

La evaluación oficial del área de lenguaje

Guillermo Bustamante Zamudio
**

Fabio Jurado Valencia

Presentación

El Ministerio de Educación Nacional [MEN] ha resuelto evaluar la calidad de la educación colombiana. Para ello, en 1990 decidió construir instrumentos de medición del logro cognitivo en Lenguaje y Matemáticas, dado que los estudiantes parecen exhibir bajo rendimiento académico en estas áreas y que se reconoce como básico el papel que ellas tienen durante todo el proceso educativo. No obstante, no

* Investigador de la Universidad Pedagógica Nacional Bogotá

** Lingüista Universidad Nacional

son explícitos los criterios con los que se detectó ese bajo rendimiento, ni con los que se consideraron básicos al lenguaje y a la matemática en la educación.

Uno de los autores del presente artículo trabajó en el equipo conformado para elaborar los mencionados instrumentos, específicamente el del área de lenguaje para 5^o grado. Esta elaboración requirió construir un marco teórico interdisciplinario para hacer una propuesta alternativa a los instrumentos ya existentes, pues al analizarlos se encontró su debilidad en el contexto de una propuesta de calidad de la educación.

Algunos resultados del trabajo, que empezó con 3^o y 5^o de primaria, y que ya se extendió a 7^o y 9^o, son públicos, no así el proceso, ni, particularmente, la conceptualización implicada. En consecuencia, estamos desarrollando una investigación ("Hacia el mejoramiento de la calidad educativa: forma alternativa de evaluación en el área de lenguaje") que no sólo pretende reconstruir la elaboración de la prueba de 5^o, sino también profundizar en el análisis de aspectos que en esa ocasión se despacharon de manera rápida —debido al plazo existente para elaborar el instrumento— o que no se tocaron. Así mismo, pretendemos contribuir a la apertura de un espacio para investigar, tanto la prueba realizada, como el asunto del mejoramiento de la calidad de la educación colombiana.

El presente artículo, que constituye una revisión de la parte inicial del segundo avance de investigación entregado a la Universidad Pedagógica Nacional (Junio de 1993), analiza someramente algunos antecedentes de la construcción de la prueba para 5^o, y sienta algunas posiciones respecto a la evaluación en lenguaje.

Antecedentes

La evaluación tuvo varios tipos de antecedentes: de carácter institucional, evaluativo y teórico.

Institucionales

Independientemente de cualquier situación evaluativa, pedagógica o teórica, es un hecho que el MEN resolvió hacer unas pruebas para evaluar el logro cognitivo. Dicha entidad se mostró (y se muestra) empeñada en evaluar los esfuerzos realizados

para mejorar la calidad de la educación en las últimas dos décadas. Según el MEN , este empeño se debe a que:

- las evaluaciones realizadas hasta ahora no han procedido de manera rigurosa, articulada ni permanente, y que no ha habido divulgación de resultados ni aplicación de correctivos;

- varios estudios (realizados por la UNESCO en 1986; por Ospe 1980-1986, 1988; y por un equipo interdisciplinario MEN - CIUP - DTE-CEP, 1987) proponen desarrollar un sistema de evaluación permanente;

- la ley 24 de 1988 (reestructuración del MEN) crea la División de Control de Calidad, entre cuyas funciones está evaluar la calidad de la educación; y a que

- en el contrato 3010-CO [1989] entre el Banco Mundial [BM] y el MEN, mediante el cual se concedió a Colombia un préstamo con destino al sector educativo, se recomendó al país evaluar el aprovechamiento estudiantil mediante "pruebas sistemáticas".

Obsérvese que hay investigaciones recomendando la evaluación desde 1980; sin embargo, la decisión tomada no es la de hacer divulgación de los resultados, aplicación de correctivos o articulación de los trabajos, todas ellas carencias señaladas como justificación. En cambio, y a partir del contrato entre el BM y el MEN, se le dio un impulso al "Sistema nacional de evaluación de la calidad de la educación" [SINECE] y se dispuso la realización de pruebas que establecieran el logro cognitivo y, por consiguiente —al menos en los propósitos—, la calidad de la educación dada a los colombianos.

