

# La evaluación externa y la elaboración de ítemes. Algunas consideraciones sobre la construcción de pruebas

**María Elvira Rodríguez Luna**

Docente Universidad Distrital Francisco José de Caldas

## **El sentido de la evaluación: a propósito de la relación con un interlocutor internacional**

El papel asignado a la evaluación externa en la política educativa y la incidencia de los resultados de las pruebas en distintos ámbitos de la acción social, hacen que los problemas relacionados con la evaluación constituyan centros de interés para docentes de la educación básica, “media” y superior y para investigadores en las disciplinas implicadas; así como también para planeadores educativos y para profesionales especializados en la evaluación de las distintas áreas del currículo. Igualmente, estos temas interesan a los formadores de educadores y a los potenciales usuarios y destinatarios directos e indirectos del sistema nacional de evaluación.

Lo anterior indica que la evaluación externa ha pasado de ser una preocupación exclusiva de quienes están interesados en acceder a la educación superior por medio de los exámenes de estado o como un parámetro de escalafonamiento de las instituciones según los resultados logrados por sus egresados. La discusión actual indica que la evaluación externa también constituye un instrumento adecuado para impulsar procesos de cualificación del sistema educativo, en



RECIBIDO: MARZO 15 DE 2005  
ACEPTADO: MAYO 30 DE 2005

1 Magíster en lingüística española. Doctora en ciencias pedagógicas. mrodun@yahoo.com

especial en grados intermedios (prueba Saber, por ejemplo) y grados terminales de la educación Superior (como el caso de los exámenes de evaluación de la calidad de la educación superior, Ecaes), así como para determinar ejes problemáticos y temáticas fundamentales en la orientación de las acciones formativas y proyectos de investigación.

Las nuevas visiones sobre la evaluación externa amplían los ámbitos de interés e intervención de factores asociados a dicho proceso, promoviendo a su vez el surgimiento de núcleos y grupos de investigación motivados por avanzar en la comprensión y el dominio de estos problemas. Por consiguiente, se requiere fortalecer y ampliar esta participación con el fin de consolidar en el país una comunidad académica fuerte que intervenga de manera crítica y propositiva con el aporte de sus puntos de vista y sus prácticas en las tareas que demanda la transformación de la cultura de evaluación. Entre ellas, las más inmediatas se relacionan con la elaboración de un balance sobre la experiencia llevada a cabo por el Icfes en la evaluación a la educación básica, “media” y superior, la revisión de los marcos de referencia que sustentan las pruebas, la búsqueda de las articulaciones necesarias entre las políticas educativas, los lineamientos sobre la evaluación, la formación y las prácticas cotidianas en los distintos grados de la escolaridad, entre otros aspectos orientados al mejoramiento del sistema nacional de evaluación.

De allí que resulte pertinente generar espacios de reflexión y participación de las comunidades alrededor de los enfoques que sustentan la evaluación externa, el sentido que tienen los tipos de preguntas, la estructura de las pruebas, las implicaciones derivadas de los modos de interrogar y, en general, sobre los fines de la evaluación externa y sus consecuencias en los sistemas de enseñanza, así como su relación con los instrumentos utilizados y su coherencia con las políticas educativas.

En esta perspectiva la Colegiatura de Lenguaje del Icfes —en el marco del convenio suscrito con la Universidad Nacional de Colombia— propuso la realización del Seminario Evaluación externa y elaboración de ítemes. Para tal efecto fue invitada como ponente cen-

tral la profesora Ana Atorresi, investigadora de la Flacso, Universidad de Buenos Aires (Argentina), y asesora del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (Llece), de la Unesco, para la elaboración de pruebas en el área de Lengua Castellana. Este evento se llevó a cabo en

### **Es pertinente generar espacios de reflexión y participación de las comunidades alrededor de los enfoques que sustentan la evaluación externa.**

Bogotá, del 24 al 28 de enero de 2005, en las instalaciones del Icfes y la Universidad Nacional, con la participación de integrantes de las colegiaturas y los grupos de trabajo en las áreas de Ciencias Naturales, Matemáticas, Lengua Materna, Lengua Extranjera y Ciencias Sociales, profesionales especializados del Icfes, miembros de grupos de investigación en evaluación, docentes de educación básica y algunos profesionales vinculados a la elaboración de los exámenes de evaluación de la calidad de la educación superior (Ecaes).

