REVISTA DE eDUCACIÓN

EDUCACIÓN Y DEPORTE

Cómo evaluar la comprensión lectora: alternativas y limitaciones

Manuel Montanero Fernández

SEPTIEMBRE - DICIEMBRE 2004

MINISTERIO
DE EDUCACIÓN
Y CIENCIA

CÓMO EVALUAR LA COMPRENSIÓN LECTORA: ALTERNATIVAS Y LIMITACIONES

MANUEL MONTANERO FERNÁNDEZ (*)

RESUMEN. En este trabajo se hace una exhaustiva revisión teórica y empírica de los principales enfoques y técnicas que se han utilizado a lo largo de los últimos años para evaluar las capacidades de comprensión lectora, así como sus presupuestos implícitos y limitaciones, que hacen imprescindible un análisis crítico de su validez. Teniendo en cuenta el tipo de inferencias lectoras que, en la mayoría de los materiales analizados, se toma como criterio principal de evaluación, se propone una clasificación que puede ser de utilidad a investigadores y profesionales a la hora de seleccionar o confeccionar pruebas de evaluación, así como para interpretar de un modo más realista sus resultados.

ABSTRACT. In this paper, we present an exhaustive theoretical and empirical review of the main approaches and techniques used in the last few years to evaluate reading comprehension skills, as well as implicit ideas and limitations, which need a critical analysis of their validity. Taking into account the type of reading inferences that is taken as the main criteria for evaluation, in most of the materials analysed, a classification for researchers and professionals I proposed to be used when selecting or designing evaluation materials and to interpret the results in a more realistic way.

INTRODUCCIÓN

Mejorar la comprensión lectora de los alumnos que finalizan la educación obligatoria es uno de los principales retos que debe afrontar el sistema educativo. La evaluación de esta compleja capacidad se ha utilizado de hecho, reiteradamente, como uno de los índices más discriminatorios, que pone de relieve las diferencias educativas entre regiones y países. Sin embargo, se echa en falta una reflexión rigurosa sobre la validez de los instrumentos empleados para este objetivo, así como sobre su utilidad para la demanda más importante: obtener una información que nos ayude a individualizar la respuesta educativa para los alumnos que no comprenden.

Las técnicas de evaluación más utilizadas se basan en la observación y el análisis de productos inferenciales, supuestamente relacionados con los procesos cognitivos de comprensión. Existen múltiples evidencias acerca de la relación entre la comprensión lectora y la eficacia ante la demanda de inferir información de un texto. Algunos trabajos recientes aportan indicios de que esta última tarea no es simplemente un síntoma, sino incluso la principal causa que explica el nivel de comprensión que alcanzan los

(*) Universidad de Extremadura.

Fecha de entrada: 5-02-2003 Fecha de aceptación: 25-06-2003

lectores (Cain y Oakhill, 1999). Normalmente, las inferencias se explicitan mediante la respuesta a preguntas objetivas o semiobjetivas, *después* de la lectura de un texto. Por el contrario, otras técnicas más recientes se han centrado en la observación sistemática *en curso* de la actividad inferencial de los sujetos durante la lectura. Finalmente, otras tareas, como la confección de resúmenes, protocolos de recuerdo o esquemas, evalúan también otros procesos verbales, más allá de lo estrictamente inferencial.

De acuerdo con los últimos modelos cognitivos de la comprensión, podríamos establecer una distinción inicial entre dos niveles de representación sobre los que operan todos estos procesos: la base del texto y el modelo de la situación. Entre las principales operaciones cognitivas que se suelen tomar como indicadores de la comprensión de la *base del texto* podemos señalar: la extracción del significado de las palabras desconocidas a partir del contexto; la extracción de la información esencial; las relaciones sintácticas, semánticas o retóricas entre las ideas del texto (conexiones referenciales, temporales, anafóricas, temáticas, expositivas…); la detección de problemas de cohesión o coherencia...

Entre los indicadores relacionados con la *representación situacional* podríamos señalar las inferencias sobre la intención comunicativa del autor o sobre el significado de recursos estilísticos y expresiones ambiguas, las inferencias sobre la información implícita en el texto, así como las acciones implicadas en la utilización del acerbo de conocimientos previos o, en la solución de las inconsistencias que [RTF annotation: Es necesario que el autor aclare (para incorporar o no el pronombre «se») si son las acciones las que plantean inconsistencias o se establece a título impersonal.] plantean con respecto a la interpretación del contenido del texto.

