repitencia es mucho más clara en curso tercero, escuela rural graduada y mediante la prueba abierta.

A escala distrital, las cifras son las siguientes:

		Distrito			
		Repitencia			
		Baja	Media	Alta	
Desempeño	Estandarizado	Bajo	5.00	12.50	10.00
		Medio	65.00	65.00	70.00
		Alto	30.00	22.50	20.00
		p-valor	0.865		
	Competencias cerrada	Nivel 0	0.00	5.00	5.00
		L. Literal	40.00	25.00	5.00
		L. Inferencial	60.00	60.00	80.00
		L. Crítico-intert.	0.00	10.00	10.00
		p-valor	0.166		
	Abierta	Nivel 0	5.26	6.06	0.00
		C. Local	10.53	18.18	0.00
		C. Lineal	78.95	75.76	88.89
		C. Global	5.26	0.00	11.11
		p-valor	0.230		

A diferencia del análisis nacional, en el que la repitencia estaba asociada a las tres formas de interpretar el desempeño, a escala distrital no se muestra ninguna asociación. Mientras a escala nacional las cifras del análisis estandarizado y del análisis por competencias-C hablan de que el desempeño se ve afectado negativamente por la repitencia (allí donde hay más repitencia, los desempeños tienden a no evolucionar hacia los niveles superiores, y allí donde hay menos repitencia, los desempeños sí tienden a evolucionar hacia dichos niveles); y mientras el análisis por competencias-A habla de que el desempeño y la repitencia están asociados, aunque no se puede establecer si positiva o negativamente; mientras esto pasa a escala nacional, a escala distrital podría afirmarse, de acuerdo con los datos, que el desempeño no parece depender de la repitencia; de donde la pretensión de aplicar una política general, según la cual la repitencia tiene que ver con baja calidad, pierde la dimensión de lo local que, como en este caso, puede ser completamente distinta de la tendencia nacional. No es cierto, para el DC, que la repitencia tenga una incidencia negativa; pero tampoco se puede inferir que tenga una incidencia positiva; sencillamente se trata, para este caso, de que el desempeño y la repitencia son independientes.

El hecho de que, a escala nacional, en los tres casos haya asociación, y que, a escala distrital, no haya, podría advertir que “repitencia” es un objeto para el DC, y otro distinto para el país.

### **5.2.2 Deserción**

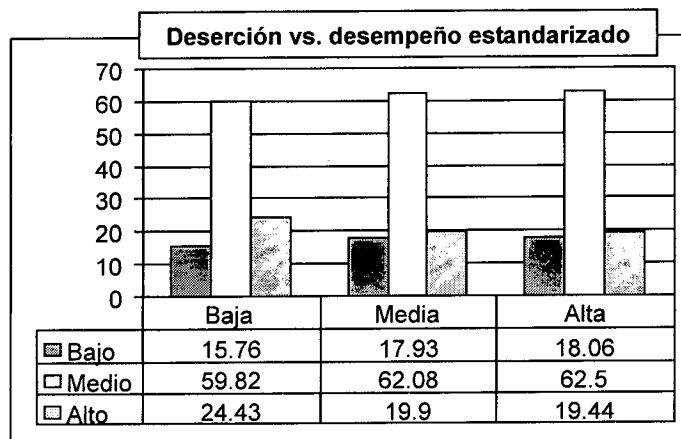
Los datos de “deserción” —tal como los de “repitencia”— también corresponden a las escuelas; no se trata de la deserción de cada niño.

La deserción que manifiestan las escuelas en los instrumentos de factores asociables se ha clasificado de la manera como aparece en la siguiente tabla:

	<b>Deserción</b>		
	Baja	Media	Alta
Número de casos	De 0 a 2	De 3 a 7	8 o más
Porcentaje	35.24	34.61	30.15

Como en el caso de la repitencia, los límites se han escogido a partir de la distribución observada en los datos (el rango de deserción va entre 0 y 130, pero de manera no consecutiva). La última categoría se escoge a partir de 8 ya que, después de ese valor, en su mayoría los porcentajes decrecen rápidamente hasta hacerse menores al 1%. Igualmente se advierte que estas categorías de deserción no son estrictamente ordinales (crecientes o decrecientes): como no todas las escuelas tienen el mismo número de estudiantes, los porcentajes no necesariamente cobijan números de deserción de escuelas de tamaños al menos similares. De todas maneras, haremos la lectura para buscar argumentos que discutan la idea —que ya se tiene construida de antemano— de que evitar la deserción mejora la calidad.

Ajustándonos a los datos que se obtuvieron durante la evaluación de impacto del PUEBP, veamos qué es posible concluir.



En la gráfica tenemos los desempeños de lenguaje, analizados desde el puntaje estandarizado. Con un p-valor de 0.001, aquí vemos que el porcentaje de desempeños altos va disminuyendo a medida que nos acercamos a las tasas más altas de deserción: la diferencia llega a ser de casi 5 puntos. E, inversamente, el porcentaje de desempeños bajos va aumentando a medida que nos acercamos a las tasas más altas de deserción. Entonces, habría una asociación altamente significativa negativa (que es la esperada por el modelo educativo, en tanto considera que la deserción afecta negativamente el desempeño).

El análisis de los desempeños en la parte cerrada de la prueba de lenguaje mediante las categorías del grupo de investigación no muestra correlación: tiene un p-valor de 0.320, valor que excede el límite que hemos puesto para asociaciones moderadas. Es decir, tiene más incertidumbre de la que hemos establecido aquí para considerar una asociación numérica como una pista de trabajo. Veamos las cifras:

		Deserción		
		Baja	Media	Alta
<b>Competencias-C</b>	Nivel 0	1.06	1.17	1.18
	L. Literal	23.68	24.61	25.00
	L. Inferencial	61.88	60.96	58.49
	L. Crít-intert.	13.38	13.27	15.33
p-valor		0.320		

No obstante, puede verse que el mayor porcentaje de desempeño crítico intertextual aparece cuando hay deserción alta; pero la tendencia no lineal hace que también el mayor porcentaje de desempeño literal esté en ese rango de deserción.

El análisis desde la categorización de los desempeños en la producción (parte abierta de la prueba de lenguaje), tiene un p-valor de 0.034. En este caso, las cifras son las siguientes:

		Deserción		
		Baja	Media	Alta
Competencias-A	Nivel 0	2.71	1.99	2.91
	C. Local	9.86	11.51	12.26
	C. Lineal	80.79	79.82	79.60
	C. Global	6.64	6.68	5.22
p-valor		0.034		

La correlación, moderadamente significativa, muestra una tendencia negativa: el aumento de la deserción favorece los desempeños más fáciles y desfavorece los desempeños más complejos (aunque la diferencia entre deserción baja y media parece no importar a la hora de la producción con coherencia global).

Desglosados por curso, los datos que aparecen para la relación entre deserción y desempeño son los siguientes:

			Grado					
			Cuarto			Tercero		
			Deserción			Deserción		
			Baja	Media	Alta	Baja	Media	Alta
Desempeño	Estandarizado	Bajo	8.85	10.11	10.74	22.72	25.65	24.78
		Medio	56.88	62.45	64.23	62.78	61.71	60.91
		Alto	34.27	27.44	25.03	14.50	12.64	14.31
		p-valor	0.001			0.397		
	Competencias cerrada	Nivel 0	1.23	1.08	0.86	0.88	1.25	1.48
		L. Literal	23.40	25.72	24.60	23.96	23.51	25.37
		L. Inferencial	62.40	60.83	57.25	61.36	61.09	59.62
		L. Crítico-intert.	12.97	12.36	17.29	13.79	14.16	13.52
		p-valor	0.026			0.820		
	Abierta	Nivel 0	2.41	1.43	2.42	3.01	2.54	3.37
		C. Local	9.46	12.42	10.70	10.26	10.62	13.68
		C. Lineal	81.63	79.94	81.59	79.94	79.70	77.79
		C. Global	6.49	6.21	5.29	6.78	7.14	5.16
p-valor		0.178			0.086			

En tercero aparece una asociación, en cuarto dos. ¿Quiere esto decir que el asunto de la deserción tiene que ver menos con el desempeño de los niños de tercero, pero que comienza a determinar su desempeño más claramente cuando pasan a cuarto?, ¿o

que la prueba de lenguaje —como hemos insistido— no mira cierta parte del desempeño que sí estaría asociada a dicho factor?

La asociación que aparece en tercero es moderadamente significativa; la dirección sería claramente negativa si no es porque el mejor desempeño global está en la deserción media. En cuarto aparecen asociaciones con el desempeño estandarizado, con el mismo p-valor y la tendencia de lo nacional. En el caso del desempeño por competencias, a diferencia de lo nacional aparece una correlación moderadamente significativa, pero no lineal: el mayor porcentaje de desempeño crítico intertextual está entre quienes tienen la mayor deserción (las otras dos deserciones parecen independientes del desempeño); pero ésta tiene el segundo porcentaje más alto de comprensión literal. Para los desempeños en producción no se presenta asociación.

Ninguno de los análisis de desempeño muestra asociaciones en ambos cursos. Pero, ¿puede un factor ser indiferente y, al año siguiente, ser importante (o viceversa)?

De todas maneras, la muestra contiene información de escuela que se traslada indiscriminadamente a los niños de tercero y cuarto.

Ahora, vista la relación desde la modalidad de escuela, se dispone de las siguientes cifras:

		Modalidad de escuela									
		Urbana			Nueva			Rural graduada			
		Deserción			Deserción			Deserción			
		Baja	Media	Alta	Baja	Media	Alta	Baja	Media	Alta	
<b>Desempeño</b>	<b>Estandarizado</b>	Bajo	14.57	17.02	16.45	18.91	21.01	18.06	16.75	14.94	20.51
		Medio	57.66	61.46	62.16	63.83	59.87	63.88	63.45	71.37	58.99
		Alto	27.77	21.52	21.39	17.26	19.12	18.06	19.80	13.69	20.51
		p-valor		0.001			0.487			0.034	
	<b>Competencias cerrada</b>	Nivel 0	1.03	1.03	1.28	0.47	1.42	1.06	1.78	1.24	1.40
		L. Literal	23.02	25.57	26.33	23.88	23.38	25.17	25.89	22.41	22.47
		L. Inferencial	61.99	59.69	57.77	62.88	63.19	58.12	60.41	62.24	60.67
		L. Crit-inter	13.95	13.71	14.63	12.77	12.01	15.66	11.93	14.11	15.45
		p-valor		0.609			0.164			0.768	
	<b>Abierta</b>	Nivel 0	2.03	1.57	2.20	2.56	2.83	3.61	5.36	2.13	2.01
		C. Local	9.08	10.59	13.57	9.97	10.65	10.93	12.60	18.72	14.08
		C. Lineal	82.12	81.02	76.25	80.31	80.20	80.52	76.41	72.34	81.90
		C. Global	6.76	6.82	7.98	7.16	6.32	4.95	5.63	6.81	2.01
		p-valor		0.097			0.642			0.002	

Como puede verse —en escuela urbana—, en el análisis estandarizado los desempeños bajos tienden a crecer a medida que aumenta la deserción; mientras que los desempeños altos disminuyen con la deserción. Es decir, tenemos una correlación negativa (para obtener mejores desempeños, sería mejor estar en escuelas con menor deserción: el desempeño más alto está en la deserción baja), altamente significativa (p-valor: 0.001); en este caso, se conserva el análisis nacional. El análisis por competencias-C muestra una tendencia parecida en la comprensión literal, pero no hay linealidad en el nivel crítico intertextual; no se muestra asociación (p-valor: 0.609), tal como en lo nacional. El análisis por competencias-A sí muestra correlación, moderadamente significativa (p-valor: 0.097), como en lo nacional; no obstante, la tendencia es no lineal: obsérvese que los desempeños, tanto locales como globales, aumentan con la deserción; es decir, ésta contribuye a tener desempeños bajos, pero también desempeños altos. Nada puede concluirse con certeza. Podría pensarse que una escuela que genera deserción (y que, en consecuencia —diría el modelo en boga—, su calidad deja qué desear), también puede generar capacidad para producir textos, lo que estaría ligado a una buena calidad.

En escuela nueva, *ningún* desempeño resulta asociado con la deserción, lo que equivale a decir que son independientes uno de la otra; en otras palabras, que los desempeños pueden ser buenos, regulares o malos, sin que la deserción de la escuela influya en ello; de otra manera: que en escuelas nuevas con alta deserción se dan desempeños de todos los tipos, y que en escuelas nuevas con baja deserción también se da todo tipo de desempeños. En el análisis estandarizado (que tiene un p-valor de 0.487), vemos que tanto los desempeños altos como los bajos son muy similares en los extremos de deserción; también en ambos casos la deserción media tiene los mayores puntajes. En el análisis por competencias-C, tanto en la comprensión literal como en la crítica intertextual, los porcentajes de los desempeños aumentan con el incremento de la deserción; en ambos casos, en la deserción media están los porcentajes más bajos. Aunque tampoco se muestra asociación (p-valor: 0.164), el comportamiento es no lineal. Y el análisis por competencias-A, pese a que los valores de la producción local crecen a medida que crece la deserción, pese a que los valores de la producción global decrecen con el aumento de la deserción, incluso pese a que la dificultad para asignar un nivel de producción (los que quedan en nivel 0) crece con la deserción —es decir, una clara tendencia negativa—, la correlación entre deserción y desempeño no es significativa: p-valor de 0.642.

Y en escuela rural graduada se presentan dos asociaciones: una alta y otra moderada. En el análisis estandarizado, los desempeños bajos, en los extremos de deserción, crecen a medida que ésta crece; sin embargo, el porcentaje más bajo está en deserción media. En los extremos, los desempeños altos son casi iguales; también en este caso la deserción media configura el porcentaje más bajo. La correlación es modera-



damente significativa, con un p-valor de 0.034. En el análisis del desempeño por competencias-C, para la comprensión literal, es indiferente que la escuela tenga deserción media o alta; pero la más baja deserción tiene los porcentajes más altos de desempeños literales. La comprensión crítico intertextual se muestra favorecida por la deserción. Es decir, tenemos una tendencia positiva, pese a que no se muestra correlación (p-valor: 0.768). En el análisis del desempeño por competencias-A tenemos una asociación altamente significativa (p-valor: 0.002). En los extremos de deserción, el porcentaje de desempeño local aumenta con la deserción; y disminuye para el porcentaje de desempeño global. Es decir, tenemos una tendencia negativa; sin embargo, la deserción media tiene los más altos porcentajes en ambos casos. Esta manera de analizar el desempeño es la que muestra más casos por fuera de las categorías de producción (incluso más casos que algunos de los desempeños globales); esto mostraría que el instrumento para calificar esta modalidad (abierta) en la prueba es un tanto incapaz de interpretar la producción de los estudiantes.

En general, cuando se desglosa por modalidad de escuela, la correlación nacional en el análisis estandarizado se mantiene en escuela urbana y en rural graduada. El análisis por competencias-C continúa sin mostrar correlación. Y la asociación nacional en el análisis por competencias-A se mantiene en escuela urbana y en rural graduada.

Para el DC, las cifras son como sigue:

			Distrito	
			Deserción	
			Baja	Media
<b>Desempeño</b>	<b>Estandarizado</b>	Bajo	12.50	7.50
		Medio	67.50	65.00
		Alto	20.00	27.50
		p-valor	0.609	
	<b>Competencias cerrada</b>	Nivel 0	5.00	2.50
		L. Literal	42.50	5.00
		L. Inferencial	50.00	80.00
		L. Crít.-intertext.	2.50	12.50
		p-valor	0.0013	
	<b>Abierta</b>	Nivel 0	5.88	2.78
		C. Local	14.71	8.33
		C. Lineal	76.47	83.33
		C. Global	2.94	5.56
		p-valor	0.706	

La escala distrital está recategorizada porque la muestra no presenta “deserción alta” (es decir, en la muestra del DC no se presentan más de 7 casos de repitencia por es-

cuela). Pese a esto, con ánimo descriptivo podemos ofrecer una comparación con la escala nacional, atendiendo a que las modalidades conservan el orden en las dos escalas.

Mientras a escala nacional los análisis del desempeño que resultan asociados muestran tendencias negativas —es decir, que la deserción incidiría negativamente en el desempeño—, a escala distrital el único análisis del desempeño que resulta asociado (competencias-C) muestra una tendencia positiva: como se ve en la tabla anterior, en las escuelas que presentan más deserción, los desempeños tienden a evolucionar hacia los niveles más altos; y en las escuelas que presentan menos deserción, los desempeños tienden a quedarse en los niveles bajos. De nuevo (como en el caso de la repitencia), la idea de aplicar una política homogéneamente no tiene lugar si se hacen estudios más puntuales. De cualquier manera, entender la educación como un conjunto complejo permite entender ambos hallazgos, en la medida en que la deserción (y la repitencia y los demás factores asociables) podría mostrar asociación pero indirectamente; es decir, si otro factor (o conjunto de factores) muestra asociación y está relacionado con el que se comenta, éste muestra asociación, pero no directamente. De tal manera, la política educativa no puede pensarse como la aplicación directa de lo que asociaciones numéricas pueden mostrar (y entonces se habla de evitar la deserción), sino como la investigación del conjunto producido por los factores en juego.

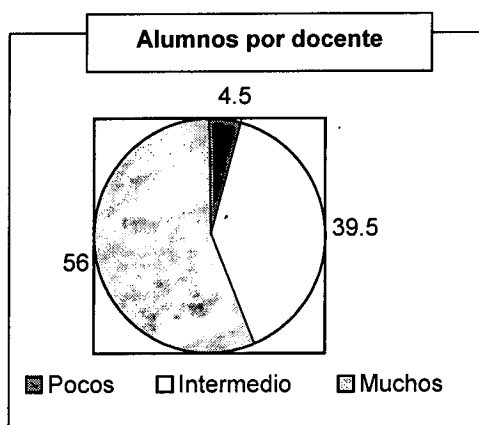
Curiosamente, los análisis del desempeño que muestran asociación a escala nacional son los que a escala distrital no resultan asociados, y viceversa. Esto podría llamar la atención en el sentido de pensar si “deserción” es un objeto (por ejemplo, por tener causas y condicionamientos diferentes) para el DC, y otro para el país.

### *5.2.3 Número de alumnos por docente*

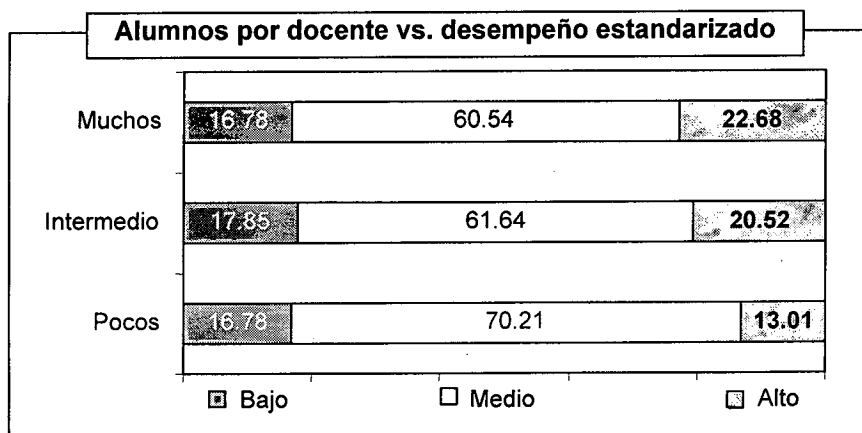
En el asunto de la tasa de estudiantes por docente, como hemos anotado más atrás, no hay solamente hipótesis “pedagógicas” (como las de los maestros, para quienes la posibilidad de brindar una mejor educación es inversamente proporcional a la cantidad de estudiantes que cada uno tenga), sino también *económicas* (como la del gobierno, para el cual una cosa no tiene que ver con la otra). Si la premisa de los docentes fuera cierta, habría que disminuir el número de estudiantes por docente, para que la calidad mejorara; si la premisa de los economistas fuera cierta, se podría aumentar la cobertura sin detrimento de la calidad ni, sobre todo, del presupuesto.

Pues bien, en esta investigación hemos determinado tres rangos de proporción entre docente y alumnos: “pocos”, cuando un docente tiene hasta 10 estudiantes; “inter-

medio”, cuando un docente tiene entre 11 y 24 estudiantes; y “muchos”, cuando tiene más de 24 estudiantes. Escogimos estos límites sobre la base del último promedio estimado por el Ministerio de Hacienda y Planeación Nacional para el número de alumnos por docente en Colombia: 23.2 para el sector oficial (para el sector privado, se estima en 19.0) [MINHACIENDA-DNP, 1999:685]. Vale la pena anotar la discrepancia con los datos de la evaluación de la calidad, que en su último informe estima un promedio de 23 alumnos por docente en la zona rural oficial, 27 en la urbana oficial<sup>74</sup> y 28 en la urbana privada [MEN, 1997b:34], en las instituciones encuestadas (entonces, ¿algún problema con el promedio del gobierno, con la muestra de la evaluación de la calidad, con ambas?). Así se distribuyen por porcentajes los grupos definidos:



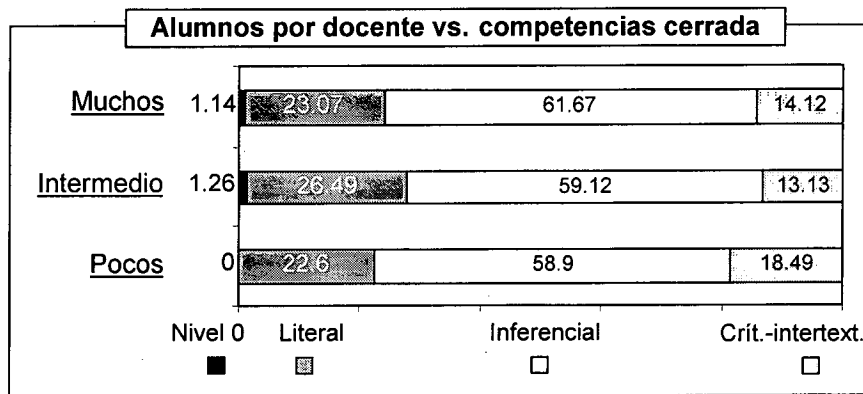
Si hacemos el cruce de esta variable con el desempeño de los estudiantes visto desde los puntajes estandarizados, tenemos un comportamiento estadístico que muestra una correlación positiva entre ambos aspectos (p-valor 0.001):



<sup>74</sup> Esta cifra es muy aproximada a la que la SED [1998:10] establece para el DC: un promedio de 26.55 alumnos por docente en las localidades urbanas

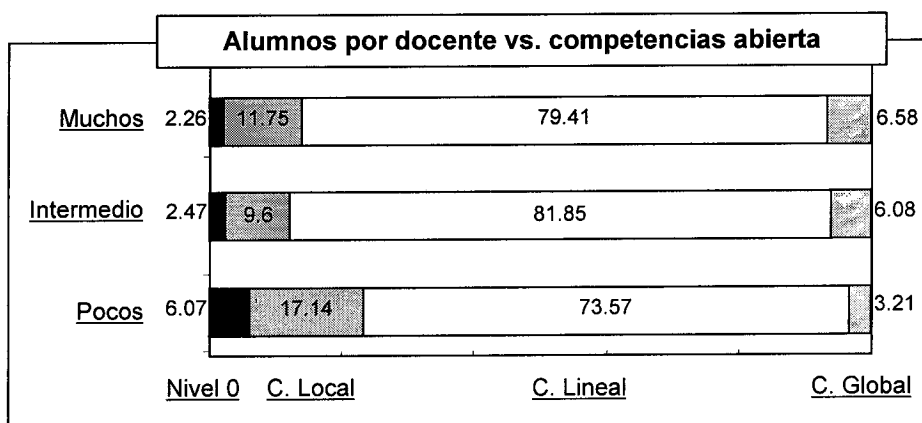
Los desempeños bajos son casi idénticos en los tres casos de número de alumnos por profesor; para alcanzar desempeños medios, es más favorable tener pocos estudiantes; y para llegar a desempeños altos, es más favorable tener alto número de alumnos por docente a tener un número intermedio de alumnos por docente (en 7.51%), y en 9.67% es mejor tener muchos estudiantes que tener pocos. La asociación es altamente significativa y positiva, más allá incluso de la expectativa gubernamental, según la cual una cosa no dependería de la otra.

En cambio, cuando hacemos el cruce de esta variable con el desempeño de los estudiantes en la parte cerrada de la prueba, visto desde las categorías que definió el grupo, vemos:



Los desempeños literal e inferencial son ligeramente superiores entre quienes tienen muchos estudiantes, frente a los que tienen pocos estudiantes (con diferencias de 0.47% y de 2.77%, respectivamente); y el desempeño crítico-intertextual es superior entre quienes tienen pocos estudiantes. Es decir, la correlación se realiza en cualquiera de los sentidos que se lean los datos: por los desempeños bajos y por los altos. La correlación es altamente significativa (p-valor de 0.005), y negativa: a menos estudiantes, mejores desempeños.

Desde el desempeño en la parte abierta de la prueba, los resultados son similares a los del análisis por competencias-C. Con una asociación altamente significativa (p-valor: 0.001) y positiva, los desempeños globales, en el caso de quienes tienen muchos alumnos por docente, son el doble de quienes tiene pocos; y los desempeños locales son las dos terceras partes entre quienes tienen muchos alumnos por docente.



Veamos si el desglose por curso permite avanzar en este sentido. Las cifras son las siguientes:

		Grado						
		Cuarto			Tercero			
		Alumnos por docente			Alumnos por docente			
		Pocos	Intermedio	Muchos	Pocos	Intermedio	Muchos	
<b>Desempeño</b>	Estandarizado	Bajo	10.14	10.72	9.20	23.61	24.73	24.17
		Medio	74.32	61.92	59.20	65.97	61.36	61.84
		Alto	15.54	27.36	31.59	10.42	13.91	14.00
		p-valor	0.001+			0.766		
	Competencias cerrada	Nivel 0	0.00	1.36	0.95	0.00	1.16	1.31
		L. Literal	22.97	27.44	22.67	22.22	25.58	23.46
		L. Inferencial	58.11	57.28	62.68	59.72	60.90	60.69
		L. Crítico-intert.	18.92	13.92	13.69	18.06	12.36	14.54
		p-valor	0.013-			0.225		
	Abierta	Nivel 0	3.50	1.70	2.21	8.76	3.23	2.31
		C. Local	17.48	10.96	10.22	16.79	8.27	13.24
		C. Lineal	74.13	81.73	81.12	72.99	81.97	77.75
		C. Global	4.90	5.61	6.45	1.46	6.53	6.71
		p-valor	0.094+			0.001+		

Tercero está correlacionado una sola vez (contra prueba abierta). En cuarto, en cambio, el número de alumnos por docente muestra asociaciones contra cualquier forma de mirar el desempeño. Para la parte cerrada de la prueba (tanto estandarizada como por competencias), el desglose por grado muestra que la correlación nacional oculta el hecho de que la sensibilidad al asunto del número de estudiantes por profesor sólo se muestra en cuarto; en otras palabras, según las cifras, el asunto es indiferente para los desempeños de los niños de tercero. En consecuencia, no se puede pensar en este factor como incidiendo en la calidad *independientemente* del curso —quizá de la

edad—. En cambio, en relación con la parte abierta de la prueba sí aparece una sensibilidad para ambos cursos (se conserva la tendencia nacional). En el análisis estandarizado para cuarto, la correlación, es altamente significativa (p-valor: 0.001) y positiva: a más estudiantes por profesor, mejores desempeños; tal como en lo nacional. En el análisis por competencias-C, la asociación es moderadamente significativa (p-valor: 0.013) y negativa: los porcentajes de los desempeños crítico intertextuales disminuyen a medida que aumenta el número de estudiantes por profesor (se tiene mejores desempeños allí donde los profesores tienen pocos estudiantes); para los desempeños literales, en los extremos de número de alumnos por docente no hay diferencia en el desempeño. En el análisis por competencias-A, la asociación es moderadamente significativa (p-valor: 0.094) y positiva: a medida que aumenta el número de estudiantes por docente, disminuyen los desempeños locales y aumentan los globales. Esta es la única manera en la que el desempeño se muestra dependiente del factor en mención en grado tercero, donde la correlación también es altamente significativa y de tendencia positiva.

Ahora, veamos las cifras de alumno por docente contra desempeño en el área de lenguaje, desglosada por modalidad de escuela:

		Modalidad de escuela									
		Urbana			Nueva			Rural graduada			
		Alumnos por docente			Alumnos por docente			Alumnos por docente			
		Pocos	Inter-medio	Muchos	Pocos	Inter-medio	Muchos	Pocos	Inter-medio	Muchos	
<b>Desempeño</b>	<b>Estandarizado</b>	Bajo	17.00	14.52	16.40	11.02	20.95	18.02	25.68	17.71	16.19
		Medio	76.00	58.59	59.80	71.19	63.33	60.77	60.81	63.98	64.05
		Alto	7.00	26.89	23.81	17.80	15.72	21.21	13.51	18.31	19.76
		p-valor	0.001(+)			0.002±			0.322		
	<b>Competencias cerrada</b>	Nivel 0	0.00	1.33	1.01	0.00	1.22	0.99	0.00	1.21	2.14
		L. Literal	31.00	26.07	23.67	24.58	26.75	21.54	8.11	26.76	23.10
		L. Inferencial	56.00	59.92	60.76	57.63	57.44	64.73	64.86	61.17	60.00
		L. Crít-intert.	13.00	12.68	14.56	17.80	14.59	12.75	27.03	10.87	14.76
		p-valor	0.295			0.029±			0.001-		
	<b>Abierta</b>	Nivel 0	7.29	2.46	1.39	9.01	2.28	3.44	0.00	2.89	4.51
		C. Local	22.92	8.60	10.60	9.01	10.13	11.49	21.92	10.33	18.55
		C. Lineal	66.67	82.57	80.61	77.48	81.73	79.15	76.71	80.79	73.43
		C. Global	3.13	6.37	7.40	4.50	5.86	5.92	1.37	5.99	3.51
		p-valor	0.001+			0.011+			0.001(+)		

La prueba abierta es la única que muestra asociaciones en todas las modalidades de escuela. En escuela nueva la asociación es moderadamente significativa, mientras en las otras dos es altamente significativa. Se conserva la tendencia nacional del de-

sempeño en abierta, que es positiva, con la excepción de escuela rural graduada, donde el puntaje más alto del desempeño global está en el número intermedio de alumnos por docente. Escuela nueva es la única modalidad que muestra asociaciones entre todos los análisis del desempeño y número de estudiantes por docente.

El análisis por competencias-C muestra una correlación moderadamente significativa, pero no lineal: los porcentajes más altos, tanto del desempeño literal como del crítico intertextual, están entre quienes tienen pocos estudiantes por profesor (favorece el desempeño elemental, pero también el más complejo).

En escuela rural graduada, en cambio, la asociación —altamente significativa— conserva la tendencia nacional: negativa.

El análisis estandarizado conserva la tendencia positiva en una correlación altamente significativa en escuela urbana (salvo que el mayor porcentaje de desempeños bajo y alto está en el número intermedio de alumnos por docente). En escuela nueva, también es altamente significativa la asociación, pero es no lineal.

A escala distrital, las cifras son como sigue:

		Distrito		
		Alumnos por doc		
		Interm.	Muchos	
<b>Desempeño</b>	Estandarizado	Bajo	5.00	11.67
		Medio	65.00	66.67
		Alto	30.00	21.67
		p-valor	0.574	
	Competencias cerrada	Nivel 0	0.00	5.00
		L. Literal	40.00	18.33
		L. Inferencial	60.00	66.67
		L. Crít.-intertext.	0.00	10.00
		p-valor	0.108	
	Abierta	Nivel 0	5.26	3.92
		C. Local	10.53	11.76
		C. Lineal	78.95	80.39
		C. Global	5.26	3.92
		p-valor	0.987	

En este caso, la escala distrital también está recategorizada porque la muestra no presenta casos en la categoría “pocos alumnos por docente” (es decir, en la muestra del DC habría más alumnos por docente que en el promedio nacional). Pese a esto,

con intención descriptiva, podemos presentar una comparación con la escala nacional, atendiendo a que las modalidades conservan el orden en las dos escalas.

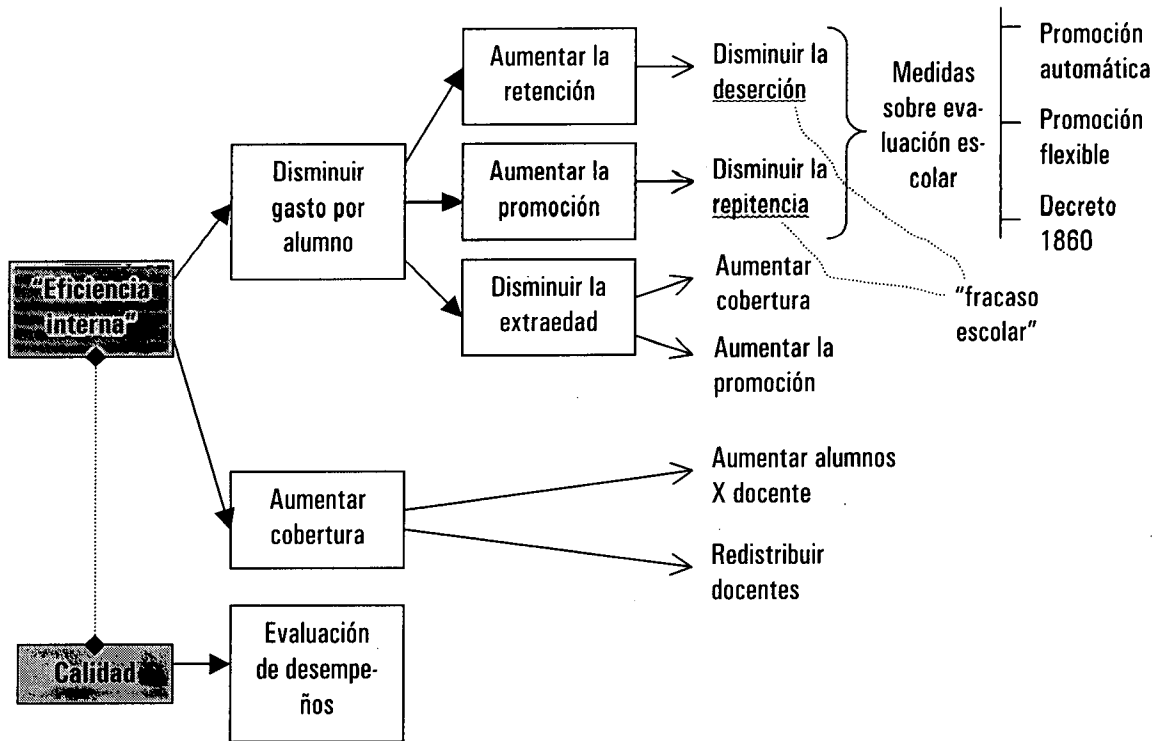
Mientras a escala nacional se presentaban asociaciones para los cruces entre alumnos por docente y desempeño, visto éste desde las tres formas de análisis adoptadas, a escala distrital ninguno de los análisis del desempeño muestra asociación con el factor en mención. Es decir, mientras a escala nacional se puede afirmar que de acuerdo con las cifras es posible hablar de una asociación (en dos casos positiva, en el otro negativa), a escala distrital habría que decir que el desempeño de un estudiante en la evaluación del PUEBP es independiente del número de compañeros con los que comparte la clase de lenguaje. Así, si a escala nacional podría pensarse que el hallazgo amerita decisiones educativas (aumentar o disminuir el número de estudiantes por docente, de acuerdo con cada una de las tendencias de las asociaciones), a escala distrital se justificaría la idea de que se puede ampliar la cobertura sin aumentar el número de docentes, pues no se evidencia asociación; en cuyo caso no sería, en principio, perjudicial ni benéfico *para el desempeño* (aunque lo sea —según el modelo de política educativa vigente— para el valor per cápita imputado al estudiante).

Como veremos, conclusiones como éstas no se pueden sostener. De nuevo, se trata de analizar un sistema complejo y no concluir linealmente; por ejemplo, pensar si la idea del número de estudiantes con los que un profesor debe trabajar es distinta en Bogotá que en el resto del país. Si esto fuera así, ¿hasta dónde las acciones tendrían que provenir de análisis locales?

### **5.3 Resumen parcial**

5.3.1 A continuación hacemos un esquema de la concepción que aparece de “eficiencia interna” en §5.1:





### 5.3.2

Factores	Desempeño					
	Estandarizado		Competencias-C		Competencias-A	
	p-v	Asociación	p-v	Asociación	p-v	Asociación
Repitencia	0.001	Alta-negativa	0.003	Alta-(negativa)	0.001	Alta-no lineal
Deserción	0.001	Alta-negativa	0.320	No	0.034	Moderada-negativa
Alumnos X docente	0.001	Alta-positiva	0.005	Alta-negativa	0.001	Alta-positiva

5.3.3 A continuación se comparan las cifras del análisis nacional con los factores grado y modalidad de escuela en relación con la eficiencia interna.