Obsérvese también (Cfr. Nota 1) que el ciclo de foros que enmarca esta puesta en acción de la evaluación tiene que ver con el *Plan de Apertura Educativa*. Pues bien, en la "Introducción" de dicho documento se plantea el propósito de llevar la economía colombiana "por la senda del crecimiento sostenido y la eficacia", lo que "a largo plazo" [sic] permitirá a un mayor número de personas disfrutar de las oportunidades del entorno económico. En esta suerte de acción merlincsa, dicen que la educación juega un papel primordial: "mayor educación para la población

* Cfr. MEN. "Ciclo de foros sobre el Plan de Apertura Educativa. Tema: la evaluación de la calidad de la educación". Mimeo, 1991.

** Departamento Nacional de Planeación, 1991.

trabajadora representa mayor capacidad productiva para todo el sistema económico". En tal sentido, según el documento, pasamos a llamarnos "capital humano" y competimos con la acumulación tradicional de capital físico. Los países orientales son puestos como ejemplo del funcionamiento exitoso de tales ideas, aunque no se menciona la pérdida de identidad de estas culturas, ni el alto costo social que pagan por la aplicación de tales mediadas neoliberales.

En el aparte sobre "Problemas del sector educativo" (p. 9), el *Plan* habla de estudios sobre la calidad de la educación que han detectado el bajo nivel de la educación primaria, aunque no los menciona. Cuando en el mismo aparte se refiere a la educación secundaria (p. 12), habla de la mala calidad en determinadas regiones y en las áreas de matemáticas y lenguaje; en este caso, sí menciona la fuente que le permite hacer esa afirmación: los llamados "Exámenes de Estado" (Ley 80/80), aplicados por el ICFES.

Estas y otras consideraciones llevan al gobierno a plantear una "Política educativa para el periodo 1991-1994", en la cual está contemplado el propósito de "elevar la calidad de la educación" mediante once medidas (ps. 25 a 28), entre las cuales está:

[...] establecer un sistema nacional de evaluación de la calidad de la educación primaria, secundaria y superior, con base en pruebas cognitivas a los estudiantes, aplicadas anualmente.

Así, las pruebas aparecen como un instrumento fundamental del mejoramiento de la calidad de la educación, programa prioritario del sector educativo (p. 36):

Para mantener información permanente sobre la calidad de la educación, se establecerá en el MEN un sistema de pruebas para la evaluación de los logros educativos de los niños y jóvenes de la primaria y la secundaria.

En respuesta pues, a la mencionada recomendación del BM, el programa de la División de Evaluación estableció un sistema de pruebas para evaluar los logros de los alumnos de primaria y secundaria, en todas las áreas. Así, el SINECE se propuso evaluar las áreas de español y matemáticas, en los grados 3º, 5º, V y 9º, paulatinamente. Después de dicha evaluación, se pasaría al diseño y aplicación de pruebas en las otras áreas... (Cfr. "Ciclo de foros"). Los anteriores propósitos son iguales a algunos de los concebidos en el *Plan de Apertura Educativa*:

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. Los resultados de la evaluación serán públicos y soportarán los programas de mejoramiento de la calidad de la educación básica y las actividades de supervisión y asesoría pedagógica de los docentes (p. 36).

La División de Evaluación se dedicó, en consecuencia, a elaborar los instrumentos para evaluar. Durante el primer semestre de 1990, el SINECE aplicó diversos instrumentos en la región del Viejo Caldas, comprometiendo a autoridades locales, regionales y a maestros, y capacitando a los docentes.

Pero vamos atrás un momento: al MEN le pareció inconsistente el trabajo investigativo y evaluativo existente hasta ese momento; sin embargo, decidió mejorarla calidad. ¿Cómo sabía que la calidad no era buena? Dos son las respuestas que da el *Plan*: la primera, considerar los índices estadísticos de repitencia, deserción y cobertura; y, la otra, tomar como base los resultados de las pruebas de Estado.

No obstante, la primera respuesta es cuantitativa y no cualitativa. Nadie niega que entre repitencia, deserción y cobertura, por un lado, y calidad, por el otro, haya relación; pero como esa relación no es una directa proporcionalidad (a más de lo uno, entonces más de lo otro), esta primera respuesta no sirve para sacar conclusiones inmediatas sobre la calidad de la educación.

La segunda respuesta es la que parece tener carácter cualitativo; pero no es tan claro que lo tenga, por una parte, porque si bien los exámenes de Estado que aplica el ICFES se basan en el currículo oficial, según el mismo *Plan de Apertura Educativa* (p. 9), tiene poca relación con la capacitación que se ha dado a los docentes. Y, por la otra, porque se hacen con arreglo a la Norma; es decir, necesitan mantener una curva para ser confiables estadísticamente, lo cual es un índice de

* Están próximos a publicarse los resultados de la "Consulta Nacional" del MEN; consulta que se proponía, entre otros, establecer el conocimiento real de los maestros sobre la Reforma Curricular. ¿Con qué base, entonces, el *Plan* afirma esto?

tipo cuantitativo. Por lo tanto, tampoco se puede concluir con mucho tino sobre la calidad de la educación, a partir de dichos exámenes.