Casi todos estos criterios han sido explícita o implícitamente suscritos por el MEC (1996) para la evaluación psicopedagógica de las capacidades cognitivas implicadas en la comprensión lectora. Algunos tests consideran también las habilidades léxicas y subléxicas que intervienen en la segmentación fonológica, la discriminación visual de letras, palabras o pseudopalabras o la velocidad lectora como variables de comprensión. Sin embargo, el énfasis de la evaluación debe recaer en aquellos indicadores que permiten inducir si el lector es capaz de desarrollar activa y estratégicamente operaciones para inferir significados progresivamente más elaborados a partir de sus conocimientos previos. Esta información es, de hecho, la que puede resultar más útil para orientar la intervención psicopedagógica sobre la comprensión. En lo que sigue, vamos a exponer una revisión más detallada acerca de las principales alternativas de evaluación y cómo pueden contribuir al análisis de los diferentes aspectos de la capacidad inferencial y estratégica del lector.

EVALUACIÓN DE LA COMPRENSIÓN LOCAL Y REFERENCIAL

La actuación estratégica del lector comienza propiamente en la fase de conexión proposicional que da lugar a la microestructura de la base del texto. En este proceso nuestra memoria de trabajo debe mantener activa una selección de las proposiciones procesadas en ciclos previos para establecer el «hilo conductor» entre las mismas. Dicha *conexión referencial* debe inferirse en tres instancias, cada una de ellas, progresivamente más complejas.

En primer lugar, el sujeto debe inferir una primera conexión sintáctica entre las diferentes cláusulas de las oraciones. Los esquemas proposicionales y sintácticos del lector estipulan qué argumentos puede admitir el predicado de cada proposición y permiten, por otro lado, decidir la estructura sintáctica más plausible.

Posteriormente, el sujeto debe integrar las nuevas proposiciones y las activadas en la memoria operativa en cada ciclo de procesamiento. Para ello es necesario producir inferencias referenciales más complejas, *anafóricas, temporales o causales,* ayudándose de ciertas señalizaciones lingüísticas. La dificultad para comprender relaciones anafóricas depende del tipo de anáfora (los pronombres demostrativos y los sustantivos anafóricos plantean más dificultades), así como de lo alejado que se encuentre el antecedente. Otro subtipo de inferencias referenciales, las llamadas «conectivas», son aquellas que permiten conectar una parte temática que se viene tratando y otra que es comentario de la misma, es decir, lo nuevo que se predica del tema.

Pruebas objetivas y semiobjetivas

La mayoría de los instrumentos dirigidos, explícita o implícitamente, a evaluar este primer nivel de comprensión, se fundamentan en la formulación de *preguntas objetivas* o *semiobjetivas,* después de la lectura del texto. Se trata de preguntas breves cuya respuesta se ofrece cerrada (acompañada de distractores) o limitada a un par de frases, que se valoran en función de los anteriores criterios inferenciales.

El emparejamiento entre frases y dibujos que representan diferentes estructuras sintácticas es una de las tareas más simples que aparecen en el PROLEC (Cuetos, Rodríguez y Ruano, 1996). Para evaluar la comprensión referencial, el CL-4 (Alonso Tapia y Corrales, 1994) formula preguntas dirigidas a detectar si el sujeto es capaz de seleccionar correctamente el antecedente de una proposición conectada referencialmente con otra anterior, es decir, si es capaz de establecer una conexión anafórica o causal entre ambas, durante la lectura.

Las preguntas más habituales de este nivel de comprensión son, no obstante, las que demandan que el sujeto parafrasee una oración, con objeto de valorar si la recuerda como parte del texto o si comprende la relación entre dos proposiciones en función de un esquema causal o medio-fin más o menos explícito. Este es el caso, por ejemplo, de la prueba de comprensión lectora del EVALUA (García Vidal y González Manjón, 1996), donde los alumnos de primaria deben completar frases del tipo «El sendero estaba fresco porque…», eligiendo entre varios distractores sin el texto delante.

La prueba de comprensión lectora de Lázaro (1982), por su parte, incluye diversas tareas en las que el sujeto debe reconocer entre varias oraciones (que no aparecían literalmente en le texto), aquella cuyo significado se corresponde con una idea que sí está presente explícitamente. Esta última demanda (*sentence recognition*) se considera un buen indicador de la capacidad de transcender de un nivel de comprensión meramente *superficial* del contenido lingüístico de un texto, aunque no se requiera necesariamente la comprensión global del mismo.

Finalmente, existen evidencias de que los lectores expertos suprimen mejor y más rápidamente los significados irrelevantes aprovechando la información semántica que proporciona el contexto inmediato. En este sentido, las tareas de emparejamiento entre palabras relacionadas semánticamente en el texto o las preguntas sobre el significado de una palabra poco familiar o ambigua, que el sujeto debe inducir a partir del contexto inmediato de la propia frase o del párrafo, podrían considerarse también adecuadas para valorar la capacidad de comprensión lectora en este primer nivel.