	Repitencia						
	Nacional	4°	3°	Urbana	Nueva	Rural grad	Distrital
Estandarizada	0.001(-)	0.001(-)	0.013(-)	0.002(-)	0.507	0.001(+)	0.865
Competencias-C	0.003(-)	0.188	0.008(-)	0.001±	0.286	0.002(+)	0.166
Competencias-A	0.001±	0.001(-)	0.002(-)	0.044±	0.002±	0.008±	0.230

### Deserción

	Nacional	4°	3°	Urbana	Nueva	Rural grad	Distrital
Estandarizada	0.001-	0.001-	0.397	0.001-	0.487	0.034±	0.609
Competencias-C	0.320	0.026±	0.820	0.609	0.164	0.768	0.001+
Competencias-A	0.034-	0.178	0.086(-)	0.097±	0.642	0.002(-)	0.706

### Alumnos por docente

	Nacional	4°	3°	Urbana	Nueva	Rural grad	Distrital
Estandarizada	0.001+	0.001+	0.766	0.001(+)	0.002±	0.322	0.574
Competencias-C	0.005-	0.013-	0.225	0.295	0.029±	0.001-	0.108
Competencias-A	0.001+	0.094+	0.001+	0.001+	0.011+	0.001(+)	0.987

## 5.4 Un análisis parcial

- § De las posiciones expuestas sobre “eficiencia interna” [cf. §5.1] se infiere una concepción que relaciona de manera tautológica el presupuesto y las cifras de factores asociables; esta situación es idéntica en otros campos: Michel [1996:29] afirma que el enfoque en términos de indicadores cuantitativos es antiguo en los campos económico, ecológico y de la salud, pero reciente en el campo educativo.
- § Triana [1996:48] señala que a esta tendencia, iniciada en otros ámbitos sociales hace más de treinta años se ha *sumado* recientemente la educación. En cualquier caso, las diferencias entre campos exigiría una prudencia que no parece observarse en la aplicación indiscriminada que tiene lugar en nuestro país: la idea de capitación, de eficiencia interna, de más clientes por unidad de tiempo y por funcionario son comunes en muchos campos de la vida pública. Entonces, podrían no ser esos los principales indicios para comprender y valorar nuestros procesos educativos. No compartimos dicho enfoque que, además, no ha sido construido en un espacio en el que la participación tenga que ver con la decisión [tal como proponemos en §7].
- § Para este enfoque de la política educativa, el proceso educativo se puede pensar desde una óptica más o menos empresarial: que se logren los propósitos (y entonces cada vez más se habla de que sean medibles<sup>75</sup>) y que se logren en la me-

<sup>75</sup> Es el caso del PEI: se crea «para lograr la formación integral del educando», «debe responder a situaciones y necesidades de los educandos, de la comunidad local, de la región y del país» y, no obstante debe ser «concreto, factible y evaluable», de manera que pueda compararse con otros y participar en concursos cuyos ganadores tendrán premios [cf. Ley 115, art. 73].

nor cantidad de tiempo y con el menor gasto posible. Tal vez esa perspectiva sea válida en ciertos contextos (no negamos, por ejemplo, que producir más y mejores alimentos en menos tiempo y con menor inversión, sea algo deseable)<sup>76</sup>, pero se requiere demostrar que es aplicable en educación; es decir, que en principio no estaría justificada la generalización que se está haciendo. La misma UNESCO [1998:60] habla del *peligro* de «caer en la opinión de que las escuelas pueden ser organizadas como fábricas».

- § Michel [1996:22] también plantea la necesidad de estar en guardia contra una visión simplista de la aplicación de la teoría del mercado y de la competencia en el sector educativo: la teoría (en economía) de que el mercado permite alcanzar la coherencia del conjunto, descansa en ideas como competencia perfecta, transparencia de la información, regulación sólo interna y estaticidad, que no tienen pertinencia en educación, donde hay un horizonte temporal, el conjunto también es efecto de incidencias externas, y se está entre lo público (con funciones tutelares) y lo privado.
- § La perspectiva de “eficiencia interna” hace que la repitencia, la deserción y el número de alumnos por docente estén en el centro de la mira. Pero juzgar la educación por la inversión decreciente y la máxima producción, hace que resulten irrelevantes los sujetos involucrados y la cultura a la cual pertenecen. En tal contexto, “calidad” sería mostrar que se alcanzan los objetivos propuestos —*no importa si son legítimos*— y que para ello se invierte cada vez menos o que, con los mismos recursos, los objetivos mensurables se incrementan —*no importa a costa de qué*—.
- § Cuando decimos que tales medidas se toman sin tener en cuenta si los objetivos son legítimos, pensamos, por ejemplo, en las evaluaciones masivas. El DC, pongamos por caso, invierte 17 mil millones de pesos en evaluación censal (durante los tres años de la alcaldía de Peñalosa [SED, 1998: Anexo 1]), para encontrar generalidades que ya los investigadores sobre educación saben (y cuyos resultados no difieren mucho de los de la evaluación de la calidad); o sea, la evaluación censal no es del todo legítima, pero entra en una dinámica de políticas educativas que parece obligatoria; se argumenta que permitirá detectar las necesidades de capacitación, por ejemplo, pero, por otro lado, el DC capacita en planeación es-

---

<sup>76</sup> Pero ni siquiera en el paradigma que la vio nacer es completamente aplicable: John Pike, analista de la Federación de Científicos de Estados Unidos, considera que los pocos fondos y el reducido personal, efectos de una política conocida como “más barato, más rápido y mejor”, podrían haber sido factores determinantes en el reciente fracaso de «Mars Polar Lander», misión estadounidense a Marte. 320 millones de dólares se han perdido, tratando de hacer “eficiencia interna”; la misión Viking, 10 veces más costosa, sin embargo llevó con éxito dos naves a Marte en 1976 [cf. «Nasa reevalúa misiones a Marte», en: diario *El Espectador*, diciembre 9 de 1999, p.3-C].

tratégica, sin que ello se haya inferido de la evaluación censal, que fue posterior, etc.

- § Cuando decimos que no importa a costa de qué, pensamos, por ejemplo, en el *Programa nacional de incentivos*, que bajo el prurito de que competir entre pares mejora la calidad (principio neoliberal), produjo o revivió una serie de conflictos en muchas comunidades educativas; por razones como estas hubo quienes se negaron a participar de esta “rifa”, que se resolvía con un instrumento en el que una de las preguntas al maestro era si asistía cumplidamente a la escuela, con 5 opciones de respuesta [cf. §7.1.2]; pero, ¿quien respondería al patrón, por escrito, que siempre llega tarde? Ya Antanas Mockus llamaba la atención sobre los efectos perversos a que da lugar este tipo de principios<sup>77</sup>, que es uno de los riesgos que Triana [1996:42] señala para la evaluación de la educación (él pone el caso de la utilización de los resultados para hacer clasificaciones y comparaciones injustas de las instituciones educativas... ¡como se hace a partir de la evaluación censal en el DC!: es el principio con el cual se llevan a cabo las “Acciones focalizadas”, pues éstas están dirigidas a las 100 escuelas que obtuvieron los puntajes más bajos).
- § Cuando se define la eficiencia del gasto como «la calidad alcanzada (medida por las pruebas de Saber) en relación con los recursos financieros invertidos (gasto por alumno) [...]» [DNP, 1999:30], queda excluida toda variable cultural. Ni siquiera se tiene en cuenta la idea bajo la cual se introdujeron las evaluaciones masivas de la calidad: «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» [Toranzos, 1996: 65]; o sea, las evaluaciones de la calidad se anunciaron con el fin de “mejorar”, no con el de “reducir costos”.
- § Cuando se dice que la incipiente investigación sobre factores asociados en Colombia «limita a su vez el análisis costo-beneficio de los resultados obtenidos y hace difícil el diseño de una canasta educativa de calidad que permita maximizar

---

<sup>77</sup> En el Seminario «Pedagogía siglo XXI», realizado en la Universidad Pedagógica Nacional del 12 al 16 de 1999, Mockus relató las siguientes dos anécdotas: durante la guerra en Vietnam, el ejército norteamericano resolvió premiar con viajes a los EUA a los soldados más eficientes (que reportaran más muertos); esto incrementó el asesinato de niños y ancianos: las cifras “mejoraron”, pero el número de soldados contrincantes no disminuyó. El otro caso es el del Departamento de Policía de New York: con la intención de disminuir la impunidad, resolvió premiar la eficiencia (casos resueltos)... sólo que entonces los policías optaron por reportar los casos sólo cuando tuvieran certeza de solución; las cifras mejoraron, pero la criminalidad no disminuyó.

la utilización de los recursos asignados a mejorar la calidad educativa» [DNP, 1999:11], se aclara el papel de la evaluación de la calidad con factores asociables: sencillamente elevar la eficiencia interna; de aquí las ideas de “subsidio a la demanda”, “aprovechamiento de la capacidad instalada”, “nuevo colegio”, etc., que vemos en el vocabulario de los gobiernos nacional y distrital. Por eso cuando aparece la pregunta por la “calidad” de la educación en la que se retendrían y se promoverían los estudiantes, el modelo responde con la tautología de que esa calidad depende de la retención y de la promoción... siempre comparándolo contra estándares internacionales.

- § Como la educación pública ha adoptado el método de promediar el gasto que representa un alumno por año, dice que le resulta oneroso que un estudiante repita porque, supuestamente, le quita el cupo a otro y hace que el Estado deba invertir otra vez en él, lo que sube la inversión promedio per cápita. Igualmente, le resulta oneroso que un estudiante se vaya de la escuela, pues se reduce la cobertura y se “pierde” el recurso en él invertido. Como los préstamos internacionales ahora se condicionan a la reducción de esta inversión, entonces se establecen políticas que ataquen factores como repitencia y deserción, y que tiendan a aumentar el número de estudiantes por docente.
- § Esa tal vez haya sido una de las razones para tener las medidas que sobre evaluación escolar se han expedido en los últimos años: hay que promover a los niños y retenerlos en el sistema. Pero después el mismo MEN [1993c:7] se encarga de demostrar que la medida no sirvió: «la promoción automática no ha afectado la baja en las tasas de repitencia [...] formas de mantener la repitencia que se encontraron [...] los maestros, en general, no han modificado sus prácticas pedagógicas ni sus juicios evaluativos». Y más adelante [:8] «[...] los cambios en la escuela a partir de la promoción automática han sido muy pocos [...] los boletines de calificaciones, más que descriptivo-explicativos como planteaba la norma, cambiaron los números por letras, manteniendo el carácter absolutista y terminal de los juicios». Costosas medidas —para bajar los costos— que fracasan, según concluyen costosos estudios.
- § Queda claro —aunque no por “consenso”, como enuncia el ministro Bula [1999:3]— que la evaluación del rendimiento se viene manipulando como un mecanismo para modificar los indicadores numéricos, aunque a la comunidad educativa se le diga que las modificaciones dependen de asuntos pedagógicos (por ejemplo, que la repitencia baja la autoestima y las expectativas de los niños, y que no sirve para mejorar los aprendizajes [MEN. 1993c.:6]). Es decir, las medidas sobre evaluación de aula en realidad materializan una política de eficiencia. No en vano la FECODE se ha quejado reiteradamente de los efectos de la promo-

ción automática: según ellos, dar muestras a los organismos multilaterales de crédito de que hay voluntad de mejorar las cifras de retención, a costa de la mala calidad... aunque, de un lado, como vimos [cf. §5.2.1 y 5.2.2], tales medidas no han logrado dar el golpe de gracia a la repitencia y a la deserción [cf. DNP, 1991:9]; y, de otro, la investigación del mismo MEN no ha mostrado consistentemente una relación negativa entre repitencia-deserción y desempeño, o positiva entre promoción-retención y desempeño [cf. MEN, 1997b:47, 64 y 67]: por ejemplo, según el MEN, repetir “en el mismo plantel” está asociado positivamente al desempeño de lenguaje de los niños de grado 5° en el sector urbano oficial, calendario A.

- § Antes de considerar si la repitencia conlleva factores de encarecimiento de la educación —que no lo negamos— sería muy interesante saber por qué no han tenido éxito las medidas en evaluación que buscan eliminarla (hace rato están vigentes las promociones flexible y automática<sup>78</sup>). ¿Qué hay en esos buenos propósitos que no permite transformar las condiciones específicas que viven los maestros, los niños, las escuelas, los padres? Por ejemplo, todavía hay maestros para los que la repitencia es sinónimo de que en la institución educativa hay disciplina, firmeza: en la investigación del MEN [1993c.:12] sobre repitencia, los maestros respondieron en las entrevistas que «la repitencia es beneficiosa para los niños. A través de ella, esperan que los estudiantes logren los aprendizajes necesarios o la madurez y responsabilidad necesarias para estudiar». Concepto para el que no parece necesario interrogar por la génesis del conocimiento, por las relaciones en la escuela y por la naturaleza de nuestras acciones, de manera que podamos exigir una disciplina de trabajo como *efecto* de la calidad de nuestra labor y no como *condición* para trabajar (que es lo dominante). El asunto de perder o no perder podría ver disminuída su importancia frente a una co-responsabilidad de las partes, no como medida para ahorrar dinero. Una mirada desde el desarrollo de la competencia comunicativa implica identificar las diferencias en los estados del saber-hacer. Si hay repitencia, podríamos interpretar una advertencia de que los cronómetros aplicados para la administración son distintos a los de los sujetos de la educación.
- § Esto daría lugar a pensar cómo generar relaciones entre la administración —en todos sus niveles— y los docentes, con el fin de que el profesional de la educación se vea enfrentado a condiciones cada vez más desafiantes desde el punto de vista profesional; entre la administración y la comunidad educativa... En cambio de eso, se está viendo cómo salir de algunos maestros, cómo debilitar su estabilidad. Lo que pasa es que generar nuevas condiciones de relación entre los miem-

---

<sup>78</sup> El Decreto 1860 no dice de manera explícita “promoción automática”, pero tampoco la desmonta: la conserva para los grupos de grados: de 0 a 6° y de 7° a 9°.

bros de la comunidad educativa es considerado “eficiencia interna” sólo en la medida en que ello implique una gerencia administrativa y no condiciones de producción de comunidades académicas, pedagógicas y educativas.

- § “Deserción” y “repetencia” son visibles y problemáticas pero desde la mirada que espera reducir los gastos del Estado *de cierta manera* (pues no se trata de atacar todos los factores de gasto), irlo descomprometiendo de su responsabilidad con la educación; todo ello, además, se justifica en la medida en que se postula (y se investiga de manera que se demuestre, aunque ello no es condición) un nexo entre calidad, y promoción y permanencia en la institución. Es por eso que ahora la extraedad no se maneja sólo como un problema de costo, sino además como factor asociado negativamente al logro [MEN, 1992:89], o —según la SED— como dificultad para la reinserción de desertores de la escuela [Carrillo, 2000]. Así, en el programa de “Fomento a la retención escolar” de la SED [cf. §5.1] la subdirectora respectiva, argumenta que el programa se inspira en la idea de calidad de vida en la escuela; es decir, mientras por un lado se menciona la búsqueda de causas, por otro ya se presuponen: los niños desertan de la escuela porque ésta no les ofrece la calidad de vida adecuada. De ahí que el maestro finalmente asuma las funciones del Estado: «El problema de la calidad está relacionado con la propuesta de vida que tenga el docente frente a sus estudiantes; lo que haga y su entrega. Ahí es donde está la calidad de vida, la capacidad de solucionar sus propios problemas y los de su comunidad» [Carrillo, 2000:13]. Para que esto sea realizable, la SED contrata un consultor internacional que se compromete a disminuir el conflicto, a mejorar el rendimiento, a lograr más compromiso de los actores educativos y a disminuir la deserción... todo postulable mientras el culpable sea el maestro, pero imposible de cumplir si las causas de la deserción desbordan la escuela y el maestro.
- § Se reconocen dificultades para acceder a la educación en aquellos que no tienen cómo pagarla; sin embargo, la “política de equidad” no busca atacar la inequidad estructural, social, sino a ampliar la cobertura educativa. Más oportunidades para todos... pero, como decía Zuleta, cuando se prohíbe “a todos” dormir bajo los puentes, en realidad la medida va dirigida a aquellos que lo requerirían. De otro lado, entender la permanencia en la escuela como equidad y tratar de producirla desde la educación, es negarse a plantear que el asunto es estructural: los problemas que obligan, en muchos casos, a que un niño abandone la escuela tal vez no sean fundamentalmente educativos, sino sociales; y aun los factores educativos tendrían que ver con asuntos sociales. Por ignorar estas razones, se puede hacer una investigación para encontrar las causas de la repetencia (con la Universidad de Harvard como socio) ¡en la que «se excluyeron zonas con problemas de orden público y difícil acceso»! [MEN, 1993c.:6]. ¿Qué tal que los problemas ex-

cluídos al definir la muestra sean justamente las principales causas? Claro que oficialmente es más “medible y evaluable” —léase “conveniente”— un esfuerzo a nivel escolar que uno a nivel de la estructura social.

- § De ahí que se siga hablando de que el niño se va (“deserción”); sin embargo, desde otras perspectivas, se ven otras cosas:
- de un lado, como dice Galeano [2000:41], «la expulsión de los niños pobres por el sistema educativo se conoce bajo el nombre de deserción escolar»;
  - y, de otro, que el muchacho no se va de la institución, sino que hay —como dice Pellicer [1983]— una *expulsión simbólica*: actividades, espacios, tipos de saberes y actos de segregación que ponen al sujeto ante la opción de irse; la decisión del sujeto es el efecto de una relación. No se trataría entonces, de evitarla, sencillamente, como si con iguales actividades, espacios, saberes y segregaciones pudiéramos mejorar la retención.

La deserción o la repitencia, como efecto, hablan en alguna medida de la vida en la escuela; pero tal vez hay una forma de hacer la política educativa que sólo ve el síntoma y no las causas...

- § Es lógico pensar, por ejemplo, que una escuela con mucha deserción esté en una zona con problemas en la que no sólo hay niños que deben irse, sino que los que se quedan ven alterado su cronograma educativo; así, cuando llegan los evaluadores, los sorprenden con mucho menos trabajo que a otros en condiciones en las que más o menos ha podido mantenerse la programación lectiva.
- § Si los niños abandonan por razones económicas o políticas, por ejemplo, porque su contribución al ingreso familiar es necesaria, o porque viven en una zona de conflicto que los obliga a desplazarse, etc., eso parece no importar (eso se trata en otra mesa de negociaciones<sup>79</sup>), en tanto se trata de controlar la “inversión” por estudiante. Por esto se invita a hacer lo posible por mantener a los estudiantes en el sistema educativo. Incluso a nombre de la “equidad” se dice que los factores extraeducativos que pueden incidir en la calidad de la educación no los podemos cambiar y, en consecuencia, sólo nos quedan las variables de la escuela [cf. Ávila, 1999]; cuando la misma UNESCO, que se evoca en estos contextos, dice que «si se espera que la educación ayude a los pobres a salir de la pobreza, primero habrá que sacar a la educación de la pobreza en los países más necesitados».

---

<sup>79</sup> El modelo neoliberal introduce una dispersión de los espacios de resolución de los problemas relativos a los componentes de la vida social, de manera que se le exige a las asociaciones entre personas, constituidas con el fin de protestar y reclamar, que tramiten sus inquietudes independientemente.



Entonces, las relaciones construídas en el modelo se toman como relaciones de causalidad (“la calidad depende de la retención”) y de ahí se pasa, por ejemplo, a políticas sobre evaluación y promoción escolar (como la promoción automática) sin mediar una confrontación teórica.

- ❖ Dado que algunos análisis a escalas distintas difieren entre sí (por ejemplo, las cifras del DC en “eficiencia interna” son inversas a las que se presentan a escala nacional), valdría la pena pensar si esto no hace del factor en cuestión un objeto distinto en cada escenario, lo que apoyaría la idea de que los factores dependen de variables sociales; o sea, que trascienden el alcance que tienen en los instrumentos que los indagan.
- ❖ Cuando se encuentra una correlación altamente significativa y negativa (que es la esperada por el modelo educativo) entre desempeño de lenguaje (estandarizado nacional) y deserción, esto puede ser leído de varias maneras: se acostumbra, por ejemplo, derivar sencillamente una política de retención en la escuela: “¡hay que retener a los estudiantes!”. Pero también puede concluirse, como creemos justo, que en este caso *hay evidencia empírica, obtenida con ayuda de instrumentos de tales características, para señalar que ha disminuído la incertidumbre* sobre la asociación entre deserción en la escuela y desempeño de los estudiantes en el área de lenguaje, leído éste con la categorización estandarizada. O sea, que hay indicios para investigar ese factor, sin olvidar que, la deserción puede tener muchas causas que desbordan el ámbito educativo, aunque, por supuesto, la escuela debe investigar cuál es su participación en el hecho de que algunos estudiantes la abandonen; y sin olvidar que otras maneras de entender el desempeño revelan otra relación con el factor en mención.
- ❖ Las correlaciones nacionales encubren el hecho de que el desempeño en la escuela nueva no muestra asociación con la deserción, pese a haber sido pensada para eliminar ese factor, mediante una flexibilidad y un sistema de guías específicos. Tampoco muestra asociación entre desempeño y repitencia, pese a que por considerarse de antemano fuertemente relacionados estos aspectos, se creó para ella una modalidad específica de evaluación. El campo es una zona expuesta a condiciones y/o conflictos (siembra, cosecha, inundación, violencia política, etc.) que producen efectos migratorios masivos que, por ende, impactan directamente los índices de deserción (además de que tales condiciones y/o conflictos pueden alterar el funcionamiento regular de las escuelas); sin embargo, parecería que los niños encuestados producen y comprenden textos a pesar de esta situación. Es la modalidad de escuela donde las cifras tienden a mostrar que el desempeño y la repitencia, que el desempeño y la deserción, son asuntos *independientes*. Esto

nos dice, entre otras, que la escuela es más compleja que un listado de factores asociables.

- § Cuando se habla de un “ausentismo” de los estudiantes [MEN, 1997b] —cuyas causas son, por ejemplo, cuidar a los hermanos o trabajar— casi siempre asociado negativamente al desempeño en la zona rural, se homogenizan lo rural y lo urbano; es decir, el procedimiento contrario a la creación de la escuela nueva por el mismo gobierno: ésta se creó justamente para permitir, entre otros, las labores de los niños campesinos. La especificidad de la modalidad en mención se sigue desconociendo cuando se afirma [:49] que «esa socialización del conocimiento que se vive en clase, no puede ser recuperada»; es decir, que los efectos del “ausentismo” son irreversibles, siendo que la flexibilidad y el sistema de guías se inventó (y se esgrime como un gran logro) para que las ausencias fueran compensadas.
- § No hay una incidencia de la deserción en la “calidad” educativa. O sea, no sólo ella no escapa a lo social, sino que podría —como en escuela nueva— ser indiferente para los desempeños, no incidir positiva ni negativamente en ellos; la primera condición es opuesta a los instrumentos y la segunda a la interpretación que suele hacerse de sus resultados.
- § Disminuir la repitencia no puede convertirse en una consigna por efecto de una asociación estadística (si ella estuviera asociada de manera consistente al bajo desempeño de los estudiantes, que no es el caso, según hemos visto). En cambio, sí puede ser una consigna del sistema educativo en toda situación, y no necesariamente como indicador de gasto, sino como criterio pedagógico (de todas maneras, la presente investigación no ha encontrado que la no repitencia sea contraproducente). Tal parece ser la idea —pese al tono— que se publicitó cuando se trató de implementar el *Programa nacional de incentivos*:

*Una escuela es buena cuando todos los alumnos aprenden lo que tienen que aprender en el momento que lo tienen que aprender y en un ambiente de convivencia y felicidad. Los educadores debemos comprender y hacer comprender a los padres de familia y a la sociedad, que todos los niños son capaces de aprender en la escuela si juntamos nuestros esfuerzos y les damos las oportunidades necesarias [MEN, 1996b:1].*

- § La aspiración a mayor cobertura y a mayor retención, entonces, no parece justificarse en términos de “eficiencia interna”; que todos puedan estudiar, que la escuela les ofrezca alternativas que los motiven a permanecer en ella, son aspiraciones plausibles porque así lo queremos, no hay que esperar a que las cifras nos digan qué hacer (los propósitos éticos y políticos son los que crean las condicio-

nes de la acción). De la otra manera, nos volveríamos malthusianos<sup>80</sup>, camino por el cual, si somos consecuentes, terminamos afirmando que la mayor eficiencia interna es la de la segunda ley de la termodinámica y que como la vida va en contra de la entropía, entonces sobra.

- ❧ Pero todo esto no hace más que evidenciar que el problema no es de “eficiencia”, sino un asunto social complejo, y no se le puede atribuir exclusivamente a la escuela (“eficiencia *interna*”), pues allí confluyen las tensiones sociales. Que el mismo *Plan educativo para la paz* lo reconozca no sorprende, pues lo hace de manera que el modelo adoptado no resulta cuestionado: «Las causas de la inasistencia escolar son diversas. Por un lado se encuentran las restricciones de la oferta, como la falta de cupos y la distancia de la escuela a la población demandante. Por otro, las restricciones de la demanda, derivadas de la necesidad de trabajar, los costos educativos, la falta de pertinencia de los contenidos educativos y los problemas sociales causados por el conflicto armado» [DNP, 1999:7]. Lo curioso es que para sacar estas conclusiones no se usó la evaluación de la calidad, ni las cifras estadísticas; por ejemplo en el último informe nacional sobre factores asociables, en los instrumentos no se preguntaba sobre el conflicto armado, ni sobre la pertinencia de los contenidos [cf. MEN, 1997b].
- ❧ No hay una tendencia unificada, de manera que con los instrumentos y las cifras disponibles, podemos afirmar que la tasa de deserción tiene incidencia (en la mayoría de los casos), pero no podemos señalar su dirección, o sea, si incide directa o inversamente en los desempeños de los estudiantes en el área de lenguaje. No obstante, la evidencia de alguna relación sugiere la posible existencia de relaciones más complejas por las que se debe indagar: «en estos países no se podrán formular políticas globales bien fundadas sin tener en cuenta las condiciones educativas que determinan la permanencia en la escuela» [UNESCO, 1998:58]; condiciones que, no sobra decirlo, no han sido establecidas.
- ❧ Los organismos encargados de implementar la política educativa denominan “fracaso escolar” al hecho de no aprobar o de irse. Pero, cuando hay repitencia o deserción, ¿quién pierde realmente el año?, ¿quién pierde realmente la opción?: ¿el estudiante?, ¿su familia?, ¿la escuela?, ¿el sistema educativo?, ¿la sociedad? En las políticas oficiales no se considera un fracaso de la escuela o del sistema educativo, sino un fracaso de los estudiantes, un asunto de la “conducta” del estudiante. Por ejemplo, el hecho de que los padres tengan que dormir a la intemperie en las afueras de la institución escolar para buscar, en muchos casos en va-

---

<sup>80</sup> La hipótesis de Malthus era la siguiente: como la población aumenta en proporción geométrica y las subsistencias aumentan en proporción aritmética, crece la pobreza y los peligros contra la sociedad; en consecuencia, hay que restringir los nacimientos por medio de la restricción moral.

no, un cupo para sus hijos —tal como se vio en el DC para el año lectivo 2000—, no se considera un fracaso de la obligación gubernamental de dar educación a todos hasta el grado 9º, según establece la *Constitución*.

- § El nuevo lenguaje en la administración convirtió la obligación constitucional de dar educación, en dar la “oportunidad” de que la gente se eduque; el mínimo giro verbal permitirá después argumentar que si hay falta de cobertura no es porque se haya dejado de prestar el servicio, sino porque la población no aprovechó las oportunidades. Que la gente se quede sin cupo en el DC, por ejemplo, aparece entonces como un asunto de no haber madrugado más que otros, de no ser lo suficientemente competitivos... y no de que se carezca de cupos suficientes para todos.
- § Parece imposible, pero a nombre de la eficiencia interna del sistema se puede justificar, con cierto tipo de “investigaciones” de factores asociables, el recorte de personal docente: si lo que mejora la calidad es la inversión y ésta es imposible en tanto la nómina de los maestros se lleva casi todo el presupuesto para educación, habría que reducir esa nómina para ampliar el margen de inversión:

*Más del 80% de los recursos sectoriales se orienta al pago de docentes, reduciendo el margen de financiación de otros gastos recurrentes (como los textos y útiles escolares) requeridos para mejorar la calidad de la educación y a través de ésta, para mejorar la eficiencia interna del sistema [Rosas et al., 1996].*

Sin embargo este principio de “compromiso” [cf. UNESCO, 1998:58]<sup>81</sup> no tiene lugar en estos países: «la cuestión de la calidad y la pertinencia de la enseñanza y del papel del personal docente se plantea de una manera totalmente distinta a lo que sucede en los países de las regiones más desarrolladas del mundo» [:62].

- § El hecho de que los mejores desempeños de la parte abierta esté entre aquellos cuyas escuelas presentan “repitencia media”, leído literalmente, sería: “cierto grado de repitencia es necesario para obtener buenos desempeños”. En otras palabras, hay una especie de principio de incertidumbre “heisenbergiano”: si ajustas la repitencia hasta cero, se te reduce el desempeño; si ajustas el desempeño, aparece repitencia. Es decir, con los datos disponibles, no es posible identificar “eficiencia interna” con “calidad” (entendida ésta a partir de los buenos desem-

---

<sup>81</sup> Considerando el aprendizaje como un proceso de producción que requiere insumos, entre ellos un profesor, algunos economistas buscan establecer el insumo que, aumentado con un nivel dado de gasto, mejoraría en mayor medida el rendimiento escolar. «Por lo general, este argumento se ha presentado como una solución de compromiso entre el gasto adicional en maestros y el gasto adicional en material didáctico» [UNESCO: 1998:57]. En países como el nuestro, «el compromiso favorece el gasto en material didáctico adicional, en lugar de en más docentes» [:58].

peños). De otro lado, tampoco habría certeza de que los desempeños disminuyan con el aumento de la repitencia; de donde las políticas al respecto no pueden ser unilaterales.

- § Tampoco las políticas al respecto pueden ser homogéneas: el DC muestra, en todos los casos de lo que aquí se denominó “eficiencia interna”, un comportamiento *opuesto* al que aparece a escala nacional. En cualquier caso, los instrumentos no buscaban establecer —por lo menos directamente— el tipo de actividad educativa que produce repitencia, sino el *hecho* de la repitencia.
- § Pese a la tendencia a atacar la repitencia con el argumento de que sería contraproducente para el desempeño, cuando se mira la repitencia por estudiante (cuántos años ha repetido el niño que responde la prueba y no la repitencia de su escuela, como se preguntó en la evaluación de impacto del PUEBP), en las categorías “no repitencia”, “repitencia de un año escolar” y “repitencia de más de un año escolar” frente a desempeño, sólo se muestra asociación en el desempeño estandarizado, mientras que la repitencia por escuela —como vimos en §5.2.1— muestra asociación contra todos los análisis del desempeño. Así, en la categoría de repitencia *por escuela* en la que ningún niño repite, está el 20.8% de los niños; mientras que en la categoría de repitencia en la que *el niño* no ha repetido, está el 61.5% de ellos. De esta manera, el primer indicador de “eficiencia interna” resulta injusto y puede revelar cosas falsas: alta ineficiencia por alta tasa de repitencia o, recíprocamente, alta eficiencia por baja tasa de repitencia, cuando en realidad se trata de un problema de la unidad de observación: de un lado las escuelas y, de otro, el niño, el cual participa de una manera ponderada diferente en las distintas categorizaciones (grado, modalidad, etc.). ¿Hasta dónde el asunto de la repitencia es algo que se resuelve a nivel de la escuela? Sin responder esta pregunta crucial, muchas lecturas de factores asociables escogen la primera perspectiva, asignándole al niño información que en realidad es de la escuela; y como siempre habrá más niños que escuelas, los resultados podrán investirse del calificativo “altamente significativo”. Curiosamente, no se resuelven pruebas de desempeño entre todos los niños de la escuela, que sería lo justo si se procede así con las cifras de repitencia y deserción (y que sería hasta cierto punto válido desde el punto de vista de la competencia comunicativa).
- § Ante la discrepancia anotada, procedimos a hacer el mismo análisis en la óptica de número de años repetidos por el niño. En la categoría “no repitencia”, está el 61.5% de los niños; en la categoría “repitencia de un año escolar”, está el 26.5%; y en la categoría “repitencia de más de un año escolar”, está el 12.0%. Cruzado con el desempeño, a nivel nacional sólo aparece correlación con el análisis estandarizado: p-valor de 0.001 y tendencia negativa. Por grado, sólo cambia el

análisis por competencias-A, en el que aparece una asociación moderadamente significativa (p-valor: 0.093) y negativa en grado 3°. Por modalidad de escuela, el área rural no muestra correlación alguna entre repitencia del niño y desempeño, bajo los tres análisis del desempeño. En otras palabras, que el niño repita años es algo que en la educación rural (en cualquiera de sus dos modalidades) no incide positiva ni negativamente en su desempeño. En escuela urbana se mantienen las tendencias de lo nacional: altamente significativa y negativa para el desempeño estandarizado; correlación moderadamente significativa y positiva para el desempeño por competencias-A (en lo nacional había un p-valor de 0.117 con tendencia positiva, mientras en modalidad urbana aparece un p-valor de 0.073, con igual tendencia); y ausencia de asociación para el desempeño por competencias-C.

- § La razón por la que —según el discurso oficial— la deserción podría influir en el desempeño de los estudiantes, es porque una escuela de la que los estudiantes se van posiblemente no esté brindando una calidad lo suficientemente atractiva como para que el estudiante se quede<sup>82</sup>; e incluso, que el maestro no está brindando solución a los problemas propios y de la comunidad, como se piensa para el DC [cf. Carrillo, 2000:13]. Pero, ¿qué es causa de qué?: ¿los bajos desempeños se deben a la deserción? o ¿la deserción se debe a los bajos desempeños?
- § La idea de que la permanencia o no en la institución influye en el rendimiento de los estudiantes —además del componente economicista— podría basarse en una serie de prejuicios que, en aras a la comprensión del fenómeno educativo, sería importante discutir: que quien permanece en la escuela es juicioso, no ha perdido años (o está arrepentido de haberlo hecho), está más adaptado, aprende más; y viceversa: que quien trasiega es indisciplinado, pierde años, está desadaptado, aprende menos. O sea, que «La constancia vence lo que la dicha no alcanza». Desde esta perspectiva, por ejemplo, se prohíbe reingreso de alguien a la institución, debido a su “mala conducta” o a su “bajo el rendimiento”; con ello, se lo condena a cambiar de institución, y se cierra un círculo tautológico: “¿te das cuenta que eres malo? ¡Andas de aquí para allá!”. Consideraciones como estas agencian “valores” como fidelidad a la institución (“sentido de pertenencia”), repudio a la diferencia, condena al cambio, culpa...; sin embargo,
- la *fidelidad* a la institución podría constituir en algún caso infidelidad al propio proyecto (cambiante e indefinido);
  - la *diferencia* se sacrificaría en nombre del proyecto “de todos”, en el que íntimamente podría no creerse;

---

<sup>82</sup> De ahí la idea de “expulsión simbólica” [Pellicer, 1983].

- el *cambio* parecería contraproducente cuando se consideran transculturales y transhistóricos los conocimientos y los valores (algo opuesto a su propia génesis);
- la *culpa* sería bien recibida cuando el modelo educativo se edifica desde la consideración de que las cosas se hacen “por tu propio bien”.

La búsqueda de factores asociables, entonces, tal vez sirva para legitimar ciertas concepciones que subyacen en los instrumentos y en los análisis.

- § Hay dos posiciones frente a la relación entre número de alumnos por docente y desempeño; de un lado, lo que hemos llamado la “intuición docente” de que los resultados educativos son mejores si se disminuye el número actual de estudiantes por docente; y, de otro lado, el supuesto hallazgo gubernamental de que el número de estudiantes por docente no influye negativamente en el desempeño de los estudiantes. Con los datos arrojados por el presente estudio es imposible decidir entre esas opciones que, además, no cubren todo el campo de posibilidades al respecto. No obstante, esto que se ha mostrado como la confrontación entre una intuición y un resultado “técnico” trasciende nuestra parroquia: para la UNESCO [1998:56], padres y docentes estiman que este factor es algo fundamental, «mientras que para las autoridades encargadas de la educación tiene consecuencias muy importantes en los presupuestos del Estado, ya que el número de alumnos por clase determina el número de maestros necesarios para una población escolar dada, y en la mayoría de los países los sueldos de los docentes suelen representar las dos terceras partes o más del gasto total en enseñanza». Este organismo internacional piensa que el número de alumnos por clase es un *problema* que se suma a la precariedad infraestructural y de material didáctico; o sea que es un factor que afecta negativamente el desempeño.
- § Pues bien, con los datos disponibles en este estudio, ¿podemos afirmar ambas cosas!: de un lado, desde el análisis estandarizado del desempeño, se muestra con creces lo que plantean quienes diseñan la política educativa: no se muestra que no haya correlación (lo que justificaría el aumento del promedio, en tanto no sería perjudicial) y además se muestra que a más estudiantes por profesor, mejores desempeños de los estudiantes; es decir, la política planteada se justificaría con más fuerza (más que “inofensiva”, es “favorable”). No obstante, desde el análisis por competencias, ¿se evidencia lo contrario!: no sólo que no es inofensiva, sino que es desfavorable. Y cuando cambiamos de escenario, las cosas se vuelven a barajar: en el DC el número de estudiantes por docente no muestra asociaciones bajo ninguno de los análisis del desempeño...

- § ¿Qué quiere decir todo esto? En primera instancia, recordemos que, en un caso, se cuentan preguntas bien contestadas y se distribuye a los sujetos de acuerdo con su distancia a la media, en términos de más o menos una desviación estándar. Con esto, por una parte, la categorización de su desempeño se hace depender de los demás; y, por otra parte, las categorías son excluyentes: quien es medio no es alto ni bajo, etc. En el otro caso, se deciden los desempeños con base en un sistema cuyos niveles incluyen al precedente: quien está en nivel inferencial, también es del literal y quien está en crítico-intertextual también está en inferencial y en literal, en el entendido de que son órdenes de complejidad creciente [cf. §1.3]. El desempeño de cada uno depende solamente de su manera de responder frente a unos niveles preestablecidos teóricamente<sup>83</sup>. Así, cuando comparamos los factores con ambas formas de interpretar los desempeños de la parte cerrada, no hablamos de pruebas distintas —se trata del mismo instrumento de lenguaje, leído desde dos teorías—, ni de poblaciones distintas —se trata de los mismos niños—. Para las personas escogidas de la primera manera, las cifras parecen mostrar que la posibilidad de que sus desempeños sean los mejores tiene que ver con que el número de alumnos por docente sea alto; en el otro caso, bajo la otra escogencia, las cifras parecen mostrar que la posibilidad de que sus desempeños sean los mejores tienen que ver con que el número de alumnos por docente sea bajo. Esto a escala nacional; pero los mismos tipos de análisis del desempeño se comportan, en la mayoría de los casos, de manera diferente cuando pasamos a escala distrital.
- § En el último informe del MEN sobre factores [1997b:47] se habla de una asociación negativa entre desempeño y alumnos por docente en el sector urbano oficial, en grado 3°, calendario B, para el logro individual en el área de lenguaje. Para el logro promedio en grado 5°, se registra asociación negativa, pero solamente para matemáticas, en el sector urbano (privado y oficial). Para el logro promedio en grado 3°, se registran asociaciones negativas también para lenguaje. No es muy claro en el documento, pero, en consecuencia con la lógica que él mismo sostiene, si es negativa quiere decir que a más alumnos por docente, menor logro; o sea, que con datos del mismo gobierno se desmiente la posición que sostiene la política educativa oficial de eficiencia interna. En el mismo documento se dice más adelante [:54] que es un “antifactor” sobre el que se requiere una acción inmediata. No obstante, la redacción del comentario no permite entender claramente (lo transcribimos textualmente, con todo y puntuación): «lo que indica una medida en el manejo de la solución para alcanzar una mayor cobertura de la educación, haciendo, cada vez, más crítica la razón número de alumnos, por profesor». Podría pensarse que el comentario que se hace en el documento oficial es

---

<sup>83</sup> No quiere decir que en la estandarización no haya teorías preestablecidas, sino que no son ellas las que determinan la clasificación.



un llamado de atención sobre la falta de medida del gobierno, que aumenta el número de alumnos por docente para alcanzar mayor cobertura sin establecer el punto crítico de ese aumento.