Entre los temas propuestos en el seminario se destacan:

- La experiencia del Segundo Estudio Regional Comparativo de la Calidad de la Educación (Serce), impulsado por la Unesco y el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (Llece).
- La reflexión sobre problemas relacionados con la parte abierta de las pruebas.
- Posibilidades de la evaluación del carácter oral en la evaluación externa.
- Las teorías del discurso y su incidencia en la evaluación externa.
- Análisis de ítemes en relación con enfoques teóricos de la evaluación.

El seminario promovió la discusión entre los asistentes y señaló algunos puntos críticos sobre los cuales es necesario ahondar para avanzar en la definición de estrategias en el marco de la cualificación de las acciones vinculadas a la evaluación externa. Se parte de reconocer que para la conformación de una comunidad académica y científica es indispensable examinar los fundamentos teóricos

que sustentan las distintas pruebas nacionales y confrontarlos con puntos de vista que alrededor de ellas circulan en ámbitos académicos distintos del nuestro. Sin duda, la posibilidad de conocer experiencias realizadas en otros países latinoamericanos y la discusión con pares portadores de visiones semejantes o diferentes de las que orientan los procesos en las áreas en el Icfes son de utilidad para alimentar el proceso de revisión de los marcos de referencia, aclarar y precisar conceptos, validar enfoques y ajustar los instrumentos utilizados en las pruebas.

Con ese ánimo en el presente escrito se han retomado algunos de los problemas abordados, sugeridos o derivados de las reflexiones propuestas en el seminario, por considerarlos importantes en la reorientación de los fundamentos conceptuales, las estrategias de construcción de las pruebas, la validación de los instrumentos utilizados y la búsqueda de mecanismos de socialización de los resultados por parte del Icfes. Vale aclarar que no se trata de una relatoría en sentido estricto, ni de una crítica a los planteamientos de la expositora central, sino de una recontextualización de las preocupaciones y las inquietudes expresadas por algunos asistentes al evento, en función de las necesidades y las tareas que competen a los diversos grupos que conforman las colegiaturas.

### **En torno de las funciones de la evaluación**

La evaluación es un proceso complejo frente al cual se han construido fuertes prejuicios por parte de los docentes y de otros usuarios del sistema educativo, por considerarla esencialmente como un dispositivo de control y selección. Con frecuencia se destaca la asimetría existente entre lo que se hace en la escuela y lo que se pretende medir a partir de los instrumentos de evaluación, con lo cual se señala la arbitrariedad en los resultados de las pruebas y sus usos como simples indicadores de la inversión en educación. La mayor parte de las veces la evaluación se considera como un evento concreto en vez de un proceso; en esa medida se considera que se trata de una actividad que “proviene siempre de un exterior más o menos lejano cuyos intereses son ajenos a los del ámbito evaluado”. (Atorresi, 2005a).

Frente a tales representaciones que ponen de relieve aspectos de política educativa y merecen la atención de las comunidades, también es necesario advertir que la evaluación permite la “toma fundamentada de decisiones” (2005a) y, en consecuencia, cuando se trata de establecer en qué medida los alumnos alcanzan ciertos aprendizajes, desarrollan determinadas capacidades y adquieren algunos modos de valorar situaciones y problemas que se les proponen por parte de la escuela, la evaluación adquiere otros sentidos.

En cuanto se constituya en un “medio para saber si los propósitos educativos en las escuelas son realizados” (2005a); la evaluación y su pertinencia alcanzan una dimensión significativa para los docentes, los alumnos y las instituciones. En especial cuando se busca establecer el grado de correspondencia entre los propósitos de la enseñanza y los aprendizajes alcanzados, para introducir acciones formativas que garanticen mayor eficacia en los procesos educativos, se requiere obtener información sistemática que oriente la toma de decisiones en diferentes ámbitos de la escuela.

Por consiguiente, la superación de los prejuicios existentes sobre la evaluación implica cambiar el uso restringido que han tenido los resultados de la evaluación para los fines de la clasificación y la selección. Para tal efecto se considera importante promover, ampliar y consolidar acciones orientadas a:

- Trabajar sobre la resistencia de la población a ser evaluada.
- Acompañar el proceso de reflexión sobre los resultados de las pruebas para establecer los aspectos sobre los cuales es necesario profundizar en la formación.
- Proponer estrategias de socialización de los resultados de las pruebas que vayan más allá del establecimiento de la clasificación sobre la “calidad de las instituciones”.
- Fortalecer los grupos de trabajo en las instituciones educativas haciendo de los resultados de la evaluación un objeto de estudio permanente.
- Coordinar acciones de mejoramiento, en las cuales se comprometan las distintas instancias formativas, grupos y centros de investigación, para superar ciertos problemas críticos relacionados con la enseñanza.