Las pruebas objetivas plantean, no obstante, diversas objeciones en cuanto a su validez para evaluar estos indicadores. La actividad inferencial que interviene en la comprensión del «hilo conductor» se desarrolla principalmente *durante* la lectura del texto, por lo que resulta difícil de

evaluar con preguntas que se formulan una vez terminado el proceso. Así, en las preguntas que algunos tests denominan «literales» es difícil saber si se está evaluando capacidades de comprensión o simplemente de recuerdo inmediato. En otros tipos de preguntas parece difícil discriminar si el sujeto realiza las inferencias espontáneamente o si dicha actividad es más bien producto de la demanda que se plantea a posteriori, al responder a la cuestión.

En el caso de los tests en los que el sujeto puede acceder al texto mientras responde a las preguntas, por el contrario, las respuestas a preguntas referenciales o sobre información literal pueden realizarse mediante simples operaciones de *reconocimiento*, por lo que su utilidad para evaluar la comprensión ofrece muchas dudas. Algunas investigaciones han comprobado por ejemplo, que los sujetos de primaria mejoran significativamente sus resultados en este tipo de preguntas cuando se les permite volver a consultar el texto. Esta ventaja se reduce, no obstante, cuando las preguntas de la prueba hacen alusión a información global o implícita (Cain y Oakhill, 1999).

Una variante semiobjetiva que pretende soslayar esta limitación, proviene del procedimiento *cloze*. Se trata de ofrecer al lector un texto mutilado (con espacios en blanco) que el sujeto debe rellenar *durante* la misma lectura, de manera que éste sólo será capaz de restaurar las palabras que faltan si utiliza de forma estratégica las claves textuales que conducen a la representación del hilo conductor (Condemarin y Milicic, 1990). La técnica aporta la ventaja de posibilitar una cierta evaluación del proceso de la comprensión, limitando en cierta medida la contaminación derivada de la capacidad de recuerdo o de las inferencias más globales que el sujeto genera al final de la lectura. Las inferencias semánticas y «elaborativas» que posibilitan la jerarquización de las ideas del texto y la conexión con los conocimientos previos del sujeto, son difícilmente evaluables con este procedimiento. No obstante, se han obtenido correlaciones altas con pruebas clásicas, como el test de Lázaro, que avalan su validez de criterio (Artola, 1988).

Otra alternativa semiobjetiva, hasta ahora mucho menos explorada, consiste en diseñar tareas de comprensión más o menos estructuradas donde los sujetos deban realizar inferencias a partir de la información que obtienen de las señalizaciones textuales y de su propia base de conocimientos previos. Frecuentemente, la presencia de algunas palabras clave, como un pronombre, un adjetivo o un sustantivo (sinónimos de una expresión anterior), nos ayudan a deducir la relación entre una idea y su antecedente. Yuill y Oakhill (1988) demostraron que la interpretación de estas palabras en textos narrativos ayuda mucho a discriminar a los niños con problemas de comprensión. La actividad de *reordenar un texto fragmentado* permite evaluar esta capacidad de un modo muy funcional, a la vez que resulta extraordinariamente estimulante para los alumnos. La ordenación de este particular rompecabezas requiere que el lector descubra este tipo de señalizaciones y expresiones anafóricas, reconozca el hipotético antecedente, infiera el «hilo conductor» y anticipe la progresión temática, con objeto de establecer «correferencias» entre los fragmentos del texto. En un estudio reciente con una muestra de 87 chicos de 11 y 12 años, hemos comprobado que, a pesar de su sencillez, la tarea discrimina con cierta precisión entre buenos y malos lectores, ofreciendo índices de validez más elevados que otras pruebas objetivas estandarizadas, como el mismo PROLEC-SE. Al mismo tiempo, el profesor puede obtener una información cualitativa de gran interés para ofrecer las ayudas adecuadas al alumno y entrenar su actividad estratégica e inferencial (González y Montanero, 2002).