- § Para ampliar la cobertura, la SED habla de hacer un estudio con el objetivo de buscar «una adecuada relación alumno-docente»; pero a renglón seguido habla del «número adicional de niños que podrían ingresar al sistema utilizando la infraestructura existente» [SED, 1998:29-30]. Si la primera parte fuera cierta, hipotéticamente habría la posibilidad de aumentar, mantener o disminuir el número de alumnos por docente; sin embargo la decisión ya estaba tomada antes del análisis que se promete, pues se habla de número *adicional* de niños.
- § El número de estudiantes genera espacios de interacción: talleres, concursos, deportes, juegos, etc.; cosa que al equipo le parece muy importante por la posibilidad de generar —además de mayores y más variadas opciones de comunicación— mayor número de hipótesis en las discusiones. Pero, ¿hasta dónde el crecimiento del número es beneficioso?, ¿hasta dónde su decrecimiento es perjudicial? ¿Es constante el beneficio en relación con la edad de los estudiantes? ¿Es constante el beneficio en relación con el área, con las respectivas competencias? ¿Qué otros factores intervienen? Son preguntas que no nos parece que tengan respuestas como para fundar una política de incremento del número de estudiantes con los objetivos de “ahorrar profesores” (está dicho más atrás que se consideran un factor que impide invertir en educación) y de aumentar la cobertura. Pensamos que investigar el asunto que plantean las preguntas enunciadas es perentorio, pues permitiría pensar que el factor “número de alumnos por docente” no significa lo mismo en todos los casos (por ejemplo: edad, área, escala) y que por tal razón no se lo puede tomar como universal y homogéneo.

La misma UNESCO [1998:59], considerando que hay muy pocos estudios rigurosos sobre la relación entre desempeño y número de alumnos por docente (o por clase), que las conclusiones que se sacan son propias de la cultura estudiada e inaplicables en otra, y que pocos reconocen abiertamente lo problemático del concepto “número de alumnos por clase” (dadas las características variables del alumnado por causa de factores como la inmigración y la desintegración familiar [:60]), se hace preguntas pertinentes al respecto:

*¿se refiere al de hoy o al del mes pasado? ¿O bien al de hace cinco años, cuando el niño empezó a leer? ¿Cuánto tiempo tienen que pasar los niños en una clase de determinado “tamaño” antes de que se pueda evaluar de manera fiable las “consecuencias”? ¿Cómo compararlos de manera útil con otro grupo de niños diferentes de una clase en la que haya más (o menos) alumnos a cargo de otro maestro? [...] ¿En qué aspectos difiere el comportamiento de los*

*maestros y alumnos en las clases con pocos alumnos en comparación con aquellas en que son numerosos? [:58, 59].*

- ♣ La “eficiencia interna” y la evaluación masiva pueden ir en detrimento de la comprensión de la especificidad de este factor: dado que cierto estrés de docentes y alumnos, que las agresiones y que el exceso de trabajo están ligados a clases numerosas, la UNESCO [1998:59] piensa que «la cuestión del número de alumnos por clase ha de situarse en un contexto más amplio que el de los meros resultados cognoscitivos»... y, sin embargo, en estas evaluaciones de la calidad se lo cruza con desempeño. De otro lado, «Teniendo en cuenta los hechos observados hasta la fecha, sería aconsejable que los responsables de las políticas educativas y de las políticas en general favoreciesen las clases pequeñas» [:59]... y, sin embargo, lo que se busca en Colombia y, particularmente en el DC, es agrandarlas.

## 6. Disciplinas

Como dijimos en §1.2, la evaluación de impacto del PUEBP incluyó pruebas de desempeño en lenguaje, matemáticas y ciencias; para estas dos últimas áreas, se hicieron 30 preguntas cerradas de 4 opciones con única respuesta, 15 para cada área, que aparecen mezcladas en un sólo cuestionario. Se trata de preguntas que podrían haber estado en cuestionarios distintos, o haber configurado partes bien diferenciables de un instrumento, y que se intercalan por otro motivo: «Por razones de procedimiento y tiempo asignado a la prueba de matemáticas y ciencias naturales, las dos pruebas se integraron en un mismo cuestionario, intercalando las preguntas de una y otra asignatura [...]. La alternación se hizo para evitar que una de las áreas fuera perjudicada al no alcanzar algunos alumnos a contestar las últimas preguntas del cuestionario» [MEN:1997a:83].

Los puntajes de ciencias y matemáticas, correspondientes a la proporción de preguntas respondidas “correctamente”, fue estandarizado y categorizado, tal como explicamos en §1.3 (esto también se hizo con lenguaje, con el objetivo de hacerlos conmensurables, es decir, de poder comparar entre sí los puntajes de las tres áreas). De ahí, y teniendo en cuenta los reparos que se harán enseguida, intentaremos una exploración estadística como la que se ha venido haciendo hasta ahora.

### 6.1 Por qué ciencias y matemáticas

En Colombia, la evaluación de desempeño ligada a la idea de “calidad de la educación” ha sido sobre las áreas de lenguaje y matemáticas. Esto parece contradecir parcialmente «la frecuente valoración social según la cual las Ciencias Naturales y las Matemáticas se consideran de mayor importancia que las Humanidades y las Ciencias Sociales» [Carretero, 1993:101].

El *Plan de apertura educativa* estableció unos “programas prioritarios del sector educativo”, uno de los cuales (el número 4) era el “mejoramiento de la calidad de la educación básica”; en el numeral 4 de este programa se promulgan “pruebas para evaluar la calidad en primaria y secundaria”, bajo el siguiente plan: «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].

Pero, ¿por qué se empezaba con lenguaje y matemáticas? ¿Por qué después de casi una década no ha comenzado la evaluación de la calidad en otras áreas? ¿Por qué olvidó el entonces ministro Jaime Niño Díez [MEN, 1997:6], seis años después, que estaba planeado evaluar otras áreas?

En el documento del *Plan de apertura educativa*, antes de establecer las pruebas, se había planteado que lenguaje y matemáticas son las materias básicas [DNP, 1991:12]. En 1997, el entonces ministro Niño sigue sosteniendo lo mismo: «Son piezas centrales de un proceso de evaluación iniciado con pruebas técnicas diseñadas y realizadas con talento colombiano, para examinar cómo se enseñan y aprenden las matemáticas y el lenguaje, *esas dos herramientas básicas del saber* [...]» [MEN, 1997b:6].

Pero así como la iniciativa de hacer evaluaciones masivas no hay que buscarla solamente en Colombia —pues ello constituye una dinámica internacional, según señalamos en §1.1—, igualmente la razón de que sean esas dos áreas las que se evalúan también hay que buscarla por fuera:

- ⌚ la *Declaración mundial sobre educación para todos* [UNESCO, 1992:99], habla de “necesidades básicas de aprendizaje”, entendidas como la adquisición de “herramientas esenciales” (lectura, escritura, expresión oral, cálculo y solución de problemas), de “contenidos básicos” (conocimientos teóricos y prácticos) y de valores y actitudes. Son las “herramientas esenciales” las que se ha decidido trabajar (puede verse la palabra en la cita anterior del exministro Niño).
- ⌚ Esto también lo retoma la SED [1998:12] a propósito de las “competencias básicas”, entendidas como «saberes y destrezas básicas [que] abarcan el espectro necesario para desempeñarse como adulto autónomo, tanto en el ejercicio pleno de la ciudadanía, como de cualquier actividad productiva. Lograr estos aprendizajes es la misión esencial de las instituciones dedicadas a la educación de niños y jóvenes».
- ⌚ Cuando el *Salto educativo* estableció la creación del examen de Estado de la educación básica, el propósito era: «evaluar las aptitudes y competencias de los estudiantes en las áreas de lectura, expresión escrita, razonamiento lógico y solución de problemas» [DNP, 1994:25].

‡ Según Peña [1999:25], para pensar la educación hoy no bastan los criterios de cobertura y retención, sino que también se requiere “producir” ciudadanos competentes frente a los requerimientos básicos de aprendizaje definidos por la *Declaración* de la UNESCO citada, o sea, lenguaje (lectura, escritura, expresión oral) matemáticas (cálculo, solución de problemas) y valores (para mejorar calidad de vida y decidir con fundamento).

También dice que haber seleccionado lenguaje y matemáticas «no sólo obedece a lo que técnicamente es más fácil de evaluar<sup>84</sup>, sino que apunta a conocer el desempeño de los alumnos en aquellas áreas que constituyen herramientas para el aprendizaje posterior y sin las cuales ningún niño debería dejar la educación básica» [:27]; que «sin importar las diferencias, hay un núcleo común de competencias y conocimientos que, en un mundo cada vez más integrado, es indispensable que todos los alumnos adquieran» [:29]. No se dice por qué son *requerimientos*, por qué son *básicos*, por qué se *debería*, por qué son *indispensables*, por qué *no importarían las diferencias*, etc.

Estos enunciados pasan de un discurso a otro; una voz sin autor circula en los ámbitos educativos, todos la agenciamos y, de alguna manera, verbaliza un manejo específico que se hace de los sujetos. “Eso es lo que se dice; por lo tanto, eso es: *Cuando el río suena, piedras trae*”. Por ejemplo, en el informe sobre los factores asociables en la evaluación censal del DC [SED-Corpoeducación-DNP, 1999:iv] se dice: «Hay un acuerdo internacional en que la lectura comprensiva, la expresión oral y escrita y el razonamiento lógico-matemático, es decir, que el desarrollo de habilidades de comunicación y de pensamiento son fundamentales para seguir aprendiendo. Por ello, las evaluaciones se han centrado en el logro de las competencias en el dominio de estas dos áreas de conocimiento»; de igual manera, en el plan sectorial de educación del DC se dice [SED, 1998:4] que en su paso por las instituciones educativas, los niños deben desarrollar «unas pertenencias [sic.] básicas en áreas como matemáticas, español y ciencias», razón por la cual, como una estrategia de calidad, se evaluarán los logros en las “áreas básicas”.

En vano se buscará, en todos los casos, una explicación de las razones para considerar importantes, fundamentales, pilares, ejes, básicas... las áreas en mención.

Desde el momento en el que se hicieron públicos, los resultados de la evaluación oficial de la calidad —en matemáticas y lenguaje— son citados como verdades en los planes, en las políticas específicas, en las evaluaciones, en las investigaciones. ¿Por qué, entonces, la evaluación de impacto del PUEBP incluyó una evaluación en

---

<sup>84</sup> O sea que —según ella— estas áreas son más fáciles de evaluar, desde el punto de vista técnico. No obstante, esto no sólo no está demostrado, sino que está lejos de ser cierto.

ciencias? No es que el MEN se haya dado cuenta de que: «obviamente, la valoración de los logros alcanzados debe referirse a los diversos ámbitos del currículo y no sólo a algunos de ellos. Dicho de otro modo, la evaluación debe ser lo más completa posible» [Triana, 1996:52]. No; la evaluación de impacto del PUEBP incluyó una evaluación en ciencias tal vez por otra razón que, sin embargo, no es ajena al modelo educativo: la confrontación con los estándares internacionales: «La evaluación global del sistema constituye un instrumento fundamental de conducción desde una perspectiva política. El objetivo es, simultáneamente, dar cuenta a la Nación de la utilización de un presupuesto considerable, medir el camino recorrido en la búsqueda de grandes orientaciones y *permitir situarse respecto a los sistemas educativos extranjeros*» [Michel, 1996:29]. Veamos cómo Colombia cumple cabalmente:

- «El país no ha rehuído tampoco las evaluaciones internacionales. Fue así como participamos en una evaluación internacional [TIMSS], junto a otros 40 países, sobre la enseñanza y los aprendizajes en ciencias y en matemáticas. [...] era necesario afrontar los riesgos de compararse con países de alto desarrollo» [MEN, 1997b:7];
- «La participación de Colombia en el TIMSS ha sido muy importante para conocer el nivel de la educación en ciencias y matemáticas en el país con referencia a un estándar internacional y para identificar las principales deficiencias que aquejan al estado actual de estas asignaturas en nuestro sistema educativo [...] el TIMSS para Colombia nos deja una fuente de datos importante para investigar los principales factores que están determinando el bajo rendimiento y la baja eficiencia de la educación en ciencias y matemáticas en el país [...]» [Díaz, 1998:46];
- para la SED [1998:13] «Los resultados del TIMSS constituyen uno de los indicadores más preocupantes en cuanto al dominio de estas competencias básicas por parte de los estudiante colombianos. Más allá de la muy desfavorable comparación internacional —Colombia ocupó el penúltimo lugar entre los 41 países participantes—, preocupan las deficiencias muy marcadas en desempeños relacionados con el razonamiento y la solución de problemas, así como en la comunicación escrita, cuando se trata de preguntas abiertas»;
- en el *Plan educativo para la paz* se plantea [DNP, 1999:9]: «En el reciente estudio internacional sobre logros académicos en matemáticas y ciencias —TIMSS— los estudiantes colombianos de más alto rendimiento en dicha prueba obtuvieron puntajes por debajo del promedio internacional, especialmente en matemáticas [...] En todo caso, estas mediciones hacen necesario explicar dónde está la falla, si se encuentra en los establecimientos o en la forma como se enseña, o si los estándares en Colombia son distintos a lo que esperan, la mayor parte de los países, en matemáticas y ciencias».

Es claro, de un lado, la inserción de esta tendencia en los parámetros del cruce entre desempeño y factores; y, de otro lado, la ausencia de posicionamiento frente a los estándares internacionales (pese a que la *Constitución* [art.7] dice que «El Estado reconoce y protege la diversidad étnica y cultural de la nación colombiana»). Sin haber respondido a la pregunta de si los estándares en Colombia son distintos, se participó en la prueba internacional<sup>85</sup>.

Esta última se viene perfilando desde hace años; por ejemplo, en el *Salto educativo* [cf. DNP, 1994] y en el *Programa nacional de incentivos* [cf. MEN, 1996b] se insiste en que el tiempo anual que un niño dedica al estudio en la escuela es un factor asociado al desempeño. De ahí que entre los resultados de la aplicación del TIMSS se mencione que el desempeño está en relación con el tiempo anual de estudio, y se recomiende que para mejorar la calidad de la educación hay que ampliar dicha dedicación. Pues bien, ya lo vemos en los actuales planes para el sector educativo, tanto nacional como distrital.

## **6.2 Las disciplinas como factor asociable**

Asumimos el desempeño en matemáticas y en ciencias como factores asociables con el desempeño en lenguaje. Si bien esto no es más que una manera de aproximarse a la explicación del desempeño en lenguaje, tenemos claro que esta postura es, además de subjetiva, simétrica: se puede explicar matemáticas y ciencias en términos de lenguaje. Esta particularidad sirve para insistir en una reflexión fundamental frente a las correlaciones que aquí se han estudiado: que se observe asociación entre un factor y desempeño no es garantía de la existencia de una relación de causalidad entre ellos. Por ejemplo, pueden estar mediando otros factores no explícitos en el estudio para el investigador; característica que tiene dos materialidades distintas:

- de un lado, puede deberse a “olvidos”, “intereses”, o “conveniencias”; en este caso, se puede intentar superar, por ejemplo, mediante estudios con diferentes metodologías, y/o mediante la explicitación de los intereses y la confrontación con intereses distintos.

---

<sup>85</sup> Pero ni siquiera un organismo internacional como la UNESCO, que está promoviendo estos movimientos, es tan ciego como para no ver que «Los indicadores internacionales de la educación existentes [...] dan la falsa impresión de que todos los alumnos matriculados en un determinado nivel de enseñanza en los distintos países, según las estadísticas, están escolarizados en condiciones esencialmente similares» [UNESCO, 1998:54]. Es claro que se está diciendo que *no son comparables* porque partimos de una base desigual.

- De otro lado, puede deberse a imposibilidades estructurales, pues su superación compromete cambios paradigmáticos (relativos a la especificidad del sujeto e incluso de la sociedad que intenta pensar) que no dependen simplemente de una voluntad; es el caso del “prestigio” de la psicometría para una disciplina cuyo estatuto se estableció con arreglo a un paradigma de científicidad empirista [cf. Tort, 1979; Benedito, 1975].

### 6.2.1 Ciencias

Cuando se decide evaluar ciencias en los grados 3° y 4°, no se hace una diferenciación entre las edades de los estudiantes. En los grados más bajos se esperaría un esfuerzo por integrar los conocimientos; mientras que a medida que la persona avanza por la escala educativa, los saberes se van separando (por supuesto que desde cierta perspectiva epistemológica también se esperaría una articulación, de otro nivel; es decir, en cualquier caso parece pertinente tener en cuenta las diferencias entre los niveles de desarrollo de la persona). Esta desintegración del conocimiento es clara también en el hecho de que se haya convocado un equipo para realizar la prueba de lenguaje y una persona distinta para la de ciencias y matemáticas (que realizó las dos pruebas por estar vinculado a la prueba TIMSS, en la que la preocupación eran esas dos áreas).

En el caso del presente estudio, asociar lenguaje y ciencias implica relacionar no las especificidades de dos disciplinas, sino las especificidades de las pruebas respectivas: para lenguaje, lo hemos dicho, es la *competencia comunicativa*, y para ciencias es:

*el marco de la Renovación curricular, orientada a formar en el niño una concepción científica del mundo, a través del conocimiento de la realidad; esto quiere decir que su enseñanza no debe tener por meta transmitir a los alumnos un cuerpo de conocimientos, sino formarlos para que frente a los seres y fenómenos de la naturaleza, adopten una actitud científica y gracias a ésta sean capaces de plantear interrogantes sobre a la naturaleza e interactuar con ella, y experimentar e interpretar las respuestas que ésta le proporciona [MEN, 1997a:91].*

Desde nuestro punto de vista, este marco de la evaluación en ciencias, así planteado, presenta una serie de problemas. Cuando establecimos la *competencia comunicativa* como objeto de evaluación en lenguaje, tratamos de obrar en consecuencia con el contexto educativo, pues, en primera instancia, desde 1994, desde la Ley 115, la *Renovación curricular* ya no tiene vigencia. No obstante, en el área de lenguaje hay una relación entre los lineamientos de dicha medida y lo que definió la Ley 115: su enfoque fue el de un trabajo sobre el lenguaje que no describiera científicamente la



lengua, sino un trabajo instrumental, con el objetivo de que los estudiantes usaran su lengua como instrumento de conocimiento, comunicación y disfrute estético; estos objetivos pasaron a la nueva ley y a los indicadores de logro, razón por la que los mantuvimos como eje para la evaluación.

En cambio, vemos un enfoque completamente opuesto en ciencias: “formar en el niño una concepción científica del mundo”, cosa que, de otro lado, consideramos discutible. Si esa era la perspectiva en la *Renovación curricular* para ciencias, la nueva normatividad (¡la vigente para los estudios que habían realizado los niños que fueron evaluados!) es muy distinta. Pese a ello, el documento del MEN [1997a:91] cita como otra de sus fuentes la resolución de indicadores de logro para los grados 1º, 2º y 3º —con lo que excluye a 4º, que también se evaluó—. Ahora bien, los indicadores allí planteados, en lo que respecta a ciencias, esquemáticamente son: curiosidad, deseo de saber, predilección por un tema, ubicación crítica en el entorno, descripciones sencillas con clasificaciones, narración de eventos con la diversa participación de los objetos, formulación de preguntas, y documentación [MEN, 1996c:28], que no sólo son contradictorios con los que el mismo documento le achaca a la *Renovación curricular*, sino que no se ven materializados en la prueba.

Una pregunta como:

5. *¿Qué par de elementos forman conjuntamente el arco iris?*
- A. *La neblina y las nubes.*
  - B. *La lluvia y la nieve.*
  - C. *Las nubes y el hielo.*
  - D. *Los rayos solares y la llovizna*

no muestra una relación muy clara con la curiosidad, la predilección, la ubicación crítica, la descripción, la clasificación, la narración, la formulación de preguntas, la documentación... Parece ser exclusivamente una típica “pregunta escolar” —es decir, «una forma histórica determinada del aprendizaje que comporta en su propia estructura características de clase» [Tort, 1979:8]—, imprecisa, ambigua, sin una relación con el conocimiento científico (que supuestamente trata de medir) y, finalmente, con presupuestos dudosos y opciones de respuesta también dudosas, incluyendo la que en la prueba se considera verdadera: no sólo con rayos solares y llovizna se forman arcos iris (también con luz y un chorro de agua artificiales); tampoco el arco iris es formado “conjuntamente” entre luz y gotas de agua; se trataría, más bien, de una propiedad (hacer de prisma) de las pequeñas gotas de agua, combinada con cierta posición de un observador, para el cual se hacen visibles unas gamas de color de la luz descompuesta y otras (la mayoría) no; es decir, distintos observadores percibirían diferentes tipos de arco iris o no lo percibirían.

De otro lado, ¿para qué sirve una pregunta como esta?, ¿de qué le sirve al niño saber la respuesta *escolar* a una pregunta escolar?, ¿cómo podría hablarse de la formación de una concepción científica del mundo en el niño a partir de preguntas como la que se comenta? La prueba de ciencias, en general, es susceptible de ser analizada de esta forma, pregunta por pregunta, pero, para la presente investigación, es suficiente lo anotado —y eso que en evaluación de la calidad se ha insistido no en tener buenas preguntas, sino en tener *pruebas*— (ya que podríamos tener buenas preguntas que no configuran una prueba). En todo caso, lo que se evalúa a nivel internacional, ¿es la familiaridad de los niños con ciertos estereotipos atribuidos a la asignatura de ciencias?

A propósito, vale la pena dejar claro que la prueba de lenguaje también puede estar plagada de problemas como éstos. Es por ello que insistimos en no hablar más allá de los datos obtenidos con los instrumentos que fueron aplicados; es decir, no hablar de los niños, de sus desempeños, de la ciencia, de la educación en el país, de cómo debe ser la política educativa, etc., pues, como se ve con el ejemplo del arco iris, a veces es difícil hasta dar cuenta de los propios instrumentos...

Esta prudencia no se tiene en la evaluación de la calidad; allí es la ciencia, como tal, la que aparece en el currículo y en las pruebas: «Aunque lo académico es, en general, el conocimiento obtenido mediante estudios que, establecidos con autoridad pública, causan efectos legales, los planes de estudio, como elementos del currículo *priorizan como contenido el saber acuñado por la ciencia* en las diferentes áreas del conocimiento» [MEN, 1997b:14-15]. Y, más adelante [:19]: «si se persigue evaluar la calidad de la educación evaluando principalmente el logro cognitivo, es porque se ha conjeturado que entre el alumno y su adaptación —futura o actual— se coloca como medio el saber organizado, tal como aparece en las ciencias y tal como lo reparten de ordinario los planes de estudio».

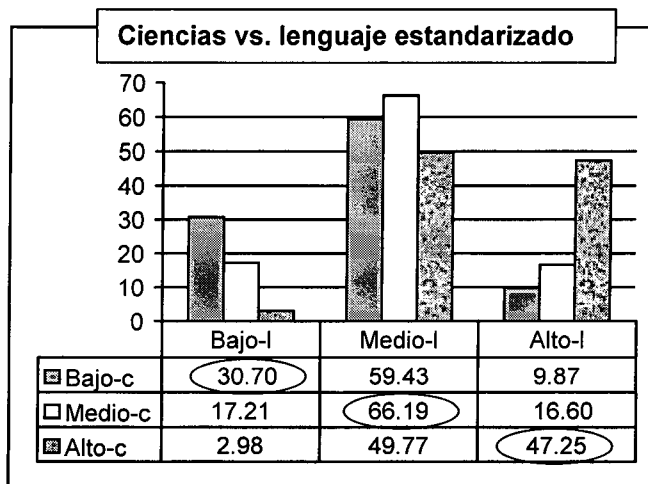
De la ciencia en la educación no falta quién hable. La *Misión ciencia, educación y desarrollo* [1997], por ejemplo, plantea:

- ‡ «La relación cercana entre ciencia y desarrollo depende de la interacción entre educación e investigación. Si se quiere que la educación forme ciudadanos con capacidad de comprender, la única vía posible es la de asegurar una estrecha relación entre la educación como proceso de aprendizaje y la investigación como proceso de generación y adaptación de conocimiento» [:171];
- ‡ «En el mundo industrializado de hoy la ciencia y la tecnología se han convertido en factores de competitividad y de acceso a mercados» [:172];

- ⌚ para reformar la educación básica, entre otros hay que «estimular a los niños, desde etapas muy tempranas, el interés y la curiosidad por la ciencia, con el manejo adecuado del lenguaje propio de las distintas disciplinas científicas, como también con el marco conceptual de las matemáticas» [:200];
- ⌚ un examen de Estado de primer ciclo general de educación preguntaría sobre los siguientes aspectos: «a. Lectura comprensiva y rápida de distintos tipos de textos, símbolos, medios y gráficos. b. En la capacidad mínima de expresión comunicativa escrita y de generación de textos. c. En las habilidades de pensamiento, como la habilidad para hacer inferencias, para razonar deductiva e inductivamente y para el pensamiento lógico-matemático» [:221].

No obstante, entre la ciencia mencionada en estos enunciados y lo que ocurre en los espacios escolares durante la asignatura “ciencias” puede haber mucha diferencia.

La asociación entre los desempeños de las áreas de lenguaje y ciencias, visto desde el análisis estandarizado del desempeño en lenguaje, se grafica de la siguiente manera:



Aunque la gráfica muestra los perfiles parciales de los desempeños de ciencias frente a los de lenguaje, para facilitar la lectura se grafican en forma inversa a la tradicional. En una lectura vertical de la tabla, se observa que los desempeños bajos de lenguaje están concentrados en los desempeños bajos de ciencias y que decrecen a medida que aumentan los puntajes de ciencias. Y, recíprocamente, los desempeños altos en lenguaje aumentan a medida que aumentan los de ciencias. Esto se corrobora con un p-valor de 0.001; es decir, hay una asociación altamente significativa y positiva. Desde esta perspectiva, se esperaría que quien tenga buenos desempeños en

ciencias también tenga buenos desempeños en lenguaje; de donde un esfuerzo por mejorar la enseñanza de las ciencias redundaría en un mejoramiento de los desempeños en lenguaje.

En cambio, desde el análisis por competencias-C, aunque el comportamiento de las cifras es más complejo, vemos una tendencia contraria en algunas modalidades:

		Ciencias		
		Bajo	Medio	Alto
<b>Competencias-C</b>	Nivel 0	1.23	1.14	0.99
	L. Literal	20.25	25.91	24.43
	L. Inferencial	63.72	59.54	60.08
	L. Crít-intertext.	14.80	13.41	14.50
p-valor		0.005		

La correlación es altamente significativa (p-valor: 0.005), pero no lineal. Los desempeños literales se distribuyen en los desempeños de ciencias de manera que el mayor porcentaje no está en el nivel bajo (como en el caso anterior) sino en el medio, y el alto retiene más que el bajo. Por su parte, los desempeños crítico-intertextuales tienen el mayor porcentaje de niños en el nivel bajo de ciencias. En contraposición al análisis desde la estandarización del puntaje de lenguaje, en este caso se esperaría que a quien le vaya bien en ciencias, sus desempeños en las competencias superiores de lenguaje no necesariamente son altos; de donde los esfuerzos por mejorar la enseñanza en un área se convertirían en un debilitamiento de otra. ¿No será esta paradoja un efecto de separar en pruebas distintas asuntos relativos a una misma competencia?

Con estos elementos, nada podemos afirmar sobre la existencia o no de un favorecimiento del desempeño en ciencias, con relación al desempeño en lenguaje; no obstante, la asociación mostrada por las cifras sugiere el desarrollo de investigaciones sobre la relación en mención. El grupo espera una asociación positiva, pues considera que el mejoramiento del desempeño en ciencias estaría ligado a la comprensión-producción de textos, a la capacidad argumentativa, asuntos todos relativos a la competencia comunicativa. Lo que hemos encontrado no nos disuade, en tanto la prueba de ciencias no necesariamente está en relación con efectos como los mencionados. Cuando sugerimos otras investigaciones, nos referimos también a la reflexión sobre las maneras de indagar en ciencias naturales, de manera que —como dicen los mismos indicadores de logro que se mencionan en la justificación de la prueba— se perciba la curiosidad, la crítica, la narrativa, etc. (que, como se ve, no están distantes de la competencia comunicativa).

De otro lado, el análisis por competencias-A no muestra asociación (p-valor: 0.198); no obstante, de manera descriptiva se insinuaría una tendencia positiva.

		Ciencias		
		Bajo	Medio	Alto
Competencias-A	Nivel 0	2.13	2.74	2.30
	C. Local	11.80	11.56	9.27
	C. Lineal	80.29	79.56	81.46
	C. Global	5.78	6.14	6.97
p-valor		0.198		

Cuando desglosamos los resultados por grado, encontramos:

			Grado					
			Cuarto			Tercero		
			Ciencias			Ciencias		
			Bajo	Medio	Alto	Bajo	Medio	Alto
Desempeño	Estandarizado	Bajo	20.05	11.74	2.07	35.04	22.31	5.54
		Medio	63.91	68.19	46.23	57.61	64.32	59.77
		Alto	16.04	20.07	51.71	7.35	13.37	34.69
		p-valor	0:001+			0:001+		
	Competencias cerrada	Nivel 0	1.00	1.10	1.03	1.33	1.18	0.87
		L. Literal	17.29	26.07	24.72	21.45	25.76	23.62
		L. Inferencial	63.16	59.81	60.19	63.94	59.28	59.77
		L. Crít-intert.	18.55	13.01	14.06	13.28	13.78	15.74
		p-valor	0:006(±)			0.173		
	Abierta	Nivel 0	2.11	2.27	1.67	2.14	3.19	4.06
		C. Local	10.29	11.31	10.23	12.41	11.81	6.56
		C. Lineal	81.00	80.94	81.20	80.00	78.25	82.19
		C. Global	6.60	5.48	6.90	5.45	6.75	7.19
		p-valor	0.679			0:030+		

El análisis estandarizado mantiene las correlaciones y sus tendencias en ambos cursos. El análisis por competencias-C, en cambio, distingue entre tercero y cuarto: en cuarto se mantiene la asociación y la tendencia (con la aclaración de que el porcentaje más bajo en el intertextual no está en el desempeño alto de ciencias, sino en el medio; y que el porcentaje más alto del literal no está en el desempeño alto de ciencias, sino en el medio), mientras que para tercero no se evidencia correlación.

El análisis por competencias-A (que no mostraba asociación en lo nacional) es inverso al anterior frente a los cursos: cuarto no muestra asociación, mientras tercero aparece con una asociación moderadamente significativa, de tendencia positiva: el

puntaje más alto de desempeño local está en el nivel bajo de ciencias, y el puntaje más alto de desempeño global está en el nivel alto de ciencias. Una curiosidad: los porcentajes de niños que quedan clasificados en el nivel 0 son mayores a medida que aumentan los desempeños en ciencias; como este nivel incluye a los niños que no alcanzan el desempeño local, podríamos estar diciendo que un niño que ni siquiera es capaz de escribir una oración con concordancia sujeto/verbo, puede tener altos desempeños en ciencias (descrito su desempeño desde la estandarización).

Las cifras por modalidad de escuela son las siguientes:

		Modalidad de escuela									
		Urbana			Nueva			Rural graduada			
		Ciencias			Ciencias			Ciencias			
		Bajo	Medio	Alto	Bajo	Medio	Alto	Bajo	Medio	Alto	
<b>Desempeño</b>	<b>Estandarizado</b>	Bajo	31.31	15.75	2.22	31.31	19.17	4.74	27.80	17.72	2.45
		Medio	56.71	66.09	47.59	60.36	66.46	52.63	65.56	65.93	53.37
		Alto	11.98	18.17	50.20	8.33	14.38	42.63	6.64	16.35	44.17
		p-valor	0.001+			0.001+			0.001+		
	<b>Competencias cerrada</b>	Nivel 0	0.87	1.21	0.91	1.58	0.94	0.79	1.66	1.36	1.84
		L. Literal	20.35	26.54	23.60	20.05	25.14	26.84	20.33	25.55	22.70
		L. Inferencial	63.35	58.82	61.54	64.41	60.17	57.63	63.49	60.48	58.90
		L. Crít-intert.	15.44	13.43	13.95	13.96	13.75	14.74	14.52	12.61	16.56
		p-valor	0.051+			0.232			0.638		
	<b>Abierta</b>	Nivel 0	1.37	2.19	1.53	3.77	2.85	3.48	1.28	4.23	3.23
		C. Local	11.59	10.86	8.21	11.79	10.73	8.99	12.39	15.52	14.84
		C. Lineal	80.95	80.25	81.78	79.48	80.39	81.45	79.91	75.66	80.00
		C. Global	6.10	6.70	8.48	4.95	6.04	6.09	6.41	4.59	1.94
		p-valor	0.112			0.786			0.123		

Como podemos observar, los p-valores de lo nacional se conservan para todas las modalidades de escuela, cuando se cotejan los desempeños en ciencias y lenguaje, visto éste desde la estandarización. El sentido más claro de las correlaciones en todo el estudio nacional está, hasta ahora, en este caso de la relación entre lenguaje y ciencias: las diagonales principales (bajo-lenguaje con bajo-ciencias, medio-l con medio-c, y alto-l con alto-c) son las que tienen los más altos puntajes de las filas *en todos los casos* del análisis estandarizado.

En cambio, desde las categorías por competencia-C de la prueba de lenguaje, la alta asociación de lo nacional se convierte, visto desde las modalidades de escuela, en una asociación moderada para escuela urbana y ausencia de correlación para la escuela rural; se conserva la tendencia negativa (con el intermedio teniendo los valores más altos y más bajos). En otras palabras, desde este análisis, esa correlación tiene

que ver más con la diferencia entre grados que con las modalidades de escuela; particularmente la modalidad de escuela nueva no mostraría ventaja alguna para la producción de mejores desempeños en ciencias y en lenguaje.

Por su parte, desde las categorías por competencias-A de la prueba de lenguaje, no se muestra asociación (tal como en lo nacional) cuando desglosamos por modalidad de escuela (no obstante, en atención a la proximidad de sus p-valores a los valores que hemos considerado asociaciones moderadas, vemos tendencias contrarias: positiva en urbana y negativa en escuela rural graduada). En escuela nueva, es donde se muestra más independencia entre ciencias y lenguaje.

Para el DC, las cifras son las siguientes:

			Distrito		
			Ciencias		
			Bajo	Medio	Alto
<b>Desempeño</b>	Estandarizado	Bajo	33.33	10.64	0.00
		Medio	55.56	78.72	45.83
		Alto	11.11	10.64	54.17
		p-valor	0.001±		
	Competencias cerrada	Nivel 0	11.11	4.26	0.00
		L. Literal	22.22	23.40	25.00
		L. Inferencial	55.56	72.34	54.17
		L. Crítico-intert.	11.11	0.00	20.83
		p-valor	0.050±		
	Abierta	Nivel 0	0.00	2.50	9.09
		C. Local	37.50	12.50	0.00
		C. Lineal	62.50	77.50	90.91
		C. Global	0.00	7.50	0.00
		p-valor	0.061±		

Como puede observarse, las cifras del desempeño en lenguaje, analizadas mediante la estandarización, se distribuyen a lo largo del desempeño en ciencias de una manera muy clara: la mayoría de los “bajos” de lenguaje se concentran en los “bajos” de ciencias; la mayor parte de los “medios” de lenguaje se concentran en los “medios” de ciencias; y la mayor parte de los “altos” de lenguaje se concentran en los “altos” de ciencias. De ahí que la asociación sea positiva y altamente significativa (p-valor: 0.001).

Por su parte, el análisis por competencias-C muestra una distribución casi uniforme de los puntajes de lectura literal en las categorías estandarizadas del desempeño de ciencias; en cambio, los puntajes de lectura inferencial se concentran en el desempe-

ño “medio” de ciencias, tal como se esperaría; y los puntajes de lectura crítico-intertextual se concentran en el desempeño “alto” de ciencias, también tal como se esperaría. Como resultado, tenemos asociación, pero no lineal (en el sentido en que hemos estado leyendo los resultados, pues “neutralizada” como está la lectura literal, la asociación sería positiva; algo que confirmaría el nivel 0: los de nivel más alto en ciencias quedan todos clasificados y la mayoría que no alcanzan el nivel literal están en el nivel más bajo de ciencias).

Por último, en el análisis por competencias-A la mayoría de los puntajes de coherencia local se concentran en en los “bajos” de ciencias, tal como se esperaría; en cambio, la mayoría de los puntajes de coherencia lineal se concentran en en los “altos” de ciencias, como si en el DC, para este caso, la coherencia lineal tuviera una dificultad a la altura de la que presenta el nivel “alto” de ciencias; esto tendría que ver con el hecho de que en total solamente un 7.5% de la población pasa a coherencia global... lo curioso es que este grupo está concentrado en los desempeños “medios” de ciencias. En este caso, el nivel 0 se comporta de manera opuesta al análisis por competencias-C: los de nivel bajo en ciencias quedan todos clasificados, mientras que los no clasificados crecen a medida que avanzamos a niveles altos en ciencias. Es como si el análisis de la comprensión constituyera un objeto al menos parcialmente distinto del que construye el análisis de la producción. Para este caso, también la asociación, moderadamente significativa, se considerará no lineal.

A escala distrital, se replica exactamente el resultado nacional, para el análisis estandarizado del desempeño; en ambos casos tenemos una asociación positiva, altamente significativa. Pero mientras a escala nacional se muestran dos asociaciones (desempeño en ciencias con desempeño en lenguaje, desde las perspectivas estandarizada y competencias-C), a escala distrital las tres maneras de analizar el desempeño muestran asociación; ahora bien, las asociaciones a escala distrital contra el desempeño por competencias (abierta y cerrada) son moderadamente significativas.

