

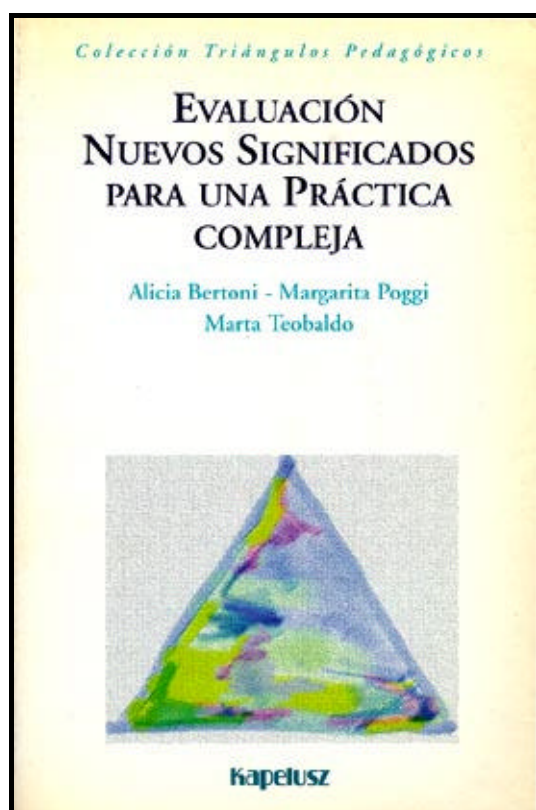
EVALUACIÓN NUEVOS SIGNIFICADOS PARA UNA PRÁCTICA COMPLEJA

Alicia Bertoni – Margarita Poggi
– Marta Teobaldo

Colección Triángulos Pedagógicos

Dirección: Martha Güerzoni de García Lanz

Directora de colección: Dra. Graciela Frigerio



KAPELUSZ editora s.a.,
1996, Buenos Aires,
Libro de edición argentina.

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

INDICE

Prólogo	5
Introducción	7
1. Los significados de la evaluación educativa: alternativas teóricas.....	11
La articulación del campo educativo con otros campos: implicancias para la evaluación	
Los usos sociales de la evaluación	
Algunas observaciones sobre la acción de evaluar	
Los múltiples significados de la palabra "evaluación"	
2. Reflexiones, dilemas y posiciones metodológicas	31
El contexto de las decisiones teórico-metodológicas	
¿Evaluación externa o evaluación interna?	
¿Evaluación cuantitativa o evaluación cualitativa?	
Los niveles para gestar un sistema de evaluación	
3. La evaluación de los "saberes" aprendidos	49
Los grupos intervinientes en la construcción de las pruebas de evaluación	
¿Qué se evalúa a partir de los instrumentos de evaluación?	
Las claves de corrección y el esquema de valoración	
Los criterios en la administración y corrección de las pruebas	
4. La evaluación de las pruebas de evaluación	69
La confiabilidad de las pruebas de evaluación	
El concepto de validez	
Otras pretensiones para definir la validez en los enfoques cualitativos	
La búsqueda de equivalencia en los controles metodológicos	
5. La función de la devolución	85
El informe de evaluación	
Concepción de evaluación y modalidades de devolución	
La devolución como instancia de la evaluación formativa	
Sobre el referente	
Sobre la concepción de enseñanza y de aprendizaje y las metodologías de enseñanza	
La evaluación formativa en la práctica docente	
Sistema de evaluación y evaluación formativa	
6. Quince premisas en relación con la evaluación educativa	97
Bibliografía	105

3. LA EVALUACIÓN DE LOS "SABERES" APRENDIDOS.

Los grupos intervinientes en la construcción de las pruebas de evaluación

La evaluación de lo que "saben" los alumnos, como resultado de los procesos de enseñanza y aprendizaje que les son pertinentes, nos conduce al análisis de los problemas relativos a la construcción, administración y corrección de las pruebas de evaluación.

La descripción del proyecto realizado en el ámbito de la MCBA da a conocer cómo fueron abordados y resueltos estos aspectos en cuanto a lo metodológico y lo procedimental.

El primer interrogante que se plantea con respecto a la construcción de pruebas es: ¿quiénes las construyen?

La incorporación de grupos de docentes en la elaboración de las pruebas

La incorporación de los maestros de grado en la etapa de la construcción de las pruebas constituye un reaseguro de la correspondencia entre "lo enseñado" y "lo evaluado". En nuestro caso se formaron cuatro comisiones con docentes previamente reconocidos por supervisores, directores y colegas, como profesionales que contaban con experiencia, trabajos especiales y/o formación académica en Lengua, Matemática, Ciencias Sociales y Ciencias Naturales. Como coordinadores del equipo técnico de la DIE sostuvimos reuniones con estas comisiones para la discusión de las características del trabajo por realizar y el análisis de las orientaciones y criterios para llevarlo a cabo.

Se acordó que cada comisión, en cada área curricular, elaborara el mayor número posible de ítems, tomando como referencia los siguientes criterios:

- representatividad de los propósitos y contenidos que hubieran explicitado el 80% de las escuelas, o más, como parte del currículum enseñado en ellas;
 - validez de contenido de los ítems;
 - inclusión de diferentes tipos de ítems (variedad);
 - presencia de ítems con distinto grado de dificultad;
 - claridad en la formulación de consignas y enunciados;
- presencia de diferentes tipos de contenidos.