Pensamiento en voz alta y registros de observación

La evaluación de proceso de la comprensión resulta difícil si no se tiene en cuenta información cualitativa que puede obtener el evaluador demandando inferencias concretas o incluso, suministrando ayudas para observar los comportamientos estratégicos del sujeto. El análisis de los protocolos orales de *pensamiento en voz alta* (*thinking aloud*) constituye una opción que permite fundamentalmente valorar inferencias referenciales durante la lectura. El procedimiento consiste en grabar (magnetofónicamente o en vídeo) y analizar posteriormente las reflexiones en voz alta que los sujetos realizan durante la lectura de un texto; razón por la cual algunos autores consideran más bien esta última técnica como un caso particular de autoinforme. Al margen de las dificultades de cuantificación objetiva, propias de un procedimiento tan poco estructurado, el método presenta otras dificultades, como son la necesidad de un entrenamiento previo del sujeto o la posibilidad de que el propio procedimiento interfiera, al realizarse simultáneamente y en condiciones de demanda excesiva, en los procesos y estrategias que se pretenden evaluar. Los sujetos son estimulados para hablar acerca de la relación de cada frase con lo leído anteriormente, así como para anticipar la información subsiguiente, por lo que existe también el riesgo de que manifiesten más inferencias de las que harían en condiciones naturales. Beishuizen y Cols. (1999) encontraron que los resultados de este tipo de evaluación apenas «correlacionaba» con las puntuaciones obtenidas en pruebas objetivas que demandan inferencias locales y globales, aunque se utilicen los mismos textos. Sin embargo, no cabe duda de que el *pensamiento en voz alta* permite vislumbrar ciertos procesos que los sujetos desarrollan estratégicamente en su interacción con el texto.

EVALUACIÓN DE LA COMPRENSIÓN GLOBAL Y ESTRUCTURAL

Un segundo nivel de comprensión se fundamenta en la actividad mental que el sujeto desarrolla para asignar un significado coherente de carácter global al texto leído. Este proceso implica, en gran medida, inferencias aplazadas al final de la lectura del texto o de sus unidades temáticas, con objeto de construir la denominada *macroestructura* del texto, es decir, una representación mental, más o menos jerarquizada de las *macroproposiciones* del contenido de las proposiciones (Van Dijk y Kintsch, 1983).

Pruebas objetivas y semiobjetivas

Los materiales más utilizados para evaluar este nivel de comprensión se basan de nuevo en la formulación de preguntas que requieren la realización de inferencias sobre el significado global del texto o la distinción entre ideas jerárquicamente diferentes. En algunos estudios dichas preguntas se dirigen a suscitar la aplicación de las clásicas *macrorreglas* del modelo de Kintsch y Van Dijk , *supresión*, *generalización* e *integración*. Se trata de averiguar si el lector es capaz de reconocer información redundante o ideas-detalle; o bien construir una idea más general a partir de información más específica, recogida en el texto. En concreto, la inducción de *títulos* se considera como un buen indicador de la *comprensión temática*, es decir, del significado más global que da sentido al texto (León y Martín, 1993). Los distractores en las preguntas objetivas sobre el título deberían reflejar respuestas erróneas de cuatro tipos: títulos imprecisos, demasiado generales, demasiado particulares o alejados de la verdadera temática. Igualmente, algunas preguntas que requieren reconocer o parafrasear una idea esencial (*sentence recognition*) pueden también constituir un indicador de este

segundo nivel de comprensión. En algunos tests, como la Batería SURCO o el ya citado CL-4, se solicita explícitamente al sujeto, que exprese o reconozca la *idea principal* del texto o la información más relevante que se deduce directamente del mismo (*shadow questions*).

Análisis del contenido de resúmenes y protocolos de recuerdo

El análisis de contenido de *productos de la actividad inferencial* es una de las modalidades de evaluación más utilizada en los trabajos de investigación sobre comprensión de textos. Las técnicas más habituales se basan en el análisis de resúmenes y protocolos de recuerdo libre, inmediato o demorado. El objetivo principal es evaluar el producto de la representación semántica generada por el sujeto, utilizando diversos métodos de análisis del contenido proposicional del resumen que elabora, sin tener acceso al texto que se leyó. Algunos de estos métodos se encuentran altamente formalizados, de modo que permiten establecer comparaciones cuantitativas con la representación proposicional del contenido original. Para ello, es necesario fraccionar los protocolos en unidades proposicionales, de modo que se puedan comparar con un texto-modelo, aunque tengan una expresión lingüística diferente. El método de (Bovair y Kieras, 1985) y, sobre todo, el de (Meyer, 1985) son los más conocidos.

El primero se basa en la aplicación de algoritmos basados en el modelo ya referenciado de Kintsch y Van Dijk. La regla fundamental para reconocer en un protocolo una proposición del texto original es que aparezca el predicado de dicha proposición más uno de sus argumentos (teniendo en cuenta la sinonimia y que el sentido general de la expresión tenga sentido). El método, sin embargo, tiene la limitación de no reflejar el establecimiento de niveles jerárquicos entre las proposiciones.