En este caso, a diferencia de los que hemos analizado hasta ahora cuando comparamos el DC con el país, el desglose a escala distrital muestra más correlacionados el desempeño en lenguaje y el factor en cuestión. A este nivel, entonces, el análisis responde más a lo esperado: el tipo de desempeño en lenguaje coadyuva al tipo de desempeño en ciencias y recíprocamente (ya que se trata de dos desempeños); es decir, los que tienden a avanzar a mejores desempeños en lenguaje, también tienden a avanzar a mejores desempeños en ciencias, y los que tienden a quedarse en desempeños bajos o intermedios en lenguaje, también tienden a ello en ciencias<sup>86</sup>.

---

<sup>86</sup> Sobre esto, podemos hipotetizar que en la escala nacional el tamaño de muestra es tan grande que permite introducir “ruido” en las asociones; en cambio, la muestra distrital tiene una información un poco más depurada con relación a este tipo de desempeños.



## 6.2.2 Matemáticas

Ya vimos [§6.1] que el área de matemáticas cuenta con una doble justificación: por el lado de las “competencias básicas” y por el lado de los “estándares internacionales”.

Con la prueba de matemáticas aplicada en la evaluación de impacto del PUEBP se quiere tener idea acerca del desempeño de los niños en áreas tales como: valor posicional, operaciones y relaciones de orden en los números naturales, algunos lugares geométricos y algunos sistemas de medición (lineal, de superficie y de volumen). En la justificación se plantea con acierto [MEN, 1997a:81] que el actual marco legal no impone enfoques de desarrollo curricular para ninguna de las áreas; lástima que esta justa apreciación no haya sido aplicada en ciencias, que cumple la misma condición.

Se considera las matemáticas como “ciencia exacta cuantitativa” [:81], definición que se contradice con la adoptada para la prueba de ciencias, donde se trataba de algo relacionado con una realidad empírica, natural —y eso que ambas pruebas fueron realizadas por la misma persona—. No obstante recordar que «cada institución tiene autonomía para formar su propio currículo» [:81], «para efecto del diseño de la prueba se adoptó el enfoque de la Renovación curricular (*Marcos generales* y programas curriculares)» [:81], que el autor fecha en 1990, pero que existe desde una década antes y que fueron suspendidos por la nueva legislación en educación que se produjo como reglamentación de la *Constitución* de 1991. De todas maneras, se aplicó la prueba, pues la formulación de las preguntas parte «de un banco de preguntas, experimentadas y validadas para el TIMSS, de donde se seleccionaron algunas, del conjunto, de las que interceptaron con el Marco curricular definido para los grados 3° y 4°» [:83].

La prueba se ampara en un marco de referencia que surge de una matriz en la que se cruzan “contenidos” (por filas, en número de 9) y “habilidades” (por columnas, en número de 10) [cf. MEN, 1997a:81-82]. Pese a que de tal forma se esperarían 90 cruces, la prueba sólo tiene 15 preguntas. No olvidemos, además, que para preguntar sobre cada uno de los cruces, desde la psicometría se esperaría tener más de un ítem; por eso, en la prueba de lenguaje hay 5 preguntas para los desempeños de comprensión literal, 9 para los de comprensión inferencial y 5 para los de desempeño crítico-intertextual [cf. §1.3.1]. Como en la prueba de matemáticas no se establecieron niveles de logro, sus resultados se estandarizaron (lo mismo que para ciencias).

Veamos una de las preguntas de la prueba:

21. *¿Cuál dígito está en el lugar de las centenas en el número 2345?*

- A. 2
- B. 3
- C. 4
- D. 5

Aunque la respuesta correcta es “B”, el hecho de que un niño la seleccione no garantiza que tenga las “habilidades y destrezas” en el “sistema posicional” (en términos de la sustentación de la prueba), pues igual resultado se obtiene si el niño asigna a los números de la cifra

- “unidad”, “decena” y “centena”, comenzando de derecha a izquierda, o
- “decena”, “unidad”, “centena”, en la misma dirección, o
- “decena” y “centena” comenzando por la dirección contraria, de izquierda a derecha,
- etc.

Además, aplicar la frase «unidad, decena, centena y un puntico que indica mil...» da cuenta de la buena memoria y de la mecanización, pero no forzosamente de una competencia en el sistema posicional (desde Piaget se sabe que los niños pueden aprender a contar antes de tener el concepto de cantidad).

El autor de la prueba, entonces, ha hecho una selección de aspectos en el marco de la disciplina a evaluar. Pero esta selección no es la única posible; por ejemplo, en la evaluación de competencias básicas en lenguaje y matemática del DC, «los criterios de evaluación proponen valorar, por ejemplo, hasta qué grado el estudiante ha integrado a su hacer el conocimiento matemático, le ha dado sentido y significado al poderlo aplicar a situaciones que requieren para su solución razonamiento y modelación matemática» [García y Acevedo, 1999:95].

Desde una perspectiva similar a la tratada para explorar asociaciones entre ciencias y lenguaje, emprendemos el ejercicio de encontrar y explicar relaciones entre matemáticas y lenguaje.

A escala nacional, los desempeños estandarizados de lenguaje y matemáticas muestran una asociación altamente significativa y positiva. Como en ciencias, se esperaría que a quien le vaya bien en matemáticas, le vaya bien en lenguaje, y viceversa. Se ve en la tabla cómo los mayores porcentajes están en las celdas donde se cruzan

idénticas categorías estandarizadas (bajo-bajo, medio-medio, alto-alto). Es decir, a medida que aumenta uno, aumenta el otro.

		Matemáticas		
		Bajo	Medio	Alto
Desempeño estandarizado	Bajo	35.42	17.50	4.24
	Medio	58.07	65.02	50.53
	Alto	6.51	17.48	45.22
p-valor		0.001+		

O sea que, desde la perspectiva estandarizada, el desempeño en lenguaje es dependiente del desempeño en matemáticas. De donde las políticas para mejorar o apoyar la enseñanza de las matemáticas, redundarían en un desarrollo de la competencia comunicativa.

En cambio, analizando los desempeños desde las categorías por competencias (tanto abierta como cerrada), no se muestra correlación:

		Matemáticas		
		Bajo	Medio	Alto
Competencias-C	Nivel 0	0.75	1.15	1.31
	L. Literal	22.28	24.44	25.63
	L. Inferencial	62.95	60.66	58.53
	L. Crít-intertext.	14.02	13.75	14.53
	p-valor	0.465		
Competencias-A	Nivel 0	2.49	2.66	2.04
	C. Local	10.99	11.80	8.87
	C. Lineal	80.37	79.42	82.45
	C. Global	6.15	6.13	6.65
	p-valor	0.140		

En la parte cerrada, el p-valor es de 0.465, y en la parte abierta es de 0.140. En otras palabras, los desempeños en cuestión son independientes: si un niño tiene buenos desempeños en matemáticas no debemos esperar que indefectiblemente tenga buenos desempeños en lenguaje.

El desglose por cursos arroja las siguientes cifras:

			Grado					
			Cuarto			Tercero		
			Matemáticas			Matemáticas		
			Bajo	Medio	Alto	Bajo	Medio	Alto
<b>Desempeño</b>	Estandarizado	Bajo	23.86	11.41	2.96	39.20	23.03	7.51
		Medio	65.99	65.87	48.12	55.48	64.25	56.65
		Alto	10.15	22.72	48.92	5.32	12.72	35.84
		p-valor	0.001+			0.001+		
	Competencias cerrada	Nivel 0	1.02	0.86	1.59	0.66	1.42	0.58
		L. Literal	19.80	24.76	25.14	23.09	24.15	26.88
		L. Inferencial	60.41	60.69	59.50	63.79	60.63	56.07
		L. Crítico-intert.	18.78	13.69	13.77	12.46	13.80	16.47
		p-valor	0.193			0.143		
	Abierta	Nivel 0	2.12	2.20	1.73	2.61	3.07	2.80
		C. Local	7.94	11.66	9.54	12.00	11.93	7.17
		C. Lineal	82.54	80.03	83.15	79.65	78.85	80.69
		C. Global	7.41	6.11	5.58	5.74	6.15	9.35
		p-valor	0.401			0.083+		

Se muestra una permanencia de la tendencia y la asociación del análisis estandarizado. De la misma manera, el análisis por competencias-C no muestra correlación en ninguno de los cursos (aunque en tercero las cifras van en un sentido negativo, mientras en cuarto van en un sentido no lineal). La correlación buscada no es algo que se vea afectado por la diferencia que hay entre pertenecer a cuarto o a tercero. Para el análisis por competencias-A, mientras cuarto mantiene la tendencia nacional (no asociación; p-valor de 0.401), tercero muestra una correlación moderadamente significativa (p-valor: 0.083), con tendencia positiva. Es decir, que en ese grado las áreas en mención dependen mutuamente.

Desglosado por modalidad de escuela, las cifras son las siguientes:

		Modalidad de escuela									
		Urbana			Nueva			Rural graduada			
		Matemáticas			Matemáticas			Matemáticas			
		Bajo	Medio	Alto	Bajo	Medio	Alto	Bajo	Medio	Alto	
<b>Desempeño</b>	<b>Estandarizado</b>	Bajo	34.78	16.33	3.99	34.48	19.75	5.21	39.01	16.79	3.03
		Medio	57.07	64.10	47.19	59.66	64.93	56.51	57.45	68.32	50.30
		Alto	8.15	19.57	48.82	5.86	15.32	38.28	3.55	14.89	46.67
		p-valor	0.001+			0.001+			0.001+		
	<b>Competencias cerrada</b>	Nivel 0	0.27	1.04	1.63	1.03	1.12	0.78	1.42	1.61	1.21
		L. Literal	20.92	25.14	24.70	23.10	24.17	26.04	24.11	22.63	28.48
		L. Inferencial	63.59	60.04	59.76	62.07	61.14	57.55	63.12	61.75	55.76
		L. Crít-intert.	15.22	13.78	13.91	13.79	13.56	15.63	11.35	14.01	14.55
		p-valor	0.264			0.848			0.720		
	<b>Abierta</b>	Nivel 0	1.41	1.98	1.76	3.27	3.43	2.03	3.73	3.32	3.14
		C. Local	10.14	11.09	8.17	11.27	10.88	9.28	12.69	15.99	10.69
		C. Lineal	81.69	79.94	83.01	81.45	79.88	81.45	74.63	76.77	82.39
		C. Global	6.76	6.99	7.05	4.00	5.81	7.25	8.96	3.92	3.77
p-valor		0.517			0.494			0.123			

El desglose por modalidad de escuela no hace mayor aporte; es decir, que la búsqueda de asociación entre desempeños de lenguaje y matemáticas no es sensible a la diferencia entre modalidades de escuela. Para el análisis estandarizado, se mantiene, como en los casos anteriores, la tendencia y la significatividad de la asociación. El análisis por competencias (pruebas cerrada y abierta) no muestra asociaciones; algo semejante a lo observado a escala nacional.

Para el DC, las cifras son las siguientes:

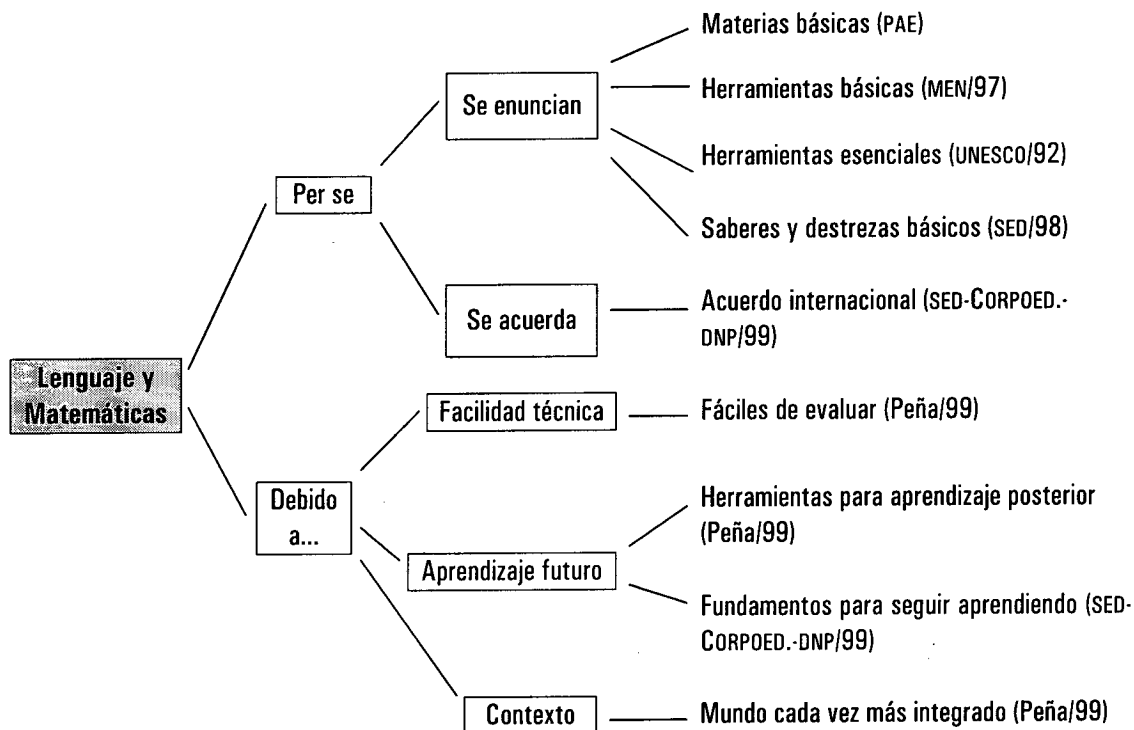
			Distrito		
			Matemáticas		
			Bajo	Medio	Alto
<b>Desempeño</b>	Estandarizado	Bajo	40.00	6.78	0.00
		Medio	60.00	66.10	72.73
		Alto	0.00	27.12	27.27
		p-valor	0.009		
	Competencias cerrada	Nivel 0	10.00	3.39	0.00
		L. Literal	20.00	25.42	18.18
		L. Inferencial	60.00	66.10	63.64
		L. Crítico-intert.	10.00	5.08	18.18
		p-valor	0.675		
	Abierta	Nivel 0	0.00	3.85	10.00
		C. Local	25.00	7.69	20.00
		C. Lineal	75.00	82.69	70.00
		C. Global	0.00	5.77	0.00
		p-valor	0.553		

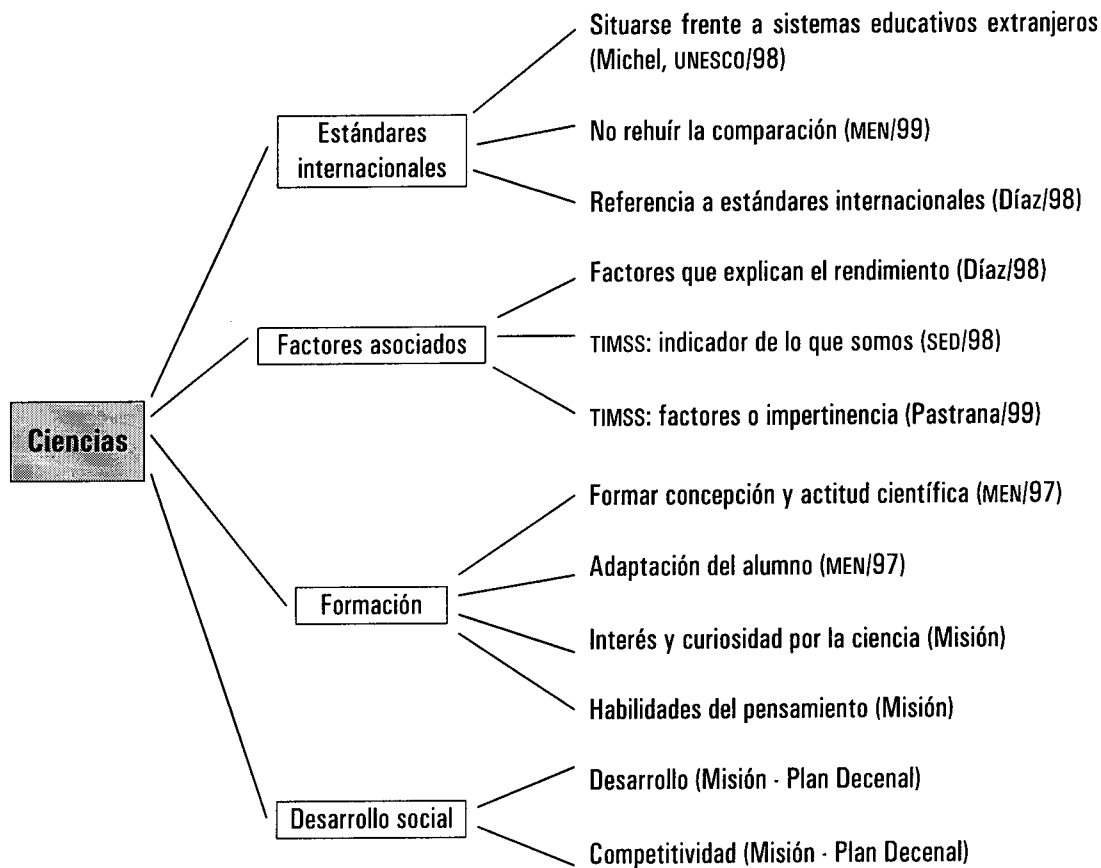
Como puede verse, el análisis estandarizado de las cifras del desempeño en lenguaje, se distribuye a lo largo del desempeño en matemáticas de una manera menos clara que en ciencias: la mayoría de los “bajos” de lenguaje se concentran en los “bajos” de matemáticas, tal como se esperaría; sin embargo, la mayoría de los “medios” de lenguaje no está en los “medios” de matemáticas, sino en los “altos”: el desempeño “medio” de lenguaje se distribuye de manera creciente hacia los desempeños “altos” de matemáticas; y los desempeños “altos” de lenguaje no presentan casos en los desempeños “bajos” de ciencias, como se esperaría, pero se distribuyen casi uniformemente en los desempeños “medio” y “alto” de matemáticas. La pequeña ventaja del último nos inclina a postular una asociación positiva, altamente significativa (p-valor:0.009), como a escala nacional.

Por su parte, los análisis por competencias-C y por competencias-A no muestran asociación, tal como a escala nacional. Aquí, pues, el desglose no permite precisar algo más de lo que el análisis nacional logró. Aún así, es curioso que se replica exactamente el comportamiento de las cifras de los no clasificados (nivel 0) en ciencias: cuando se miran desde la comprensión, los puntajes disminuyen a medida que mejoran los desempeños de matemáticas; y lo contrario, cuando se miran desde la producción: aumentan a medida que mejoran los desempeños de matemáticas.

## 6.3 Resumen parcial

6.3.1 A partir de las referencias [§6.1 y 6.2] a la necesidad de evaluar, de un lado, lenguaje y matemáticas y, de otro, ciencias, hacemos las siguientes clasificaciones:

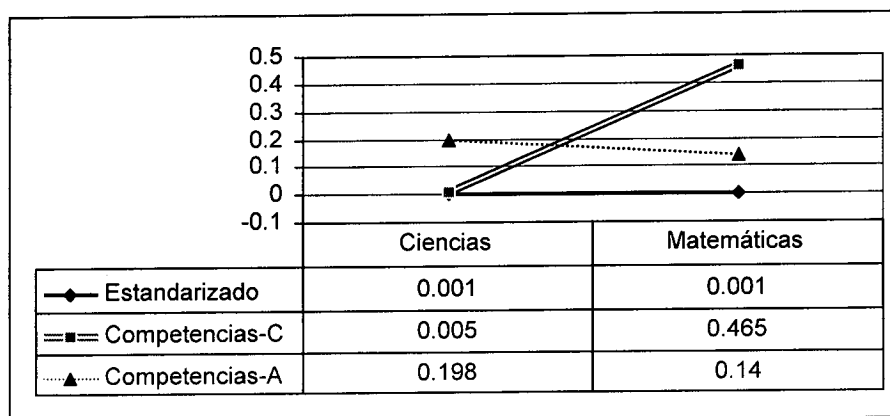




6.3.2

Factores Disciplinas	Desempeño					
	Estandarizado		Competencias-C		Competencias-A	
	p-v	Asociación	p-v	Asociación	p-v	Asociación
Ciencias	0.001	Alta-positiva	0.005	Alta-no lineal	0.198	No
Matemáticas	0.001	Alta-positiva	0.465	No	0.140	No





6.3.3 A continuación resumimos los p-valores para la relación entre desempeños en el área de lenguaje y desempeño en las áreas de ciencias y matemáticas, en ambas escalas, en ambos grados y en las modalidades de escuela:

#### Ciencias

	Nacional	4°	3°	Urbana	Nueva	Rural grad	Distrital
Estandarizada	0.001+	0.001+	0.001+	0.001+	0.001+	0.001+	0.001+
Competencias-C	0.005±	0.006(-)	0.173	0.051±	0.232	0.638	0.050±
Competencias-A	0.198	0.679	0.030+	0.112	0.786	0.123	0.061±

#### Matemáticas

	Nacional	4°	3°	Urbana	Nueva	Rural grad	Distrital
Estandarizada	0.001+	0.001+	0.001+	0.001+	0.001+	0.001+	0.009+
Competencias-C	0.465	0.193	0.143	0.264	0.848	0.720	0.675
Competencias-A	0.140	0.401	0.083+	0.517	0.494	0.123	0.553

Hay un comportamiento idéntico en lo que respecta a las disciplinas, visto el desempeño de lenguaje desde el análisis estandarizado. Siempre hay correlación, positiva, altamente significativa, con el p-valor más pequeño posible; en los desgloses por grado y por modalidad de escuela no se alteran estas tendencias. Sospechamos que la alta asociación encontrada entre las pruebas podría obedecer al criterio común de calificación (número de respuestas correctas), de estandarización y categorización (alto, medio, bajo).

De otro lado, la mirada que trata de aprehender la comprensión y la producción de textos sólo resulta asociada en lo nacional una vez; el desglose por curso produce asociaciones que se alternan entre los grados (de modo que no podemos concluir

acerca de una asociación en función de ese factor), más en ciencias que en matemáticas; el desglose por modalidad de escuela, para las dos disciplinas, sólo marca asociación en un caso (moderada); de manera que podría afirmarse que el asunto es indiferente a ese aspecto en el sector rural.

Y, finalmente, en lo que al DC respecta, las pruebas de competencias sólo muestran una sensibilidad mayor en el caso de ciencias; para matemáticas se replica el análisis nacional. ¿Proviene esto del hecho de que las ciencias estarían más ligadas al manejo de textos, mientras las matemáticas se concentran más en las fórmulas? En cualquier caso, esa parecería ser la tendencia en la enseñanza. A esto podríamos objetar que las fórmulas son algoritmos de una “prosa” que las antecede y, en consecuencia, si la enseñanza ha tomado este camino, es por asumir equivocadamente las fórmulas; no obstante, del lado de los discursos más “prosaicos” y menos “formulistas” (el lenguaje, en primer lugar, y las ciencias en segundo lugar, en tanto ellas también son “formulistas” en alguna medida), valdría interrogar si proceden también por el camino de la *formalización*, que no es ajeno a disciplina alguna y que no niega la naturaleza de un estudio. En otras palabras: ¿hasta dónde la enseñanza del lenguaje y de las ciencias —que aquí se mostrarían más próximas— procede en el sentido de la ciencia?

#### **6.4 Un análisis parcial**

- ‡ Para las referencias traídas a cuento en §6.1, que son representativas del contexto de evaluación de la calidad, la “importancia” del lenguaje y las matemáticas está solamente supuesta, no está demostrada ni fundamentada en una teoría explícita, por ejemplo, de la formación del sujeto. Se explica que estas áreas se evalúan por ser materias, herramientas, saberes, o destrezas “básicas”, “esenciales”, pero esa jerarquía implícita en relación con otras áreas no queda demostrada. En alguna ocasión se hace una tautología: es importante porque hay un acuerdo internacional en que es importante (pero las razones del acuerdo no se mencionan). Cuando aparecen explicaciones, son a propósito de aspectos colaterales: por ejemplo, que se escogen estas áreas porque son más fáciles de evaluar (lo cual, a su vez, no se demuestra), porque son necesarias para aprendizajes futuros (verdad de Perogrullo que tampoco se demuestra), porque el contexto así lo exige (afirmación que no establece el contexto sino que lo califica: “mundo cada vez más integrado”). Esto tratamos de esquematizarlo en el primer gráfico de §6.3.1.
- ‡ Sin embargo, creemos que cuando se están poniendo de por medio los resultados de investigaciones nacionales e internacionales no basta con “creer” que son áreas fundamentales. Recordemos, de un lado, que la exigencia de evaluaciones

masivas [cf. §1.1] era justamente contra las inversiones basadas en las “intuiciones” (los que lo esgrimen para ciertas cosas, lo olvidan para otras); y, de otro lado, que hasta ahora en el presente estudio se ha visto que las asociaciones estadísticas pueden tener sentidos contrarios a los esperados.

⌘ En relación con el área de ciencias, la justificación de su aparición en la evaluación también deja mucho que desear:

- Se justifica por la necesidad no demostrada de comparar los desempeños colombianos con los “estándares internacionales”, sin establecer explícitamente la relación que se busca entre lo particular (no sólo nacional, sino también regional) y lo general; lo que, a su vez pasaría por definir esas especificidades... cosa que no se hace.
- De otro lado, se considera necesario involucrar el área de ciencias puesto que la incluye la evaluación internacional, participación que serviría para establecer los factores asociados a nuestros bajos desempeños (comparados con los de los otros países), que los resultados representan una fotografía del estado de nuestra educación. No se entiende, entonces, para qué lleva el país 10 años haciendo evaluación de la calidad.
- También se justifica en términos del papel de la ciencia en la formación, entendida ésta como hacer del niño un científico, como adaptación, como producción del interés y la curiosidad, y como producción de habilidades del pensamiento; es decir, ¡al menos cuatro teorías, excluyentes entre sí, a propósito de la formación e, incluso, de la misma ciencia!
- Y, finalmente, la inclusión del área de ciencias en la evaluación se justifica por la relación que habría entre ciencia y desarrollo social, en términos del crecimiento económico o del incremento de la “competitividad”. Pues bien, ese paso de la asignatura “ciencias”, a la ciencia, como si fueran lo mismo, no está explicado, como tampoco la perspectiva desde la cual ciencia y técnica se muestran como factores de competitividad y desarrollo económico; pero el “desarrollo” —decir así ya es una posición: sería preferible hablar de la “transformación social”— también puede ser pensado como un efecto complejo en el que no solamente ni principalmente ciencia y técnica intervenirían.

Todo esto tratamos de esquematizarlo en el segundo gráfico de §6.3.1.

⌘ De las preguntas de ciencias y matemáticas decimos que están “mezcladas” pues no se trata de un cuestionario *integrado*, con preguntas comunes que pudieran ser

leídas desde las dos disciplinas, o con problemas comunes para ser resueltos desde la especificidad de cada una, tal como el mismo discurso oficial piensa que debe ser la contribución de las disciplinas en los primeros grados, y como se ha pretendido hacer en la evaluación censal del DC —mediante la utilización de una narrativa icónico-textual común, salvo que, finalmente, las preguntas son independientes—.

- ⌘ La búsqueda de dependencia entre el desempeño en lenguaje, por un lado, y el desempeño en ciencias y matemáticas, por otro, analizados los instrumentos respectivos, es una búsqueda de dependencia entre factores cuya naturaleza se define de antemano como incompatible: quien hizo las pruebas de ciencias y matemáticas posiblemente desearía cruzar sus resultados con *otra* prueba de lenguaje, que tuviera un concepto de “cientificidad” más acorde con el suyo; y quienes hicimos la prueba de lenguaje vemos que sería más deseable cruzar los resultados del área con *otras* pruebas de ciencias y matemáticas, hechas con base en una concepción menos desintegrada del conocimiento.
- ⌘ En relación con las disciplinas está lejos de ser una evidencia su importancia; no olvidemos que se está preguntando no por los conocimientos, sino por lo que unos instrumentos portan, y que las evaluaciones no controlan las fuentes de “saber” que hacen a alguien capaz de responder “cuál parte de la planta chupa más agua”<sup>87</sup>. Por ejemplo, el niño puede acertar —sin entender bien lo que se le pregunta— si asocia el hecho de que el agua está “abajo” y, en consecuencia, la parte que más chupa son las raíces, en tanto en el suelo se almacena el agua. Pero también puede pensarse que las opciones de la pregunta son incorrectas, en tanto “chupar”, propiamente, es un evento que requiere un agente y la planta no es agente de chupar, ni de absorber; la capilaridad no es un asunto de “voluntad”. Ya Bachelard [1979: §IV] mostraba que la idea de que “la esponja absorbe” es un *obstáculo epistemológico*.
- ⌘ Lenguaje y matemáticas entran a la evaluación por una razón ligada a las orientaciones de la UNESCO, y a la tradición de la “evaluación de la calidad”; mientras que ciencias entra por otra razón: la comparación con los estándares internacionales.
- ⌘ La sustentación teórica de las áreas (ciencias y matemáticas), no se somete a la discusión: en el último informe del MEN sobre factores [1997b], cuando se promete dar cuenta de los enfoques para evaluar matemáticas y lenguaje, es con la pretensión de «preparar al lector para la mirada e interpretación de los resultados

---

<sup>87</sup> Pregunta 4 del cuestionario de ciencias y matemáticas de la evaluación de impacto del PUEBP.

que se presentan» [:10]. Es decir, se trata al lector como quien no sabe mirar ni interpretar resultados, razón por la cual hay que prepararlo; y los marcos teóricos no se ofrecen a discusión sino a *divulgación*. De otro lado, no es el *proceso* el que se pone en lo público, sino los resultados; o sea, no puede ponerse un sólo interrogante sobre la *especificidad* de esos resultados.

- ⌘ Cuando se trata de asociar “lenguaje” y “disciplinas (ciencias y matemáticas), se relacionan no las especificidades de dos disciplinas cada vez, sino las especificidades de las pruebas respectivas.
- ⌘ La concepción de ciencia que se exhibe en la justificación de la prueba respectiva en la evaluación de impacto del PUEBP es completamente positivista: ciencia sería sólo aquello que describe la naturaleza, la cual da respuestas; pensamos que hoy las discusiones epistemológicas permiten afirmar que la ciencia se refiere a muchos tipos de objeto (incluso se afirma que la posibilidad de lograr convertir en objeto de la ciencia un objeto “natural” es un proceso social: «la ciencia depende del símbolo para producir sus efectos» [Forrester, 1995:159]); y la naturaleza no “da respuestas”, sino que las respuestas están en el constructo conceptual desde y con el cual se le mira.
- ⌘ La prueba de ciencias, en general, es susceptible de ser analizada de esta forma, pregunta por pregunta: ¿para qué sirve el tipo de pregunta que se hace en la prueba de ciencias?, ¿de qué le sirve al niño saber la respuesta *escolar* a una pregunta escolar?, ¿cómo podría hablarse de la formación de una concepción científica del mundo en el niño a partir de ese tipo de preguntas?
- ⌘ Dada la relación entre la prueba TIMSS y la evaluación de impacto del PUEBP, cabe preguntar si lo que se evalúa a nivel internacional es la familiaridad de los niños con ciertos estereotipos atribuidos a la asignatura de ciencias.
- ⌘ Desde la perspectiva estandarizada, se espera que quien tenga buenos desempeños en ciencias y matemáticas también tenga buenos desempeños en lenguaje. Así, un esfuerzo por mejorar la enseñanza de las ciencias y de las matemáticas o por apoyar su realización, redundaría en un mejoramiento de los desempeños en el área de lenguaje.
- ⌘ Desde la perspectiva del análisis por competencias-C, nada puede afirmarse sobre un favorecimiento del desempeño en ciencias con relación al desempeño en lenguaje. Lo que hemos encontrado no nos disuade de la idea de que mejorar el desempeño en ciencias estaría ligado a comprender textos, a producir textos y a argumentar, asuntos relativos a la competencia comunicativa. Esto porque la

prueba de ciencias no necesariamente está en relación con ese tipo de efectos (juicio del cual tampoco la prueba de lenguaje está eximida). La asociación estadística encontrada sugiere el desarrollo de investigaciones sobre la relación en mención, en el sentido de reflexionar sobre las maneras de indagar en ciencias naturales, de manera que —como dicen los mismos indicadores de logro que se mencionan en la justificación de la prueba— se perciba la curiosidad, la crítica, la narrativa, etc. (que, como se ve, no están distantes de la competencia comunicativa).

- ⌘ Para los niños de tercero el desempeño en lenguaje muestra ser independiente del desempeño en ciencias (algo muy consistente con hacer pruebas distintas para cada área), mientras que para cuarto las cifras sugieren que un avance en una de las áreas sería un retroceso en la otra. ¿Cambian tan radicalmente las cosas con el paso de un año lectivo?, ¿o tenemos una curiosidad de instrumentos y cifras?
- ⌘ Habría que analizar si en ciencias sólo se están realizando ejercicios como el que se hizo para obtener la información del área que aquí estamos considerando; si el asunto de argumentar se considera propio de otra área (lenguaje), porque en ciencias se trataría de “saber” que es la raíz la que más agua absorbe. Igualmente podríamos dudar de si en la prueba de lenguaje efectivamente se está dando cuenta de la comprensión y de la producción de textos.
- ⌘ ¿Podría afirmarse que una condición válida para tercero desaparece completamente cuando se pasa a cuarto?, ¿que algo que no mostraba relación alguna comienza a estar “altamente asociado” cuando un niño pasa al grado siguiente? ¿No será, más bien, que podemos dudar de la acción misma de “medir”, de la idea de buscar causalidad?
- ⌘ En matemáticas, ¿tiene pertinencia una prueba que se construye con arreglo a unos programas que los estudiantes evaluados no han visto? Es más, ¿tiene pertinencia una evaluación de desempeños, en cualquier área, en un país que ha reglamentado la autonomía curricular de las instituciones escolares?
- ⌘ La justificación de las preguntas que conformaron la prueba en matemáticas procede no dando lugar a la discusión, en tanto por haber sido “probadas” en otros contextos, se asumen como válidas. Pero, ¿acaso esa otra validación no podría interrogarse a propósito de esta evaluación? [cf. §7.1.4].
- ⌘ Aplicar ciertas rutinas memorísticas (como la frase: «unidad, decena, centena y un puntico que indica mil...») podría dar cuenta de la buena memoria o de la me-

canización, pero no forzosamente de una competencia en el aspecto que se quiere indagar.

- ⌘ A la hora de construir la prueba de matemáticas, se hace evidente que, aún tratándose de la misma disciplina, podrían construirse innumerables objetos de evaluación sobre ella. Así, cada conclusión no sería comparable con las otras, sino se discute y se concerta sobre el fundamento epistemológico que hay en la decisión. Habrá tantos perfiles de sujeto como posibles entradas se puedan asignar a la matriz: objetos de evaluación –vs.– criterios de evaluación.
- ⌘ Desde las competencias, la no asociación parece consistente con el hecho de que las prácticas pedagógicas de matemáticas y de lenguaje no son homogéneas, y con el hecho de que ni siquiera en una de las áreas tenemos comunidades académicas para afirmar que hay cierto control teórico de las prácticas y, en consecuencia, no tenemos prácticas pedagógicas interdisciplinarias controladas.
- ⌘ Entre lo que se juega cognitivamente en lenguaje y lo que se juega cognitivamente en matemáticas hay una correlación; pero existe un abismo entre eso y el hecho de que entre las prácticas pedagógicas correspondientes haya una asociación.
- ⌘ Las diferencias entre ciencias y matemáticas, ¿proviene de que las ciencias estarían más ligadas al manejo de textos, mientras las matemáticas se concentran más en las fórmulas?; pues tal parecería ser la tendencia en las prácticas pedagógicas. Pero si la enseñanza ha tomado este camino, es por asumir equivocadamente las fórmulas, pues ellas son algoritmos de una “prosa” que las antecede. Y del lado de los discursos menos “formulistas” (el lenguaje, en primer lugar, y las ciencias en segundo lugar, en tanto ellas también son “formulistas” en alguna medida), valdría interrogar si proceden también por el camino de la *formalización*, que no es ajeno a disciplina alguna. En otras palabras: ¿hasta dónde la enseñanza del lenguaje y de las ciencias —que aquí se mostrarían más próximas— procede en el sentido de la ciencia?

**CAPÍTULO III:  
PARA CONCLUIR... Y SEGUIR**



## 7. Políticas educativas y “evaluación de la calidad”

La evaluación es una forma distinta y nueva de hacer política\*

Haremos en este punto una reflexión sobre la relación que percibimos, una vez realizada la investigación, entre la llamada “evaluación de la calidad” y la determinación de las políticas educativas. Se explica de manera crítica algunos componentes de la evaluación de la calidad (que son idénticos en la evaluación de impacto del PUEBP, objeto específico de nuestro trabajo), la interacción entre ellos y la distancia que todo eso puede tener con una evaluación del sistema educativo (o de alguna de sus partes), con la toma de decisiones, con la definición de la política educativa. Al final se hace una propuesta que no constituye otro “deber-ser” de la educación, sino unas implicaciones prácticas del estudio realizado, si se es consecuente con sus presupuestos.

### 7.1 Algunos componentes de la evaluación masiva

En Colombia, es posible interrogar la sistematicidad del proceso de evaluación masiva, tal como la palabra “sistema” promete, en los nombres “*Sistema* nacional de evaluación de la calidad de la educación”, antes de la Ley 115, y “*Sistema* nacional de evaluación de la educación”, según establece el artículo 80 de dicha Ley. Haciendo referencia a algunos componentes de la evaluación masiva —instrumentos para medir desempeño estudiantil, instrumentos para establecer factores asociables, búsqueda de asociaciones, investigación y estadística—, a continuación enunciamos algunos argumentos a favor de la idea de que *sistema* [de evaluación], 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. Pero incluso podría ponerse en duda si ha habido evaluación de la educación (por ejemplo, si hubo evaluación de impacto del PUEBP) y si “evaluar la calidad de la educación” sería o no lo que está en juego.

---

\* OEI, 1994:5.

### **7.1.1 Evaluación del desempeño**

Durante la evaluación de la calidad educativa en Colombia, lo indagado en los estudiantes ha tenido varias denominaciones: “desempeño” ha sido la denominación clásica; “logro cognitivo” o “académico” se empezó a decir en la época de Gaviria; “competencias” también se dice actualmente (por ejemplo, “competencias básicas” en la evaluación censal del DC). Pese a que estas denominaciones puedan tener algún correlato conceptual y, en consecuencia, diferencias teóricas entre sí [cf. MEN, 1997b.], han terminado utilizándose como sinónimos, pero con la connotación que esté de moda.