El nivel de productividad de las comisiones docentes permitió contar con el material apropiado para efectuar el análisis de los contenidos que los maestros desarrollan, las conceptualizaciones en que los sustentan y los enfoques de los que parten en la enseñanza de cada área.

La incorporación de especialistas de contenido en la revisión de ítems y selección de la forma final.

Como venimos sosteniendo, la construcción de un sistema de evaluación para el nivel primario implica un proceso gradual si se pretende que contribuya a mejorar la calidad de la educación. A medida que se lo lleve a cabo, deberá promover instancias para que la actualización de la práctica docente incluya la apropiación adecuada de principios y técnicas de evaluación.

La ejecución de este primer proyecto asociado a la construcción de tal sistema promueve el acercamiento entre los maestros y los especialistas. Este acercamiento desborda, por supuesto, la problemática de la evaluación, pero a través de él, se visualiza la conexión con el importante tema de la transposición didáctica, por lo que debemos señalar la pertinencia de esta última con respecto al problema central del presente proyecto.

Al comenzar la investigación, se informó a los docentes que la construcción de las pruebas de evaluación se iniciaba desde su misma práctica. Pero también se les informó desde el principio que las tareas de construcción no se limitaban a ese único aporte.

Por otra parte, en el diseño de los instrumentos se requiere el juicio de los especialistas para cumplimentar las exigencias técnicas que son complementarias, y no opuestas, de las que deben satisfacer los maestros al redactar los ítems.

Los especialistas consultados de las áreas de Lengua, Matemática, Ciencias Sociales y Ciencias Naturales leyeron la producción docente correspondiente a su área profesional. Recomendaron la inclusión de aquellos ítems que, a su juicio, satisficieran los siguientes criterios:

- validez de contenido;
- consistencia interna (ítems referidos a un mismo propósito de logro);
- posibilidad de gradación del orden de dificultad;
- promoción de la producción de resultados comparables;
- diferenciación e integración de diferentes tipos de contenidos;
- posibilidad de lograr el *equilibrio* de la prueba (este criterio tiende a que todas las partes de la prueba contribuyan a resultados que estén de acuerdo con los propósitos generales).

En síntesis, la elaboración de las pruebas ha contado con el concurso de tres grupos, a cada uno de los cuales le correspondieron actividades y responsabilidades específicas:

1. El grupo formado por docentes de grado, quienes compusieron cuatro comisiones (una por cada área curricular evaluada). Estuvieron encargados de elaborar los ítems para componer la forma inicial de la prueba. Se estableció como requisito que los docentes que participaran de las comisiones no pertenecieran a las escuelas donde se iban a aplicar los instrumentos de evaluación.

2. El grupo de especialistas, que fueron consultores del proyecto, y que asumieron la responsabilidad de analizar los ítems incluidos en la forma inicial y hacer una selección destinada a componer la forma final de las pruebas.

3. El grupo de investigadores de la DIE, encargados del primer análisis de los ítems construidos por los docentes, de la coordinación de la tarea con los especialistas, y responsables ante las escuelas de la forma final de cada una de las cuatro pruebas construidas.

¿Qué se evalúa a partir de los instrumentos de evaluación?

Como paso previo a la definición de los principios y criterios de carácter pedagógico que orientan la construcción de instrumentos de evaluación, conviene presentar una perspectiva más amplia en la que se enmarca esta cuestión.

Es interesante plantearse las siguientes preguntas en relación con la construcción de un sistema de evaluación:

- ¿la evaluación se centra en cada alumno en forma individual, a partir de la presentación de sus conocimientos sobre una disciplina o área?, o, por el contrario,
- ¿se centra en la estimación de los conocimientos de una disciplina o área a partir de la presentación que alumnos y alumnas hacen de sus conocimientos? (Kvale, 1992).

No se trata en este caso de un simple juego de palabras, sino que el modo en que se formula la pregunta y las respuestas que de ella se derivan adquieren implicancias desde la misma definición de qué es evalúa y para qué se evalúa un sistema educativo, así como implicancias pedagógicas a la hora de diseñar los instrumentos de evaluación.

Frente a la primera pregunta, aparecen en un primer plano las funciones de calificación, de clasificación y de selección que tradicionalmente han sido adjudicadas a las evaluaciones. A partir de allí, se legitima el establecimiento de distinciones entre los individuos, la construcción de escalas normativizadas en las que se los sitúa y, derivado de esta posición, el análisis de la influencia educativa de los exámenes en la motivación del aprendizaje de los alumnos, es decir, lo que algunos autores denominan "estudiar para el examen", con la consiguiente preocupación que esto también produce en los docentes.

Por el contrario, el segundo interrogante, aquel que parte de considerar la evaluación como una estimación de los conocimientos comunes de una disciplina o área, nos permite contribuir al análisis de los conocimientos válidos en los diferentes campos del saber. De este modo, "los exámenes son una valoración de los conocimientos sobre una disciplina, la comprobación y la confirmación anual de la calidad de los conocimientos impartidos y adquiridos en un ámbito determinado"¹. Desde esta perspectiva, la evaluación misma implica creación de conocimientos, y no sólo la "ilustración", en el momento de la prueba, de contenidos transmitidos en una disciplina. Popkewitz afirma que la evaluación supone "un problema de

¹ S. Kvale en *Revista de Educación* N° 299, 1992, p 122.

producción social que implica relaciones sociales y dimensiones de poder"². En otros términos, la cuestión de la evaluación entraña conocer y comprender mejor las tensiones, las luchas, los dilemas y los conflictos que subyacen en el mejoramiento de la calidad y en las condiciones en las que se produce y genera el conocimiento en las escuelas.