- Disminuir la distancia entre los expertos en la evaluación y los formadores, mediante nexos construidos a partir de la investigación sobre los procesos y el fortalecimiento de la participación y crítica de la comunidad educativa.

### **Sobre la utilidad de los resultados de la evaluación**

Si se consolida en nuestro contexto una cultura de evaluación en la cual se supere la concepción de la evaluación como simple mecanismo para la selección y el control, los resultados de las pruebas servirán para mucho más que “calificar o certificar el cumplimiento de requisitos” (2005a). En buena medida ese cambio depende del uso de la información y de lo significativa que ella resulte para los afectados; de lo contrario, los prejuicios existentes alrededor de las evaluaciones masivas se convierten en realidades inobjetables.

Lo deseable es que las evaluaciones masivas entreguen informaciones útiles a maestros, alumnos e instituciones, que les permitan identificar necesidades en la formación, así como ponderar la eficacia de los enfoques y las dificultades presentes en los procesos de enseñanza, y establecer prioridades en las acciones educativas. En esa dirección es importante que los resultados de la evaluación constituyan herramientas para la política educativa. De este modo, como se señala desde otras experiencias, puede esperarse que:

El acceso al sistema educativo se traduzca efectivamente en acceso al conocimiento. Para ello es necesario desarrollar sistemas de información que permitan “ver” qué ocurre en materia de aprendizajes, poner de manifiesto eventuales procesos de vaciamiento de contenidos y/o de segmentación interna del sistema, localizar las situaciones más problemáticas, analizar alternativas de acción y enriquecer los procesos de toma de decisiones. Permite, así mismo, contar con evidencia empírica a la hora de evaluar el impacto de programas e intervenciones específicas (Administración Nacional de Educación Pública, 2000, p. 13).

Para introducir cambios en el uso social de los resultados de la evaluación, puede ser relevante la realización de investigaciones y la puesta en marcha de acciones dirigidas a:

- Identificar las ideas existentes entre los profesores, los alumnos y las instituciones sobre la evaluación y su utilidad.
- Empezar acciones de sensibilización respecto a los beneficios y los impactos de la evaluación en cuanto a la introducción de transformaciones en los procesos educativos.
- Hacer visibles los ajustes a los procesos de formación y toma de decisiones con base en los resultados obtenidos en las pruebas.

### **Algunas pautas relacionadas con los propósitos y los criterios de la evaluación**

Es fundamental que la población evaluada conozca con anterioridad a la aplicación de las pruebas los propósitos y los criterios a partir de los cuales éstas se organizan, de modo que tales orientaciones les permitan la realización adecuada de las tareas escolares y la evaluación de sus resultados. Con esto se hace claridad respecto a lo que se espera de los procesos de formación y se facilita la toma de decisiones respecto a los posibles problemas, avances y acciones necesarios para dar cumplimiento a las expectativas trazadas.

En cumplimiento de tales fines, el Icfes ha impulsado en el país acciones de socialización sustentadas en talleres regionales, publicación de documentos seriados, uso de la página web y seminarios, entre otros medios. La búsqueda de nuevos mecanismos que faciliten la discusión permanente sobre los fundamentos, los problemas y la apropiación de enfoques y orientaciones de las pruebas, entre ellas los foros virtuales, el trabajo en red y la conformación de nodos por medio de los cuales se hagan visibles y circulen las propuestas y los trabajos desarrollados en el país, son de gran valor para garantizar la eficacia de las políticas<sup>1</sup>.

<sup>1</sup> La conformación de una Red Nacional de Investigación en Evaluación (proyecto impulsado por el Icfes en convenio con la Universidad Nacional, la Universidad Pedagógica Nacional y la Universidad Distrital Francisco José de Caldas, con el aval de Colciencias) aspira a constituirse en una estrategia básica para avanzar en estos propósitos.