Si las pruebas de Estado sirven para diagnosticarla calidad, ¿por qué el MEN propuso coonstruir un nuevo tipo de instrumento? Fuera de eso, las pruebas de calidad no hacen más que concluirlo que ya se dijo para justificarlas: mala calidad, regiones deprimidas, áreas de conocimiento débiles, diferencia entre la educación oficial y la educación privada de los estratos sociales más altos ; y las medidas también ya estaban tomadas antes de hacer las pruebas: dotación de textos, cobertura, grado cero... Afortunadamente, todo está calculado:

El costo de estas acciones, \$1.000 millones, se financiará con recursos del crédito del BM para primaria y con fondos de cooperación técnica internacional (p. 36).

El BM había otorgado un plazo de tres años para elaborar las pruebas. Próximo este término, se presentaron los trabajos de los equipos respectivos. Sin embargo, el BM tuvo dos objeciones fundamentales sobre el instrumento elaborado para el área de lenguaje:

Por una parte, no le pareció que se comprometiera con niveles de logro cognitivo claramente delimitados; es decir, que no definía los niveles de exigencia desde donde el niño debe partir, ni aquellos hasta donde debe llegar, así como tampoco logros mínimos por nivel. O sea, el instrumento no cumplía, a juicio del BM, algunas condiciones de las pruebas con arreglo a criterio .

La otra objeción era más de tipo operativo: el BM juzgó que la prueba no se prestaba a una codificación en grandes volúmenes, dado el carácter abierto de la mayoría de sus preguntas, vale decir, preguntas no estructuradas: ejercicios que no se contestan en una hoja de respuestas (para lectura electrónica) y cuya "verdad" se establece de antemano, sino que el niño los redacta con alguna libertad. La lectura de miles de estas respuestas implica —dada la carencia de programas computari-

* Cfr. MEN. *Saber*. Bogotá, 1992.

** El lector puede hacer su propio juicio: Cfr. Charria, María Elvira (coordinadora).

"Estudio sobre la evaluación del aprendizaje del lenguaje" (documento de trabajo). Bogotá: MEN, Noviembre de 1990.

zados para leer manuscritos— una enorme cantidad de tiempo; y, por otro lado, un problema de confiabilidad, al tener que recurrir a un amplio número de personas para la lectura y la calificación

Ante esta situación, el BM puso a disposición de Colombia la información de que el "Educational Testing Service" [ETS] tenía elaboradas unas pruebas adecuadas al propósito buscado en Colombia. En otras palabras, del préstamo — recibido o no— ya se había invertido una parte en el pago de la investigación conducente a la elaboración del instrumento rechazado, y suponemos que otra buena parte se habría gastado en la compra y aplicación masiva de la prueba norteamericana.

Pero la prueba no se compró al ETS: en 1991, el MEN, contando con unos pocos meses todavía, estableció una serie de equipos de trabajo para intentarlo de nuevo: para las pruebas de logro cognitivo de 5^o de primaria se responsabilizó al ICFES, a través de su "Servicio Nacional de Pruebas" [SNP]; de las pruebas de factores asociados, se hizo cargo el Instituto SER de Investigación; de las pruebas de desarrollo personal y social, se hizo cargo el SINECE; para la aplicación de las pruebas y el procesamiento de la información, se responsabilizó al Instituto SER de Investigación; y, por último, el ICFES se encargaría del análisis de los resultados.

Pese a tener pruebas construidas en las áreas a evaluar, el ICFES resolvió que se elaboraran instrumentos diferentes a los que usualmente él mismo aplica, pues no se trataba de una prueba con arreglo a la Norma —como las que tal institución hace regularmente— sino con arreglo a Criterio, ya que se indagaría calidad. Para ello, integró un equipo con especialistas suyos (sicólogos, sicómetras), con personas del MEN que trabajan las áreas curriculares correspondientes y con maestros en ejercicio o investigadores de distintos niveles educativos .