A partir de la evidencia de que la problemática del conocimiento es central en la construcción de un sistema de evaluación, deben las ciertas cuestiones en la elaboración de los instrumentos de evaluación. Las mismas se encuentran estrechamente relacionadas y vinculadas entre sí, como veremos a continuación.

En primer lugar, la **problemática de los contenidos** lleva a conceptualizar qué se entiende por contenidos escolares, cuestión que, por otra parte, alcanza a otros aspectos de la acción educativa que exceden a la evaluación y son objeto de tratamiento en distintos ámbitos de un sistema educativo, como la definición y concreción de los contenidos pertinentes en un diseño curricular, su traducción en un proyecto pedagógico institucional, o su abordaje en la acción educativa cotidiana.

En segundo lugar, es necesario especificar **qué contenidos serán seleccionados y priorizados** para la organización de los ítems que integran el instrumento de evaluación.

Para la resolución de esta cuestión se ha optado por realizar un relevamiento de los objetivos y contenidos elegidos y priorizados, en los establecimientos educativos, por los maestros de las cuatro áreas sobre las que se elaboran posteriormente las pruebas.

En tercer lugar, es importante definir **qué tipo de instrumento de evaluación** será elaborado para indagar sobre las características y el nivel de aprendizaje de los contenidos seleccionados para integrar la prueba.

Al respecto, conviene reseñar alguna de las críticas, cada vez más extendidas, que se formulan a las pruebas denominadas "objetivas", por lo menos en cuanto a las implicancias y derivaciones con respecto a la cuestión del conocimiento y de los contenidos. Entre ellas podemos mencionar las siguientes cuestiones:

- Los conocimientos tienden a descontextualizarse a partir de las respuestas breves o de la selección entre las opciones proporcionadas en el instrumento de evaluación.
- Este tipo de pruebas produce información sobre el volumen cuantitativo de conocimiento, pero no siempre informa o da cuenta de los matices y las diferencias cualitativas.
- En ocasiones se priorizan las reglas formales: de este modo el conocimiento aparecería compuesto por hechos, datos, informaciones, que deberían relacionarse mediante la aplicación de reglas lógicas.
- Este tipo de instrumento puede conducir a una simplificación de los conocimientos en lugar de favorecer el desarrollo de procesos cognitivos complejos.
- Se produce lo que algunos teóricos denominan "la instrumentalización del aprendizaje": el conocimiento adquiere un valor de uso e intercambio, del que se rinde cuenta en el momento de la prueba.
- No se deja espacio para la ambigüedad y la complicitad de los conocimientos y mucho menos para cuestionar el conocimiento mismo.

Tanto las pruebas de completamiento como aquellas que requieren respuestas breves y prácticamente unívocas, como las pruebas de opción múltiple, proporcionan una información muy empobrecida sobre el conocimiento de los alumnos porque apelan a procedimientos intelectuales poco complejos. Elliot (1990) ha señalado que este tipo de pruebas opera sobre bases muy reduccionistas, equiparando las definiciones formales que se dan como respuesta a la comprensión de su significado. Se produce así una reducción de la comprensión de los conceptos a la comprensión de los términos y de ésta "a la de recuerdo de las definiciones formales"³.

No obstante los aspectos señalados como críticas a este tipo de instrumentos, puede afirmarse que proveen indicadores muy simples para empezar a pensar sobre los resultados de aprendizaje de los alumnos.

² Popkewitz en *Revista de Educación* N° 299, 1992, p 98.

³ J. Elliot, *La investigación-acción en educación*, Madrid, Morata, 1990.

Con el propósito de superar parte de las críticas señaladas a las pruebas de tipo "objetivo", el proyecto de investigación nos permitió avanzar en la construcción de instrumentos que progresivamente incluyeran ítems que apuntaran a respuestas de mayor grado de complejidad, desde el punto de vista de los procesos cognitivos implicados. Vamos a analizar más detalladamente las distintas cuestiones mencionadas en este apartado.

Los contenidos escolares en los instrumentos de evaluación, su selección para la evaluación y el tipo de instrumento elaborado

La cuestión de los contenidos escolares nos confronta como educadores con diversos aspectos, los cuales pueden ser abordados desde algunos de los siguientes interrogantes: ¿qué enseñamos en la escuela? ¿es significativo lo que enseñamos?, ¿qué evaluamos cuando realizamos evaluaciones?, ¿cómo evaluamos aquello que queremos evaluar? Estas preguntas, y otras que podemos seguir formulándonos, encuentran algunas respuestas, aunque sea provisionarias, desde la experiencia que dio origen a estas reflexiones.

Para la definición de **contenidos escolares**, C. Coll plantea que "se entiende por contenidos el conjunto de formas culturales y de saberes seleccionados para formar parte de las distintas áreas curriculares en función de los Objetivos Generales de la Educación"⁴. A su vez, diferencia entre contenidos conceptuales, procedimentales y actitudinales.