Para conocer el desempeño (hemos adoptado este término como genérico a lo largo del documento), en la evaluación de la calidad se diseñan pruebas específicas; no se utiliza el examen de Estado, de un lado, porque se hacen aplicaciones a lo largo de toda la educación básica, grados 3°, 5°, 7° y 9°, cosa que dicho examen no permite, pues se aplica solamente en grado 11° (hubo un examen del ICFES para grado 5° en el DC, que ya quedó abolido). Pero, entonces, ¿por qué no ampliar el examen de Estado a estos cursos? Peña [1999: 26] lo explica como sigue,

*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.*

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 evaluación de la calidad. Ante esto, cabrían preguntas tales como: los conocimientos “adquiridos” por los bachilleres, ¿acaso no son exigidos por la sociedad al sistema educativo?; los resultados que la sociedad le exige al sistema educativo, ¿acaso no son parte de los conocimientos adquiridos por los bachilleres? Si las respuestas fueran afirmativas, la justificación citada no sería admisible. Pero, aún así, la actual promoción del examen de Estado como una prueba “de competencias” invalida los argumentos de Peña y muestra que dicho examen no tiene “propósitos distintos”, como dice ella.

Otra justificación se hizo con el nacimiento del SINECE: se dijo que 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 SNP 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).

Por su parte el exministro y exdirector del ICFES, Jaime Niño Díez, plantea que el examen de Estado está diseñado más «para acceder a la educación superior que para evaluar la calidad de la educación» [“Carta a los maestros”, en: MEN, 1997b:6]. Agrega que «De esta forma también se evitará que dicho examen, que es por lo demás excelente para su cometido específico, y que se perfecciona año tras año, se preste para distorsionar los procesos de enseñanza y aprendizaje de la educación básica primaria y secundaria, procesos que no se agotan con el paso a otro nivel de educación» [:6]<sup>88</sup>.

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, a veces parece no haber trascendido más allá de los equipos técnicos que desarrollaron las evaluaciones; es decir, no ha habido ni una justificación pública de por qué no se lo tomó para este propósito, ni una apropiación de dicha justificación. Esto ha facilitado que, con la misma naturalidad, en algunos momentos se lo deseché y en otros 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 MEN, en otra se diseñaba el Programa Institucional de Mejoramiento<sup>89</sup>, en el que el examen de Estado era el instrumento para establecer los

---

<sup>88</sup> Este extraño reconocimiento de que el examen de Estado ha distorsionado las prácticas educativas (ahora no importa aprender sino aquello que va a “salir” en dicha prueba) debería llamar la atención sobre la pragmática de la evaluación, de manera que es lícito preguntar si las evaluaciones “de la calidad” no producen también, a su manera, una distorsión: prepararse para quedar bien en dichas evaluaciones, siendo que —parafraseando al exministro— “los procesos de enseñanza y aprendizaje de la educación básica primaria y secundaria no se agotan en la llamada evaluación de la calidad”.

<sup>89</sup> Formato para aspirar a recursos del «Programa de Ampliación de la Cobertura en Educación Secundaria» - PACES.

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

El segundo caso es el de la Ley 115, 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.*

La instauración de estos estímulos (independientemente de que sean demandables jurídicamente por no ser equitativos) indica que el legislador cree —después de varios años de estar consolidando el SINECE— 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 cambió la prueba de Estado, su plan de desarrollo para el sector educativo valida la que había estado vigente hasta hoy, con propósitos de informar sobre la calidad de la educación: «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*» [DNP, 1999:9]; las cifras del examen de Estado comparadas son las que van de 1986 a 1995 [:10], es decir, que abarca 5 años de implementación del SINECE, el cual rechazó dicha prueba como indicio de calidad.

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.*

De ahí que la evaluación oficial de la calidad argumente extensamente a favor de una indagación de “logro cognitivo” y no de “desempeño”, a favor de pruebas con referencia a criterio y no a la norma, y, sin embargo, termine aceptando lo segundo. De un lado:

*Estas “pruebas con referencia a criterio” se diseñan dentro de una evaluación que da significado al nivel obtenido por cada estudiante, con referencia a qué ha aprendido, qué ha aprovechado, qué ha logrado en su experiencia educativa, en nuestro caso, en las áreas de matemáticas y lenguaje. En cambio, en las pruebas conocidas como de referencia a la “norma”, utilizadas fundamentalmente con fines de selección, como sucede en el examen de Estado para ingreso a la educación superior, los resultados de cada estudiante, al compararse con los de los otros estudiantes, se interpretan o expresan, en función del rendimiento promedio del grupo y de las diferencias individuales [MEN, 1997b:11].*

Y, de otro, el análisis se hace con arreglo al segundo concepto: en una nota de pie de página se aclara que «El instituto SER se refiere a desempeño y a rendimiento escolar porque ajustó los resultados de logro suministrados por el ICFES a una escala que permitiera la aplicación del modelo multicorrelacional que utilizó» [:29]; lo que queda aclarado en el aparte llamado “El procedimiento”: «La transformación de los resultados de “logro cognitivo” (suministrados por el ICFES en escalas de naturaleza categórica ordenada) a puntajes teniendo en cuenta el número de preguntas totales correctas y obviando el nivel al que corresponde cada pregunta» [:39]. Insistimos: para el MEN —por lo menos en ese documento— la evaluación de desempeño no es evaluación de la calidad y, sin embargo, acepta que el análisis, cuya síntesis representa el documento en mención, se haga con arreglo a esa noción. Igual ocurre en el DC a la hora de cruzar los desempeños con factores asociables: promedios de rendimiento sobre escalas estandarizadas [SED-Corpoeducación-DNP, 1999]. Es necesario insistir: *interesa una clasificación del desempeño, independientemente de cómo se obtenga.*

Con todo, en atención a que las pruebas de desempeño en la evaluación de la calidad han sido un reciente objeto de investigación, no han mantenido una misma conceptualización a lo largo del proceso, de manera que sus resultados son hasta cierto punto incomparables. Es decir, a que a la fecha no se dispone 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...<sup>90</sup> incluso reconociendo a renglón seguido que «estas pruebas no se han estabilizado de manera que permitan hacer comparaciones en el tiempo de factores identificados como estratégicos para el mejoramiento de la calidad. El sistema de evaluación de la calidad de la educación es aun débil y no está articulado al diseño de la política educativa» [DNP, 1999:11].

---

<sup>90</sup> Es como medir la estatura de una persona, cada vez con unidades distintas, sin tablas de convergencia entre sistemas de medición, y sin embargo comparar los resultados; esto llevaría a afirmar, por ejemplo, que la persona rebajó su estatura (“hace un momento medía 180 y ahora mide 70.8”... claro: en un caso se usan centímetros y en el otro pulgadas).

Otra cosa es pensarlo desde el ángulo de la investigación del desempeño: no se sabe todo sobre los procesos de conocimiento y, particularmente, en las diversas áreas. Esto plantea una tensión: o aparente 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 niños saben (es decir, *investigación*) y, en consecuencia, de un lado —para quienes se mantengan en ese tipo de indagación—, cifras con distintos grados de “incomparabilidad” que ninguna prueba, por “buena” que sea, eliminaría; y, de otro lado, indagaciones desde múltiples perspectivas<sup>91</sup>.

En otras palabras, en este momento en Colombia se cree tener certeza y por eso se habla de la educación y de los niños, y se toman medidas; pero podría tenerse incertidumbre y hablar del aprendizaje que hemos ido teniendo, de la distancia, tanto con los procesos cognitivos de los niños, como con los fenómenos educativos, con lo que se tomarían medidas *modestas*<sup>92</sup>, que explicitan sus condiciones de posibilidad, que hacen historia y que se reconocen como limitadas; en tal sentido, compartimos algunas de las ideas de Michel [1996:18], aunque su objetivo sea defender una “conducción” del sistema educativo: «la complejidad [...] puede conducir, en la esfera educativa, a la prudencia en las recomendaciones y a orientaciones estratégicas que privilegien acciones de mayor duración y a experimentos, antes que a intempestivas reformas basadas en análisis apresurados o parciales»; más adelante dice: «la conciencia de la complejidad del sistema educativo ha de llevar a *renunciar a esquemas demasiado centralizados y simplistas de conducción* y a tener mucho más en cuenta quiénes son las fuerzas en presencia» [:33, destacado del autor]<sup>93</sup>.

A veces resultan incompatibles incluso los *objetos de evaluación* entre pruebas de la misma área. Veamos el caso de dos pruebas realizadas en el área de lenguaje para el grado tercero: la primera utilizaba 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), etc. En este caso, el objeto de evaluación explícito era «el manejo de los aspectos formales de la lengua y la adquisición de las habilidades necesarias (hablar, escu-

---

<sup>91</sup> Es a lo que hemos hecho un llamado a lo largo del presente trabajo, lo cual no hace incompatibles ambas perspectivas, sino inviábiles las alternativas unilaterales.

<sup>92</sup> Triana [1996:60] coincide, a su manera, en la idea de la modestia: no intentar cubrirlo todo, no hacer sistemas muy sofisticados ni complejos, no generar amplia contestación. Pero, como se ve en sus criterios, podría tratarse de una renuncia a la investigación y de una evitación del conflicto social.

<sup>93</sup> El autor retoma [:16-17] de Lesourne [*Education et société demain*] los atributos de la complejidad educativa: objeto conformado por seres humanos, diversos y autónomos; dimensiones muy grandes; multiplicidad formal; conjuga centralización y autonomía; sistema cerrado, pero abierto a la sociedad; sus efectos abarcan grandes periodos; limitaciones económicas vs. una eficacia difícil de definir y de medir; causas indiscriminables; multiplicidad e imprecisión de objetivos; lugar de conflictos e intereses divergentes entre sus actores.

char, leer y escribir) para comunicarse» [MEN, 1993b: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 [cf. Baena, 1976], 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 no equivale a “las cuatro habilidades” (y menos al “manejo de los aspectos formales de la lengua”), pues en la enseñanza de la lengua éstas, 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 había definido como objeto de la evaluación el concepto de competencia comunicativa explicado.

El otro ejemplo que puede exponerse es el que hemos materializado en la presente investigación: la misma prueba, interpretada con categorizaciones diferentes, arroja resultados distintos (y más interesante resulta la idea de que incluso el orden de las preguntas sea determinante<sup>94</sup>). Distintas pruebas (o interpretaciones) 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 (tampoco se justifica, entonces, decir que se evaluó el impacto del PUEBP). 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.

### ***7.1.2 Indagación de factores asociables***

En §1.1 vimos las menciones a los factores asociables. Comienza a ser una especie de obligación hablar en términos de: «Un sistema evaluativo que realmente pretenda

---

<sup>94</sup> Oliverio [1998] trae la noticia del experimento de unos psicólogos norteamericanos: hacen dos encuestas a muestras de adultos en las que indagan la postura de sus compatriotas frente a la Unión Soviética. Se hacen dos preguntas, pero a una muestra se le presentan en un orden, y a la otra en el orden inverso. Cuando se pregunta primero: «Un país comunista como la URSS, ¿debería permitir a los periodistas norteamericanos difundir en Norteamérica aquellas noticias que consideren oportunas sin censurarlas?», el 82% de los encuestados responde que sí y, al abordar la siguiente pregunta: «¿Deberían permitir los EEUU a los periodistas rusos difundir las noticias tal y como las ven y juzgan bajo su óptica particular», aproximadamente el 75% responde que sí. Pero cuando se invierte el orden de las preguntas, el 75% baja a 55% y el 82% baja a 65%.

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 a cabo una búsqueda de soluciones pertinentes y diferenciadas» [MEN, 1992:25]. Se entiende, pues hay claros condicionamientos de parte de las instancias que prestan el dinero para educación, que es necesario justificar en qué aspectos se invertirá el dinero, y no parece haber otra manera de hacerlo más que la “investigación” estadística: «factor (de), en general, es todo aquello que produce un efecto (en) si puede verificarse empíricamente y que, al utilizar un modelo estadístico para identificar los factores (de), se busca una explicación de carácter probabilístico» [MEN, 1997b:28]. Cualquier otro tipo de investigación, como se ve, queda descartada... y eso que “investigación estadística” no hay; hay, más bien, investigación que *utiliza la estadística* [esto se amplía en §7.1.5].

Entre los factores asociables sobre los que se ha indagado en la década del 90, se han venido incluyendo aquellos indicadores tradicionales que se utilizaban en décadas anteriores, pero también una serie de aspectos que, sin embargo, varía de una aplicación a otra, aparentemente de manera inconsistente. 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 (con muchas más opciones cada ítem); el que aplicó la evaluación de impacto del PUEBP tenía 11... De esto no suele darse 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 al mismo nivel [cf. §7.1.4].
- 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 explícita que organice la indagación;



- copia de otras indagaciones, poco o nada explicitadas, cuyas condiciones de realización pueden diferir radicalmente;
- exclusión de ciertos factores que podrían ser fundamentales, pero que tocan a las instancias que determinan la realización de las evaluaciones: «Como parte de un sistema global juegan un papel muy importante en el proceso: la nación, las entidades territoriales y las entidades locales» [MEN, 1997b:42], pero, aclaran entre paréntesis: ese aspecto no forma parte del estudio (claro, ¿qué tal que el MEN saliera asociado negativamente al logro?).

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 (o se mencionan) para hacer comparaciones, 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 el sistema legaliza una política educativa no del todo legítima, mientras que con ayuda de una investigación más amplia 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 estudiantes de grado 3°:

En 1992, durante la evaluación de la calidad del SINECE, se pregunta<sup>95</sup>:

10. En tu casa, lees libros distintos de los textos escolares:	
Muchas veces?	<input type="checkbox"/>
Rara vez?	<input type="checkbox"/>
Nunca?	<input type="checkbox"/>

En la evaluación de impacto del PUEBP [MEN, 1997a] esto se indaga de la siguiente manera<sup>96</sup>:

<sup>95</sup> El cuestionario va encabezado con la leyenda «Ministerio de Educación Nacional. Sistema nacional de evaluación de la calidad de la educación. Colombia, 1992», y corresponde al «Cuestionario para el alumno».

<sup>96</sup> El instrumento se denomina: «Evaluación de impacto, Plan de Universalización. Prueba de matemáticas y ciencias naturales». La pregunta está en la página 3 y fue analizada en §4.2.1.

5. ¿Cuántos libros hay en tu casa? (No incluyas las revistas ni los periódicos)	
<input type="checkbox"/> A	Solamente los textos escolares
<input type="checkbox"/> B	Los textos escolares y menos de 10 libros
<input type="checkbox"/> C	Los textos escolares y más de 10 y menos de 50
<input type="checkbox"/> D	Los textos escolares y más de 50 libros

Y en la evaluación censal del Distrito Capital, correspondiente a 1998<sup>97</sup>, en relación con el mismo aspecto se hace la siguiente pregunta:

20. Para ayudarte a hacer las tareas en tu casa tienes:		
Enciclopedia	Sí <input type="checkbox"/>	No <input type="checkbox"/>
Libros de literatura o cuentos	Sí <input type="checkbox"/>	No <input type="checkbox"/>
Diccionario	Sí <input type="checkbox"/>	No <input type="checkbox"/>
Otros libros	Sí <input type="checkbox"/>	No <input type="checkbox"/>

Como puede verse, sin ir más allá de la literalidad de las preguntas, parece haber más de una concepción sobre el material impreso:

- En la evaluación de impacto y en la evaluación del SINECE, se hace una diferencia, que la evaluación censal no hace, entre “texto escolar” y “libros”.
- La segunda hace una jerarquización, que la tercera no hace, entre materiales impresos, pues solicita excluir revistas y periódicos (la primera tiene una pregunta sobre los periódicos que se refiere a la frecuencia de lectura de tal material *en la casa del niño*, no por parte del niño).
- La primera y la segunda indagan desde cierta “unanimidad” del material impreso; a la tercera, le interesa la diversidad de textos.
- La segunda interroga por la tenencia de los libros, la tercera indaga por una de las posibilidades de su uso, y la primera por la frecuencia de lectura.

No parece haber claridad ni obrarse en consecuencia con lo que se define como objeto de evaluación. Veamos, por ejemplo, las siguientes declaraciones [MEN, 1997b:19]:

---

<sup>97</sup> El instrumento se denomina: «Evaluación de competencias y saberes básicos en lenguaje y matemática». La pregunta está en la página 17.

*podríamos explorar sobre los “determinantes” del conocimiento concernientes a la estructura cognitiva del sujeto e irnos por una vía de investigación experimental que resultaría adecuada para propuestas de investigación universitaria en escuelas de postgrado. Podríamos por supuesto, explorar los factores de transformación y desarrollo del conocimiento, lo que nos llevaría a una historia del conocimiento científico. Así que lo más adecuado para los propósitos de la evaluación de la calidad educativa, a nivel global, es asumir que los hechos que la afectan se dan en el contexto escolar, comunitario, cultural..., y mediando o “desencadenando” los logros cognitivos.*

En otras palabras, la investigación que propiamente tendría lugar, la científica, no es la que se hace, sino una más fácil, más asequible, más comprensible por los miembros de la comunidad educativa (de antemano considerados incapaces de comprender cierto tipo de investigación) y por quienes toman decisiones de política educativa. La “adecuación” no es a principios epistemológicos, es a “los propósitos de la evaluación de la calidad”. ¡Una perfecta tautología!: los objetivos de la evaluación de la calidad están determinados por los propósitos de la evaluación de la calidad. No se desecha teóricamente la opción de que el desempeño tenga que ver fundamentalmente con la estructura cognitiva del sujeto, con la transformación y desarrollo del conocimiento, pero se toma, eso sí, una decisión política: se va a evaluar el desempeño de los estudiantes *como si* tuviera que ver principalmente con los contextos escolar, comunitario, cultural (su propia estructura se desecha). ¿Y si lo desechado fuera cierto, al menos parcialmente?

Que no parezca que se obra de esa manera porque estamos en Colombia; desde España se dice: «La evaluación es conceptual y empíricamente difícil, por eso desplaza la valoración de los resultados deseables a los *efectivamente medibles*» [Triana, 1996:53]; y desde Francia: «En este campo resulta preciso realizar un esfuerzo especial de conceptualización para construir algunos indicadores susceptibles de iluminar las decisiones a adoptar y los términos de un consenso social» [Michel, 1996:30]. No se trata de *conocer* los factores asociados, sino de *producir* los factores en función de las “decisiones a adoptar” (es decir, los factores anteceden a la búsqueda) y en función del “consenso social” (no es que nos consulten a todos, sino que existen los conflictos sociales y se trataría de no exacerbarlos). De ahí que las informaciones que se recojan

- tienen que estar en «un idioma común...» [:23];

es decir, sin problematización conceptual, que es la única forma de hacer consensos e “idiomas comunes”.

- «... para poder ser agregadas y comparables» [:23];

es decir, la comparabilidad no depende de la consistencia de la información, sino de su “aplanamiento” a una nomenclatura de validez universal: transcultural, transcontextual; es decir, para que haya tal grado de validez, tal posibilidad de comparación con culturas y contextos diversos, debió ocurrir una reducción epistémica al mundo de la medida, una homogenización de procesos heterogéneos.

- Además, «estas informaciones, en la medida de lo posible, deberán estar cuantificadas...» [:23];

es decir, no se trata realmente de un “idioma común”, sino del idioma que cierto modelo quisiera que fuera común: los números; no obstante, como no todo conocimiento es reductible a cifras, esta declaración deja ver claramente los tipos de “conocimiento” y de “investigación” a los que se refiere.

- «... y ser puestas a disposición rápidamente» [:23];

es decir, estamos bajo la égida de las leyes del mercado y del consumo que hace todo precedero, no bajo la especificidad de lo educativo, que tal vez no podría reducirse a la categoría de “producto desechable”. En otras palabras, se trata de buscar una actuación «en tiempo real» [:24]: que los resultados de las evaluaciones no lleguen cuando las condiciones hayan cambiado, sino que constituyan una fotografía del aquí y el ahora para poder incidir en la educación real; es decir, se trata de la ilusión referencial de los mass media: lo que se transmite en directo es conocimiento “en tiempo real”, como si la inmediatez fuera sinónimo de estatuto científico; sabemos que los enunciados de los medios no se juzgan *verdaderos o falsos*, sino *verosímiles o inverosímiles* y en ese caso la impresión de objetividad que causa la transmisión en directo<sup>98</sup> es un efecto de sentido;

- De otro lado, del hecho de que las necesidades de información sean «inmensas y muy diversas en todos los niveles de responsabilidad» [:23], «*deriva el problema de escoger las informaciones más pertinentes* en función de una jerarquía de las necesidades, puesto que una abundancia excesiva de datos sería fuente de ineficacia» [:24, destacado del autor];

es decir, la política comienza a imponer una manera de hacer las cosas que implica obtención de información (por ejemplo, los sistemas de evaluación masiva) y después se refiere a esos montos de información como el punto de partida de un problema en el que la política no tiene responsabilidad, pero sabe que es fuente de ine-

---

<sup>98</sup> Por ejemplo, de noticias o de la vida de alguien durante las 24 horas del día, tal como en los noticieros permanentes se mueve incesantemente una franja de la transformación de los valores de la bolsa.

ficacia, cosa que ella ataca, porque está para eso; el criterio, entonces, es escoger la información más pertinente (lo que quiere decir que el sistema reconoce que produce información impertinente), salvo que la pertinencia, de nuevo, no está en función de los conceptos (que es como la nueva política se presenta a sí misma: técnica, no política [cf. §7.1.4]), sino de la jerarquía de las necesidades... salvo que las necesidades cambian de acuerdo con el modelo. Lo hemos visto en este estudio, por ejemplo en el caso de los textos: para el modelo es una necesidad inobjetable, mientras que para otras miradas se trata de lo que se hace con los textos.

- «*la información también ha de ser fiable*, lo que implica protocolos de acceso estrictos y procedimientos de control de la calidad de los datos» [:23, destacado del autor];

es decir, de nuevo aparece el criterio técnico, cuando los términos de consenso y de necesidad los han excluido; en todo caso, se trata de “protocolos estrictos”, es decir, de acceso denegado para los que participan del consenso.

- «Todas estas características explican el desarrollo de la informática en el seno de la Educación Nacional y la *creación de nuevos sistemas de información*» [:24, destacado del autor];

es decir, la causa se pone después como solución: se impone una manera de hacer las cosas que *requiere* la informática (por ejemplo, que las pruebas deban ser masivas, para que el procesamiento de los datos sea imposible hacerlo manualmente) y después se la pone como algo que nos ayuda, que nos saca del problema.

- En definitiva, «Se trata, pues, de construir un conjunto de indicadores que den una *medida sintética, y por ello aproximativa, de los costes, de la actividad y de los resultados del sistema educativo*. Este conjunto constituye, en definitiva, un esbozo de “modelo” del sistema, con una finalidad más operativa que explicativa» [:24, destacado del autor]

... con una finalidad más operativa que explicativa: ahí cae la máscara cientificista.

De ahí la posición a la que se reduce a la comunidad educativa: en la presentación al documento del MEN [1997b] que publica los resultados de la aplicación de instrumentos de “factores asociados” a estudiantes de 3º, 5º entre 1993 y 1995, el entonces ministro de educación habla de los “factores asociados al logro” como un tema *de suma importancia* «cuyo análisis permitirá a la comunidad educativa y a las autoridades locales y nacionales introducir innovaciones pedagógicas, curriculares, de materiales educativos o de políticas de apoyo, tanto en educación como en otros asuntos, para que el esfuerzo del maestro y del alumno conduzca de una manera más

segura a una educación de calidad para todos» [:7]. Es decir, según los funcionarios educativos, los propios agentes del proceso educativo no pueden pensar, investigar, decidir ni innovar por su propia cuenta, deben esperar a que los “expertos” apliquen sus instrumentos y les digan qué hacer<sup>99</sup>. Esa pretensión es consistente con el cuerpo del documento: cuando se promete dar cuenta de los enfoques para evaluar las áreas (matemáticas y lenguaje), se pretende «preparar al lector para la mirada e interpretación de los resultados que se presentan» [:10]. No es para discutir los datos, pues éstos están por fuera de toda duda, es para poner a los no preparados a tono con el lenguaje de los expertos (y eso que el documento en mención es una “vulgarización” en 67 páginas de una investigación que tiene 6 tomos).

Pero no es esta una posición esporádica, accidental; es una perspectiva que le da razón a la “evaluación de la calidad”. En el DC se encuentra lo mismo, pues se dice que las evaluaciones tendrán cobertura censal con dos objetivos, uno de los cuales es «que cada institución reciba la información sobre el desempeño de sus alumnos, con el fin de que pueda orientar sus planes de mejoramiento» [SED-Corpoeducación-DNP, 1999:ii].

Lo curioso es que a esta pretensión se le responde de manera tal que se cierra el círculo. Los instrumentos están constituidos por un conjunto de enunciados ante los cuales quienes responden se posicionan en forma diversa [Velásquez, 1995a]: por ejemplo, preguntarse qué querrá el otro y, en consecuencia, responder sobre la base de esa expectativa. Esto —que en otro estudio llamamos *pragmática de la evaluación* [Caicedo y Bustamante, 1998]— se deduce de una propiedad de la comunicación; no obstante, la nueva política educativa necesita hacer evaluaciones masivas para averiguarlo: «A menudo se tiene la sensación de que se contesta en los cuestionarios lo que los sujetos creen que debe ser respondido» [Cabrera, 1999:82]. Lo cual, como el mismo Cabrera afirma [:82], le resta credibilidad a los procesos de evaluación<sup>100</sup>. Pero las muchas otras actitudes que son posibles (entre ellas está la “ideal” presupuesta por los instrumentos: responder como si el encuestador tuviera necesidad real de esa información y como si sus intenciones estuvieran claras a lo largo del cuestionario) *no van a ser tenidas en cuenta*, pues, para efectos de la “investigación” estadística, las posiciones frente al evaluador se consideran homogéneas (de lo contrario, la información obtenida con el instrumento *no sería confiable* [cf. Simiand, 1994]).

---

<sup>99</sup> Algo parecido a lo que ocurría con el diseño instruccional: el docente era considerado una especie de auxiliar del currículo diseñado por los expertos del MEN. Lo curioso es que en este momento simultáneamente se le asigna al maestro un perfil distinto: investigador, constructor de PEI, etc.

<sup>100</sup> Según Triana [1999:58], «la experiencia internacional demuestra que la aparición de efectos no deseados ni deseables en este ámbito es más frecuente de lo que debiera».

Sin embargo, desde otro punto de vista, es posible pensar que esas posiciones son determinantes. Por ejemplo, cuando están en juego los dineros de los incentivos [cf. MEN, 1996b] y al docente le preguntan si en su trabajo es puntual:

Siempre	5	Muchas veces	4	Algunas veces	3	Pocas veces	2	Nunca	0
---------	---	--------------	---	---------------	---	-------------	---	-------	---

¿Quién respondería otra cosa que no sea “Muchas veces” o “Siempre”? En cuyo caso lo único que se estaría “midiendo” es la capacidad del encuestado para enfrentar la pragmática del instrumento<sup>101</sup>.

Los formularios de factores asociables siempre van encabezados con algún nombre como MEN, SINECE, SNP, SED, etc. Quien diligencia el formulario puede pensar que estos organismos tienen una relación con el asunto de la calidad educativa y, en consecuencia, que *todo* lo dicho en el instrumento apunta a eso. El que contesta puede creer que contribuye a la preocupación por la calidad de la educación, pues, además, lo interpelan como individuo: en todos los instrumentos se pide a los estudiantes identificarse. Sin embargo, esta interpelación por aquello que supuestamente constituye al sujeto es aparente: por una parte, en realidad el instrumento sólo necesita del sujeto una marca que permita sacar diversos tipos de conclusiones producto del “cruce de información” (perfectamente podría pedírsele una identificación aleatoria pero que se mantuviera sistemáticamente en todo el sistema); y, por otra parte, el efecto del diligenciamiento del formulario es en contra de la individualidad apelada: con él se obtendrá una estandarización de la población, una identificación de muchas personas sobre la base de los rasgos que el sistema de información hace aparecer como comunes, no sobre la base de sus rasgos específicos, sino a la escala en que se quiere mirar<sup>102</sup>: «Los alumnos de los grados tercero y quinto de la educación básica primaria no muestran, en su mayoría, logros que reflejen el nivel educativo en el que están ubicados» [MEN, 1992:97].

### 7.1.3 Asociaciones o correlaciones

Fiel al principio de la psicometría de asegurar la validez vía “expertos” —principio que amerita una discusión epistemológica que siempre se evita—, los aspectos cuya

<sup>101</sup> Lo mismo subyace a las posiciones de estudiantes hispanos y norteamericanos en E.U. cuando son entrevistados: Stubbs [1984] señala que, para los primeros, la situación de entrevista está ligada a problemas con las autoridades (sobre todo con inmigración).

<sup>102</sup> Procrusto era un bandido de Ática que poseía un lecho pequeño y uno grande, en los que obligaba a los viajeros a echarse: a los altos, en el pequeño y, a los bajos, en el grande. Si sobrepasaban la longitud del lecho, les cortaba el sobrante; de lo contrario, los estiraba hasta que alcanzaran la medida.

asociación se busca en la evaluación de la calidad son tratados independientemente, incluso por equipos distintos. Por ejemplo, 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 con independencia 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].*

... sin embargo, cuando ha habido tiempo, tampoco se han relacionado los trabajos. En el caso de la evaluación de impacto del PUEBP, a las reuniones de trabajo asistían todos los grupos, pero en ningún momento se permitía transgredir el proverbio según el cual “entre bomberos, no se pisan las mangueras”.

Pero si, como creemos, en los instrumentos de desempeño hay alguna o algunas concepciones —así sean implícitas— 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 alguna o algunas concepciones —así sean implícitas— sobre desempeño, entonces a la hora de cruzar las informaciones que vienen de instrumentos aislados, se pueden estar cruzando medidas realizadas sobre objetos que pertenecen, al menos parcialmente, a clases diferentes, es decir, hasta cierto punto incompatibles (como hemos intentado mostrar en este estudio), y pueden verse asociaciones debidas más a la accidentalidad o, en el mejor de los casos, al azar<sup>103</sup>. De donde las correlaciones encontradas, en cualquier magnitud y sentido, pueden ser tan azarosas como las decisiones tomadas en política educativa a partir de esa información o usándola como legitimador.

Veamos un caso en relación con el factor *textos*. En la prueba de desempeño en lenguaje para la evaluación de impacto del PUEBP 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. Jurado et al., 1998: Anexo I] que enfatiza sobre la diversidad textual y sobre los usos de los textos. 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” [cf. §4.3.1], frente al cual resulta relevante la cantidad y resultan irrelevantes la diversidad y el

---

<sup>103</sup> No discutimos la validez estadística (o psicométrica) que puedan exhibir tales instrumentos.



uso. Entonces, cuando se cruzan informaciones recogidas con ideas opuestas sobre el mismo objeto, ¿qué tan confiables pueden ser las conclusiones?

El “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 (parece como si el objetivo fuera acumular entropía). 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, 1993c:7]. También es el caso de la evaluación de impacto del PUEBP: hubo un equipo distinto para el diseño de la muestra y para cada una de las siguientes pruebas: matemáticas-ciencias, valores-actitudes infantiles, lenguaje, gestión, costo-efectividad; y otro equipo, de estadísticos, hizo el análisis “multivariado”. 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).

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 PUEBP, el de la evaluación censal en el DC, el estudio sobre repitencia en Colombia, el TIMSS... Claro que el desconocimiento de lo hecho incluso puede ser explícito: después de casi diez años de estar gastando empréstitos internacionales<sup>104</sup> 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» [DNP, 1999:11].

A continuación se ve cómo cada plan de desarrollo educativo “usa” este mecanismo. En primer lugar, 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 asociaciones estadísticas entre indicadores para justificar la política educativa: «Las características del plantel y las de los

---

<sup>104</sup> 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.

procesos pedagógicos son las variables que más pesan<sup>105</sup> 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» [DNP, 1994:11]. Y, en segundo lugar, 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 [DNP, 1999:8-9]. A partir del cruce con factores asociables, plantea que las variables escolares explican hasta el 30% del rendimiento escolar [:10] y que factores como textos, bibliotecas, medios pedagógicos, promoción de la lectura y tiempo efectivo de aprendizaje son factores *decisivos* —pero, recordemos que sólo explican el 30% (¿no serán decisivos los que explican el 70%?)— 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

- que las diferencias de desempeño tienden a coincidir con las diferencias sociales:
  - a) En su investigación sobre repitencia, el MEN [1993c.:8] dice que «los resultados de los análisis estadísticos muestran que los niños de los sectores más pobres de la población —urbano-marginales y rurales— tienden [...] a tener menos logros»; y más adelante: «las escuelas situadas en los sectores más pobres son las que presentan las peores condiciones materiales y académicas y por lo tanto, son aquellas en las que los niños tienen menos oportunidades de aprender» [:12].
  - b) Por su parte, el DNP [1999:20] afirma: «Cuando se analizan los factores asociados de las pruebas de calidad (Saber), es evidente que buena parte de las ventajas que tienen los niños que estudian en colegios privados, no se debe a la calidad del establecimiento, sino a soportes que las familias de más altos ingresos dan a su educación».
  - c) Se une al coro la *Misión de educación ciencia y desarrollo* [1997:141], diciendo en qué estratos está la baja calidad de la educación primaria: «sobre todo la oficial y la de gran parte de colegios privados de estrato medio y medio-bajo».
  - d) El mismo MEN [1997b:40] afirma: «Se puede suponer que, a

---

<sup>105</sup> Lo del "peso" se refiere a la significatividad estadística.

mayor nivel de satisfacción de necesidades básicas, corresponden mejores condiciones iniciales de partida para el aprendizaje» (lo cual justifica que se hable de «democratizar la calidad» [:9]); e) SED-Corpoeducación-DNP [1999:2] encontraron lo mismo: «hay una fuerte asociación positiva entre logro y nivel socioeconómico»; un poco más atrás dicen [:2]: «En general, los planteles privados, los de mayor nivel socioeconómico, los no mixtos, los de jornada única y los que ofrecen todos los niveles educativos, son los planteles de mayor logro académico promedio»; y

- 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, pero que sí puede incidir sobre un porcentaje del desempeño, el debido a factores propiamente educativos. A partir de esto se habla de *equidad*, en un sentido compensatorio: «el maestro puede jugar un papel fundamental en el aprendizaje, a pesar de las condiciones desfavorables del medio en el que viven los niños» [MEN, 1993c.:5]; «[...] estas condiciones desfavorables sean *compensadas* con un buen ambiente escolar» [:8]. Y es que, según Michel [1996:37], «una actuación evaluativa constituye el resultado de una triple exigencia: eficacia, transparencia democrática y equidad»; así mismo, para Triana [1996:52], la reducción de las desigualdades ocupa un lugar prioritario entre los objetivos de la evaluación de la educación. Pero veamos qué es “democracia”: «Desde una lógica democrática, parece razonable que la toma de decisiones acerca de cuestiones que tanto y tan íntimamente afectan a la sociedad...» aquí espera uno de qué manera es que los afectados van a participar, a discutir, a tomar decisiones, a controlar... pero no; en cambio, la frase continúa: «... se realice a través de una evaluación cuidadosa y sistemática de los efectos producidos en las situaciones de cambio» [:50].

Mientras para el *Salto educativo* estos principios indican 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%. Por su parte, el plan sectorial en educación de la SED [1998:3] dice: «La educación es una condición indispensable para lograr una sociedad más igualitaria, y puede compensar en parte las condiciones de pobreza a las que están sometidos muchos de nuestros conciudadanos»; y más adelante [:3-4]: «no tenemos desde el Estado la capacidad para hacer óptimo el hábitat individual de cada uno de los niños del Distrito, pero sí estamos en capacidad de ofrecerles ambientes educativos adecuados y placenteros. Buscaremos que la infraestructura educativa de los más pobres esté a la altura de la que tienen los mejores colegios» (o sea, ¿de los ricos?). Para sustentar 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 educativo 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
			información oportuna
		recursos asignados con eficiencia	

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... Y eso que el modelo que impone las evaluaciones masivas dice que «la coherencia entre la política de evaluación y la política educativa en general es la única condición capaz de asegurar una evaluación realista y justa de las reformas educativas» [Triana, 1996:59].

De otro lado, es notable que la manera de concluir del DNP tenga la marca del modelo que impulsa para toda la política nacional: la administración. En otras palabras, para el DNP, el asunto educativo es más de índole administrativa; repasemos su lenguaje: “desarrollo institucional”, “gestión”, “oportunidad”, “eficiencia”, “recursos”. Pero claro: «Es este un estilo de funcionamiento frecuente en las organizaciones productivas, que ha ido arraigando también entre las administraciones públicas. Parece ser que la educación no constituye una excepción a lo que va siendo una regla cada vez más extendida en la gestión de los servicios públicos» [Triana, 1996:42].

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

- Sobre número de alumnos por docente, por ejemplo, en la evaluación de impacto del PUEBP 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 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* [DNP, 1994:11] afirma que la calificación de los docentes —entre cuatro factores— está asociada positiva y significativamente con el logro de los estudiantes.
- 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, 1997b: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.
- Mientras se justifican las evaluaciones de la calidad porque la escuela, aspecto en el que sí podría actuarse, explicaría un porcentaje importante del desempeño, el último informe del MEN sobre factores asociables encontró que «el mayor número de factores de logro del estudiante se registra, y no sólo estadísticamente hablando, en el “estudiante y su hogar”» [MEN, 1997b:56].

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 juega la interpretación de las cifras. No se expli-

citan 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 extinto principio musulmán gracias al cual quedaba demostrado que, excepto *El Corán*, se podían quemar todos los libros: si decían lo mismo que éste, sobraban, pues estaban repitiendo; si decían cosa distinta, debían ser falsos, pues sólo aquél dice la verdad. Entonces, queda claro que es el interés el que decide qué es y qué no es verdad, y no la “verdad” la que determina el interés.