- "Servicio de Pruebas en Educación", de los Estados Unidos.
- Carlos Augusto Hernández (Universidad Nacional); Margarita de Angarita (MEN); Jaime Riveros (INEM); Elizabeth Forero (Escuela Nueva Delhi - Bogotá); Carlos Ordóñez (ICFES - SNP); Patricia Duarte (ICFES - SNP); y Guillermo Bustamante Z. (entonces. Universidad Pedagógica y Tecnológica de Colombia). En los informes del SINECE - *Saber* y al colección *Avances de saber*— casi todas estas personas aparecen asignadas al SNP, pasando por alto la fundamental composición interinstitucional e interniveles de los grupos de trabajo.

En 1991, se terminaron dos grupos de instrumentos: por un lado, factores asociados y, por otro, logro cognitivo, en sus dimensiones "académica" y "de desarrollo personal-social". Finalmente, el BM aceptó las pruebas elaboradas por el nuevo equipo. Después de aplicarlas en 1991, y una vez articulado el trabajo con el de los otros equipos (factores asociados, análisis de las pruebas, etc.), el MEN tomó la determinación de preparar nuevos instrumentos para aplicar en 7^o y 9^o grados, a finales de 1992; en esta ocasión, todas las fases corrieron por cuenta del SNP.

Eval uativos

Este tipo de antecedente tiene tres vertientes: a) estudios sobre evaluación del logro cognitivo en el área de lengua materna, o que la impliquen; b) estudios que han tenido que aplicar instrumentos de evaluación para el área; y c) evaluaciones en el área.

El primero, es el caso de trabajos como el de Mabel Condemarín y Chadwick , el de Triado y Forns ; o el de disposiciones sobre evaluación, como en el caso de la "Promoción flexible" (Escuela Nueva) y de la "Promoción automática". El segundo es el caso de investigaciones (Centros, pregrado, postgrado) que han aplicado instrumentos de evaluación terminales; por ejemplo, el trabajo de Camacho y otros , que utiliza y anexa pruebas de comprensión de lectura y de aptitud verbal; o de investigaciones que han desarrollado concepciones evaluativas por procesos; por ejemplo, el trabajo de Gloria Rincón ; o el caso de evaluaciones del "rendimiento" en el área de español para Escuela Nueva (como parte de una evaluación a esa modalidad de estudios), etc. El tercero es el caso de un amplio abanico de evaluaciones que va desde las oficiales (como la que el ICFES —por

* *La escritura creativa y formal*. Santiago: Andrés Bello, 1986.

** *La evaluación del lenguaje*. Barcelona: Anthropos, 1989.

*** "Competencia comunicativa y competencia matemática: análisis a partir de la filosofía de Wittgenstein". Tesis de grado. Maestría en Educación. Universidad de la Sabana. Bogotá, Julio de 1991. *****

"Las interacciones lecto-escritas con función comunicativa y el proceso de construcción de este objeto de conocimiento en un contexto escolar". Tesis de grado. Magister en Lingüística y Español. Universidad del Valle, 1987.

resolución 14584 de Agosto de 1979— aplicaba a los niños que aspiraban a ser admitidos en los colegios oficiales de secundaria de Bogotá), hasta las evaluaciones que cotidianamente hacen los maestros en la escuela.

Teóricos

Como la prueba debía, implícitamente, hacerse sobre el currículo vigente, fue necesario echar una mirada sobre este aspecto del área y sobre la bibliografía que analiza la Reforma Curricular para español. Además, este antecedente se tuvo en cuenta por la necesidad de ser consistentes con la idea de que las formas evaluativas están relacionadas —directa o indirectamente— con la concepción del lenguaje manejada, con el currículo vigente y con la relación maestro/alumno establecida. Si bien la relación entre estos factores no es uno a uno, parece cierto que las formas evaluativas permitirían percibir una imbricación en la que aparecen diferentes

a) tipos de modelos: concepciones del lenguaje y otros como modelos prácticos, pedagógicos, etc.; b) especificidades: complementarios o excluyentes entre sí; c) niveles de asunción: desde asumidos hasta simplemente enunciados; d) tipos de presencia: explícitos o implícitos; e) formas de articulación: buscada deliberadamente o resultante de la realización del proceso educativo.