Dentro de los límites que supone un instrumento como una prueba que podría ser administrada a una muestra importante de alumnos de las escuelas primarias, se trató en las diferentes áreas de incluir tanto contenidos conceptuales (hechos, conceptos y principios) como procedimentales y actitudinales (valores, normas y actitudes).

No se desconoce que la evaluación de procedimientos y actitudes más sistemática y con mayor confiabilidad requeriría de otros instrumentos diferentes de los seleccionados en este proyecto, pero no obstante se trató de no reducir los ítems elaborados a la simple constatación de la presencia o ausencia de determinados hechos o conceptos.

En consecuencia, incluiremos algunas consideraciones sobre los contenidos actitudinales y procedimentales por la importancia que reviste integrarlos, en la medida en que sea posible, en un instrumento de evaluación.

Existen, en relación con los contenidos, tendencias predominantes y contrapuestas. Una de ellas, presente en las concepciones más tradicionales, reduce los contenidos casi exclusivamente a hechos y conceptos. Como contrapartida de esta posición, puede encontrarse otra tendencia que centra su formulación en valores o actitudes muy genéricos, incluso no siempre considerados como contenidos escolares sino como grandes fines o propósitos de la acción educativa. Estos valores generales presentan dificultades en el momento de ser evaluados tanto en el nivel de un proyecto curricular como en el de los logros de aprendizaje de los alumnos. Esta última posición encuentra además otras dificultades derivadas del hecho de que algunas de las actitudes deseadas en los alumnos no son de incumbencia exclusiva de la institución escolar, sino que la familia, el medio social, los medios de comunicación, etc., pueden contribuir a su adquisición.

También en relación con los **contenidos actitudinales**, más allá de la tendencia mencionada antes de considerarlos como grandes marcos referenciales, puede agregarse que si no son explicitados como contenidos pasan a formar parte del currículum oculto o no escrito (Dreeben).

La conceptualización de Coll sobre los contenidos permite abordar con mejores elementos técnicos esta cuestión tan controvertida, que, si bien se evidencia en el momento de la evaluación, se halla presente en toda práctica educativa.

Una cuestión que conviene recordar respecto de esta categorización es que se realiza con fines didácticos, es arbitraria (como toda otra clasificación) y eventualmente es reformulable.

Afirmar que se realiza a partir de un propósito centralmente didáctico significa destacar que, en la práctica, los contenidos escolares no pueden ser estrictamente separados en contenidos conceptuales, procedimentales y actitudinales. Esto significa, entonces, que, cuando enseñamos contenidos, enseñamos contenidos que hacen referencia a conceptos, procedimientos y actitudes simultáneamente. Por esta razón la clasificación es arbitraria y reformulable, pues constituye un modo de focalizar los contenidos y mirar aquello que enseñamos⁵.

⁴C. Coll, *Psicología y Currículum*, México, Paidós, 1991, p. 138.

⁵ Este punto puede ser ampliado en M Poggi(comp.), *Apuntes y aportes para la gestión curricular*, Buenos Aires, Kapelusz, 1995.

Avancemos con algunos ejemplos que pongan en evidencia lo que afirmamos ejemplificando con ítems que se elaboraron para las pruebas de 7° y 5° grado.

Ejemplos de ítems

CIENCIAS SOCIALES

Ítem 19. Pensá y recordá todas las guerras que estudiaste en tu escuela primaria o sobre las que escuchaste y viste algo en la televisión. Te damos ejemplos de algunas de ellas: las luchas por la independencia de las colonias americanas, las guerras mundiales, la guerra de Vietnam, la guerra del Golfo, la guerra de Medio Oriente, la guerra de las Malvinas.

¿Podés explicar qué significa esta frase: "En una guerra todos pierden"?

CIENCIAS NATURALES

Ítem 4. La relación armónica que existe entre los componentes de la naturaleza se llama *equilibrio ecológico*. Nombrá, por lo menos, *tres* factores con que el hombre modifica el medio perturbando el equilibrio natural.

Item 5. ¿Qué medidas adoptarías para evitar la destrucción del equilibrio ecológico? Nombrá tres medidas.

En los ejemplos citados no se encuentran involucrados solo contenidos actitudinales; de hecho suponen también el conocimiento de hechos y conceptos, pero implican, además, la construcción de normas y valores, y reflejan el nivel de respeto en la asunción de los valores a los que se alude. En ambos casos, ellos se evidencian con las claves de corrección presentadas.

Respecto de los contenidos procedimentales⁶, es importante señalar diversas cuestiones. La primera de ellas se refiere a que no implican necesaria ni prioritariamente manipulación de instrumentos, sino que implican procedimientos cognitivos.

Dichos contenidos tampoco indican, en sentido estricto, la metodología por enseñar en el aula, sino conocimientos que los alumnos adquirirán. Por supuesto que la construcción de ciertos procedimientos requiere de la implementación de metodologías adecuadas para facilitar esa construcción, pero no hay un isomorfismo entre métodos didácticos y contenidos procedimentales.

La segunda cuestión se refiere al hecho de que los procedimientos han sido, y siguen siendo, enseñados en las escuelas, pero en muchas ocasiones son planteados como propósitos educativos de un gran nivel de generalidad.