Meyer, por el contrario, distingue entre relaciones retóricas y léxicas. Las primeras se basan en la taxonomía de superestructuras propuesta por la autora, mientras que las segundas se inspiran en la gramática de casos de (Fillmore, 1968). Sobre este análisis, se pueden realizar hasta nueve tipos de medida: recuerdo reproductivo, recuerdo sustantivo de los contenidos tópicos, relaciones de rol, predicados léxicos, predicados retóricos, unidades de contenido que no estaban en el texto, relaciones lógicas que no estaban en el texto, confusiones intratextuales y puntuaciones de la estructura de alto nivel. Este tipo de evaluación aporta una interesante información, tanto cuantitativa como cualitativa, sobre la actividad estratégica e inferencial que lleva al sujeto a reconstruir en un protocolo de resumen (con recuerdo inmediato o demorado) el significado original del texto. Como contrapartida debemos citar la considerable complejidad del procedimiento, que repercute en su fiabilidad. Por ello muchos investigadores han optado por utilizar métodos mucho más simplificados y con criterios de puntuación menos estrictos.

En este mismo sentido, para simplificar la valoración de «resúmenes» en tareas de recuerdo inmediato, (Sánchez, 1989) utilizó un sencillo método de análisis cualitativo, que ha sido posteriormente replicado en sucesivas investigaciones. Los criterios de categorización del resumen pretenden evaluar la utilización de estrategias semánticas de forma eficiente con textos breves, de carácter expositivo. La categoría C (correcto) se asigna sólo a los resúmenes que reflejan con precisión y de forma económica las macroproposiciones más importantes. La categoría A (abstracción) incluye las respuestas que incorporan el significado global del texto pero no las características de precisión y economía. Se puntúan así resúmenes que incluyen ideas demasiado generales o que presenten demasiados detalles que podrían haber

sido sustituidos por proposiciones más generales. En las categorías D (detalle) y S (secuencia) el resumen estaría construido a partir de uno de los detalles del párrafo, con pérdida de información relevante, o como un simple listado de los mismos.

Análisis del contenido de esquemas y gráficos

La mayoría de los anteriores métodos pueden adaptarse para evaluar la comprensión de textos narrativos o expositivos. Otro tipo de tareas, como la clasificación y jerarquización de conceptos (*shorting of concepts*) o la construcción de un esquema que organice gráficamente las ideas del texto, se centran en cambio, en evaluar el uso de estrategias semánticas y estructurales con textos de naturaleza expositiva.

Uno de los principales antecedente puede verse en el diseño y evaluación de *mapas conceptuales* (Novak y Gowin, 1984). La construcción de mapas puede constituir en este caso un método para valorar cómo los sujetos representan la *macroestructura* de textos expositivos. Además de una valoración cualitativa, la cuantificación puede hacerse teniendo en cuenta tres criterios: la cantidad y calidad de las relaciones jerárquicas que el alumno ha sido capaz de elaborar, comparándolo con un mapa-modelo; el número de niveles de la representación jerárquica (los mapas excesivamente horizontales suelen revelar una clara dificultad para emplear estrategias semánticas de alto nivel); y el número de nexos transversales que representen las relaciones semánticas entre diferentes ramas de conceptos.

En diversas investigaciones Alonso Tapia y han utilizado una alternativa más simplificada que consiste en solicitar a los sujetos que elijan el gráfico más adecuado entre varias opciones que representan las ideas conectadas por flechas. Sin embargo, esta opción no permite discriminar si el alumno es capaz de generar, o simplemente reconocer, la organización retórica del texto (especialmente si no se controla el sesgo del azar).

Una alternativa, semejante al procedimiento *cloze* para este segundo nivel de comprensión, consiste en ofrecer un esquema mutilado que el sujeto debe rellenar con conceptos de un texto descriptivo (Ramos y Cuetos, 1999). De esta manera, se puede cuantificar si el sujeto es capaz de utilizar las claves jerárquicas del esquema para recordar y clasificar la información esencial que conforma la macroestructura del texto.

Sin embargo, hay que reseñar que ninguna de estas opciones evalúan realmente la capacidad del sujeto para reconocer o transformar la *superestructura retórica* que vincula dichas ideas, más allá de relaciones estrictamente semánticas que caracterizan a los textos descriptivos. Una tarea mucho más exigente en este sentido consiste en ofrecer un esquema, un diagrama o un cuadro sinóptico vacío, de manera que el sujeto seleccione y rellene uno de ellos con un listado de las ideas de un texto académico. El procedimiento ofrece la posibilidad de conseguir información cuantitativa o cualitativa. Para cuantificar la capacidad del sujeto para representar la organización de las ideas del texto basta con contabilizar el número de relaciones descriptivas, comparativas o lógico-causales que el sujeto es capaz de reflejar gráficamente. Este tipo de tareas presenta, no obstante, el riesgo de resultar poco discriminatoria debido a su elevada dificultad, como se ha demostrado en algunas investigaciones recientes (Montanero, 2002).