Es indudable, como lo señala Ana Atorresi, que:

La evaluación modula el currículum; no lo reemplaza ni constituye un currículum paralelo, sino que genera un sistema de jerarquías y énfasis que da una forma especial al contenido de la enseñanza. Los temas evaluados se vuelven más importantes y los tipos de producción exigidos, más relevantes. Así, por ejemplo, como en el área de lenguaje se observa un predominio de pruebas de lectura por sobre las de escritura, ya que la corrección resulta relativamente más sencilla, suelen enfatizarse la enseñanza y el aprendizaje de las competencias lectoras. La influencia de la evaluación como moduladora del currículum puede ser valorada como una indeseable desviación de las normas oficiales sobre el contenido escolar, también como un dispositivo complementario que actualiza esas normas, si el sistema procede al respecto de un modo más o menos consistente (2005b).

La interdependencia entre las orientaciones y los criterios de la evaluación y los propósitos de la formación escolar hace necesaria la articulación entre las instancias comprometidas. En tal sentido, la información que se dé al respecto por parte de la entidad que evalúa requiere estar acompañada de estrategias de trabajo y de comunicación (talleres, reuniones de grupos pequeños, documentos que circulan en la escuela, página web, entre otros elementos) que permitan a los evaluados y a las instituciones educativas tener claridad sobre el sentido de la evaluación, los contenidos, las habilidades y las competencias, al igual que sobre el tipo de instrumentos y parámetros desde los cuales serán interpretados los resultados.

Los destinatarios directos e indirectos de la evaluación podrán formularse preguntas relacionadas con el empleo de la información, entre las cuales Ana Atorresi destaca estos interrogantes:

- ¿Cuáles son las pautas que deberán cumplir los evaluados?
- ¿Deben dominar todo el programa de estudios?
- ¿Deben dominar todo el programa por igual?
- ¿Se evaluará el aprendizaje de contenidos integrales, globales, básicos o mínimos?
- ¿Se evaluará el aprendizaje de contenidos útiles en la vida cotidiana o profesional o de contenidos difícilmente transferibles a otras situaciones?

- ¿Se tratará de contenidos interesantes o motivadores para los alumnos?
- ¿Qué tipo de desempeño deben tener?
- ¿Qué tareas deberán ser capaces de resolver? (2005b).

Este proceso de información previa exige de los equipos de trabajo la toma de decisiones de orden teórico, metodológico y pedagógico, así como la planeación de estrategias de información que garanticen a los destinatarios de las pruebas el conocimiento y la apropiación de los fundamentos, los criterios y los instrumentos que se utilizarán. En esa discusión es bueno diferenciar la naturaleza de la información pertinente para los alumnos y para los docentes, proponer documentos ágiles utilizables independientemente de la prueba como soportes para la enseñanza y el aprendizaje que promuevan la reflexión, evitando así el riesgo de “adiestrar” para responder de manera exitosa una prueba o reducir las actividades de clase a “aprender a responder preguntas” para obtener buenos puntajes en las pruebas.

Dicho riesgo puede disminuirse si se logra que la evaluación se asuma como un proceso inherente a la actividad escolar, que da información importante para los evaluadores y los evaluados, también si se conforman equipos de trabajo en las instituciones y se investiga sobre los resultados y los factores asociados a los mismos. En síntesis, si se consigue promover nuevas visiones sobre la evaluación externa y su importancia para impulsar transformaciones en la cultura escolar.

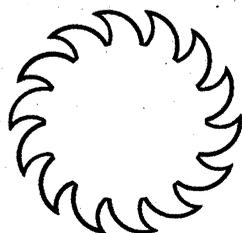
### Aspectos relevantes en el proceso de composición de la prueba

La elaboración de la prueba implica la toma de decisiones importantes respecto al enfoque

que se adopta y desde el cual se organizan los componentes y las especificaciones de la prueba. La consistencia del enfoque y de las categorías establecidas contribuye a la claridad por parte de quienes elaboran las pruebas para evitar las distorsiones y las incongruencias entre los propósitos declarados y lo que efectivamente se solicita en la evaluación. Así mismo, la precisión que se alcance en el fundamento teórico, las categorías, los dominios y los grados evaluados en las pruebas permite eliminar la ambigüedad en las posibles respuestas y la interpretación de los resultados.