El punto central sobre el que es necesario reflexionar aquí se vincula en ocasiones con que el aprendizaje de los contenidos procedimentales parece haber quedado a cargo de los alumnos, lo cual hace que algunos de ellos puedan reconocer, por ejemplo, en un problema matemático cuál es el procedimiento más económico para su resolución (ya sea porque "ahorra" operaciones, porque conlleva una menor posibilidad de error, etc.), o que otros alumnos no "descubran" ese procedimiento como el más adecuado para resolver el problema si no media la intervención activa del docente en cuanto a enseñarlo explícitamente como contenido.

La tercera cuestión se refiere al hecho de que existirían procedimientos particulares, propios y específicos de ciertas áreas del saber y otros más generales, que pueden ser enseñados, y en consecuencia evaluados, a través de distintas disciplinas (por ejemplo, los procesos de búsqueda de información, análisis y síntesis de ella según distintos recursos, las estrategias de resolución de situaciones problemáticas).

En este sentido, contenidos procedimentales generales atraviesan las diferentes pruebas, tal como se ejemplificará más detalladamente con los siguientes ítems.

Ejemplos de Ítems

CIENCIAS SOCIALES

⁶ Coll define un procedimiento como "conjunto de acciones ordenadas y finalizadas, es decir, orientadas a la consecución de una meta", 1991, p. 139.

Te presentamos el siguiente cuadro:

EXPORTACIÓN DE ALGUNOS PRODUCTOS AGRÍCOLAS Y GANADEROS
(En pesos oro)

	AÑO 1890	AÑO 1900	1910
Carnes congeladas:			
vacuna	53.029	3.436.955	29.550.914
ovina	1.633.105	6.887.211	10.667.541
Lanas	26.144.053	25.884.928	58.847.699
Trigo	9.951.281	48.627.653	72.202.260

Fuente: *Historia Integral Argentina*, Buenos Aires, Centro Editor e América Latina, Tomo 5, 1971.

12. ¿Cuál fue el producto que aportó más dinero al país en 1890, en 1900 y en 1910?

En 1890:

En 1900:

En 1910:

13. Varios historiadores han afirmado que los productos agrícolas fueron una importantísima fuente de ingresos para el país y que a partir de 1904 ocuparon generalmente los primeros puestos en relación con los productos ganaderos. ¿Qué datos del cuadro que te presentamos permiten afirmar que el país era considerado agroexportador?

Como puede constatarse, se ejemplifican aquí algunos ítems de las pruebas de 7º grado, que en las áreas de Ciencias Sociales indagan sobre contenidos procedimentales como el análisis de cuadros con información estadística, pero además suponen contenidos conceptuales articulados con estos procedimientos, los que son específicos de cada una de las áreas mencionadas.

No obstante afirmar que existirían algunos procedimientos que atraviesan diferentes áreas, es necesario insistir en que, por lo menos desde algunas investigaciones, se ofrece evidencia de que estos procedimientos cognitivos ni se construirían al margen de los contenidos conceptuales de las disciplinas ni se generalizarían a cualquier campo disciplinario, independientemente de los contenidos conceptuales.

En este sentido, en lo que hace a la enseñanza de los contenidos procedimentales, debemos afirmar, como lo hicimos anteriormente, que su aprendizaje no se realiza al margen del abordaje de los contenidos conceptuales propios de las disciplinas o áreas escolares.

Un último aspecto que deseamos abordar respecto del tratamiento de los contenidos en la elaboración de las pruebas se refiere a **la articulación de los distintos ítems para construir situaciones de carácter más abarcador** que contextualizan los contenidos por evaluar.

Desde diferentes enfoques se viene enfatizando la importancia de un tratamiento de los contenidos de manera articulada e integral. Quienes trabajan sobre las redes semánticas o mapas conceptuales (G. Posner, J. Novak, B. Gowin) coinciden en enfatizar que la construcción y la apropiación del conocimiento se facilita a partir de propuestas didácticas centradas en la integración de los conocimientos. Éstos conforman redes y mapas conceptuales, que a su vez contextualizan nuevos conocimientos, posibilitan enfoques globalizadores, favorecen la comprensión y la retención de nueva información y, en definitiva, amplían los esquemas cognitivos y enriquecen las estructuras cognoscitivas de los alumnos.

Por ello, y a fin de no abordar conceptos en forma fragmentada y aislada, se construyeron, en la medida de lo posible, ítems que formaran parte de situaciones más abarcativas e integradoras. Éste fue también, en consecuencia, un criterio organizador de cada una de las pruebas, en particular en las de Ciencias Sociales y Ciencias Naturales .

En las pruebas de Matemática y Lengua, también se presentaron algunas situaciones integradoras, pero se optó por abarcar más contenidos, aunque en ese caso la integración se encuentra relacionada con las operaciones cognitivas involucradas para la resolución de los ejercicios propuestos.

Cabe aclarar que la inclusión de los ítems en situaciones que los relaciona no implica anular la independencia de cada ítem. Esto significa que la posibilidad de no poder resolver correctamente o resolver parcialmente un ítem no afecta la posibilidad de seguir respondiendo al resto de los ítems que integran la situación elaborada.

En **síntesis**, hasta ahora hemos avanzado en fundamentar los siguientes criterios priorizados para la elaboración de los instrumentos de evaluación:

- la inclusión de ítems que se centren en contenidos no sólo conceptuales sino que también incorporen contenidos procedimentales y actitudinales, aunque en el marco de los límites del tipo de pruebas realizadas;
- la construcción de situaciones que integren diversos ítems que contextualicen los contenidos seleccionados, aunque sin eliminar la independencia de los ítems en el interior de cada situación.