Los últimos resultados de evaluación de la calidad se exponen en un orden muy particular: se habla de la científicidad de la investigación, se exponen las asociaciones y, como si fuera algo a agregar, se plantea que «cerca del 50% de las variables exploradas no registraron asociación con el logro cognitivo y que, las que sí lo hicieron, tuvieron poca densidad: no lo hicieron para los tres sectores» [MEN, 1997b:51]. Pensamos que lo más importante de la cita es el fracaso de la indagación; sin embargo, en todo el documento se habla dando respuestas o planteando preguntas *sin poner en duda los instrumentos ni la metodología*; se dice incluso que «nos aseguramos que nuestras fuentes sean válidas»<sup>106</sup> [:31], que la información «debe recogerse en situación confiable que permita, con un nivel conocido de confianza, considerar los hallazgos de la investigación como factores que realmente lo son» [:31]<sup>107</sup>. Valdría la pena darle crédito al mismo documento cuando plantea [:50] que «todos sabemos que una dicotomía 50% vs. 50% es atribuible al azar», con lo que el resultado de la investigación puede calificarse de azaroso. Después de tanta validez, confiabilidad, confianza y realidad mencionadas, ¿qué pasó? Como dice el mismo documento más adelante [:52]: «Próximas investigaciones, posiblemente aclararán el problema».

---

<sup>106</sup> Claro que para salir del paso en un aprieto se puede decir: «cabría la pregunta de si es exacta la información sobre el porcentaje de docentes que ha tomado los cursos [de capacitación]» [:55]. O sea, que no sólo no es cierto que se hubieran asegurado de que las fuentes fueran válidas, sino que es posible dudar de *toda* la información recibida.

<sup>107</sup> Es la misma posición en todo el mundo; Triana [1996:46], por ejemplo, dice: «Parece suficientemente demostrada la capacidad de la evaluación para generar conocimiento *válido, fiable y relevante* acerca de la situación y el estado de la educación [...]». Así mismo, pone la tendencia mundial de las evaluaciones masivas en términos de «La demanda de respuestas *objetivas y fiables* a las cuestiones suscitadas acerca del sistema educativo estadounidense [...]» [:37]. Y más adelante [:48] sostiene que los modernos sistemas de indicadores interesan porque proporcionan información *relevante*, permiten comparaciones *objetivas*, permiten estudiar las tendencias y enfocarse en *puntos críticos de la realidad*.

Pero, ¿se justifica que el país diga que está evaluando la calidad de la educación, invierta no pocos recursos en ello y, al final, se haga tal afirmación? Si próximas investigaciones aclararán el problema, ¿por qué no invertir de una vez en ellas? ¿Acaso ya no se invalidó toda pesquisa que no tuviera piso en el análisis estadístico?

Esto sería contraproducente para la educación si la política se diseñara efectivamente con base en estos resultados, tal como reza el credo del actual modelo: «Parece suficientemente demostrada la capacidad de la evaluación para generar conocimiento válido, fiable y relevante acerca de la situación y el estado de la educación, como proceso previo a la toma de decisiones susceptibles de producir una mejora de su calidad» [Triana, 1996:47]. Pero si a la hora de definir la política educativa fueran determinantes otros asuntos, incongruencias como estas resultarían secundarias...

Por ejemplo, pese a que el documento que da cuenta de los factores asociables en la evaluación del SINECE de los años 1993 a 1995 reconoce que no funcionó más del 50% de las variables exploradas sobre el director y su gestión [MEN, 1997b:51], se aventura a preguntar si la gestión tendría que estar centrada más en el desarrollo humano que en técnicas, dotaciones y operaciones [:53], de donde «el diagnóstico guía para las decisiones en política educativa es que falta gestión de lo intrínsecamente humano» [:53]. Extraña conclusión para un documento que no lo ha posibilitado y que se ha preocupado todo el tiempo por mostrar su científicidad. No obstante, nos sirve para mostrar las flagrantes contradicciones entre la política educativa (para la que la gestión es de orden administrativo) y lo que supuestamente le da sustento (para esa “investigación”, la gestión es de orden “humano”).

#### **7.1.4 Investigación**

De un lado, la búsqueda de correlaciones se postula a sí misma como una investigación científica: Triana [1999:53] afirma que «El conocimiento *objetivo*, y la valoración *rigurosa* de los resultados de la educación, generalmente efectuados a partir de su comparación a través del tiempo y del espacio, constituye una base sólida para la toma de decisiones encaminadas a la mejora». Más adelante plantea que «movidos por ese propósito de desvelar la realidad educativa tal y como se presenta en su encarnación escolar, los mecanismos de evaluación permiten arrojar luz sobre el interior de la mencionada caja negra que son los centros y comprender cómo la educación en abstracto se convierte en relación educativa concreta» [:54]. Se trata de un modelo político que se presenta a sí mismo como eficaz, como técnico, como desprovisto de atavismos politiqueros. Sin embargo, tal modelo de investigación y conocimiento no se sostiene ni en el mismo documento. Más adelante, Triana

[1999:56] plantea que «la acción evaluadora está forzosamente sometida a una tensión interna entre conocimiento e interpretación, que no siempre resulta fácil de resolver» [:56].

De otro lado, hay una “investigación” de otro orden: aquella que viene después de la investigación que se identifica con evaluación. Según el MEN [1992:26], los estudios que indagan factores asociables dan lugar a la investigación cuando es necesario explicar la interacción de variables «no completamente conceptualizadas»; o sea, la investigación comenzaría cuando el cruce de información (de desempeño y de factores) revela hechos especiales. Esto equivaldría a plantear que durante la elaboración de los instrumentos de factores asociables no mediarían teorías, salvo las del manejo estadístico. Tal vez por eso, en el cuestionario de factores asociados del profesor, durante la evaluación de impacto del PUEBP, pueden estar preguntas como: «¿Es usted hombre o mujer?» y «¿Con qué frecuencia se reúne con otros profesores de su escuela o escuelas similares, para discutir y planear enfoques curriculares o didácticos?»; también por eso la tabla de factores pone estos aspectos al mismo nivel, sin jerarquías entre ellos, como si fueran del mismo nivel la diferencia sexual anatómica y la construcción de comunidades académicas y pedagógicas.

Esta posición no ha cambiado con el curso de las aplicaciones; en 1997 se definen como factores asociables al logro cognitivo: «a los componentes que por hacer parte del sistema educativo mismo inciden de manera directa en el proceso educativo y aparecen simultáneamente con el buen o el mal desempeño de los alumnos, sin por ello explicar necesariamente el resultado encontrado, *ya que se requiere de mayor investigación para poder afirmar conclusiones al respecto*» [MEN, 1997b:29].

Esta manera de proceder se asemeja a la situación comentada por Bachelard [1979:257] sobre el abuso de las determinaciones recíprocas; dice que para un espíritu precientífico «todas las variables características del fenómeno están en interacción y el fenómeno se considera como igualmente sensibilizado en todas sus variaciones. Ahora bien, aun cuando las variables estén ligadas, su sensibilidad no es recíproca. Hay que hacer de cada investigación un caso especial». Por eso es tan fácil hacer listados: cualquier persona puede listar los factores que él “cree” que afectan el desempeño de los niños. Lo difícil, en cambio, es establecer relaciones...

Se ha dado por hecho que *hacer un listado* de factores es el resultado de “observar objetivamente” la realidad (particularmente la realidad educativa); recordemos que, para el MEN [1997b:28], una variable es factor de algo cuando produce un efecto que puede verificarse *empíricamente*. De esta manera, las discusiones teóricas parecen secundarias, tal vez inconvenientes, sobre todo cuando el “investigador” que cruza factores y desempeño se pone en la posición de «representar a los colombianos que



desean saber sobre la calidad de la educación» [:31] (así, los que quieren discutir automáticamente quedan del lado de los que no quieren saber sobre la calidad, ¡concepto que el mismo MEN ha dicho que es discutible!). En el último informe sobre factores son notables, de un lado, el esfuerzo por hacer una sustentación teórica y, de otro, el papel secundario que ella ocupa realmente. Se habla de que hay dos puntos de partida para proponer variables asociables al logro: una aproximación sociológica (de cuyas fuentes y principios no se dice nada) y una indagación bibliográfica sobre factores ligados al proceso de aprendizaje; es decir estudios anteriores que ya hicieron búsquedas parecidas... y cuyos fundamentos no se discuten ni se explicitan, pues lo importante son los factores que ya se buscaron, que van a servir para nutrir los instrumentos. Prueba del papel secundario de la teoría es que el documento se expone en el segundo punto de partida: cita indagaciones realizadas en países latinoamericanos y publicaciones de la UNESCO<sup>108</sup> y cuando va a definir los factores, se olvida que había enunciado un primer punto de partida y afirma: «Tratando de agrupar los factores hallados en la literatura científica revisada, el Instituto SER encontró sugerencias teóricas que le autorizan a considerar factores ligados indirectamente al desempeño escolar [...] y factores ligados directamente al desempeño escolar», y a continuación agrega una lista.

Recordemos que tal actitud también se muestra con ocasión de los instrumentos de logro: las preguntas de la evaluación del PUEBP en ciencias y matemáticas se sacaron «de un banco de preguntas, *experimentadas y validadas* para el TIMSS» [MEN, 1997a:83]. De igual manera, cuando el Banco Mundial rechazó las primeras pruebas realizadas en Colombia, puso a disposición del país pruebas hechas en Estados Unidos que, no obstante, servirían para los propósitos de evaluar la calidad de la educación en Colombia, en tanto eran pruebas *validadas* por el Educational Testing Service de los Estados Unidos.

Al punto vienen, entonces, las anotaciones de Tort [1979:5]:

*Si se le acusa, esgrime las armas propias de su competencia: comparación, muestreo, contrastación [...] al no haber sido instruido en los arcanos de la estadística, no [se] tiene más remedio que admitir estos argumentos como moneda de curso legal [...]. Cuando hace preguntas, no tiene que hacerse preguntas acerca de las preguntas (eso es asunto del autor del test), pues las preguntas han sido comparadas, contrastadas, muestreadas, patentadas [...]. La cifra, además de la autoridad que se*

---

<sup>108</sup> En este punto recordamos que durante un Seminario sobre investigación en educación (Biblioteca «Luis Ángel Arango», 1993), uno de los estudios concluía que mejorar la calidad de la educación requería cambiar la composición femenina del magisterio de primaria, a favor de una composición masculina. Ante semejante afirmación, una de las preguntas a los ponentes fue por el marco epistemológico de su investigación, por las teorías que los «autorizaban» a escoger los factores asociables que interrogaron; ante esto, respondieron que los factores sobre los que obtuvieron información venían de un listado *validado* por la UNESCO.

*confiere a la cantidad como tal, cuenta para sí con el marchamo de la "Ciencia", es decir, que es la medida de la "Ciencia" y no la palabra de nadie.*

De ahí que suelen tomarse los factores de listas elaboradas en otras investigaciones<sup>109</sup>. Es lo que anuncia el MEN en su último informe y lo que efectivamente hace [1997b:30]. De nada sirve la "aproximación sociológica" que anunció. Pensamos que se pueden validar pruebas sobre objetos de conocimiento con muchas deudas frente a lo científico [cf. §6.2.1], que se pueden validar instrumentos con fundamentos discutibles [cf. §4], que se pueden encontrar asociaciones estadísticas entre cifras sin que eso represente una relación de causalidad [cf. §7.1.5] y, lo que es menos visible, que los instrumentos reafirman la circulación de ciertas nociones o las promueven a la existencia en el ámbito educativo... como en el caso siguiente: en el instrumento de factores asociables aplicado en 1993, se le pide a los estudiantes decir si ven y oyen bien (preguntas 3 y 4). Esta pregunta posiblemente se origine en la apreciación de que los sentidos son los principales instrumentos del conocimiento: "ver para creer", "¿no entiende?, ¿es que no me oye?", «El niño no puede aprender qué es un imán, si no manipula un imán» [Colbert y Mogollón, 1980:168]). Pero el proceso cognitivo humano también ha sido interpretado de maneras para las que los órganos de los sentidos se vinculan de otra manera y, en consecuencia, las limitaciones en relación con ellos requieren explicaciones complejas. Hay, por supuesto, un soporte orgánico para los fenómenos cognitivos, una condición de posibilidad; pero éste no es determinante. Conceptos como el de "obstáculo epistemológico", en filosofía, o el de "resistencias", en psicoanálisis, o el de la "producción social de la realidad", en sociología y ciencias del lenguaje, hablan más de la naturaleza propiamente humana del conocimiento.

Ciertas limitaciones en los órganos de los sentidos ni siquiera son percibidas por la persona que las sufre (¿cómo se puede saber si se es ciego cuando nunca se ha visto?); éstas aparecen como problema sólo a partir de la interacción con otros seres humanos. Razón de más para pensar en la importancia de esa zona de interacción, en detrimento de la que podría atribuirse a los órganos de los sentidos. Beethoven era sordo, pero oía la música, por eso podía componer: se trata de la producción, mediada por el lenguaje, de *sentidos culturales*. Por eso decía Marx que *los sentidos han devenido sentidos humanos*: «el ojo se ha hecho un ojo humano, así como su objeto se ha hecho un objeto social, humano, creado por el hombre para el hombre. Los sentidos se han hecho así inmediatamente *teóricos* en su práctica [...] estos órganos inmediatos se constituyen así en órganos *sociales*, en la *forma* de la sociedad» [1989:148-149]. Investigaciones en óptica revelan [Foerster, 1994], por ejemplo, que nuestro campo visual también está integrado por puntos ciegos, pero que

---

<sup>109</sup> Es el objetivo de la Organización para la cooperación y el desarrollo económico [OCDE]: construir indicadores internacionales de la educación [cf. Triana, 1996:38].

nosotros no los percibimos; es decir, *no vemos que no vemos*. Estos estudios concluyen —contra el adagio empirista “ver para creer”— que no vemos algo si no lo creemos: aquello de “ojos que no ven, corazón que no siente” resulta bastante extraño para la lógica del celoso, cuyo corazón siente aun en ausencia de un hallazgo visual.

El formulario, entonces, valida un modelo empirista de conocimiento (“nada hay en el entendimiento que no pase por los sentidos”). La dificultad del maestro para explicar determinadas diferencias de ritmos de aprendizaje, proviene de este modelo: o bien piensa que la “incomprensión” del estudiante debería provenir de deficiencias en los órganos de los sentidos, ya que supuestamente son los instrumentos del conocimiento; o bien, no entiende por qué el estudiante no entiende [Bachelard, 1979:20], siendo que sus órganos de los sentidos están en buen estado. La perspectiva materializada en el cuestionario, legitimada por él, promovida por él a la existencia, es inherente no sólo a una mirada sobre el conocimiento desde lo disciplinar, sino a ciertas condiciones de entender a los sujetos de acuerdo con perspectivas históricas precisas.

Por todo esto, la investigación tal vez comenzaría con el análisis de la concepción que da lugar a los listados de factores, no con su transformación en instrumentos con los cuales tener después una coartada a la hora de trazar la política educativa. Mientras de un lado se convierte la “confusión” del mundo circundante en una lista de factores (lo que incluye las asignaturas sobre las cuales se pregunta) que se seleccionan con arreglo a criterios no siempre explícitos<sup>110</sup>, de otro lado se intenta convertir esa misma “confusión” en *complejidad*, es decir en una red de relaciones que jerarquiza el papel de los elementos y sus interrelaciones en los efectos y, en consecuencia, la organización y selección pretende obedecer a criterios explícitos.

Si se tiene en cuenta que es la interpretación racional la que ubica los hechos (y no el experimento o la estadística), incluso en ciencias consideradas “duras”, la ausencia de explicación que parece haber en estos estudios es significativa. «Sólo la razón dinamiza a la investigación, pues sólo ella sugiere, más allá de la experiencia común (inmediata y especiosa), la experiencia científica (indirecta y fecunda)» [Bachelard, 1979:19]. ¿Cuál es, en el caso de la investigación de factores asociables, la “interpretación racional”, la razón más allá de la experiencia común, la que da consistencia a esos listados? Para poner un par de ejemplos, en el listado del MEN [1997b:30] se enuncian las “condiciones climáticas” dentro de los factores “relativos al medio

---

<sup>110</sup> Tal como ocurre a veces con diversos programas en educación: pueden tener propósitos opuestos. Es el caso de una de las ideas asociada a la propuesta de PEI (trabajo colectivo) -vs- el *Programa nacional de incentivos* (premios individuales). En la SED, es el caso de la evaluación de competencias (para la que la diversidad de material impreso es fundamental) -vs- la dotación de textos, que se hace con una concepción amarrada a los libros de texto [cf. SED, 1999].

social”; y siendo que hay factores “de carácter educativo”, la “repetencia” y el “alto desarrollo educativo del distrito escolar” también quedan dentro de los factores “relativos al medio social”.

Para encontrar estas cosas no se necesita perspicacia, se puede buscar en la literalidad de los enunciados ministeriales: «la lista de factores de desempeño escolar verificada por estudios adelantados en diferentes países [...]» [:31]. Tampoco es difícil encontrar que, más allá de toda justificación teórica hay otro interés: la preparación de los instrumentos de factores se hizo «con el ánimo de identificar algunas de las variables que posiblemente se asocien al rendimiento escolar, *especialmente de aquellas que puedan ser objeto de intervención de la política educativa*» [:31]; para lo cual no se necesitan “estudios sociológicos”: recordemos que la política educativa tiene muchos puntos en común con otros países...

A la hora de establecer y de evaluar los desempeños hay al menos una concepción de escuela<sup>111</sup>; es decir, las ideas sobre el desempeño en las disciplinas (lenguaje, matemáticas y ciencias, en el caso de la evaluación de impacto del PUEBP) o en los valores ya comportan ideas sobre los factores asociables, tal vez de manera implícita; y también en la indagación de factores, así esto se trate de ocultar detrás de un listado, detrás de una instancia cuya autoridad nos impida preguntar, o detrás de la “validación” previa. La incertidumbre, entonces, aparece en estos estudios reemplazando una explicación que no se da, pero que se podría dar, pues sí median teorías: las concepciones no pueden no mediar.

La investigación sobre asociación entre desempeño y factores asociables parece obedecer a la siguiente descripción de Bachelard [1979:13-14]: «La experiencia común no está en verdad compuesta, a lo sumo está hecha con observaciones yuxtapuestas [...] la experimentación debe apartarse de las condiciones ordinarias de la observación». La prueba es que los factores son listados y que se prestan de una investigación a otra, se toman fragmentariamente, pueden incluir pocas o muchas preguntas (como en el caso de la evaluación del SINECE o la evaluación de impacto del PUEBP, respectivamente), pocos o muchos cruces, etc.

De otro lado, si las preguntas vienen después de aplicados los instrumentos y si se trata de explicar la interacción de variables «no completamente conceptualizadas», ¿por qué puede afirmarse que la televisión —al salir asociada positivamente— es un factor «contrario a la predicción del modelo» [MEN, 1992:90-91]? Si esto es así para la televisión, querría decir que el modelo puede predecir *todos* los demás factores; o

---

<sup>111</sup> De hecho, el grupo que elaboró la primera prueba de logro en lenguaje que aplicó el SINECE en grado 5°, empleó más tiempo en discutir sobre la especificidad de la escuela y su relación con la comunicación, que en la formulación de preguntas para el instrumento.

sea, sí están conceptualizados de alguna manera, y la investigación se justifica sólo en relación con lo que es contrario a la predicción del modelo; el resto, el modelo ya lo sabe. La tautología se completa cuando, además, se rechaza lo que no sirve para la política educativa.

Cuando en el cuestionario de factores asociables del SINECE se pregunta a los estudiantes si están de acuerdo con la frase *Las matemáticas son más para los hombres que para las mujeres*, ¿acaso la pregunta ya no presupone la correlación buscada?, ¿cuál es el descubrimiento?, ¿qué se está entendiendo por “investigación”? ¿No hay acaso una concepción sobre mujeres y hombres cuando se pregunta de manera tal que se “encuentra” que las niñas sirven menos para las matemáticas?; a esto se le agrega la idea —no sustentada, por supuesto— de que las diferencias entre niños y niñas puede explicarse con bases neurofisiológicas<sup>112</sup>; y, finalmente, se dice que «Esta variable requerirá de mayor investigación» [MEN, 1992:92]. Pero todo el esfuerzo de la evaluación masiva justificaría, además de una recomendación de investigar, unas líneas de pertinencia para esas investigaciones, así como una discusión académica, pública, sobre los marcos teóricos que están a la base de dichas evaluaciones. Si con esa concepción sobre hombres y mujeres se hubieran elaborado instrumentos para establecer el desempeño de los estudiantes, se habrían hecho cosas muy distintas, como haber diferenciado los textos según si el estudiante fuera niño o niña, o cosas por el estilo; y si quienes elaboraron instrumentos para establecer logro hubieran hecho los instrumentos de factores, nunca habrían aparecido preguntas así<sup>113</sup>.

Se genera desorden (entropía) y después se dice que allí hay cosas por investigar, que puede introducirse un orden. Esta “incertidumbre” que supuestamente conduce a la investigación —y que aparece cuando se buscan correlaciones entre desempeño y factores asociables—, más bien da cuenta de la ausencia de vinculación conceptual entre los equipos que se encargan de diseñar y producir los instrumentos respectivos [Duarte et al., 1996:47]. «Los argumentos racionales que afectan a la experiencia son ya momentos de esta experiencia [...]. Ya no se trata de confrontar un espíritu solitario y un universo indiferente» [Bachelard, 1973:136-137].

Para Bachelard [:135], en disciplinas fuertemente constituídas tienen lugar búsquedas bien planteadas que, sin embargo, no obtienen resultados; pero no pueden ocurrir búsquedas como las que analizamos: ocasionales, experiencias improvisadas,

---

<sup>112</sup> Recordemos que la estadística también sirvió para probar que el coeficiente intelectual de los negros era inferior al de los blancos... Un premio Nobel, William Shockley, lo “demostró” para proponer que el gobierno los esterilizara, a cambio de una indemnización [Tort, 1979:21].

<sup>113</sup> En el estudio sobre evaluación de impacto del PUEBP, por ejemplo, el equipo de lenguaje se opuso a que se introdujera la variable “género”.

aunque muestren “resultados” y largos “antecedentes”. Las primeras tienen valor epistemológico, las segundas no: «Ningún físico malgastaría “sus créditos” para construir un instrumento sin destino teórico» [:135].

### **7.1.5 El uso de la estadística**

Como disciplina, la estadística tiene su especificidad; esto quiere decir que —dentro de la matemática— tiene su objeto, sus métodos, etc. Y aunque en casos como la detección de asociaciones, ella se ofrece como una ayuda para procesar una información cuya naturaleza y sentido son establecidos por otros campos conceptuales, no se puede desdeñar la implicación recíproca, en el sentido que diferentes campos del saber han obligado, de un lado, a que la estadística transforme su estructura conceptual (de ahí la aparición de distintas escuelas, como la no-paramétrica, la robusta, la bayesiana, etc.) y, de otro, a que esos campos del saber también tengan que repensar algunos de sus componentes.

Claro que, para Triana [1996:48], la notable transformación que los indicadores educativos han introducido en la estadística tradicional, «se debe más a su significación y a la utilización que de ellos se hace, que a sus fundamentos técnicos y su procedimiento de obtención. En realidad, desde el punto de vista de su cálculo no son muchas las novedades que introducen, aunque no ocurre lo mismo en lo que se refiere a su definición y, sobre todo, su uso».

De cualquier forma, la manera como se ha utilizado la estadística en la evaluación de la calidad educativa, en el establecimiento de la política educativa y en la evaluación de determinados programas (como en la evaluación de impacto del PUEBP) ha implicado —al menos— dos tipos de problema:

De un lado, que los criterios técnicos condicionen en gran parte los criterios teóricos, pedagógicos y educativos (de donde la educación cede su especificidad a aquella de la estadística); es uno de los riesgos que Triana [:42] señala para la evaluación de la educación: el hecho de reducirla a los elementos más asequibles, *al margen de su relevancia*<sup>114</sup>. Ejemplos:

---

<sup>114</sup> Como en el chiste de quien busca algo en un sitio distinto de aquel en el que lo perdió, con el argumento de que allí hay más luz.

- es la dificultad de procesar la información venida de pruebas abiertas (confiabilidad) la que restringe su uso, y no la naturaleza del estudio<sup>115</sup>;
- se ha pretendido tener pruebas de desempeño estandarizadas, “probadas”, antes de tener suficiente investigación sobre el conocimiento en las diferentes áreas objeto de análisis, y antes de tener suficiente discusión pedagógica sobre la evaluación; de ahí que a veces haya que tratar de “teorizar” *a posteriori* los resultados estadísticos;
- y se ha omitido la discusión sobre qué factores asociables 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, de las siguientes cuatro maneras:

- a. *Por la selección de los resultados que, a juicio del funcionario de turno, parezcan acordes con las expectativas previas, con la continuidad de la política educativa, o con la promoción de una nueva política educativa.*

Por ejemplo, después de que mediante herramientas estadísticas se hizo 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 las que presentan signos contrarios, *¿por qué no se descartan las que presentan signos favorables, 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];

---

<sup>115</sup> Y, no obstante, en la presente 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; y en “eficiencia interna” muestra tantas correlaciones como el análisis estandarizado.

- b. *Por el uso indebido de algunos procedimientos en los cuales los datos no se corresponden con los supuestos teóricos que su teoría demanda.*

Este caso se presenta, por ejemplo, cuando se hacen regresiones incluyendo factores que son asociables a unidades de observación jerarquizadas. El MEN [1992; 1997b], buscando asociaciones entre el desempeño promedio de los alumnos y las características de los docentes, el plantel y el director, adopta un modelo de regresión lineal múltiple. No obstante, esto no es adecuado, pues de una parte algunas de las variables explicativas son de naturaleza categórica, y de otra, las unidades de observación son de tipo jerárquico<sup>116</sup>, ya que, por ejemplo, los alumnos están incluidos dentro de clases (variables del profesor), las clases están incluidas dentro del plantel (información del plantel incluyendo a la del respectivo director). De manera que:

- la aplicación de la metodología llamada «análisis de varianza ponderado en una dirección» [MEN, 1992:54], si bien puede servir para hallar diferencias, por separado, entre desempeños por niveles de la variable categórica asumida, esta metodología no tiene en cuenta la posible asociación que existe entre los factores;
- la aplicación de la metodología descrita [MEN, 1992:53-56; 1997b:27-31] trata a las variables como si todas fuesen de tipo cuantitativo y deja de lado mucha información contenida en la variabilidad de los datos;
- la técnica empleada para la “generación” de factores al parecer es el análisis factorial clásico, la cual no es pertinente para estos casos [MEN, 1997b:27-31]; en su lugar, está el análisis de correspondencias, el cual trata de explicar la información contenida en variables categóricas en una serie de “nuevos” factores [Díaz, 1999a:§11];
- para establecer no sólo la posible asociación entre los factores y el desempeño, el análisis de regresión [cf. MEN, 1992; 1997b] no es adecuado, hay disponibles otras técnicas, tales como: los modelos log-lineales, la regresión logística, el análisis de correspondencias, entre otras;
- valerse de un modelo de regresión lineal [cf. MEN, 1992; 1997b] implica estar seguro del cumplimiento de una serie de supuestos estadísticos sobre los cuales se construye y valida la relación que se propone para las variables en juego. Supuestos tales como: no asociación entre las variables explicativas (no colinealidad), varianza constante de los errores, independencia entre las

---

<sup>116</sup> Para estos casos están los “modelos jerárquicos” [Goldstein, 1987].



observaciones, entre otras, “avalan” las posibles asociaciones, que a través de la regresión lineal, se puedan evidenciar entre los datos disponibles. De lo contrario, estos modelos pueden mostrar relaciones que sólo se encuentran en la realidad de los cálculos, pero, muy posiblemente, contrarias a la realidad del marco conceptual con el que supuestamente se obtuvo tal información;

- no se explicita el nivel de incertidumbre asumido [cf. MEN, 1992];
- la metodología para la selección de variables tipo “backward” [cf. MEN, 1992] debería haber sido complementada con la metodología “paso a paso” (step-wise) [sí se hizo en MEN, 1997b] o la tipo regresión hacia adelante (forward), para hacer una especie de “panel de metodologías”, que junto con la mirada conceptual con la cual se está indagando la realidad, permita tomar decisiones sobre la exclusión (o la inclusión) de las variables que formarán parte del modelo que se quiere construir.

c. *Por la aplicación de un modelo a un objeto cuya naturaleza no aceptaría esa aplicación.*

Este caso se ilustra con el informe sobre factores “asociados” al logro en el DC [SED-Corpoeducación-DNP, 1999]:

- Siendo “matemáticas” y “lenguaje” dos áreas que tienen su propio objeto de trabajo y fundamentación conceptual, no deben compararse sus resultados; tampoco se trata de la misma prueba ni de la misma escala de medición. No obstante, se habla de un “logro total” (suma de puntajes de matemáticas y lenguaje). Cada disciplina se evalúa sobre la base del concepto de “competencias”, pero cada competencia obedece a una concepción particular y distinta de desempeño; así, considerar un puntaje global a la hora de hacer el cruce con los factores asociados, neutraliza las categorías de cada equipo, con lo que más valdría no haber hecho pruebas con arreglo a criterio, sino haber contado número de preguntas correctas. Se hacen inversiones económicas y esfuerzos teóricos para construir y concertar teorías sobre los objetos de evaluación, pero al final todo se aplana bajo un solo criterio de naturaleza técnica. *Hablar de medias y compararlas para cada uno de los factores “asociados” resulta inadecuado.*
- Por su parte, las variables de los alumnos, las características de los docentes y de los planteles en su mayoría son de tipo nominal y ordinal, de manera que el coeficiente de Pearson no es adecuado para buscar la correlación de estas variables con el logro académico; no obstante, si así fue, las correlaciones

mostradas son muy bajas; otra cosa es que, como se comprueba matemáticamente, por los grandes tamaños de muestra las correlaciones se evidencien significativas.

- Buscar factores “asociados” al logro desde el análisis de regresión jerárquica, es atribuir a la mera relación de contención conjuntista —por el hecho de “compartir” un salón, una escuela, unos docentes, unas normas— capacidad para dar cuenta de la afectación (aporte) que tienen los factores asociables sobre el desempeño. Pero tal vez no sea suficiente<sup>117</sup>. Atribuir el desempeño a los otros estudiantes, a los docentes, a la escuela, y por último, como una consecuencia de todo lo anterior (“sumando” en este orden), al sistema educativo, sobre la base de la agrupación jerárquica —anidada—, tal vez equivalga a dar explicaciones desde un punto de vista puramente geográfico. Lo pertinente sería indagar y buscar las variables o factores que estarían determinando la asociación entre las diferentes unidades de observación y que atravesasen el eje desempeño construido.
- Por último el informe se pregunta si los resultados obtenidos por los alumnos del grado 5° se deben al desempeño académico efectivo de los estudiantes o a la prueba misma; no obstante, este gran vacío, advertido por los mismos investigadores<sup>118</sup>, el trabajo fue desarrollado...

*d. Finalmente, por la asignación de causalidades en la realidad educativa a partir de correlaciones de cifras, sin establecer las mediaciones entre ambas.*

En este caso, quienes diseñan los cuestionarios los tienen como instrumentos objetivos de aprehensión de lo real. Para lograr esa supuesta “objetividad”, se recurre a todas las técnicas de elaboración de instrumentos psicométricos, con respaldo matemático, y validación mediante pruebas piloto. Después de esto, parecería infundado dudar sobre la capacidad predictiva de los instrumentos (aunque ellos conduzcan a absurdos como afirmar que una persona con la cabeza entre un horno y con los pies entre un congelador, en promedio, se siente bien). Se cree hablar desde la verdad misma, con lo que «el pensamiento del otro sólo puede ser un error o mala fe y el hecho mismo de su diferencia con nuestra verdad es prueba contundente de su falsedad, sin que se requiera ninguna otra»... es la concepción paranoide de la verdad de la que hablaba Zuleta [1994:13]. Pero *dos eventos pueden mostrar asocia-*

---

<sup>117</sup> Un ejemplo es la familia: las condiciones de la familia no afectan de la misma manera a todos los hermanos...

<sup>118</sup> Como decía Churchill, a veces el hombre se tropieza con la verdad, pero casi siempre se levanta y sigue adelante.

*ción estadística no solamente porque sean una causa del otro sino, por ejemplo, porque media un tercero que no está considerado.*

Lo curioso es que efectivamente se encuentran correlaciones estadísticas... Pues bien, para quienes ven en los instrumentos ventanas hacia el mundo, estas asociaciones se constituyen en prueba de la existencia de relaciones causales en la vida social; no obstante, hay cantidad de ejemplos que involucran factores altamente asociados estadísticamente, pero que obedecen a una mirada parcial o accidental de los hechos; por ser evidentemente contradictorios o absurdos, queda por lo menos interrogado el hecho de que una dependencia entre cifras equivalga, sin otra mediación, a una correlación entre factores de la vida social. Veamos un par de ejemplos: la asociación entre el número de cigüeñas y el número de nacimientos... entre el número de emisoras radiales y el número de enfermos mentales en una ciudad... entre el cáncer de pulmón y el hábito de fumar.

Casos como éstos aparecen continuamente en la “investigación” de factores asociables al desempeño de los estudiantes; por ejemplo:

- el tipo de escuela en la educación pública rural: es un factor asociado negativamente con el logro de matemáticas en 3º, ¡pero se asocia positivamente en grado 5º! [MEN, 1992:91]. ¿Cómo es posible creer que el tipo de escuela incida negativamente en el rendimiento de los estudiantes de 3º y que dos años después para los mismos niños esa incidencia sea positiva?
- Otro caso: «En los planteles de calendario A la asociación es negativa: a mayor número de salones de clase, menor logro. En los planteles de calendario B esta asociación es positiva: a mayor número de salones de clase, mayor logro» [MEN, 1997b:45]. Igualmente, ¿cómo es posible creer que el número de salones de clase incida negativamente en el rendimiento de los estudiantes de un calendario y que en el otro incida positivamente (y eso que en una misma ciudad pueden funcionar ambos calendarios)?
- Es también el caso de considerar que idéntico factor es contraproducente y es favorable para la calidad: de un lado, una investigación citada busca mejorar la calidad cambiando la composición femenina del magisterio a favor de una composición masculina; y, por otro lado, Ávila [1999:38], leyendo los resultados de instrumentos similares, dice: «los docentes de sexo femenino obtienen mayores logros de sus alumnos».
- Es, finalmente, lo que se ve en varios casos del presente trabajo: vimos que se pueden hacer afirmaciones contrarias sobre el mismo aspecto, con sólo variar el ángulo de análisis del desempeño; puede ilustrarse con el factor “número de

alumnos por docente”: desde el análisis estandarizado, el desempeño presenta asociación altamente significativa y positiva con dicho factor; pero visto desde el análisis por competencias de la parte cerrada, se muestra asociación negativa altamente significativa [cf. §5.2.3].

Casos como éstos o no se explicitan o se ponen al lado de los otros como curiosidades. Pero, eso sí, la política educativa *escoge* los que le sirven, es decir, los que no contradicen sus propósitos:

*El método utilizado para la selección de variables que pasan a la siguiente fase corresponde al método Backward con un grado de confianza del 90%. Partiendo de este modelo se descartaron las siguientes variables:*

- *Aquellas que no cumplieran con las exigencias de significancia requeridas.*
- *Las que presentaran signos contrarios a la predicción teórica.*

*De estas variables descartadas se exceptuaron aquellas que presentaban especial importancia en relación con el diseño de políticas educativas [MEN, 1992:54-55].*

Pero si todos estos datos se producen con las mismas herramientas metodológicas, ¿por qué escoger unos y no otros? Es claro que o todos sirven (y habría que considerar las contradicciones lógicas o los absurdos enumerados como una manera de ser de lo social) o ninguno sirve (sabemos que a los niños no los traen las cigüeñas, que la radio no produce la locura, que muchos fumadores no padecen de cáncer pulmonar...). Es decir, las correlaciones estadísticas, primero, no muestran una relación causa-efecto y, segundo, pueden provenir de terceros factores no considerados o, incluso, ser fortuitas. En el caso del número de nacimientos y el número de cigüeñas, un estudio estadístico encontrará asociación entre ambos —lo que autorizaría, a quien establece inmediatamente relaciones causales, a concluir que a los niños los trae la cigüeña— porque 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, y más nacimientos.

Las expresiones estadísticas requieren una base que presente cierta homogeneidad o también una extensión apropiada, oportuna. Sin embargo, esto es una abstracción, pues si hubiera homogeneidad, sobraría la estadística [Simiand, 1994:234]. Por eso, la elección de las abstracciones estadísticas es arbitraria, pero debe esforzarse por ser fundada, «porque hay buenas y malas medias, medias que tienen sentido y medias que no, desconfiemos de las medias, controlemos, cotejemos las indicaciones de medias de un tipo con las de otro, con otros índices, con datos complementarios y conservemos solamente aquellas que, después de estas pruebas, nos presenten una

verdadera consistencia y respondan a alguna realidad colectiva» [Simiand, 1994:235]. Por supuesto que hay un límite, pero mientras en ciencias positivas la mala abstracción puede llevar a una imposibilidad material, una mala abstracción estadística no necesariamente salta a la vista, «la no realidad de la relación percibida no puede manifestarse de manera material» [Simiand, 1994:234], pues las cifras se dejan combinar entre sí.

De ahí que, para no engañar con abstracciones estadísticas, Simiand recomienda [1994:232-233] que las expresiones de hechos complejos, las medias, los índices, los coeficientes no resulten simplemente de computar, de combinar cifras de manera arbitraria, «sino que se modelen a partir de la complejidad concreta, respeten las articulaciones de lo real, expresen algo a la vez distinto y verdadero respecto a la multiplicidad de los casos individuales a los que corresponden [...] lo que de hecho a menudo nos despista cuando empleamos abstracciones estadísticas, *no es que sean abstracciones, sino que son malas abstracciones*». El asunto no es sencillo (como “escoger” las cifras que se adapten a una política educativa preestablecida), es un problema epistemológico: ¿cómo saber que algo está “modelado” a partir de la complejidad concreta?, ¿cómo saber qué modelo respeta las articulaciones de lo real?, ¿cómo distinguir aquello que expresa algo distinto y verdadero respecto de la multiplicidad, de aquello que no? Por lo menos queda claro que es necesario tener como mínimo una conceptualización sobre aquello que se quiere investigar, sobre su complejidad. Cuando los elementos son heterogéneos o variables en el tiempo, los indicadores disimularán el rasgo esencial del elemento estudiado, en vez de ponerlo en evidencia [Simiand, 1994:233].