Ya desde los Marcos Generales del currículo del MEN, por ejemplo, se ve claramente la presencia de dos teorías de la lengua, unidas a dos modelos de enseñanza de la lengua, que son excluyentes: el modelo normativo, basado en una teoría estandarizada de la lengua, y el modelo vinculado a las funciones del lenguaje, basado en la teoría denominada "lógico-semántica". Además, en los desarrollos por programas y en la elaboración de los textos escolares, aunque de palabra tenemos una concepción "moderna" y expresa sobre el lenguaje, el modelo anterior no ha perdido su fuerza. La evaluación obedece, pues, a la imbricación de modelos en

* Cfr. Arboleda, Rubén. "Sobre la reforma curricular: el caso de Español y literatura". En: *Revista colombiana de educación*. N° 18. Bogotá: CIUP-UPN, II Semestre de 1986. También: Rincón et. al.: "La lengua materna en la escuela y la reforma curricular". *Revista Educación y cultura* N° 9. Bogotá: FECODE, Septiembre de 1986.

juego; en consecuencia, para que haya innovación en los modelos pedagógicos (que involucren distintas formas evaluativas), el nuevo marco teórico requiere ser incorporado racionalmente.

Una de las conclusiones fue: no se puede preguntar a los niños por algo sobre lo que se ignora si están informados homogéneamente. En la búsqueda y estudio de los antecedentes, pues está gran parte de lo que se puede ofrecer como modelo teórico alternativo a la evaluación del área de español; pero también encontramos ahí una serie de formas evaluativas cuyos presupuestos teóricos están dominados por el instrumento; es decir, encontramos una advertencia de que la consistencia del modelo teórico no garantiza una solidaridad con el engranaje de la prueba.

Por ejemplo, el equipo debió estudiar pruebas oficiales chilenas, norteamericanas y colombianas, cuyos logros instrumentales parecían estar por encima de sus marcos teóricos. Cuando se llegó la hora de construir preguntas para nuestra prueba, vimos que, instrumentalmente, podíamos estar más atrás que pruebas con fundamentos teóricos más débiles: por un lado, porque lo que es claro para el evaluador no necesariamente lo es para el evaluado (la forma como se hace la pregunta, por ejemplo ; y, por otro, porque es necesario que los distractores —en las preguntas de escogencia múltiple— vayan siendo menos atractivos para los niños con mayor número de respuestas correctas.

Por qué evaluar el área

Si por evaluar se entiende una mirada analítica retrospectiva que busca someter los modelos (pedagógicos, teóricos) a las transformaciones necesarias y modificar la acción racionalmente orientada, sí es necesario evaluar el área. Pero, en consecuencia, es una evaluación que depende del momento del proyecto individual de trabajo. La evaluación masiva queda, con este principio, seriamente cuestionada. El lenguaje no se debe evaluar solamente en la escuela y, menos aún, en el aula de clase . Pretender

* Ante la pregunta por la diferencia entre la forma como ocurre el evento /correr/, cuando es ejecutado por una tortuga y por una ardilla, un niño nos respondió: "No era una ardilla, era una liebre; y perdió por quedarse dormida". El evaluador quería que se le respondiera lenta y rápidamente.

** Cfr. Michael Stubbs. *Lenguaje y escuela*. Madrid: Cincel-Kapelusz, 1984.

esa evaluación en la escuela encarna una mirada muy restringida de las funciones reales del lenguaje. La prueba masiva parece reñir con una concepción del lenguaje como significación y no como mera codificación formal.

Sin embargo, como las determinaciones estaban tomadas, la realización de la prueba masiva era ineludible; fue un paso más en el proceso. Pero sí resultaría equivocado, a todas luces, perpetuar ese tipo de evaluación.

Curiosamente, muchos comparten la idea de la necesidad de evaluar el área de lenguaje. No obstante, cuando se hace la encuesta entre maestros, son muy dominantes las justificaciones normativas para la evaluación del área: "es que la cultura de una persona se expresa en su forma de hablar". En lugar de esto, escogimos una perspectiva de evaluación centrada en el desarrollo de la capacidad para producir y entender distintos tipos de mensajes en contextos comunicativos específicos ("Competencia Comunicativa").

Como se ve, la actual iniciativa de evaluar la calidad de nuestra educación proviene, en muy baja proporción, de una dinámica propia de los maestros (y, menos aún, de los alumnos). Podría caracterizarse la iniciativa como una "heteroevaluación", con todas las implicaciones negativas que tiene, al no generar "autoevaluación" ni "coevaluación".

En realidad, esta medida, contemplada en el *Plan de Apertura Educativa*, no es más que una parte del plan de "Revolución Pacífica"; es decir, la forma que ha tomado el neoliberalismo en nuestro país y que implica el adelgazamiento de la nómina estatal, con el consecuente desprendimiento paulatino del Estado de su compromiso económico con la educación. Y sabemos que iguales procesos (evaluación de la calidad de la educación incluida) vivió hace unos años Chile, en similar aplicación del neoliberalismo. ¿Qué balance se ha hecho en Colombia de la experiencia chilena? ¿O se trata de hacer creer que es una estrategia nuestra, original?