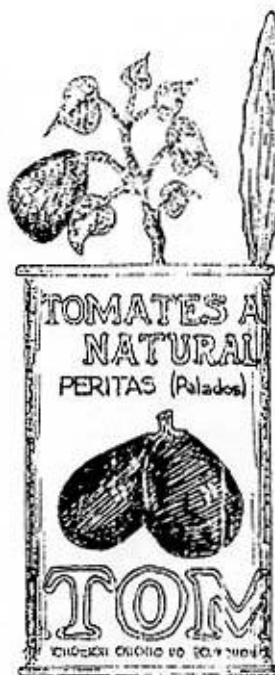
Ahora bien, los criterios para determinar qué tipo de contenidos ofrecen mayor grado de significatividad para integrar los instrumentos de evaluación deben combinarse, a su vez, con aquellos establecidos para seleccionar el tipo de instrumento por elaborar.

Como se desprende de las críticas que anteriormente se enunciaron para las pruebas de evaluación llamadas "objetivas", se intento avanzar en la construcción de instrumentos que progresivamente incorporaran ítems que involucraran procesos cognitivos de carácter más complejo, los cuales requieren mayor trabajo de los alumnos en la estructuración de las respuestas. Las pruebas integradas por este tipo de ítems habitualmente se consideran pruebas "abiertas". Para ilustrar lo dicho, podemos analizar más detalladamente el siguiente ejemplo:

Ejemplos de ítems y de claves de corrección

LENGUA

Ítems 17 a 21. Elaborá una publicidad promocionando este producto. ¡Atención! El dibujo sugiere algunas cualidades básicas del producto. Vos tenés que explicarlas con palabras en tu publicidad



Clave de corrección que corresponde a esos ítems:

En este ejercicio se discriminan 5 ítems, a cada uno de los cuales se le adjudicará un código, para dar cuenta de la calidad de la producción de los alumnos. Estos cinco códigos se colocarán en el margen derecho de la hoja con marcador de color, como en los otros ejercicios.

El código 0 corresponde a los 5 ítems si el alumno no hubiera respondido este ejercicio.

ÍTEM 17. Adecuación entre texto e imagen

3. Adecuada relación entre el texto y la totalidad de la imagen.
2. Considera el dibujo y/o el texto del cuerpo de la lata.
1. No considera las características de la imagen para elaborar el mensaje.

3. La publicidad alude a lo natural y/o fresca del producto, características que pueden inferirse de la planta que emerge de la lata. Ej.: Tienen el gusto de los tomates recién sacados de la planta. / Frescos como recién cortados. / Con un fresco y tentador aroma a tomates maduros. / Totalmente natural, sin aditivos ni especias.

2. Ej.: Pruebe exquisitos tomates al natural TOM. / Estos tomates al natural son exquisitos. / Nada igual a los tomates al natural TOM.

ÍTEM 18. Claridad en la producción del texto

3. Hay coherencia en la sucesión de ideas. Mensaje bien organizado.

2. Algunas frases incoherentes. Algunas frases incompletas. El mensaje se entiende pero no está suficientemente organizado.

1. Es muy incoherente. Repeticiones innecesarias. Frases incompletas. Mensaje confuso. No se entiende.

ÍTEM 19. Vocabulario

3. Riqueza de términos bien aplicados.

2. Aceptable empleo y variedad de términos.

1. Confusión de significación.

3. Ej.: Con un fresco y tentador aroma a tomates maduros. Totalmente naturales, sin aditivos ni especias. / Tal vez, probarlos le provoque tal satisfacción que ya no pueda dejar de comerlos. / Como un rico y sabroso tomate. Total gusto y color. Todos los detalles cuidados. Quizás comerlo provoque que toda su familia lo coma diciendo: - Mamá: ¡te felicito! O quizás, una alegría tan grande que lo va a comprar toda la vida. TOM va a pasar a ser un bien de familia.

2. Ej.: Compre los exquisitos tomates TOM. Naturales y sabrosos. Cómprelos a buen precio

ÍTEM 20. Puntuación.

3. Emplea adecuadamente los signos de puntuación.

2. Omite el empleo de algunos signos de puntuación.

1. Descuida en general el uso de signos importantes. No usa mayúsculas.

ÍTEM 21. Legibilidad

3. La letra no afecta la comprensión del contenido.

2. Algunas palabras no se entienden.

1. La letra impide que se pueda leer.

ÍTEMS 22 a 31. Vamos a contar la historia de una máquina inventada, la máquina de aprobar exámenes. Éstos son los hechos más importantes de la historia.

Los chicos inventaron la máquina de aprobar exámenes.

La probaron una vez.

Les fue mal.

La probaron otra vez.

Tuvieron éxito.

La máquina desapareció.

La buscaron.

(Falta el final)

La historia *tiene que "crecer"*; para ello, *tenés que*:

- Poner nombres a los personajes, decir cómo son, ubicarlos en tiempo y lugar.
- Inventar situaciones: fracaso de la primera prueba, éxito de la segunda, desaparición de la máquina, búsqueda.
- Darle un final a la historia.

Además podés:

- Incorporar otros personajes.
- Hacerlos hablar.