EVALUACIÓN DE LA COMPRENSIÓN PRAGMÁTICA Y «ELABORATIVA»

El proceso de comprensión no se agota ni mucho menos en la representación mental de la *base del texto*. En paralelo, el lector

enriquece todas estas representaciones a partir de su acerbo experiencial. Este tercer nivel lleva por tanto aparejado una dimensión pragmática y *situacional*: la construcción de todo un mundo más allá del texto en sí, en el que se recogen las inferencias e incluso las vivencias, que el lector activa para interpretar la nueva información. Algunas inferencias causales o intencionales que permiten conectar localmente dos proposiciones pueden considerarse también como elaborativas. Sin embargo, las inferencias más representativas de los niveles profundos de la comprensión (al menos en los textos expositivos) serían aquellas que operan sobre representaciones semánticas más globales al final de la lectura.

Pruebas objetivas y semiobjetivas

La comprensión pragmática se relaciona en parte con la capacidad de articular enunciados en unidades lingüísticas muy amplias, propias de diferentes «superestructuras» y géneros textuales. Esta capacidad comienza a desarrollarse desde los 5 ó 6 años, cuando algunos niños son ya capaces de identificar un tipo de texto o detectar contradicciones que afectan a su comprensión. De nuevo, la alternativa más utilizada se corresponde con materiales que recogen preguntas breves, formuladas con posterioridad a la lectura. Algunos tests presentan ítems acerca del género del documento, las intenciones comunicativas o los recursos estilísticos del autor en diferentes tipos de textos. En los textos de género no expositivo (narrativos, descriptivos, periodísticos o literarios) las preguntas suelen centrarse además en clarificar el escenario situacional en que transcurre la acción o las intenciones más o menos implícitas de los personajes; inferencias que resultan realmente difíciles de extrapolar de la actividad que hemos descrito en los niveles anteriores de comprensión. Otras preguntas más «profundas» pueden incluso solicitar la interpretación de una metáfora o la moraleja de un suceso (véase, por ejemplo, Lázaro, 1982). Este tipo de inferencias discriminan bastante bien la capacidad del lector para utilizar sus conocimientos previos y razonar sobre el contenido implícito del texto.

En los pasajes expositivos no es tan difícil deslindar la evaluación de las capacidades que intervienen en la construcción de un *modelo situacional* de aquellas que posibilitan la representación de la base del texto. Las preguntas de comprensión profunda que requieren la participación de conocimientos temáticos (*deep questions*) o conectar información implícita más o menos alejadas en el texto-base (*bridging inference question*), son las tareas más representativas de este nivel. Todas ellas tienen en común la necesidad de desplegar inferencias «elaborativas» más allá del texto-base. Algunos trabajos recientes están basados en diversos tipos de preguntas con esta orientación, (Orrantia y cols., 1998), por ejemplo, utilizaron preguntas «creativas» para evaluar el uso personal de la información en la construcción de modelos situacionales más allá del contenido explícito del texto. Se demandaban esencialmente respuestas explicativas, que requisieran desentrañar la trama causal de un fenómeno. La evaluación de inferencias referenciales, semánticas o estructurales no tiene por qué ser redundante con este tipo de cuestiones. En principio, para reconocer la relación causal entre dos eventos es suficiente con delimitar las causas y las consecuencias a partir de la representación semántica de la base del texto, sin que ello presuponga que se comprende por qué esas causas generan esos efectos. La llamada «trampa narrativa» (Duthie, 1987) es una de los principales riesgos en este sentido, al menos para un tipo de evaluación que pretenda abordar las comprensión profunda de textos expositivos.

En uno de los estudios anteriores hemos desarrollado también una propues-

ta para operativizar este tipo de preguntas en función de la estructura organizativa de textos (Montanero y cols., 2002). En un texto causal las preguntas deberían solicitar información implícita acerca del *nexo*, las acciones o condiciones intermedias que permiten explicar por qué un determinado episodio puede considerarse un factor explicativo de otro suceso. Los sujetos que alcancen un nivel más profundo en la comprensión de un texto expositivo, deberían ser capaces de inferir o justificar los *criterios* que permiten vincular las características de diversos conceptos entre sí; así como las *decisiones* que justifican que una determinada respuesta se presente como posible solución a un problema si el texto presenta más bien una estructura procedimental. Cuando se trata de un texto argumental, finalmente, el lector debería poder inferir por qué unas *razones* fundamentan determinadas *deducciones* o en qué medida, las razones expuestas apoyan la conclusión a la que llega el autor (explorando por ejemplo, si se han tenido en cuenta los contraargumentos que apoyan otras decisiones o posiciones diferentes).