En estos términos, elaborar una prueba para evaluar logro cognitivo en el área de español es una forma de ejercicio del poder, el ejercicio de la llamada "modernización" del Estado, que no es otra cosa que la nueva apariencia del ejercicio del capitalismo internacional. No es extraño que la evaluación esté ligada al poder, pues ha sido pensada y aplicada como justificación de la persecución, de la sanción y de medidas que afectan a miles de personas: capacitación, dotación de materiales, producción de material impreso, endeudamiento externo, etc.

No obstante, en el diseño de un instrumento podrían darse las bases para ganar condiciones distintas de negociar nuestras especificidades educativas y, por qué no, de ganar un espacio para buscar la posibilidad de negociar nuestros procesos sociales.

La evaluación podría ser una oportunidad para luchar por mayor autonomía para el maestro, así como por responsabilidades asumidas desde proyectos propios. Es decir, partimos de un Sistema Nacional de Evaluación que debería diluirse en la creación de Sistemas Regionales de Evaluación que, a su vez, se diluyan en una cultura de la evaluación basada en la autonomía real del maestro, ejercida en proyectos de innovación y en procesos de investigación. Ya el hecho de que equipos conformados por maestros en ejercicio (además de personal del MEN y del ICFES) tuviera en sus manos parte de la elaboración de la prueba, establece un comienzo de reconocimiento a las innovaciones pedagógicas y al protagonismo del maestro.

En consecuencia con estas ideas, durante el trabajo inicial, el grupo planteó taxativamente que debería propenderse por pruebas con características tales que no fuera posible hacer selecciones de "buenos" o "malos" niños, sino que permitieran expresar la proporción en la que la educación colombiana había sido capaz de llevar a los niños a los logros esperados. Desafortunadamente, la forma como se ha entendido que los resultados de esta evaluación eran "públicos" ha sido otra; aún así, el país puede exigir que lo público sea el proceso, los resultados, las medidas que se tomen y las investigaciones que se hagan.

De la misma forma, el grupo planteó que no deberían utilizarse unos mapas, provenientes de los puntajes obtenidos, para asignar recursos sacados del préstamo internacional. Que dicho préstamo no debería destinarse a perpetuar aspectos considerados problemáticos a la hora de hacer la prueba; por ejemplo, si el modelo de dar y recibir clase sobre normas de gramática ha conducido a una incompetencia comunicativa, no debería buscarse una "capacitación" de maestros en la que ahora ellos se sienten a recibir clase de otros y, así, continuar con el mismo esquema. Como alternativa, proponíamos proyectos, no de capacitación, sino de "autoinformación", vinculados con el establecimiento de actitudes investigativas en los maestros, quienes podrían llegar a construir proyectos acordes con sus contextos

* Cfr. *Saber* y la colección *Avances de saber*, N^o 1, 2 y 3.

de situación reales, con la escritura como herramienta fundamental. La investigación debe producir las pruebas; también ella podrá sugerir las medidas a tomar.

En tal sentido, también se justificó asumir la elaboración de la prueba para dejar una información tal que posibilitara futuras investigaciones, ya que, en las pruebas que han dominado hasta ahora, las miradas son fundamentalmente numéricas y el material que resulta de su aplicación, a lo sumo, sirve después para diagnosticar a los evaluadores. Tratamos de que esto se superara, por una parte, mediante la explicitación, en la medida de lo posible, de los criterios y consideraciones utilizados, para facilitar una futura aproximación crítica; y, por otra, dejando preguntas abiertas para que miradas absolutamente distintas puedan utilizarlas más adelante como información básica, menos contaminada de la conceptualización de quienes diseñaron los instrumentos. Esta propuesta es consistente con lo planteado por el *Plan de apertura educativa*, por los "Foros" del MEN sobre el tema, y por los documentos del SINECE: todos ellos anuncian que el Sistema de Evaluación en funcionamiento incluye una línea de investigación.

* No sólo conscientemente: el análisis de la aplicación piloto y de las dos aplicaciones definitivas (calendarios A y B) muestra que hay principios subyacentes funcionando en el instrumento.

** Y no "descontaminada", pues una mirada sobre las respuestas de los niños revela que la forma de hacer las preguntas abiertas es definitiva.