ATENCIÓN: Podés usar un borrador y después pasar en limpio tu redacción. Te pedimos que, junto con la prueba, entregues el borrador.

Clave de corrección que corresponde a estos ítems:

En esta situación se discriminarán 7 ítems, a cada uno dé los cuales se le adjudicará un código, para dar cuenta de la calidad de la producción de los alumnos. Estos siete códigos se colocarán en el margen derecho de la hoja con marcador de color, como en los otros ejercicios.

El código 0 corresponde a los 7 ítems si el alumno no hubiera respondido este ejercicio.

ÍTEM 22.

- 3. Incluye nombres y caracterización de los personajes.
- 0. No lo hace.

ÍTEM 23.

- 3. Inventa situaciones en torno de los núcleos propuestos.
- 0. No lo hace.

ÍTEM 24

- 3. Da final.
- 0. No cierra la historia.

ÍTEM 25. Creatividad

- 3. Reúne por lo menos tres de los requisitos considerados para estimar la creatividad.
- 2. Reúne uno o dos de los requisitos considerados para estimar la creatividad.
- 1. Revela un tratamiento rutinario, que refiere a situaciones cotidianas y corrientes, con un desarrollo lineal del texto.

Se consideran parámetros de referencia o requisitos para este ítem:

- a- Caracterización imaginativa y pertinente de los personajes. Ej.: "Los chicos, Pedro (un gordito que siempre tenía alguna excusa para faltar a las pruebas), Juan (un vago total, que coleccionaba 'unos') y Walter (el intelectual de la clase), inventaron una máquina de aprobar exámenes..."
- b- Presentación de situaciones inéditas o novedosas.
- c- Desarrollo original del texto, que puede evidenciar una lógica no convencional.
- d- Diálogos ingeniosos entre los personajes.
- e- Búsqueda de finales imprevistos, pero congruentes con el texto elaborado.

ÍTEM 26. Claridad en la producción del texto

- 3. Hay coherencia en la sucesión de ideas. Mensaje bien organizado.
- 2. Algunas frases incoherentes. Algunas frases incompletas. El mensaje se entiende pero no está suficientemente organizado.

1. Es muy incoherente. Repeticiones innecesarias. Frases incompletas. Mensaje confuso. No se entiende.

ÍTEM 27. Estructura de las oraciones

3. Correcta coordinación y subordinación de oraciones.
2. Ligeros errores sintácticos.
1. Numerosos errores de coordinación y subordinación de oraciones.

ÍTEM 28. Vocabulario.

3. Riqueza de términos bien aplicados.
2. Aceptable empleo y variedad de términos.
1. Confusión de significación.

ÍTEM 29. Puntuación

3. Emplea adecuadamente los signos de puntuación.
2. Omite el empleo de algunos signos de puntuación.
1. Descuida en general el uso de signos importantes. No usa mayúsculas.

ÍTEM 30. Legibilidad

3. La letra no afecta la comprensión del contenido.
2. Algunas palabras no se entienden.
1. La letra impide que se pueda leer.

ÍTEM 31. Ortografía

3. Ningún error ortográfico.
2. De *uno a cinco* errores ortográficos.
1. *Seis* errores ortográficos o más.

Como se comprende a partir de los ejemplos presentados, la elaboración de las claves de corrección es un elemento central e imprescindible en la elaboración de pruebas que incluyan ítems con poco grado de estructuración de las respuestas y, en consecuencia, con un amplio margen de libertad para los alumnos que elaboran sus respuestas. Sobre esta cuestión avanzaremos en el siguiente apartado.

Las claves de corrección y el esquema de valoración

Abordar más detenidamente la cuestión de la construcción de las claves de corrección implica avanzar en la elaboración de una base de referencia común para asignar puntajes a los ejercicios y situaciones que integran una prueba. Asegurar claves de corrección para todos los ítems que integran una prueba coadyuva para que evaluadores distintos asignen puntajes equivalentes a un mismo ítem, o que el mismo evaluador asigne el mismo puntaje a un ítem momentos diferentes. De este modo se asegura, en parte, la confiabilidad del instrumento de evaluación.

Pero, además, las claves de corrección aseguran que la evaluación misma se convierta en un vehículo de información hacia los docentes, y de éstos hacia sus alumnos, respecto de cuáles son las variedades de respuestas solicitadas, en relación con los criterios que se fijan previamente, para asignar los puntajes a las respuestas de los alumnos.

La búsqueda de consistencia en la valoración de las respuestas de los alumnos tiene relación con la homogeneidad en la utilización de claves de corrección de los ítems y atribución de puntajes o categorías para la calificación de la realización que suscita el ítem.

El procedimiento para asignar un valor determinado a los resultados de una prueba no debe conspirar contra su validez. En todos los casos, la atribución de "notas", "puntajes", o medidas equivalentes, debe adecuarse a las condiciones de relevancia y consistencia de la valoración, independientemente del método de puntuación adoptado.

Por lo tanto, quienes evalúan las pruebas tienen que tener claras las especificaciones que corresponden a las realizaciones esperadas de quienes las responden y la valoración que será atribuida a tal realización.