En tal sentido, frente a la evaluación por competencias se señala la necesidad de concretar aún más las definiciones empleadas con el fin de evitar confusiones derivadas de la superposición de teorías que pueden orientarse hacia fines distintos de los señalados en el marco teórico de la evaluación externa. Igualmente, resulta perentorio establecer de manera explícita la articulación del enfoque de competencias, declarado en las pruebas, con los fundamentos expresados en los estándares, los indicadores de logro y los lineamientos curriculares. La convergencia y la armonización de planteamientos expresados en estos documentos oficiales, así como la posibilidad de hacer visibles tales articulaciones ante la comunidad educativa, pueden contribuir a clarificar las dudas existentes y a evitar el fraccionamiento y la dispersión de las políticas educativas. Se trata así de destacar los aspectos comunes en el desarrollo de aprendizajes significativos para los alumnos en todos los grados de la escolaridad y la centralidad del fortalecimiento de habilidades para la vida, en la formación del ciudadano.

Según Ana Atorresi (2005a), justamente el enfoque de las pruebas surge del análisis de los elementos anteriores. En tal sentido, “los requerimientos del estado del arte a la educación lingüística o matemática para la formación de ciudadanos, da cuenta de los avances de la lingüística y de la didáctica de la lengua, y de la matemática y de la didáctica de la matemática en relación con el avance de las sociedades”. El elemento referido a los programas de estudio “da cuenta del currículum prescripto”, y los libros de texto “ofrecen una aproximación al currículum implementado”. También es importante incluir un cuarto elemento, relacionado con la investigación de lo que efectivamente ocurre



## La superación de los prejuicios existentes sobre la evaluación implica cambiar el uso restringido que han tenido sus resultados.

en el aula, aunque aún son poco divulgados los trabajos que aportan informaciones sistemáticas al respecto<sup>2</sup>.

Es importante resaltar que la adopción del enfoque por competencias y/o habilidades para la vida, supone, entre otros propósitos, "usar información escrita y oral necesaria para la participación solidaria, crítica y creativa del ciudadano en la sociedad, apropiarse de información significativa para el propio desarrollo y la autonomía y aportar al desarrollo de la sociedad mediante las propias capacidades" (Atorresi, 2005a). A estos fines en la formación, agrega, les subyacen concepciones sobre la matemática y el lenguaje tomadas como herramientas para conocer mejor el mundo que nos rodea y actuar en él en forma consciente, creativa y autónoma. En consecuencia, para que el aprendizaje sea significativo los conocimientos que se imparten al interior de las aulas o que se proponen como objeto de la evaluación tienen que estar vinculados con la realidad y el mundo cotidiano. Igualmente, se requiere que los conceptos sobre los cuales descansa la estructura de las disciplinas escolares tengan énfasis especial tanto en las prácticas de enseñanza como en la evaluación.

En el caso del lenguaje, las habilidades o competencias que el usuario de una lengua requiere dominar para poder comunicarse eficazmente son las de hablar, escuchar, leer y escribir. Sin embargo, dadas las condiciones de las evaluaciones masivas, éstas con frecuencia se limitan al campo de la lectura debido a dificultades inherentes a la validez y la confiabilidad de los resultados frente a las demás habilidades. Para ser coherente con este enfoque respecto a la lectura, se requiere la inclusión de textos auténticos, es decir, los que circulan en la sociedad, de modo que la prueba permita establecer en qué medida la formación impartida prepara a los alumnos para enfrentarse a situaciones de lectura en contextos efectivos de comunicación.

Los lineamientos curriculares, los estándares y los diversos materiales que circulan en el aula, permiten definir, con mayor grado de comprensión sobre la realidad educativa de los sujetos evaluados, la

tabla de especificaciones de las pruebas, y obtener un balance adecuado entre lo declarado en el currículo y lo que efectivamente se promueve en las aulas.

### Consideraciones sobre la selección de textos

Según lo expuesto, en el caso de la lengua castellana se requiere contar, para la elaboración de los ítems, con soportes textuales auténticos que permitan establecer el dominio de estructuras como la narración, la argumentación, la descripción y la explicación y su expresión en diversos géneros discursivos, como el cuento, la leyenda, la receta, la biografía, entre otros. Esto significa que los ítems cuentan con un contexto específico de referencia y que ninguno de ellos debe aparecer como pregunta aislada, sino articulados a una situación de comunicación propuesta. De allí que resulte tan importante la selección de un texto alrededor del cual se organiza la prueba, debido a su relación con los objetivos del currículo, su interés para los evaluados y su relevancia para la cultura de un país o una región.