La consistencia, entonces, no viene de los números, «las determinaciones numéricas en ningún caso deben sobrepasar en exactitud a los medios de investigación» [Bachelard, 1979:254]:

*La precisión numérica es frecuentemente un motín de cifras [...] Puede verse en ella uno de los signos más claros de un espíritu no científico en el instante mismo en que ese espíritu pretende la objetividad científica [...] la precisión de una medida debe referirse constantemente a la sensibilidad del método de medida y ha de tener en cuenta las condiciones de permanencia del objeto medido. Medir exactamente un objeto fugaz o indeterminado, medir exactamente un objeto fijo y bien determinado con un instrumento burdo, he ahí dos tipos de ocupación vana que la disciplina científica rechaza [:250].*

Dos posiciones se perfilan al respecto: de un lado, puede creerse que se describe mejor la realidad mientras más decimales tenga la medida, o mientras la muestra sea más grande (y, entonces, el ideal es la prueba censal); y, de otro, pueden discutirse las condiciones del estudio, determinarse la sensibilidad y el alcance de los instrumentos. Caso en el que hay más preocupación por describir el método de medida

que el objeto mismo, con lo cual éste termina siendo «un grado particular de la aproximación del método de medida» [Bachelard, 1979:250]. El tiempo se midió con ayuda de la caída de la arena, con ayuda de la sombra... ¿era un tiempo “preciso”, era *el* tiempo? Y luego, cuando se midió con un muelle elástico, ante lo cual las mediciones anteriores resultaban burdas, ¿allí sí era un tiempo “preciso”, era *el* tiempo? Y ahora, que se mide usando emisiones atómicas (relojes de cuarzo, relojes atómicos), ahora que podemos medir las diferencias entre eventos que para los relojes anteriores eran simultáneos, ¿ahora sí hemos capturado *el* tiempo? ¡No!, cada vez lo *producido* era distinto:

*El objeto puede variar de naturaleza cuando cambia el grado de aproximación. Pretender agotar de golpe la determinación cuantitativa es dejar escapar las relaciones del objeto. Cuanto más numerosas son las relaciones del objeto con otros objetos, tanto más instructivo es su estudio. Pero en cuanto las relaciones son numerosas, ellas están sometidas a interferencias y de pronto la búsqueda discursiva de las aproximaciones se convierte en una necesidad metodológica. Entonces la objetividad se afirma como método discursivo, más acá de la medida, y no más allá de la medida, como intuición directa de un objeto. Hay que reflexionar para medir y no medir para reflexionar [Bachelard, 1979:251].*

O sea, en la evaluación de la calidad hemos hecho lo contrario: creemos que es cada vez el mismo objeto el que se precisa con aproximaciones sucesivas; hemos reducido las relaciones a cifras, cuando éstas en realidad tratan de ser una aproximación a aquéllas; la búsqueda discursiva de las aproximaciones está ausente o a manera de comentario de tablas y gráficas; nuestra “objetividad” ha estado más allá de la medida.

De tal manera, con sus respuestas, los estudiantes hacen existir una clasificación en la que podrían no haber estado antes de la prueba, pero en la que estarán después de que los resultados comiencen a entretenerse en lo social. Desrosières [1995:22-23] plantea que las clasificaciones se nos muestran de tal manera que los números parecen expresar cosas que existen independientemente de las condiciones que las establecen, cuando en realidad «una codificación es una decisión convencional de construir una clase de equivalencia entre objetos diversos, considerada la ‘clase’ como más “general” que cualquiera de los objetos particulares que incluye. Precondición para ello es la asunción de que esos objetos puedan ser *comparados*». Pero la cuestión no es si son realmente equivalentes los objetos, sino —como dice Desrosières— *quién decide tratarlos como equivalentes y con qué fines*.

Sin embargo, en Colombia quien pregunta por los intereses en juego durante estas campañas de evaluación masiva, se lo mira como un impertinente. No obstante, sabemos que el actual modelo aplicado a la educación tiene objetivos claros (los fines que Desrosières sitúa como determinantes). Y aquí es donde la aparente incerti-

dumbre y la necesidad de “investigar” estadísticamente adquieren otro carisma. Cuando se pone en entredicho la relación entre la serie *números* y la serie *conocimiento* [Cf. Benedito, 1975], las conclusiones ya no pueden ser en relación con lo que los entrevistados saben, sino en relación con los intereses en juego.

Por ejemplo, no sería necio atar unos cuantos cabos a propósito de los cursos de capacitación: por un lado, el Banco Mundial —apurado por imponer el ajuste fiscal— dice que la capacitación de docentes no aporta al mejoramiento; por otro lado, la investigación en nuestro país —hecha a imagen y semejanza de la del Banco Mundial— encuentra que aquí tampoco; y, por último, aparece una serie de medidas: los Centros Experimentales Piloto —tradicionalmente encargados de la “capacitación”— han sido desestimulados, los departamentos que los quieran mantener, lo harán asumiendo los costos; la capacitación, que comenzó siendo gratuita, cada vez más corre por cuenta de los mismos docentes; y la oferta de capacitación va quedando en manos del sector privado... Todo consistente con el “ajuste fiscal”. No obstante, nada se dice del tipo de capacitación que ha habido como producto de las políticas educativas del MEN, qué otras posibilidades hay de pensar la formación permanente de los docentes en ejercicio, ni qué responsabilidades tiene el MEN en el fracaso de la capacitación que hemos tenido (si es que ha sido un fracaso...).

Desde otra mirada, se podría decir —incluso con base en los mismos resultados de la evaluación— que *esa* capacitación, la bien llamada feria del crédito, no está asociada a la calidad, pero que *otra* capacitación, pensada desde un concepto de escuela y de formación, y no desde la idea de reglamentar los compromisos con los docentes al menor precio, sí podría estar asociada.

Otro caso sería el de los textos: pese a que la Ley 115 establece que «Los textos escolares que se adquieran deberán ser definidos de acuerdo con el Proyecto Educativo Institucional [...]» [art.102], el MEN entrega *el mismo libro* (por ejemplo, cartillas de escuela nueva) a miles de escuelas [cf. Samper, 1998], sin que haya sido definido por los PEI de cada una de ellas<sup>119</sup>; pero es que hay políticas internacionales, que nos toca aplicar, para las que los textos están asociados positivamente al logro [Cf. Heyneman et al., 1991]. No es casual, entonces, que nuestros instrumentos de factores asociables busquen en relación con dicho aspecto, de una manera que encuentren lo que ya el Banco Mundial ha encontrado en todos los países donde condiciona los préstamos a que se hagan esos “hallazgos”.

Miremos, sin embargo, la forma como se obtiene este “conocimiento” sobre los textos, sabido desde antes sin gastar todo el dinero que implica poner en marcha un

---

<sup>119</sup> Claro que siempre podrá argumentarse que esa condición no se le puede poner a los textos *regalados*.

Sistema de Evaluación: al docente se le pregunta: *Durante este año, siguió para su clase ¿uno o varios textos?* No se interroga sobre la *utilización* de material impreso diverso, sino sobre el *seguimiento* de textos, concepto aplicable solamente a libros de carácter instruccional (el llamado “libro de texto”). Al estudiante se le pregunta *En tu casa, ¿lee libros distintos de los textos escolares?* Con lo que habría podido recomendarse la dotación de impresos distintos a los libros de texto, en el caso de que las respuestas afirmativas estuvieran asociadas a buenos logros, o de que no se hubiera respondido, caso en el que podría tenerse en cuenta la importancia de la diversidad de textos... pero sobre el destino de esta pregunta nada se dice en los informes. Preguntas como éstas no serían suficientes para justificar políticas a favor de determinado tipo de texto; a lo sumo, a favor de la importancia del material escrito —cosa evidente—, cuyas diferencias y cuya utilización no se exploran en la evaluación [cf. Jurado y Bustamante, 1994], asunto que, paradójicamente, sí resulta prioritario para el Banco Mundial: «No conocemos las condiciones exactas bajo las cuales se producen los efectos de los libros de texto [...] Sería de utilidad que se realizaran esfuerzos para obtener más información a este respecto» [Heyneman et al., 1991:134].

Sobre la estadística, quisiéramos resaltar un último punto. Ciertas decisiones que se toman en su cuerpo teórico son arbitrarias e implican efectos en los “resultados” de su aplicación. Por ejemplo, la estandarización en la evaluación de impacto del PUEBP se hizo, como hemos explicado [cf. §1.3.1], mediante distribuciones homogéneas de media 50 y una desviación estándar para determinar tres grupos: “alto”, “medio” y “bajo”. Pero, ¿por qué es *una* desviación estándar la que se suma y se resta a la media para determinar estos tres grupos? Cualquier otro valor mueve las cifras de una manera distinta y arroja resultados distintos. A continuación hacemos el ejercicio de aplicar media y una desviación estándar a las mismas cifras; se obtienen grupos distintos de “alto”, “medio” y “bajo”, como se ve en la siguiente tabla:

		Desviaciones estándar		
		0.5	1	1.5
Desempeño estandarizado	Bajo (%)	34.6	17.2	5.6
	Medio (%)	36.2	61.4	86.7
	Alto (%)	29.2	21.4	7.8

Como se ve, los grupos cada vez son muy distintos. En este caso, la clasificación que produce grupos más homogéneos sería la obtenida no con una desviación estándar (la que se aplica tradicionalmente), sino con media; cuando se aplica una-y-media desviación estándar, las diferencias entre los grupos es aun mayor que cuando se aplica una. A continuación hacemos el ejercicio de cruzar el desempeño de los



niños en lenguaje con los factores-libro analizados en §4, cada vez con las categorías estandarizadas obtenidas aplicando media, una y una-y-media desviaciones estándar:

		P-valores		
		Desviaciones estándar		
		0.5	1	1.5
Libros	En casa	0.997	0.130	0.390
	Recibidos x la escuela	0.161	0.368	0.139
	Guías de escuela nueva	0.796	0.002	0.592
	Para la biblioteca	0.116	0.187	0.501
	Textos escolares	0.060	0.042	0.741
	TE de lenguaje	0.066	0.048	0.844

Como se ve, con media desviación estándar aparecen dos asociaciones; mientras que con una desviación estándar aparecen tres asociaciones; y con una-y-media no aparecen. En el caso de las guías de escuela nueva, aplicando una desviación estándar habría que decir que las guías coadyuvan al desempeño en lenguaje (lo que, leído literalmente, apuntalaría una política de reedición de guías); pero aplicando 1.5 o 0.5 desviaciones estándar no se muestra asociación, de donde sólo podría afirmarse que las guías son independientes del desempeño (y la política tendría que ser completamente opuesta).

Decimos que la decisión de con qué cifra se establecen los grupos es “arbitraria” en tanto no se obtiene de deducción frente al desempeño, frente a los factores asociables o frente a su posible asociación; es una decisión en otro rango de especificidad. Esto se puede ilustrar con los siguientes dos casos:

- recientemente circuló la noticia sobre el desacuerdo entre el DANE y la ANDI en relación con el crecimiento de la economía colombiana: mientras para el DANE la economía había decrecido en un 11.18%, para la ANDI había crecido en un 11.7%. ¡No sólo diferían en la cifra sino, fundamentalmente, en la dirección! El DANE se desdijo después, a favor de un crecimiento del 5.37%, argumentando que se había aplicado el deflactor (índice de precios) nuevo que entregó el Banco de la República, a la clasificación CIUU (cifra que entrega la ONU para hacer comparables los países) antigua: «Es decir, que se empezó a implementar una cifra para deflactar antes de haber ajustado del todo la muestra a la nueva guía internacional»<sup>120</sup>. O sea, ¡estamos hablando de cifras! La diferencia entre si la economía crece o no depende de si se aplica o no determinado número.
- En las islas británicas, entre 1979 y 1997, las autoridades han cambiado *treinta y dos veces* los criterios estadísticos para definir las cifras de desempleo. En este momento, el criterio es: «no está desempleado quien trabaja más de una hora por

<sup>120</sup> *El oso del DANE*. En: Diario El Tiempo, domingo 26 de Marzo de 2000, p.19A.

semana» [Galeano, 2000:177]. De manera que el porcentaje de personas empleadas no depende solamente del empleo, sino también de lo que se defina por “estar empleado”.

Ahora bien, la “arbitrariedad” va todavía más allá: el grupo que se clasifica en “alto”, “medio” y “bajo” son los 6447 niños que respondieron acertadamente entre 0 y 19 preguntas; ahora bien, antes de producir estos tres grupos, ya estarían configurados 20 grupos: los que no responden bien ninguna pregunta, los que responden bien una pregunta... hasta los que responden bien todas las preguntas. Este grupo “natural”, era de esperarse, debería tener una alta asociación con la clasificación estandarizada; no obstante, ¡no se asocian!: el p-valor es de 0.202. En cambio, sí se asocia con los análisis por competencia (tanto de la parte cerrada como de la parte abierta de la prueba de lenguaje): asociación altamente significativa: p-valor de 0.001:

		p-valor
<b>Escala “continua” vs...</b>	Estandarizada	0.202
	Competencias-C	0.001
	Competencias-A	0.001

Pese a que toda clasificación es arbitraria, estas asociaciones estarían diciendo que parecen hablar más del desempeño de los niños las clasificaciones por competencia que la categorización estandarizada. Ante la expectativa de encontrar asociaciones con los factores asociados, parecía más propicia la categorización estandarizada, con ayuda de la cual aparecían más correlaciones; sin embargo, la confrontación que venimos de hacer parecería permitir una idea distinta. Pero, ¿para qué continuar manteniendo esa expectativa —que podría ser eterna: infinitas investigaciones demostrando una vez que la manera más objetiva es la una, otras que es la otra—, dadas las características de los instrumentos?

## 7.2 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 de interpretación, es porque aceptamos el hecho de estar limita-

dos para describir la realidad educativa tal cual es. De tal manera, ninguna propuesta puede postularse como la que muestra 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 con 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 hasta qué punto se pueden hacer equivalencias?

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 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 tal vez 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” (o las autoridades educativas y los técnicos) 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 (excluidas del saber del sabio y del hacer del político). 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 en esa dirección 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, dinámi-

ca (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; que las formas de investigar no estén técnicamente condicionadas... Todo esto sólo puede ser un efecto, no es una condición del trabajo.

Todo el mundo está de acuerdo en que debe mejorarse la calidad de la educación; sin embargo, en medio de tanta diversidad de intereses, nos parece sospechoso el hecho de que haya semejante acuerdo. En la relación con la evaluación de la calidad de la educación hemos aprendido que donde no hay homogeneidad es fácil encontrar diferencias (claro, ¡porque es una tautología!). En lugar de eso, hemos hecho la presente investigación en la 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.

En el ámbito gubernamental se entiende la calidad —cuyo mejoramiento dice buscarse— como «el grado de cercanía entre el ideal humano de una sociedad dada y su expresión educativa. Más específicamente, se la considera como el grado de aproximación entre lo establecido entre los fines del sistema educativo nacional y el logro de la población estudiantil» [MEN, 1992:24]. Este tradicional concepto de “evaluación” (confrontar objetivos y resultados) está impregnado de ideas discutibles: que una sociedad dada, tiene *un* ideal humano, que deberíamos medirnos frente al ideal, que los fines del sistema rigen efectivamente las acciones educativas, que no hay objetivos implícitos, que el desempeño de los estudiantes es visible sin que las ideologías, las teorías y los instrumentos “interfieran”. Ideas como estas no se quieren discutir, pero las evaluaciones masivas se imponen...

Cuando se intenta establecer qué quiere decir “calidad de la educación”, resulta que es algo distinto, dependiendo de los propósitos de los que es portador quien responde. Entenderíamos un “mejoramiento” si se construyera 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.

## 8. A manera de conclusiones

¿Quién puede atreverse a soñar que ha cazado la imagen fugitiva de la verdad en toda su aterradora multiplicidad?\*

► *Hay una nueva manera de diseñar la política educativa cuyo lenguaje es internacional, pero que cada uno avala a su manera*

Durante la última década del siglo pasado se introduce un cambio en el diseño de la política educativa: ahora ésta se determina mencionando los estudios que relacionan desempeño estudiantil (obtenido en evaluaciones masivas) y factores asociables a ese desempeño; una manera presentada como “más técnica”, como más allá de caprichos o compromisos personales o politiqueros. Se supone que en dichos informes técnicos quien habla es la medida de la Ciencia, y no la palabra de alguien (la ciencia a prueba de sujeto, paradigma cartesiano que domina en la cultura occidental). Por ello, según los funcionarios, los propios agentes del proceso educativo no pueden pensar, investigar, decidir ni innovar por su propia cuenta; deben esperar a que los “expertos” apliquen sus instrumentos y les digan qué hacer (cosa que no funciona tal cual si la educación es lo que sus agentes del día a día son capaces de hacer, capacidad que ha sido producida en condiciones específicas y que sólo puede transformarse mediante la producción de otras condiciones).

Este lenguaje tiene un correlato —aunque no homogéneo— con el lenguaje internacional sobre educación, de manera que parece estar codificado en la misma jerga. Como frente a estos asuntos el país hace como si las iniciativas fueran propias, las fuentes y los compromisos, cuando se citan, se enuncian como verdades por fuera de discusión. Esto aparece mezclado con otras nociones (“globalización”, por ejemplo) que, pese a no explicarse, terminan afirmando que nuestra educación es de mala calidad porque no contribuye a desarrollar ciertas competencias —relacionadas con las expectativas de los empleadores—, lo que a su vez explicaría nuestro estado de desarrollo social. Con todo, estas políticas se abren paso en Colombia y ganan cierta particularidad gracias a un aval que todos los que tenemos que ver con el ámbito educativo damos con nuestras acciones (incluyendo las discursivas) desde nuestro ángulo, conscientemente o no, y de manera heterogénea.

---

\* Durrell, 1979:141.

- ▶ *Sin que sea específico del sector educativo, el enfoque de indicadores cuantitativos se implementa allí con ayuda de la “evaluación de la calidad”*

Los “nuevos” propósitos (calidad, equidad, eficiencia) encarnados en las recientes políticas educativas tienen cierta constancia, pese a la aparente movilidad legislativa y a los cambios de gobierno; el cambio estaría más bien en la composición entre ingerencia y evaluación, que parece obedecer a una proporcionalidad inversa: a menos ingerencia en el diseño curricular y la gestión institucional, más evaluaciones masivas (“rendición de cuentas”), y viceversa; evaluaciones que relacionan tautológicamente presupuesto y cifras de factores asociables. La especificidad de la educación frente al resto de servicios públicos —donde este enfoque de indicadores cuantitativos es anterior— no ha impedido que se lo aplique de forma indiscriminada, ni ha ameritado una demostración de que es aplicable allí, siendo que las condiciones de ese enfoque (“competencia perfecta”, transparencia de la información, regulación interna y estaticidad), no se cumplen en educación, donde más bien hay un horizonte temporal, unas incidencias externas, y una compleja articulación público-privado.

- ▶ *Como esta evaluación busca elevar la eficiencia interna, hace indagaciones estereotipadas cuyo primer efecto no es describir algo, sino autovalidarse*

Como se cruzan evaluación de desempeños con factores asociables para buscar la manera de elevar la “eficiencia interna” —dadas las exigencias de las entidades crediticias—, más que dar cuenta de una realidad social, preocupa una “consistencia” de cifras; juzgar la educación sólo por la inversión decreciente y la máxima producción, hace que los sujetos involucrados y la cultura a la cual pertenecen resulten poco relevantes. De ahí que no sea infrecuente que en Colombia se busquen las mismas asociaciones que se buscan en todo el planeta y que se encuentre lo que se ha encontrado en todos los países donde la asignación de préstamos se condiciona a la obtención de esos “hallazgos”. El efecto de esto es una falta de legitimidad de la evaluación, que se percibe en el hecho de que sus resultados son prácticamente predecibles y muchas veces no permiten saber más acerca de los desempeños o de la escuela, de lo que ya otro tipo de investigaciones ha encontrado.

A veces los instrumentos preguntan de una forma que sobre cierto aspecto del objeto en cuestión la información resulta más relevante; con lo que se procede de una manera en la que se termina “encontrando” lo que se presupone. También se cruzan las informaciones, pese a que las sustentaciones de los respectivos instrumentos sean excluyentes, pues las cifras se dejan operar con arreglo al conjunto de números de donde proceden. De tal manera, los factores que se busca asociar al desempeño de los estudiantes dependen de lo que es socialmente perceptible en una época dada. Es

decir, el sistema educativo pregunta sobre aquello que él mismo produce —como política educativa, como teoría de la educación, como legislación, etc.—, y/o concibe, presentándolo como lo único existente y posible. La búsqueda de factores asociables, entonces, tal vez sirva para legitimar ciertas concepciones que subyacen en los instrumentos y en los análisis, para legitimar la necesidad de investigar ciertos temas y de cierta manera que podemos esquematizar así: idioma “común”, cifras agregadas y comparables, puestas a disposición rápidamente, seleccionables por su pertinencia, fiables, enmarcadas en sistemas de información, que den cuenta de los costes, que midan los resultados del sistema educativo, y que privilegien una finalidad operativa.

► *Los exámenes de Estado y la evaluación de la calidad en realidad no difieren, en tanto ésta sólo requiere discriminar desempeños*

La idea de que el examen de Estado del ICFES no busca medir la calidad de la educación y que por ello no sirve para tomar decisiones en política educativa, no trasciende los equipos técnicos, pues se siguió usando para tal propósito. Con su actual inclusión en la jerga de las competencias, se muestra que el examen de Estado en realidad no tiene propósitos distintos de la evaluación de la calidad (con lo que las evaluaciones de la calidad se pueden tomar como exámenes de Estado) y que ésta adoptó en su momento una forma opuesta a la de aquél, pese a que en el fondo lo único que requiere producir es una *clasificación del desempeño, independientemente de la manera que se use para obtenerla*, con tal de tener una base diferenciada para contrastar con los factores asociables. Ahora, si bien el sistema pide que todos aprueben, no pide que lo hagan de la misma manera, pues la evaluación de los desempeños requiere *discriminar* o, de lo contrario, no habría cómo buscar asociaciones; de una prueba en la que todos ganan o todos pierden, se dice que no sirve. Para la psicometría, todo grupo humano es susceptible de ser repartido en categorías como alto, medio y bajo, y todo buen psicómetra debe ser capaz de diseñar el instrumento que lo pruebe. Finalmente, la diferencia entre “norma” y “criterio” no ha sido tan relevante.

► *Es problemático el supuesto de homogeneidad de la población frente a lo preguntado*

Una de las condiciones que la psicometría exige es que la población haya estado expuesta de forma homogénea a aquello sobre lo que se la va a evaluar (de lo contrario, la evaluación no es pertinente). Una enseñanza igual para todos neutralizaría cierta variabilidad, salvo el desempeño (y por eso se hacen pruebas de desempeño)



que depende de la persona (por eso las pruebas son individuales) o de cierta condición que se quiere estudiar (por eso se promedian puntajes por escuela, por región, por localidad). Otras perspectivas, en cambio, perciben lo educativo como heterogéneo; para ellas, lo social impide realizar dicha condición de la evaluación que autoriza a preguntar a todos lo mismo, y por eso las diferencias de desempeño que se encuentran suelen coincidir con las diferencias sociales. Por ejemplo, los problemas políticos del país —que no están distribuidos de forma homogénea— expulsan a unos niños y alteran el cronograma educativo de los que se quedan (disminuyen su cantidad de trabajo escolar frente a los que no tienen tales inconvenientes). Este tipo de evidencia bastaría para suspender las evaluaciones masivas hasta que los niños cumplieran la condición de homogeneidad, si no fuera porque ella no sólo es requerida por tales estudios, sino que es producida en cuanto los resultados se toman como explicativos de algo. La objeción teórica resulta entonces insuficiente; el asunto es si la intervención de otras perspectivas puede producir el ámbito educativo como heterogéneo.

►► *La evaluación masiva no detecta clasificaciones, sino que las produce*

Para evaluar masivamente, se decide de forma convencional construir clases en las que objetos diversos ahora serán equivalentes. Esto presupone que, con determinados fines, alguien decide que tales objetos son comparables y después los trata como equivalentes. En otras palabras, con la evaluación masiva se obtienen diferenciaciones y estandarizaciones de la población, una identificación de muchas personas sobre la base de los rasgos que el sistema de información hace aparecer como comunes. Por eso, los formularios no interpelan realmente al sujeto (aunque le pidan algunos datos personales), pues sólo se necesita de él una marca permanente mientras se “cruza” la información; para luego amalgamarlo en una clase. De donde es el interés el que decide cómo queda clasificada la población; una clasificación “dada en lo real” y “percibida” por el investigador no es más que un supuesto de algunas miradas. Así, con sus respuestas, los estudiantes hacen existir clasificaciones: no es que pertenecieran a ellas antes de la prueba y que ésta detectara tal pertenencia, sino que estarán ahí después de que los resultados comiencen a entretenerse en lo social. Los perfiles de la población escolar, obtenidos de tal manera, no constituyen entonces una mera descripción.

Entonces, cruzar desempeños estudiantiles y factores asociables exige realizar clasificaciones que no provienen de la realidad a secas, sino, de un lado, del saber asignado a ciertos roles (por ejemplo investigativos), y del poder asignado a ciertos cargos; es decir, la comunidad educativa requiere de alguna manera esas clasificaciones para representarse. Y, de otro lado, las clasificaciones provienen del ejercicio de ru-

tinias técnicas: crear, colapsar, o anular otras clasificaciones cuya relación con el asunto evaluado no tiene que ser teórica, pero sí debe aparecer como evidencia empírica.

► *La estadística no genera evidencia empírica, sólo verifica bajo sus propias condiciones*

La asociación o dependencia estadística está lejos de constituir una evidencia, pues pasa por decidir si se puede asegurar, con una baja incertidumbre —arbitrariamente acordada desde el ámbito de investigación—, que las diferencias observadas no son atribuibles a causas no identificables. Es decir, no prueba, sino que verifica bajo sus condiciones: cuando no se rechaza, bajo el amparo de unos datos particulares, una hipótesis de partida que niega la asociación, no quiere decir que se acepte la validez de dicha hipótesis (no negar no es forzosamente aceptar). También se requiere elegir cifras con las cuales un continuo de respuestas se separa en grupos. Esta decisión también es “arbitraria” (por ejemplo al categorizar por estandarización) cuando no se obtiene de una deducción frente al desempeño, frente a los factores asociables o frente a su posible asociación, sino en otro rango de especificidad.

No es que las herramientas teóricas (como la estadística de Ji-cuadrado que arroja los p-valores) no se pongan a prueba permanentemente, sino que requieren decisiones que también son portadoras de intereses y de incertidumbres propias del hecho de que estamos imposibilitados de ver la realidad y juzgar si un modelo la describe como es (si lo contrario fuera posible, sobrarían las teorías).

► *Como las cifras que deciden una asociación dependen del campo de trabajo, se requiere una conceptualización propia para relacionar educación y cifras obtenidas al evaluar*

Un nivel de incertidumbre considerado “alto” en ciertos contextos, en otros sirve para animar una búsqueda (“una luz en donde antes sólo había oscuridad”); y un nivel de incertidumbre considerado “bajo” en ciertos contextos, en otros no obliga a abandonar el escepticismo encerrado en esa pequeña cifra (“una oscuridad en donde parece haber luz”). Pero, si la elección de las abstracciones estadísticas es arbitraria, por lo menos debe esforzarse por ser fundada: el problema no es que sean abstracciones, sino que sean malas abstracciones. Con el agravante de que una mala abstracción estadística no necesariamente salta a la vista (pues su “irrealidad” no se manifiesta de manera material), pero sí puede disimular rasgos esenciales del elemento estudiado, sobre todo cuando es heterogéneo o variable en el tiempo.

De tal manera, a partir de los estudios estadísticos en educación, donde los factores no considerados (que podrían ser los más importantes) a veces se encargan de introducir suficiente “ruido” como para que no se pueda concluir, o donde los factores considerados se hacen impertinentes, a) una falta de asociación numérica no desecha decisiones de política educativa, b) una asociación negativa no elimina una hipótesis de correlación positiva y c) una asociación positiva no valida una política en la dirección encontrada entre las cifras... es una conceptualización propia la que permite tomar la decisión (así se diga que fueron las cifras).

Cuando las relaciones entre cifras se hacen equivalentes a relaciones causa-efecto en el objeto que se pretende describir, las asociaciones pueden tomar de sorpresa al analista de datos, el cual se apresura a recomendar más investigación; claro: no tiene presente la concepción de escuela que maneja, y/o no hay vinculación conceptual entre los equipos (son “multidisciplinarios”, no interdisciplinarios), tal vez en el entendido de que lo producido por el método científico está hecho a prueba de sujetos y de interpretación. Pero, además de una recomendación de investigar, el esfuerzo de la evaluación masiva justificaría unas líneas de pertinencia para los estudios recomendados, así como una discusión académica, pública, sobre los marcos teóricos que están a la base de dichas evaluaciones. Las cifras pueden ser advertencias para cotejar con investigaciones más puntuales, siempre y cuando provengan de una demanda explícita de la teoría y no de un lance de dados.

► *La incertidumbre no se puede eliminar*

A veces se observa que hay un grado de incertidumbre que no se puede eliminar: a) cuando un factor “indeseable” está en “cero”, no aparece el mejor desempeño; y en el mejor desempeño, aparece cierta frecuencia de dicho factor... una especie de principio de incertidumbre heisenbergiano; b) los desgloses (por curso, modalidad de escuela, DC) no eliminan el ruido a favor de una variable: entre el análisis global y el desglosado puede haber más, menos o idéntica incertidumbre, y puede cambiar o mantenerse la dirección de las asociaciones; c) las aproximaciones de los instrumentos pueden ser difusas, en tanto distan de los requerimientos de los objetos respectivos, pues se ven exigidos desde obligaciones, presiones y condicionamientos distintos a la naturaleza de dichos objetos; d) los instrumentos también arrojan resultados contradictorios y/o absurdos; e) las dependencias estadísticas pueden provenir de terceros factores no considerados o, incluso, ser fortuitas.

Así, en estos estudios se pueden estar cruzando medidas realizadas sobre objetos que pertenecen, al menos parcialmente, a clases diferentes (producidas teóricamente de

manera distinta), es decir, hasta cierto punto incompatibles. Como en el siguiente caso: la prueba de desempeño en el área de lenguaje no fue diseñada por el mismo equipo que hizo los instrumentos de factores asociables; esto puede conducir a una imposibilidad de contingencia, toda vez que si el diseño de los instrumentos de factores asociables nos hubiera correspondido, habría preguntas distintas (acordes con una concepción sobre la relación entre los factores y la competencia comunicativa, objeto de evaluación); de igual manera, si quienes diseñaron los instrumentos en mención hubieran hecho la prueba de lenguaje, posiblemente también habrían preguntado de manera distinta. Los “hallazgos”, en ambos casos, habrían sido otros.

► *En lugar de tomar la imposibilidad de eliminar el ruido como rasgo del conocimiento, quienes interpretan la evaluación proceden de manera asistemática*

El conocimiento no se produce yendo hacia lo que pueda descomponerse cada vez más; por ese camino está lo imposible de aprehender: el corpus tiende a infinito (la búsqueda del “ladrillo” más pequeño del universo llevó más bien a la proliferación de partículas “elementales” —ya van más de 200— y a una paradoja en la que es imposible saber cuál compone a cuál<sup>121</sup>). De ahí que la exactitud de los estudios no sea directamente proporcional a la cantidad de decimales que se le ponen a las cifras. Pero tampoco el conocimiento se produce yendo hacia lo más general; por ese camino está lo indiscernible.

En lugar de girar hacia tal principio epistemológico, muchos estudios violentan la especificidad de la estadística: usos indebidos de sus procedimientos, selección arbitraria de sus resultados, y asignación de causalidades del objeto indagado a partir de correlaciones de cifras (lo cual no pocas veces lleva a absurdos que no interrogan los estudios, sino que se omiten o se ponen al lado de los otros como curiosidades). Esta escogencia de las cifras estadísticas hasta obtener los factores que obedezcan a un interés preestablecido (por ejemplo, descartando las asociaciones de signos contrarios a lo esperado, o seleccionando en términos de su relación con las políticas educativas) es proceder de forma asistemática, pues si todos los datos se producen con igual método, todos serían escogibles o rechazables. Así, cuando las distintas administraciones requieren decir qué está efectivamente asociado, aparecen y desaparecen factores, pese a que esas decisiones dicen basarse en los mismos estudios o en estudios que, por principio, tendrían que ser equivalentes.

Con esto, se muestra que el diseño de la política educativa no necesita las cifras, pero ha puesto sus criterios, limitaciones o requerimientos técnicos por encima de las

---

<sup>121</sup> Cf. Pajares, 1975:119: «¿son elementales las partículas elementales?»; Capra, 1996:104: «cada partícula está compuesta de todas las demás partículas».

condiciones propiamente educativas de los estudios. O sea, la estadística se ha usado para justificar cierta política educativa y para desdeñar o subordinar otras investigaciones sobre los objetos de evaluación en juego, sin dar cuenta: de los estudios en los que dice apoyarse, de la responsabilidad de los instrumentos en los hallazgos, ni del papel que juega la interpretación de las cifras.

►► *Las asociaciones dependen de la perspectiva que se asuma*

La asociación entre factores y desempeño no depende de las propiedades respectivas de dichos objetos, sino de la perspectiva que se asuma, lo que coincide con un principio epistemológico según el cual el punto de vista crea el objeto. Por eso, distintas elecciones frente a los mismos datos dan resultados distintos: a) la misma prueba, interpretada con categorizaciones diferentes, arroja resultados distintos; en este estudio, por ejemplo, los enfoques del desempeño en el área de lenguaje (“por estandarización” y “por competencias”) a veces muestran resultados opuestos y, sin embargo, pueden sustentar su propia validez; b) distintas pruebas (o interpretaciones) promueven a la existencia distintos objetos, perfiles distintos de estudiante (a veces resultan incompatibles incluso los objetos de evaluación entre pruebas de la misma área, cuando son hechas por equipos distintos); c) las mismas cifras, vistas desde horizontes distintos, verán evoluciones de otra naturaleza y otras direcciones podrían asignar a las asociaciones (en el presente estudio, por ejemplo, se decidió mirar los extremos del desempeño, pero quien piense que allí no está representada una proporción importante de la población, podría preocuparse por el medio; o, considerando que nuestra escuela no tiene las condiciones para el nivel más alto, se podría centrar en los primeros niveles; o la atención podría estar en relación con los mejores desempeños, en tanto indicadores de horizontes pedagógicos...).

►► *Hay que reflexionar para medir y no medir para reflexionar*

Ni la consistencia ni la exhaustividad vienen de los números. Nuestro estudio permite afirmar que si uno va armado de ciertas preguntas, el trabajo estadístico por sí solo no resuelve la complejidad del asunto educativo. Si las demandas que se le hacen tienen origen en preocupaciones de otra especificidad, la estadística puede ser un interlocutor que ayuda a interrogar sobre la validez de la teoría, sobre la confiabilidad de los instrumentos o sobre la pertinencia del estudio. Cuando los números pretenden sobrepasar en exactitud a los medios de investigación, quedamos subordinados a ellos, a sus leyes; pero si el objeto se considera un grado particular de la aproximación del método de medida, ya resulta ingenuo pensar que dicho método es independiente de la descripción del objeto mismo. De tal manera, la precisión de los

instrumentos puede aducirse para introducir el prejuicio social de no interrogar sobre dicha exactitud y, con ello, permitir que ciertos intereses se realicen. Si la serie números y las series desempeño-factores no son isomórficas, la “medición” de la calidad no sería legítima.

De ahí que en el presente estudio, lo encontrado no se plantea como único, permanente o de lectura unidimensional. Los datos objeto de nuestra mirada representan un levantamiento de información inscrito en un espacio, un tiempo y al menos una postura. Otra muestra, otros datos, podrían mostrar asociación donde antes no se evidenciaba, y recíprocamente. Las conclusiones están subordinadas a los datos disponibles, obtenidos con los instrumentos que se utilizaron y con ayuda de la interpretación hecha. Entonces, los factores que se dicen “asociados” lo son desde ciertos puntos de vista, y luego de ciertas decisiones técnicas que la estadística toma.

► *Así no haya habido sistema de evaluación, se ha procedido como si hubiera*

Lo más evidente es que si fuera *sistema* articularía de manera compleja muchos factores y estamentos; de manera que una indagación por fuera de su articulación con otros factores, no habla del sistema, del conjunto. Evaluar partes del sistema educativo (como en este caso), no conduce a tener sistema el día de mañana como resultado de sumar las partes. Lo menos evidente —pero tal vez más interesante— es que nos movemos entre dos extremos virtuales:

- un horizonte de certeza, en el que las pruebas permanecen, y las cifras aparecen como comparables (es decir, *sistema*), y entonces se habla de la educación y de los niños, y se toman medidas cuya arrogancia es la medida de la expectativa que tienen los miembros de una comunidad educativa considerada como homogénea. De ahí que los “hallazgos” —o lo que ocupa su lugar— se puedan tomar como totalizantes y concluyentes, incluso cuando se abandonan para tomar nuevos resultados o nuevas posiciones como totalizantes y concluyentes.
- Y, en el otro extremo, un horizonte de incertidumbre, en el que los mecanismos de indagación de lo que el niño sabe se transforman de manera permanente, y las cifras se hacen hasta cierto punto incomparables (es decir, *investigación*), y entonces se habla del aprendizaje que ha ido teniendo lugar, de la distancia, tanto con los procesos cognitivos de los niños, como con los fenómenos educativos, y se toman medidas cuya modestia es la medida de la capacidad de una comunidad de trabajo sin fisuras para explicitar las condiciones de posibilidad, que hacen historia (evita empezar cada vez de cero, hace uso del conocimiento producido

por los expertos y por el sistema de evaluación), y para reconocerse como limitada...

Extremos inexistentes, claro está, pero que podrían postularse como las fronteras del ámbito de las transformaciones en el que se juegan los intereses de todos los que se relacionan con la educación.

Por eso, cuando se trata de fundamentar una política en relación con un factor cualquiera, una discusión podría ser qué tanto está o no asociado estadísticamente al desempeño, y otra podría ser sobre la o las teorías que más complejidades y riquezas permitan ver para la educación. Es claro que ninguna de las dos funciona tal cual; ciertas condiciones podrían hacer que las teorías que requieren los análisis estadísticos queden subordinadas a éstos (como parece ser hoy), y cómo otras condiciones podrían producir el efecto contrario. De ahí que para controlar ciertos factores no haya sido necesario esperar las asociaciones estadísticas ni el final de las investigaciones: los agentes educativos siempre tienen una consigna para cada situación, como indicador de gasto, como criterio pedagógico, como proyecto ético... pero estas perspectivas poco dialogan.