Es decir, para evaluar hay que disponer de un conjunto de reglas que se refieran a la realización de cada ítem y su correspondiente valoración. Este conjunto de reglas está constituido por las claves de corrección y un "esquema de valoración". En los ejemplos presentados se muestra el esquema construido para algunos ítems de las pruebas de las áreas curriculares evaluadas, y la adecuación del mismo a las claves de corrección de cada ítem. Este esquema incluye la consideración de distintas contingencias, como la de que el alumno no conteste la pregunta o que sus respuestas sean en parte correctas y en parte incorrectas. Las claves de corrección especifican la amplitud y variedad del significado de las respuestas de los alumnos.

Las dificultades para elaborar las claves y el esquema de valoración son más notorias en la medida en que los ítems promueven la "libre expresión", la originalidad, la creatividad. Pero siempre es posible encontrar criterios para valorar estas características en las realizaciones de los alumnos.

Cuando es un solo examinador el que corrige las pruebas, el conjunto de reglas a que hemos aludido es necesario, pero se torna imprescindible cuando las pruebas son corregidas por numerosos examinadores, a los que hay que facilitar su tarea eliminando, en todo lo posible, la ambigüedad en la atribución de la valoración correspondiente.

La tarea hace a los parámetros que se deben consensuar para tener un marco de referencia desde el cual juzgar las realizaciones de los alumnos, de modo que la interpretación pedagógica de los resultados, en el marco de las condiciones institucionales en que se producen, sea mucho más inequívoca.

Los criterios en la administración corrección de las pruebas

Se sostiene a menudo que el evaluador externo contribuye fuertemente a la objetividad de la evaluación. Suele agregarse, además, que aquella característica puede sesgarse fácilmente cuando la evaluación es interna y establece que sea el propio maestro el que califique a sus alumnos.

El dilema planteado en términos de adjudicación "objetiva" de puntajes versus el peligro de una adjudicación "subjetiva" roza una cuestión ética más que una cuestión técnica.

Desde el punto de vista técnico, "puntuación objetiva" significa que "un evaluador asignará los mismos puntajes cuando vuelve a asignar una calificación a un test o examen, o bien que diferentes examinadores, frente al mismo conjunto de ítems, les otorgarán las mismas notas. Pero esto queda cubierto por las reglas que supone la condición de consistencia"⁷.

Desde el punto de vista ético, la cuestión nos plantea cómo puede la sociedad depositar confianza en la tarea cotidiana del maestro, incluido el acto educativo de evaluar a sus alumnos, si estamos dispuestos a retirar tal confianza cuando se trata de construir un sistema de evaluación para depositarla solamente en las virtudes concedidas a la evaluación externa. No es que esta última por el hecho de estar "fuera" de la escuela impida que el maestro la realice en forma competente. Pero una y otra evaluación tienen finalidades y procesos diferentes.

"La aplicación externa es llevada a cabo por examinadores que no pertenecen al establecimiento. La información obtenida tiene el carácter de una evaluación sumativa y sus resultados pueden ser muy útiles si son difundidos a la comunidad más amplia. La aplicación interna, en cambio, pretende entregar información acerca de los avances de los niños. La aplicación de las pruebas puede ser responsabilidad de los mismos profesores y la información obtenida estar destinada principalmente a ellos"⁸.

Lo importante es que se pueda otorgar credibilidad a ambas y esto sucede cuando los propósitos, principios y procedimientos que cada una sustenta están definidos claramente y no muestran contradicciones entre estos aspectos. Desde la perspectiva de la evaluación interna no se trata sólo de la confianza en la capacidad del maestro de evaluar con honestidad, prudencia y justicia. Se trata, en este caso, de **la confianza en la capacidad del maestro de compartir el control riguroso que supone un proceso de investigación evaluativa.**

⁷J. M. Thyne, *op. cit.*, 1978.

⁸ UNESCO, *Medición de la calidad de la educación básica: ¿Por qué, cómo y para qué? Una propuesta para el mejoramiento de la calidad y las nuevas demandas de información*, Santiago de Chile, 1992, p. 52.

La comunicación establecida entre investigadores y la escuela permite que los primeros expongan las características de este control y que "estas reglas de juego" sean aceptadas por las instituciones escolares y sus docentes.

El procedimiento adecuado indica que los maestros, asistidos por un miembro del nivel de conducción de la escuela y/o un supervisor y/o un miembro del equipo técnico, puedan estar encargados de corregir las pruebas administradas en las cuatro áreas. Por otra parte y antes de empezar la corrección, mediante un sorteo al azar, habrá de seleccionarse una muestra de pruebas para ser corregidas en los Distritos Escolares por los responsables del proyecto.

La forma de hacer el sorteo debe garantizar que la probabilidad de incluir la prueba de un alumno en un área dada sea independiente de la probabilidad de elección de la prueba respondida por el mismo alumno en otra área.

Contar con una muestra al azar de pruebas corregidas "externamente" a la escuela permite una comparación estadística con los resultados obtenidos al corregirlas dentro de las escuelas. ¿Qué se ha probado hasta ahora, en los tres años que hemos seguido este procedimiento? Que el grupo de maestros que evalúa las pruebas en la institución escolar puede efectuar una adjudicación de puntajes autoconsistente y consistente con la evaluación realizada por el equipo técnico. Si la valoración hecha por cada uno de los grupos intervinientes no revela inconsistencias significativas dentro y entre ellos mismos, estaremos más seguros de que pueden tenderse puentes entre la evaluación externa y la interna. Y será mayor la convicción de que ellos son necesarios cuando nos mueven propósitos de mejoramiento de la calidad de la educación.