Según Atorresi (2005a), puede justificarse la inclusión de un texto en la prueba, teniendo en cuenta que:

- No es especialmente conocido para un sector social.
- Tiene un tema almacenado en parte en la memoria a largo plazo de los niños, lo cual permite poner en juego fácilmente lo conocido para operar sobre lo nuevo.
- Propone un tema que coincide con el de otros textos.
- Cuenta con una extensión que permite una lectura relativamente rápida y una retención de buena parte de la información en la memoria operativa.
- Ofrece diversos elementos motivadores que responden a las características de los niños de las edades consideradas.

- En varios casos combina los códigos verbal y visual, y éstos se complementan.
- Tiene una tipología y un grado de complejidad diferentes respecto a otros textos incluidos en la prueba.
- No contiene datos matemáticos o científicos que exceden el nivel de conocimiento marcado por el análisis curricular de las otras áreas.
- Contiene una información y una estructura tales que dan origen al menos de cinco ítems adecuados al nivel.
- No presenta errores de coherencia, cohesión, gramática, ortografía y demás.

Definidos los criterios frente a éstas y otras consideraciones relevantes para la selección de los textos sobre los cuales se organiza la prueba, es importante establecer la coherencia entre los ejes temáticos seleccionados, de modo que posibiliten la integración, faciliten la interdisciplinariedad y ayuden a mantener la atención del niño en torno de un mismo referente. Es bueno resaltar que se privilegien los textos breves, a menos que capten el interés y la atención de los lectores y se ajusten al tiempo destinado para la prueba. Igualmente, es necesario establecer qué categorías, clases y géneros textuales se mantienen con referencia a pruebas anteriores, con el fin de tener parámetros de comparación frente a logros, grados de dificultad y progresión en los aprendizajes evaluados.

Preguntas como las siguientes pueden ayudar a la toma de decisiones al respecto:

- ¿Justifica el número de ítems la inclusión de un texto en la prueba?
- ¿Existen relaciones de continuidad en los ejes temáticos de la prueba que faciliten al niño mantener su atención?
- ¿Son auténticos los textos seleccionados?
- ¿A qué tipo de situaciones de comunicación están vinculados los textos?

<sup>2</sup> Buena parte de este propósito puede alcanzarse mediante la creación de una Red Nacional de Investigación en Evaluación, proyecto al cual ya se ha hecho mención en ese escrito.

- ¿Qué relación tienen las habilidades evaluadas con las habilidades para la vida?<sup>3</sup>.

Del mismo modo, también se necesita reflexionar sobre:

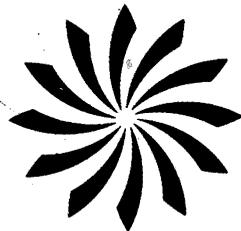
- La extensión de los textos propuestos;
- La clase de textos;
- La cantidad de datos de los textos;
- La cercanía entre los datos de los textos;
- El grado de abstracción de la información;
- El tipo de tarea solicitada (reconocer información, inferir...);
- El grado de exigencia a la memoria;
- La complejidad de las estructuras sintácticas;
- La cantidad de implícitos y suposiciones;
- El grado de empleo del lenguaje figurado;
- El tipo de relaciones lógicas presentes en el texto.

Estos criterios, aunados a las consideraciones relacionadas con las finalidades, los enfoques, los conceptos básicos y otros aspectos disciplinarios e interdisciplinarios pertinentes, según las áreas de conocimiento y los acuerdos logrados en las comunidades académicas y los grupos de trabajo, contribuyen a la selección adecuada de los textos sobre los cuales se organizan los distintos ítemes de una prueba.

### Respecto a la elaboración de ítemes

La construcción de la prueba y el procedimiento para la elaboración de ítemes toma como punto de partida la definición de criterios generales sobre la evaluación, las orientaciones específicas de las áreas por evaluar, las propuestas individuales de los miembros del colectivo de trabajo, los acuerdos globales relacionados con las orientaciones teóricas de la disciplinas y las consideraciones de orden estadístico.

Un procedimiento ágil para validar la consistencia de los ítemes puede iniciarse con la elaboración individual de las preguntas en las respectivas áreas, para proceder luego a la revisión individual por medio del intercambio entre los miembros de este colectivo. La confrontación de criterios y perspectivas adoptadas en la formulación de los ítemes con respecto a la tabla de especificaciones de la prueba contribuye a consolidar el enfoque adoptado y a eliminar ambigüedades, para proceder luego a su presentación ante el colectivo del área para afinar detalles y llegar a consensos. Finalmente, realizar la socia-



lización y la discusión ante las colegiaturas de todas las áreas mediante la estrategia de taller, facilita la superación de las lecturas fragmentadas de los textos que conducen a la adopción de visiones disciplinarias e individuales. El trabajo colectivo también favorece la identificación de errores conceptuales e imprecisiones, así como la ubicación de los detalles que pueden inducir a fallas en la coherencia global de la prueba.