► *Aunque las respuestas a los cuestionarios provengan de la pragmática de evaluación, se pueden validar instrumentos y aplicaciones*

Si se puede interrogar por la naturaleza de los instrumentos con los que se obtiene la información, aparece el asunto de la pragmática de la evaluación: hasta qué punto el encuestado responde condicionado por las preguntas, por su expectativa sobre lo que quiere el encuestador, tratando de quedar bien, escogiendo lo que socialmente es más aceptable, negándose a hacer existir una política, etc.; y si tal actitud está determinada culturalmente, se explica que haya tendencias, que se puedan obtener cifras manejables y que se hallen correlaciones. Ahora bien, como esto no es eliminable, no se trata sencillamente de mejorar paulatinamente los cuestionarios, hasta obtener instrumentos “perfectos”, sino de entender la investigación como algo comprometido con tendencias de su momento histórico (y tanto más relevante cuando permite explicitarlo). Aunque hay esfuerzos por “pulir” los instrumentos, el “ruido” no es eliminable, pues no retrocede con dichos esfuerzos sino que se introduce con cada decisión metodológica, con cada decisión teórica, con cada decisión práctica.

El efecto “perverso” de esto es que los instrumentos terminan induciendo las respuestas y entonces lo que “miden” (psico-metría) no es lo psico-, sino sino la pragmática de la evaluación, el efecto de su presencia (y de sus implicaciones) en el campo educativo.

► *Tratar de conocer la educación introduce nuevas “cosas” en el campo educativo y contribuye a producirlo y delimitarlo*

Los esfuerzos por conocer la educación introducen nuevas “cosas” en su objeto de estudio y, por lo tanto, el esfuerzo por conocer ya forma parte también del objeto; en cuyo caso, no es que la investigación educativa sencillamente describa la educación, sino que también ha transformado la educación y ha producido formas de relacionarse con ella (por ejemplo, después de la supuesta detección del conocimiento de los aspirantes a la universidad mediante el examen de Estado, la educación no volvió a ser la misma en Colombia; así mismo, después de estudiar a Piaget, un maestro no sólo pretende entender de una manera particular el proceso de conocimiento en el niño, sino que actúa y vive su trabajo de manera distinta).

Los esfuerzos por conocer la educación, además, forman parte de esa red social de límites difusos: las fronteras y las relaciones que se le atribuyen a la educación con otras prácticas sociales han cambiado en el curso de la historia; y en este cambio todos los factores (incluyendo la investigación) se relacionan, por supuesto de manera diversa. La investigación sobre la educación no ha agregado simplemente objetos y posiciones a la educación, sino que principalmente ha intervenido en el proceso mismo de hacerla existir cada vez de manera particular. De modo que se impone la pregunta de la manera como la evaluación masiva y la investigación hacen hoy existir a la educación colombiana; sin olvidar el papel que cumple la tendencia oficial hacia la estadística, que parece desdeñar otro tipo de investigaciones. *Que la escuela no se detecta sino que se produce*, no es una consigna alternativa; la actual manera de hacer política educativa lo lleva a cabo: con la idea de indagar sobre los factores asociados, *produce* los factores con sus intervenciones, privilegia la búsqueda en función de las “decisiones a adoptar”, busca cierto “consenso social” (es decir, no exacerbar los ánimos).

► *Los factores “asociados” se han buscado asistemáticamente, pero podrían buscarse desde una concepción compleja de escuela*

Entre los factores que concurren en la educación, hay unas relaciones; pero no se puede decir —con el prurito de que “todo se asocia con todo”— que basta con tener listas de factores. De ahí que el número de factores sea muy variable de un estudio a otro; que haya un ajuste de los instrumentos a los recursos económicos disponibles, a los requerimientos técnicos que implica su “medición”, y a la política educativa previamente definida; que la importancia en términos teóricos, pedagógicos y edu-



cativos quede relegada (de donde la educación cede su especificidad a aquella de la estadística); que falte jerarquía, relaciones, entre los factores indagados; que en ocasiones entre los factores listados haya algunos excluyentes entre sí o ya incluidos dentro de otros; que se copien otras indagaciones; que se excluyan ciertos factores que podrían ser fundamentales, etc.

Al tratarse de factores que determinan el logro, se espera una jerarquización entre ellos en relación con el efecto que producen, y no solamente un listado con valores estadísticos, pues no todas las variables características del fenómeno están en interacción, ni éste es sensibilizado de igual manera por todas ellas: cada investigación es un caso especial. Por eso, si se quiere sacar la investigación sobre la asociación entre factores y desempeño de sus propias suposiciones, es necesario buscar una estructura abstracta, jerárquica, compleja (con la intervención irreductible del azar) que relacione lo que tiene que ver, al menos, con las prácticas pedagógicas (no se trata de echar culpas al maestro), con el factor en cuestión y con los procesos de aprendizaje. Ahora bien, ningún estudio sobre contingencias tendría sentido para la educación si no se inscribe en los contextos específicos, pues éstos inciden (de una manera por establecer) sobre el factor, la pedagogía y el aprendizaje, los cuales, a su vez, determinan dichos contextos (complejidad de retroacción y autopoiesis).

►► *La búsqueda de la participación como factor asociable —como se interrogó en la evaluación de impacto del PUEBP—, pese a tener inconvenientes, permitió exploraciones interesantes*

De un lado, las especificaciones de escuela nueva y de la reciente legislación educativa presuponen una participación de los alumnos que puede ser mero formalismo (en tal sentido iría el hecho de que puedan cambiarse muy rápidamente), mientras la participación se produce en el seno de condiciones de la cultura escolar que trascienden la norma y los objetivos. Y, de otro lado, mientras las comunidades educativas entienden la participación en la dirección de afrontar sus condiciones específicas de violencia, pobreza, etc., en la política educativa la participación pasa de no ser necesaria a ser nominal, y luego a ser condición de aplicación de la nueva política, hasta quedar reducida a una gestión empresarial. Por su parte, las pruebas poco recogen la expectativa —tanto oficial como de la comunidad educativa—, reducen la indagación a la pertenencia a comités y al involucramiento en ciertas actividades escolares, no exploran los condicionamientos comunicativos de la participación, usan una clasificación que mezcla distintos tipos de actividades, que contiene opciones no excluyentes, etc.

- ▶ *La búsqueda de los libros como factor asociable —como se interrogó en la evaluación de impacto del PUEBP—, pese a tener inconvenientes, permitió exploraciones interesantes*

La búsqueda de los *libros* como factor asociable, tal como fue hecha en la evaluación de impacto del PUEBP, no obedece a la inserción de los actos de lectura y de escritura en una perspectiva pedagógica. Además, materializa una manera de entender los tipos de texto pensada desde una tradición para la que los libros deliberadamente instruccionales son más importantes (como si para su elaboración no se requirieran libros “auténticos”) y desde una perspectiva que se agota en la dotación. La diversidad de textos, los actos de demanda de lectura, los efectos dialógicos de los actos de lectura, la organización de la escuela en función del material impreso, la existencia de proyectos explícitos para la utilización de los libros... son aspectos que se dejaron por fuera, siendo que el objeto de evaluación del área de lenguaje era la competencia comunicativa.

- ▶ *La búsqueda de la eficiencia interna como factor asociable —como se interrogó en la evaluación de impacto del PUEBP—, pese a tener inconvenientes, permitió exploraciones interesantes*

Según la idea de “eficiencia interna”, como los factores extraeducativos que pueden incidir en la calidad de la educación no pueden cambiarse (ya cayó en desuso el discurso de que la educación era un factor de ascenso social), sólo quedan las variables de la escuela, que explican parte del desempeño. Por ejemplo, sin haber desechado teóricamente la opción de que el desempeño dependa fundamentalmente de la producción del conocimiento por parte del sujeto, se lo evalúa como si tuviera que ver principalmente con los contextos escolar, comunitario y cultural. Entonces, se busca ampliar la cobertura educativa, retener a los estudiantes desde la educación, promover medidas sobre evaluación escolar que eviten la pérdida de años, ampliar el número de alumnos por docente, etc. De tal manera, la búsqueda de la *eficiencia interna* como factor asociable ya viene marcada por este interés que renuncia a atacar la inequidad del sistema social, que sustenta las medidas sobre evaluación escolar en consideraciones pedagógicas, que desconoce los problemas sociales que obligan a un niño a abandonar la escuela (expulsión material o simbólica: la decisión del sujeto es el efecto de una relación), que no discute razones en relación con los efectos de alterar el número de alumnos por docente. No obstante, se dan correlaciones, pero de manera tal que en algunos casos las hipótesis encuentran corroboración con un análisis del desempeño, pero objeción cuando se aplica otro análisis del desempeño.

- *La búsqueda de las disciplinas como factor asociable —como se interrogó en la evaluación de impacto del PUEBP—, pese a tener inconvenientes, permitió exploraciones interesantes*

La “importancia” de las disciplinas sobre las que se evalúa el desempeño y de los factores con los que éstas se cruzan, solamente se supone, no se demuestra ni se fundamenta en una teoría explícita, por ejemplo, de la formación del sujeto. Por ello, cuando se trata de asociar disciplinas entre sí, no se relacionan las especificidades de dos disciplinas cada vez, sino las especificidades de las pruebas respectivas (las cuales pueden ser portadoras de obstáculos epistemológicos). En todo caso, las prácticas pedagógicas de las disciplinas no parecen ser homogéneas. Como es muy arriesgado afirmar que en al menos una de las áreas haya comunidades académicas de profesores —lo que autorizaría la hipótesis de cierto control teórico de las prácticas respectivas—, entonces, menos podemos hipotetizar la existencia de prácticas pedagógicas interdisciplinarias controladas. Todo esto parece estar de acuerdo con las cifras de la relación entre el análisis por competencias del desempeño y disciplinas. No obstante, el análisis estandarizado parece percibir con regularidad una relación entre disciplinas y lenguaje, más allá de las especificidades de los planes de estudio o de las prácticas de enseñanza.

- *«La elevación cuantitativa de los factores de “progreso” aumenta la intensidad cualitativa de las “regresiones”»<sup>122</sup>*

Lo humano se juega siempre con la misma fuerza, sin tener en cuenta que se empuñe un garrote o se atienda un terminal de computador; aquello que nos diferencia de cualquier otra entidad parece tener igual vigencia frente a todas las “conquistas” logradas gracias a la ciencia. Esto podría indicar que la ciencia roza la especificidad humana, pero no la agota: hay algo “no racional” presto a irrumpir en cada instante.

Según Debray [1996:55] no hay un camino de una sola vía entre la industrialización y el ateísmo, entre el mito y el saber, entre lo sagrado y lo profano; no hay una producción de lo profano mediante la progresiva expulsión de lo sagrado, pues «los territorios culturales que borra el progreso técnico se recomponen en el imaginario» [:60]. Sagrado/profano, irracional/racional, arcaico/moderno, local/mundial —factores llamados de regresión y de progreso, respectivamente—, serían figuras ideológicas opuestas variando en una relación constante: «Cada desequilibrio suscitado por un progreso técnico provoca un reequilibrio étnico» [:54], lo que explicaría los actuales entrecruzamientos entre homogenización y reivindicación de las diferencias,

---

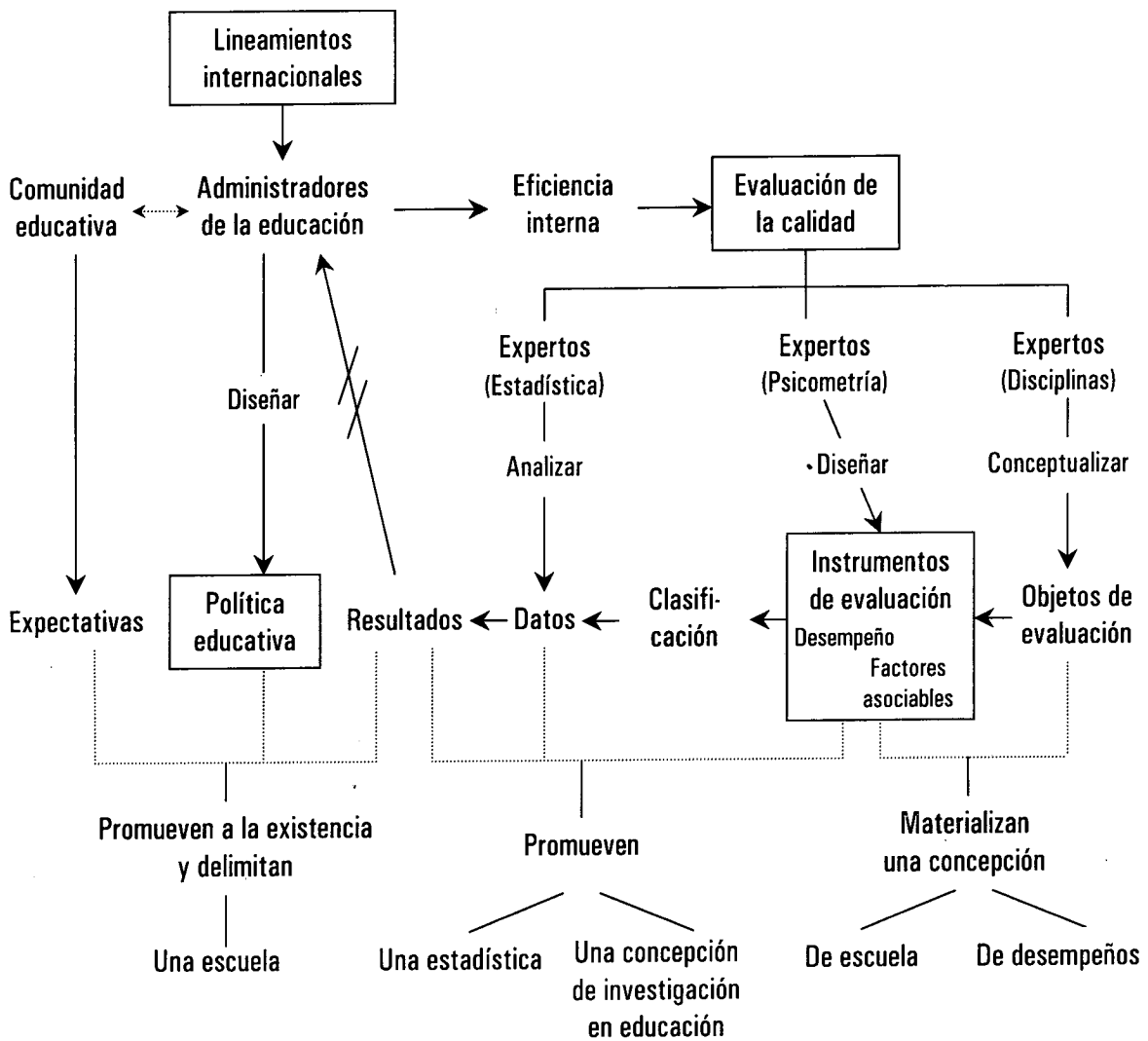
<sup>122</sup> Debray, 1996:66.

entre lo intelectual y lo afectivo, entre el imperativo económico y la necesidad religiosa [:54]. En conclusión, «cada dispositivo de desarraigo y apertura libera un mecanismo de contra-arraigo o de cierre [:51] Las ganancias en racionalidad económica parecen pagarse con pérdidas equivalentes en el orden político [:62]».

Valdría la pena, entonces, explorar si esta idea de Debray caracteriza a la evaluación masiva: de un lado, “progresos” en la medición, avances en los modelos estadísticos, amplia posibilidad de matematización, generalización de los sistemas de evaluación, ampliación de las poblaciones evaluadas (recordemos que en el DC ya se trata de un *censo*)... la pregunta sería cuáles son los ámbitos de regresión que esto genera (o si la proporción exige invocar niveles más complejos de lo social); podrían indagarse varios (por ejemplo en el orden de la respuesta tribal a la mundialización de los objetos [:67]), pero sería interesante si al menos pudiéramos relacionar tales “progresos” con una “regresión” de otro tipo de investigación educativa.

Es algo que quisiéramos seguir pensando.

El siguiente diagrama esquematiza la manera como se organizan las conclusiones:



## Bibliografía

- ABAD, Héctor. «Dulzuras y amarguras del devorador de libros». En: Revista *Universidad del Valle* N° 16. Cali: UNIVALLE, Abril de 1997.
- AGRESTI, Alan. *Introduction to categorical data analysis*. New York: John Wiley, 1996.
- ALCALDÍA MAYOR DE SANTA FE DE BOGOTÁ, Secretaría de Educación. *Evaluación de competencias básicas en lenguaje y matemáticas. Primera aplicación, Octubre-Noviembre de 1998. Resultados*. Santa fe de Bogotá: Febrero de 1999.
- ALLEN, Woody. *Sin plumas*. Barcelona: Tusquets, 1980.
- ANDRADE, Gerardo. «Niveles de conceptualización en los procesos evaluativos». En: Revista *Edu.co*, N°1. Bogotá: Enero-Marzo de 1998.
- ATO, Manuel y LÓPEZ, Juan José. *Análisis estadístico para datos categóricos*. Madrid: Síntesis, 1996.
- 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.
- \_\_\_\_\_. *Epistemología* [textos escogidos por Dominique Lecourt]. Barcelona: Anagrama, 1973.
- BAENA, Luis Ángel. «Lingüística y significación». En: Revista *Lenguaje* No. 6. Cali: Univalle, 1976.
- \_\_\_\_\_ y BUSTAMANTE, Guillermo. «El área de lenguaje y el rendimiento escolar» En: Revista *La palabra* N°1. Tunja: UPTC, 1992.
- BANCO MUNDIAL. «Staff appraisal». Bogotá: MEN (s/p), 1988.
- \_\_\_\_\_ - Instituto de desarrollo económico - División de recursos humanos, EDIHR, Fondo de las Naciones Unidas para la Infancia, UNICEF - Oficina regional para América Latina y el Caribe. *Hacia una nueva escuela: guías de formación docente en estrategias para el mejoramiento de la educación básica primaria y para el aprendizaje personalizado y grupal*. Bogotá: UNICEF, 1996.
- BENEDITO, Gloria. «El problema de la medida en psicología». En: *Psicología: ideología y ciencia*. México: Siglo XXI, 1975.

- \_\_\_\_\_. «El método de los tests e inventarios». En: *Psicología: ideología y ciencia*. México: Siglo XXI, 1975.
- BLOCH, Arthur. *Ley de Murphy II. Más razones por las que las cosas salen mal*. México: Diana, 1981.
- BODNAR, Yolanda. «La constitución y la etnoeducación, ¿una paradoja?». En revista *Educación y cultura* N° 27. Bogotá, 1992.
- BULA E., Germán A. «Carta». En: Revista *Alegría de enseñar* N° 38. Cali: FES, Enero-Abril de 1999.
- BUSTAMANTE, Guillermo. «¿Es posible evaluar “objetivamente”?». En: *Evaluación escolar, ¿resultados o procesos?* Bogotá: Magisterio, 1996.
- \_\_\_\_\_. «Editorial». En: Revista *Pedagogía y saberes* N° 11. Bogotá: Universidad Pedagógica Nacional, 1998.
- \_\_\_\_\_ y JURADO, Fabio. «La evaluación oficial del área de lenguaje». En: Revista *Opciones pedagógicas* N° 10. Bogotá: Universidad Distrital, 1993.
- CABRERA, Blas. «Evaluación del sistema educativo: ¿para qué?». En: Revista *Archipiélago* No 38. Barcelona, Otoño de 1999.
- CAICEDO, Lilian y BUSTAMANTE, Guillermo. «Construcción social de la evaluación escolar». Investigación financiada por el IDEP. Bogotá: s/p, 1998.
- CAPRA, Fritjof. *El punto crucial*. Buenos Aires: Estaciones, 1996.
- CARRETERO, Mario. *Constructivismo y educación*. Buenos Aires: Aique, 1993.
- CARRILLO, Ana Cristina. «Una escuela más humana para evitar la deserción escolar». En: *Aula Urbana* N° 18. Bogotá: IDEP, Febrero de 2000.
- CERLALC. «Proyecto de renovación de las prácticas pedagógicas en la formación de lectores y escritores». En: Revista *La Palabra* N°3. Tunja: UPTC, septiembre de 1994.
- COLBERT, Vicky y MOGOLLÓN, Oscar. *Hacia la Escuela Nueva*. Bogotá: MEN, 1980.
- \_\_\_\_\_. *Escuelas demostrativas y microcentros rurales. Programa de Escuela Nueva*. Bogotá: MEN-UNICEF, 1993.
- \_\_\_\_\_. *Manual complementario de las guías para los niños. Programa de Escuela Nueva*. Bogotá: MEN, 1981.

- DAZA, Ricardo y JIMÉNEZ, David. *El texto escolar en Colombia*. Bogotá: CERLALC, 1977.
- DEBRAY, Régis. *El arcaísmo posmoderno*. Buenos Aires: Manantial, 1996.
- 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 1991-1994*. Bogotá: DNP, 1991.
- \_\_\_\_\_. *El salto educativo. La educación, eje del desarrollo del país*. Bogotá: DNP-MEN, 1994.
- \_\_\_\_\_. *Un plan educativo para la paz*. Bogotá: UPN, 1999.
- DESROSIÈRES, Alain. «¿Cómo fabricar cosas que se sostienen entre sí? Las ciencias sociales, la estadística y el Estado». En: *Archipiélago* No 20. Barcelona, 1995.
- DÍAZ, Carlos. «Diseño, metodología y resultados del tercer estudio internacional de matemáticas y ciencias -TIMSS- en Colombia». En: *Revista Edu.co*, N°1. Bogotá: Enero-Marzo de 1998.
- DÍAZ, Luis Guillermo. *Estadística multivariada: análisis y métodos*. Bogotá: Universidad Nacional, 1999a.
- \_\_\_\_\_. *Análisis de datos categóricos*. Bogotá: XVI Coloquio distrital de matemáticas y estadística, 1999b.
- DUARTE, Patricia et al. «Posición sobre los componentes del sistema de evaluación, a propósito del área de lenguaje». En: *Revista Enunciación* Vol. I, N°1. Bogotá: Universidad Distrital «Francisco José de Caldas», Julio de 1996.
- DURRELL, Lawrence. *El cuarteto de Alejandría*. Clea. Barcelona: Edhasa, 1979.
- EVERITT, B.S. *The analysis of contingency tables*. London: Chapman & Hall, 1994.
- FECODE. «La evaluación de docentes. Debate educativo». En: revista *Educación y cultura* N° 52. Bogotá: FECODE, noviembre de 1999.
- FLÓREZ, María Victoria. «Algunos procederes prosociales de niños entre 4 y 8 años: su relación con la lectura de imágenes en movimiento» (Anteproyecto de investigación). Bogotá: Universidad Pedagógica Nacional s/p, 1997.



- VON FOERSTER, Heinz. «Visión y conocimiento: disfunciones de segundo orden». En: Fried, Dora (comp.). *Nuevos paradigmas, cultura y subjetividad*. Buenos Aires: Paidós, 1994.
- FONDO DE CONFINANCIACIÓN PARA LA INVERSIÓN SOCIAL [FIS]. *Guía para presentación de proyectos de dotación de textos. Para establecimientos oficiales de educación básica secundaria y media vocacional*. Bogotá: MEN, 1994.
- FORRESTER, John. *Seduciones del psicoanálisis: Freud, Lacan y Derrida*. México: FCE, 1995.
- FUNDACIÓN SOCIAL. *Megatendencias Siglo XXI*. Bogotá: Fundación social, 1998.
- GALEANO, Eduardo. *Patas arriba. La escuela del mundo al revés*. Bogotá: Tercer Mundo, 2000.
- GARCÍA, Gloria y ACEVEDO, Myriam. «Evaluación de las competencias en matemáticas y el currículum: un problema de coherencia y consistencia». En: *Hacia una cultura de la evaluación para el siglo XXI*. Bogotá: Universidad Nacional de Colombia, 1999.
- GOLDSTEIN, Harvey. *Multilevel models in educational and social research*. New York: Oxford University Press, 1987.
- HEDERICH, Christian y CHARRIA, María Elvira. «Impacto del proyecto “Podemos leer y escribir” en el desarrollo de la cultura escrita en las escuelas latinoamericanas». En: *Revista Folios* N° 10, Bogotá: Universidad Pedagógica Nacional, segundo semestre de 1999.
- HEYNEMAN, Stephen et al. «Libros de texto y aprovechamiento: lo que sabemos». En: Laptapí, Pablo (coord.). *Educación y escuela. I La educación formal*. México: SEP-Nueva Imagen, 1991.
- ICFES. «Cruce entre la variable logro y los factores asociables correspondientes a plantel-rector, docentes y estudiantes». Bogotá: ICFES s/p, 1996.
- \_\_\_\_\_. «Información sobre estudiantes». Formulario de factores asociados. ICFES, 1992.
- JURADO, Fabio. «Lectura, incertidumbre, escritura». En: *Los procesos de la lectura*. Bogotá: Magisterio, 1996.
- \_\_\_\_\_ y BUSTAMANTE, Guillermo. «Por una diversidad de textos en el aula de clase». En: *Revista Opciones pedagógicas* N°12. Bogotá: Universidad Distrital, 1994.
- \_\_\_\_\_ ; \_\_\_\_\_ y PÉREZ, Mauricio. *Juguemos a interpretar*. Bogotá: Plaza&Janés, 1998.

- LE, Chat. *Applied categorical data analysis*. New York: John Wiley & sons inc., 1998.
- MARX, Carlos. *Manuscritos: economía y filosofía*. Madrid: Alianza, 1989.
- MICHEL, Alain. «La conducción de un sistema complejo: la educación nacional». En: *Revista iberoamericana de educación* N° 10. Madrid: OEI, Enero-Abril de 1996.
- MINISTERIO DE EDUCACIÓN NACIONAL. *La promoción automática en educación básica primaria. Acción nacional educativo-cultural*. Bogotá: MEN, 1987.
- \_\_\_\_\_. «Programa para la Universalización de la Educación Básica Primaria». Bogotá: MEN (s/p), 1988.
- \_\_\_\_\_. «Ciclo de Foros sobre el Plan de Apertura Educativa. Tema: La evaluación de la calidad de la educación» (Documento de trabajo). Bogotá, MEN: 1991.
- \_\_\_\_\_. *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° 1*. Bogotá: MEN, 1993a.
- \_\_\_\_\_. *Avances de saber. N° 2. Los primeros resultados*. Bogotá: MEN, 1993b.
- \_\_\_\_\_. *Avances de saber. N° 3. Informe sobre repitencia*. Bogotá: MEN, 1993c.
- \_\_\_\_\_. Circular ministerial N° 28, Agosto 4 de 1994. *Política de textos*. Bogotá: MEN, 1994.
- \_\_\_\_\_. *Educación para la democracia, el desarrollo, la equidad y la convivencia. Plan Decenal de educación 1996-2005*. Bogotá: MEN, 1996a.
- \_\_\_\_\_. *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, 1996b.
- \_\_\_\_\_. *Resolución número 2343 de junio 5 de 1996*. Bogotá: MEN, 1996c.
- \_\_\_\_\_. *Evaluación de impacto Plan de Universalización Educación básica primaria. Crédito 3010-CO BIRF. Resultados*. Bogotá: MEN, 1997a.
- \_\_\_\_\_. *Factores asociados al logro cognitivo de los estudiantes. Grados 3°, 5°. 1993-1995*. Bogotá: MEN, 1997b.

- MINISTERIO DE HACIENDA Y CRÉDITO PÚBLICO - DEPARTAMENTO NACIONAL DE PLANEACIÓN. *Proyecto de presupuesto general de la nación 2000. El presupuesto de la verdad - Ministerio de educación nacional*. Bogotá: MINHACIENDA-DNP, 1999.
- MISIÓN CIENCIA, EDUCACIÓN Y DESARROLLO. *Colombia al filo de la oportunidad*. Bogotá, IDEP: 1997.
- MOCKUS, Antanas et al. *Las fronteras de la escuela*. Bogotá: Sociedad colombiana de pedagogía, 1994.
- MORIN, Edgar. *El método. Las ideas*. Madrid: Cátedra, 1992.
- OEI. *Evaluación de la calidad de la educación. Documento base*. Buenos Aires: OEI, 1994.
- OLIVERIO, Alberto. «El espectáculo de la ciencia». En: Revista *Newton*. Agosto de 1998.
- ORTIZ, Luis Augusto y BURGOS, Myriam. «La escuela en Colombia: Gobierno escolar y construcción de democracia». En: Revista *La educación* N° 122, Año XXXIX, Vol. III, 1995. Organización de los Estados Americanos.
- OTÁLORA, Blanca et al. «El uso de materiales educativos en la educación secundaria». Bogotá: Instituto Ser de investigación (s/p), 1993.
- PAJARES, Carlos. *La nueva física*. Barcelona: Salvat, 1975.
- PELLICER, Dora. «El lenguaje en la transmisión escolar de conocimientos: la clase de español en la escuela primaria». En: Revista *Nueva Antropología*, Vol. VI, No. 22, México, 1983.
- PEÑA, Daniel. *Estadística: modelos y métodos. 1 Fundamentos*. Madrid: Alianza, 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.
- PERELMAN, Chaïm. *El imperio retórico*. Bogotá: Norma, 1997.
- PÉREZ, Mauricio. «Evaluación de competencias para la producción de textos. Pruebas masivas en educación básica en Colombia». En: Revista *Alegría de enseñar* N° 38. Cali: FES, Enero-Abril de 1999.
- PIÑEROS, Luis Jaime y RODRÍGUEZ, Alberto. «Los insumos escolares en la educación secundaria y su efecto sobre el rendimiento académico de los estudiantes». Bogotá: s/p, 1999.
- POSADA, Jorge. «Notas sobre comunidad educativa». Bogotá: s/p, 1999.

- RIUS [Eduardo del Río]. *Pequeño Rius ilustrado*. México: Ediciones de cultura popular, 1971.
- ROSAS, Gabriel et al. «Racionalización del gasto público en educación». Bogotá: Departamento Nacional de Planeación, 1996.
- SAMPER, Ernesto. *Por los derechos de la gente. Informe al Congreso 1994-1998*. Bogotá: Presidencia de la República, 1998.
- SAN FABIÁN, José Luis. «EL Centro Escolar y la Comunidad Educativa: ¿Un Juego de Metáforas?». En: *Revista de Educación* N° 309. 1996.
- de SAUSSURE, Ferdinand. *Curso de lingüística general*. Madrid: Alianza, 1989.
- SECRETARÍA DE EDUCACIÓN DEL DISTRITO CAPITAL [SED]. *Plan sectorial de educación, 1998-2001*. Bogotá: SED, 1998.
- \_\_\_\_\_. *Pretextos. Reflexiones para elegir libros de texto*. Bogotá: SED, 1999a.
- \_\_\_\_\_, CORPOEDUCACIÓN, Misión social-DNP. *Evaluación de la calidad de la educación primaria en Santafé de Bogotá 1998. Factores asociados al logro. Informe final*. Bogotá, 1999.
- SIABATO, Rosa Cristina. «La escuela Nueva y su impacto en el plan de universalización de educación básica primaria». En: *Educación para la población rural: balance prospectivo*. Bogotá: MEN, 1998.
- SILVA, Luis Carlos. *Cultura estadística e investigación científica en el campo de la salud: una mirada crítica*. Madrid: Díaz de Santos, 1997.
- SIMIAND, François. «El estadístico debe saber lo que hace». En: *El oficio de sociólogo* [Bourdieu et al.]. México: Siglo XXI, 1994.
- STUBBS, Michael. *Lenguaje y escuela*. Madrid: Cincel-Kapelusz, 1984.
- TORANZOS, Lilia. «Evaluación y calidad». En: *Revista iberoamericana de educación* N° 10. Madrid: OEI, Enero-Abril de 1996.
- TORT, Michel. *El cociente intelectual*. Madrid: Siglo XXI, 1979.
- TRIANA, Alejandro. «La evaluación de los sistemas educativos». En: *Revista iberoamericana de educación* N° 10. Madrid: OEI, Enero-Abril de 1996.

ULLOA, Alejandro. «Propuesta para una teoría del discurso». En: Revista *Grafos* N°3. Cali: Universidad Santiago de Cali, Enero-Junio, 1984.

UNESCO. «Declaración mundial sobre educación para todos». En: *Educación para todos: finalidad y contextos. Monografía I*. Paris: UNESCO, 1992.

\_\_\_\_\_. *Informe mundial sobre la educación. Los docentes y la enseñanza en un mundo en mutación*. Madrid: Santillana-UNESCO, 1998.

VELÁZQUEZ, Enrique. «Descripción de los cuestionarios sobre evaluación de la calidad de la educación, aplicados por el ICFES en 1992». Bogotá: ICFES s/p, 1995a.

\_\_\_\_\_. «El dispositivo: pruebas de Estado». Bogotá: ICFES s/p, 1995b.

\_\_\_\_\_. «Pragmática de los exámenes escolares masivos». Bogotá: ICFES s/p, 1995c.

VENEGAS, Ma. Clemencia. *El texto escolar: cómo aprovecharlo*. Bogotá: MEN, 1993.

WIDDOWSON, H. G. «Manejo y uso». En: *Lenguaje y sociedad*. Cali: Univalle, 1983.

ZULETA, Estanislao. «Elogio de la dificultad». En: *Elogio de la dificultad y otros ensayos*. Cali: Fundación Estanislao Zuleta, 1994.

## Anexo I

		<i>Cerrada</i>		<i>Abierta</i>	
		Con Nivel 0	Sin Nivel 0	Con Nivel 0	Sin Nivel 0
<i>Comités</i>	Nacional	0.880	0.853	0.661	0.901
	Cuarto	0.531	0.580	0.831	0.851
	Tercero	0.976	0.887	0.746	0.794
	Urbana	0.945	0.986	0.191	0.488
	Nueva	0.864	0.944	0.721	0.549
	R-Graduada	0.192	0.172	0.451	0.274
	Distrito	0.912	0.739	0.712	0.957
<i>Gestión</i>	Nacional	0.309	0.271	0.490	0.408
	Cuarto	0.354	0.222	0.058	0.024
	Tercero	0.282	0.294	0.679	0.689
	Urbana	0.574	0.532	0.769	0.588
	Nueva	0.478	0.680	0.128	0.036
	R-Graduada	0.804	0.656	0.008	0.002
	Distrito	0.750	0.828	0.069	0.107
<i>Libros en casa</i>	Nacional	0.677	0.758	0.505	0.910
	Cuarto	0.607	0.425	0.802	0.769
	Tercero	0.148	0.136	0.340	0.463
	Urbana	0.336	0.142	0.690	0.586
	Nueva	0.592	0.710	0.369	0.411
	R-Graduada	0.141	0.150	0.014	0.144
	Distrito	0.245	0.167	0.670	0.707
<i>Libros recibidos escuela</i>	Nacional	0.662	0.514	0.190	0.022
	Cuarto	0.877	0.286	0.013	0.040
	Tercero	0.236	0.657	0.011	0.117
	Urbana	0.126	0.757	0.168	0.049
	Nueva	1.658	0.766	0.914	0.389
	R-Graduada	0.057	0.259	0.197	0.024
	Distrito	0.049	0.046	0.631	0.445
<i>Guías escuela nueva</i>	Nacional	0.826	0.478	0.002	0.404
	Cuarto	0.325	0.360	0.042	0.365
	Tercero	0.863	0.611	0.055	0.782
	Urbana	0.328	0.287	0.049	0.048
	Nueva	0.390	0.381	0.153	0.090
	R-Graduada	0.337	0.150	0.001	0.083

<i>Libros biblioteca</i>	Nacional	0.339	0.148	0.089	0.038
	Cuarto	0.348	0.176	0.030	0.013
	Tercero	0.805	0.511	0.524	0.377
	Urbana	0.294	0.145	0.202	0.385
	Nueva	0.905	0.926	0.026	0.012
	R-Graduada	0.071	0.075	0.024	0.490

<i>Textos escolares</i>	Nacional	0.026	0.009	0.033	0.115
	Cuarto	0.382	0.260	0.156	0.225
	Tercero	0.065	0.025	0.001	0.231
	Urbana	0.461	0.387	0.039	0.043
	Nueva	0.204	0.160	0.346	0.211
	R-Graduada	0.031	0.013	0.011	0.008

<i>Textos escolares lenguaje</i>	Nacional	0.154	0.068	0.007	0.019
	Cuarto	0.845	0.697	0.072	0.061
	Tercero	0.106	0.047	0.001	0.140
	Urbana	0.917	0.986	0.021	0.017
	Nueva	0.396	0.367	0.444	0.312
	R-Graduada	0.031	0.013	0.011	0.008

<i>Repitencia</i>	Nacional	0.003	0.012	0.001	0.001
	Cuarto	0.188	0.435	0.001	0.001
	Tercero	0.008	0.008	0.002	0.113
	Urbana	0.001	0.001	0.044	0.104
	Nueva	0.286	0.190	0.002	0.001
	R-Graduada	0.002	0.018	0.008	0.013
	Distrito	0.166	0.117	0.230	0.142

<i>Deserción</i>	Nacional	0.320	0.081	0.034	0.057
	Cuarto	0.026	0.003	0.178	0.222
	Tercero	0.820	0.984	0.086	0.057
	Urbana	0.609	0.303	0.097	0.085
	Nueva	0.164	0.146	0.642	0.578
	R-Graduada	0.768	0.598	0.002	0.014
	Distrito	0.001	0.001	0.706	0.626

<i>Alumnos por docente</i>	Nacional	0.005	0.005	0.001	0.001
	Cuarto	0.013	0.009	0.094	0.066
	Tercero	0.225	0.249	0.001	0.001
	Urbana	0.295	0.219	0.001	0.001
	Nueva	0.029	0.006	0.011	0.844
	R-Graduada	0.001	0.001	0.001	0.001
	Distrito	0.108	0.150	0.987	0.963

<i>Ciencias</i>	Nacional	0.005	0.001	0.198	0.205
	Cuarto	0.006	0.001	0.679	0.581
	Tercero	0.173	0.046	0.030	0.036
	Urbana	0.051	0.008	0.112	0.120
	Nueva	0.232	0.071	0.786	0.816
	R-Graduada	0.638	0.548	0.123	0.255
	Distrital	0.050	0.032	0.061	0.039

<i>Matemáticas</i>	Nacional	0.465	0.468	0.140	0.106
	Cuarto	0.193	0.231	0.401	0.283
	Tercero	0.143	0.197	0.083	0.028
	Urbana	0.264	0.495	0.517	0.324
	Nueva	0.848	0.564	0.494	0.464
	R-Graduada	0.720	0.554	0.123	0.066
	Distrital	0.675	0.495	0.553	0.454