Este último tipo de cuestiones se corresponde en cierto modo con las tareas de *análisis crítico* que sirvieron de base para la evaluación de diversas capacidades de comprensión lectora en el conocido «macroestudio» del INCE (1998). En el diseño de la investigación se intentó en concreto discriminar entre inferencias que permiten comprender el *significado literal* del texto y las que se dirigen a obtener una *interpretación personal* o realizar un *análisis crítico*. Las primeras son asimilables a la construcción de lo que hemos denominado texto-base: inferencia del significado de las palabras en su contexto, establecimiento de conexiones referenciales y comprensión global a través de la identificación del tema y la distinción entre la idea principal y las secundarias. Los dos grupos de preguntas restantes se refieren ya a inferencias que supuestamente reflejarían los niveles más profundos de comprensión. Para responder correctamente las preguntas de *interpretación personal* en textos narrativos se preguntaba sobre la identificación de personajes protagonistas y secundarios, ordenación temporal… En el *análisis crítico*, se debe determinar la función del texto y la intención del autor; clarificar la opinión, la ironía, la ambigüedad o la falacia; y valorar la línea argumental de contenidos implicativos. Este último tipo de preguntas se contestaban en presencia de textos «informativos», categoría a la que son asimilables tanto los periodísticos como el resto de los expositivos de carácter académico.

Por último, hay que señalar que el hecho de que el sujeto pueda releer o no el texto, mientras responde a estas preguntas, no condicionaría tanto los resultados como en la evaluación del primer nivel referencial. Como hemos tenido ocasión de comprobar con una muestra de alumnos de la ESO, las diferencias resultantes de responder con o sin el texto delante no resultan significativas (Montanero, 2000). Por su parte, (Cain y Oakhill, 1999) constataron que los alumnos de primaria mejoran relativamente poco, ante preguntas sobre información implícita cuando se les da la oportunidad de revisar el texto leído. Además, las diferencias entre buenos y malos lectores se mantienen prácticamente constantes.

Pensamiento en voz alta y tareas de solución de problemas con textos inconsistentes

En los niveles más profundos de la comprensión, las relaciones que los lectores deben establecer entre las ideas, tienen un carácter global: abarcan segmentos más amplios de información y una participación mayor de conocimientos previos. No es de extrañar, por tanto, que cuando introducimos información contradictoria en un texto, los chicos que no comprenden bien,

manifiesten muchos más problemas, no sólo para resolverlas, sino incluso para detectarlas. Este tipo de evaluación se basa en confeccionar textos que presenten afirmaciones contradictorias entre sí (inconsistencias internas) o con respecto a los conocimientos previos que sabemos que tiene el sujeto (inconsistencias externas). La técnica no se recoge en ninguna de las pruebas estandarizadas que hemos analizado. El procedimiento de preguntas abiertas (semiobjetivas), y sobre todo el autoinforme oral de los sujetos durante la lectura, son probablemente los que mejor se adaptan a este recurso.

EVALUACIÓN DE ESTRATEGIAS DE REGULACIÓN METACOGNITIVA Y APOYO A LA COMPRENSIÓN DE TEXTOS EXPOSITIVOS

Para terminar puede ser oportuno considerar un último ámbito de evaluación que tiene una relación indirecta con la comprensión de textos académicos. Se trata de la evaluación del uso de estrategias de apoyo al estudio, es decir, de las actividades relacionadas con la preparación psicofísica, el control de estimular la planificación de intervalos de estudio, etc., que los estudiantes deberían realizar a partir de secundaria. Estas estrategias tienen un efecto particular sobre las capacidades de atención sostenida que intervienen en la comprensión de texto expositivos. Las técnicas utilizadas para su evaluación son sensiblemente diferentes a las analizadas hasta ahora. Están sobre todo basadas en métodos de autoinforme, donde los sujetos proporcionan información sobre sus hábitos de estudio a partir de una serie de preguntas.

Autoinformes

Desde los años setenta han proliferado múltiples *cuestionarios e inventarios estandarizados* que presentan un listado de actividades, generalmente relacionadas con tareas escolares, sobre las que el sujeto debe afirmar o negar (y en algunos casos cuantificar) su uso. Casi todos introducen además ítems dirigidos, no sólo a evaluar estrategias de apoyo, sino también la utilización de algunas de las «estrategias primarias» de carácter metacognitivo y técnicas de estudio de textos expositivos. El problema reside en que algunos componentes que posibilitan el desarrollo de estas estrategias metacognitivas están en parte automatizados, y en consecuencia, no son conscientes o al menos no directamente accesibles a la propia observación del sujeto. Además, la naturaleza claramente retrospectiva de estos cuestionarios contamina la objetividad de las supuestas valoraciones de quien cumplimenta el cuestionario (aunque estén previstos ítems «residuales» como parte de una escala de sinceridad). En suma, las preguntas directamente dirigidas a evaluar si el individuo aplica estrategias de comprensión constituyen un porcentaje minoritario y difícilmente permiten discriminar entre una utilización mecánica de técnicas de estudio y la coordinación estratégica de operaciones mentales que posibilitan su eficacia.