La elaboración de los ítemes expresa la consistencia del enfoque puesto que los tipos de preguntas formuladas permiten establecer las concepciones subyacentes. Al respecto, es conveniente formularse, entre otras, preguntas como:

- ¿Sobre qué indaga un ítem?
- ¿Qué tipo de tareas cognitivas se proponen en un ítem?
- ¿Qué relación existe entre la “respuesta correcta” y las “otras opciones”?
- ¿Qué relación existe entre la selección de los textos y la construcción de ítemes con las prácticas habituales en la escuela?
- ¿Cuál es el grado de dificultad de la prueba?

Ana Atorresi (2005a) plantea que las razones técnicas para la inclusión de ítemes son de dos tipos: el juicio de los expertos a cargo y el análisis estadístico. Frente a las razones del experto, considera:

- Cualquiera de las opciones puede ser correcta y debe ser coherente en su relación con el enunciado.
- El enunciado y las opciones no presentan asociaciones verbales que puedan señalar la clave.
- Los distractores aportan diferentes informaciones.
- El ítem no presenta ambigüedad.

Del mismo modo, las razones por las cuales el análisis estadístico rechaza un ítem piloteado son las siguientes:

- Dificultad muy alta o muy baja;
- Validez baja o negativa;
- Funcionamiento excesivamente atractivo de uno de los distractores;
- no diferenciación de los sujetos con habilidades más altas, medias o más bajas;
- Estimación de la dificultad a priori;
- Estimación de la dificultad a posteriori.

### Sobre la presentación de las pruebas

Otro aspecto importante está relacionado con la organización adoptada para la presentación de los textos y la distribución de los ítemes que la componen. Es fundamental que la diagramación del cuadernillo y de las ayudas que lo acompañen (diario, mural u otro tipo de texto) sean atractivas para los evaluados, que haya un encadenamiento lógico entre las áreas que intervienen, los bloques y las temáticas seleccionadas. De igual modo, es útil tener en cuenta que las preguntas se organicen partiendo de lo general, a lo específico y lo particular para, finalmente, incorporar las de carácter correlacional que demandan una comprensión más compleja de los textos.

### Importancia de las estrategias de comunicación sobre los resultados de las pruebas

Teniendo presente el peso de los resultados de las pruebas en la toma de decisiones de orden institucional y sus efectos sociales en los beneficiarios directos e indirectos de las mismas, es fundamental definir un plan de socialización que supere la publicación de listas. En consecuencia, se necesita prever acciones para la interpretación y el análisis de resultados que contribuyan a hacer de éstos un medio para avanzar en la formación.

<sup>3</sup> Así, por ejemplo, el enfoque de evaluación impulsado por la Organización para la Cooperación y el Desarrollo Económico (OCDE) (1999) por medio del Programa Internacional para la Evaluación de Estudiantes (PISA), "no se propone evaluar el aprendizaje de contenidos curriculares específicos, sino el dominio de un conjunto de conocimientos y competencias que los jóvenes necesitarán para su desempeño en la vida como ciudadanos. El énfasis estará puesto en el dominio de procesos, la comprensión de conceptos y la capacidad para desempeñarse en diversas situaciones dentro de cada campo". Administración Nacional de Educación Pública, p. 13.

Una de ellas puede consistir en la elaboración de un documento analítico en el cual se proponga el análisis de las preguntas “acertadas” y la explicación de los “errores” y las posibles razones que los generaron. Los equipos directivos de las instituciones pueden propiciar espacios de reflexión sobre la información obtenida, con el objeto de introducir transformaciones en las prácticas que tengan mayor incidencia en los resultados alcanzados.

Así mismo, se necesita proponer desde las entidades responsables de la evaluación, de la formación, y de los equipos de trabajo e investigación sobre el currículo y la evaluación, un conjunto de recomendaciones para los docentes, los alumnos y las instituciones, que faciliten acciones de mejoramiento de acuerdo con los resultados obtenidos. Algunas de ellas estarán relacionadas con problemas de dominio disciplinario e interdisciplinario, enfoques didácticos y pedagógicos, disposiciones de los sujetos hacia la evaluación, entre otras.

Una de las recomendaciones formuladas por la profesora Atorresi respecto a los resultados de la evaluación es la de mantener presente que son los docentes, a cargo de la educación en el área, los destinatarios de los mismos, y que su finalidad es la toma de decisiones que conduzcan al mejoramiento de las prácticas pedagógicas. Aunque los informes orientados a este tipo de interlocutores cuentan con muy poca tradición, los resultados se explican, en gran parte, a partir de lo que ocurre en las escuelas. Por consiguiente, son los docentes quienes pueden protagonizar un proceso de transformación. Por ello Atorresi insiste en la necesidad de elaborar informes cuyos destinatarios y protagonistas sean los equipos de docentes que participan en la evaluación, los cuales serán complementarios de las recomendaciones metodológicas y los cursos de acción definidos por el nivel central.

### Algunas conclusiones

En general, los planteamientos formulados por la profesora Ana Atorresi durante el Seminario Evaluación externa y elaboración de ítemes, las reflexiones y las discusiones generadas a partir de sus puntos de vista, así como las consideraciones personales expuestas en estas notas, señalan la necesidad de trabajar de manera solidaria entre las distintas instancias comprometidas con los procesos de

## Resulta perentorio establecer de manera explícita la articulación del enfoque de competencias, declarado en las pruebas, con los fundamentos expresados en los estándares, los indicadores de logro y los lineamientos curriculares.

evaluación para promover cambios significativos en todos los grados de la escolaridad.

Un elemento que se destaca en este evento tiene que ver con la necesidad de avanzar en la evaluación en tanto proceso que contribuye a superar las debilidades que puedan existir en el enfoque adoptado para las pruebas nacionales. Así, mediante la discusión de pares y la confrontación de puntos de vista, sean opuestos o complementarios, se advierten aspectos sobre los cuales se requiere profundizar teóricamente con el fin de dar mayor consistencia a los instrumentos utilizados.

Para los colectivos que trabajan en el campo de la evaluación es de gran importancia conocer experiencias y visiones externas en relación con sus enfoques, sus procedimientos y sus perspectivas. Este tipo de encuentros, así como los espacios de discusión y reflexión que se generan, ayudan a validar la universalidad de los planteamientos que orientan la evaluación en nuestro medio y a asumir con rigor las particularidades de nuestra tradición académica y nuestra diversidad cultural.

El lugar natural de la investigación y de la academia se sitúa en los encuentros de pares y en el cuestionamiento permanente sobre aspectos nodales de los objetos de estudio. Por consiguiente, la formulación de interrogantes multifocales e interdisciplinarios —como ocurre en el caso de la evaluación externa— implica la reorientación tanto del sujeto como de las instituciones en una doble dirección: “hacia una nueva relación con el conocimiento y hacia una redefinición del propio objeto” (Bialakowsky, 1992).

En tal sentido, el seminario cumplió con el propósito de promover la reflexión sobre temas críticos de la evaluación externa, frente a los cuales se requiere profundizar a fin de garantizar la cualificación de las prácticas educativas y de evaluación en el país.

Igualmente, señaló la necesidad de continuar en la búsqueda de mecanismos cada vez más eficaz para lograr consolidar perspectivas interdisciplinarias en el proceso de construcción de las pruebas y la implementación de acciones creativas derivadas de sus resultados. Esto, con el fin de superar las fragmentaciones y las desarticulaciones causadas por la ausencia de diálogo entre los expertos que elaboran las pruebas, los docentes de los distintos grados de la escolaridad que las toman como referencia para el trabajo en el aula y los destinatarios directos e indirectos de la evaluación.

### Bibliografía

- Administración Nacional de Educación Pública. 2000. *Evaluaciones nacionales de aprendizajes en educación primaria en el Uruguay*. Montevideo.
- Atorresi, Ana. 2005a. Material del Seminario Evaluación externa y elaboración de ítemes. Bogotá, D. C.
- \_\_\_\_\_. 2005b. Recomendaciones para el uso de información en evaluación, Cfr. En *Boletín*, No. 1 de la Colegiatura de Lenguaje, Icfes, Bogotá, D. C.
- Bialakowsky, Alberto. 1992. La mono-multi-inter y transdisciplina. El objeto, los sectores y el sujeto. Algunas reflexiones sobre el proceso interdisciplinario. Memorias Primer Congreso Latinoamericano de Investigación Cualitativa: Medellín.