Registros de observación

Los métodos de observación constituyen de nuevo una interesante alternativa para evaluar los procesos metacognitivos que intervienen en la comprensión de un texto. Un ejemplo característico de *observación directa* se recoge en el experimento de (Garner, Wagoner y Smith, 1983), en el que se pidió a dos grupos de alumnos de diferente capacidad, que explicara a otros compañeros, el mejor método para comprender un texto. La información se grabó y analizó posteriormente. Sin embargo, al menos en este caso, difícilmente dicha valoración podría ir más allá del conocimiento declarativo de

CUADRO I

Indicadores, instrumentos y materiales para evaluar la comprensión

Nivel de comprensión	Indicadores	Tipo de prueba	Algunos materiales y referencias
Base del texto I: comprensión léxica, sintáctica y referencial local	Emparejar oraciones con dibujos	Objetivas o semiobjetivas	PROLEC (Cuetos y cols., 1996)
	Emparejar entre sí palabras del texto o inferir su significado		PCL (Lázaro, 1982); CL (Alonso Tapia y cols., 1994); EVALUA (G. Vidal y G. Manjón, 1996)
	Reconocer o parafrasear oraciones	Objetivas o semiobjetivas	PCL (Lázaro, 1982); CL (Alonso Tapia y cols., 1994); «Thinking aloud» (Pressley y Afflerbach, 1995); EVALUA; PROLECSE (Ramos y Cuetos, 1999)
	Inferir causas o intenciones	Pensamiento en voz alta	
	Inferir referencias anafóricas		
	Inferir palabras en textos mutilados	Semiobjetivas	CLT (Suárez y Meara)
	Reordenar fragmentos		TD (Montanero y González, en preparación)
Base del texto II: comprensión semántica y retórica global	Inferir o reconocer títulos o macroproposiciones	Objetivas o semiobjetivas	CL-4 (Alonso Tapia y cols., 1994); PROLEC-SE (Ramos y Cuetos, 1999); PC (Montanero, 2000)
	Reconocer o elaborar gráficos	Análisis del contenido de resúmenes	Bovair y Kieras, 1985; Meyer, 1985; Sánchez, 1987; G. Madruga y cols., 1995
	Resumir con recuerdo inmediato o demorado		
Modelo de la situación: comprensión elaborativa y pragmática	Inferir el género del documento	Objetivas o semiobjetivas	PCL (Lázaro, 1982) ;CL-4 ;TA (Pressley y Afflerbach, 1995); CL (INCE, 1998)
	Interpretar intenciones comunicativas o recursos estilísticos del autor		
	Inferir «creativamente» información contextual o implícita	Pensamiento en voz alta	
	Detectar inconsistencias (internas o externas)		
(Autorregulación estratégica)	Informar sobre el empleo de estrategias y técnicas	Autoinformes. Escalas de apreciación de allegados	IHE (Pozar, 1972); IME (Baeza, 1979); LASSI (Weinstein, Goezts y Alexander, 1983); ACRA (Román y Gallego, 1994); IDEA (Vizcarro, 1996)
	Evaluar y regular estratégicamente la actividad lectora	Pensamiento en voz alta	TA (Pressley y Afflerbach, 1995)

los sujetos, por lo que una vez más la técnica resulta limitada para valorar ciertos conocimientos estratégicos que los sujetos utilizan realmente durante el estudio de un texto. Esta limitación es más clara si cabe en el caso de la *observación indirecta* a través de *escalas de apreciación de allegados*, muy extendida en la evaluación psicopedagógica de los alumnos con necesidades especiales, en las que el profesor debe aportar esta información a partir de una valoración retrospectiva y generalmente poco sistemática.

En conclusión, la panorámica actual sobre la evaluación de la comprensión lectora ofrece una amplia diversidad de técnicas e instrumentos, que se sintetizan en el cuadro I. La mayoría de ellas tienen, como denominador común, el empeño en analizar la actividad estratégica e inferencial durante la lectura como principal indicador del nivel de comprensión que alcanza el

sujeto. A menudo, sin embargo, dichos criterios se manejan sin una reflexión previa sobre los presupuestos teóricos que los justifican. Esta limitación afecta particularmente a la evaluación de la *comprensión profunda* de textos expositivos en la Educación Secundaria, donde los lectores tienen que poner en juego capacidades y conocimientos sensiblemente diferentes a los que se necesitan en otro tipo de textos. Un requisito que hemos tratado de fundamentar, en este sentido, consiste en tratar de discriminar entre las capacidades que intervienen en diferentes niveles de comprensión e incluso entre textos de diferente género. Por otro lado, es necesario dedicar más investigación a los procedimientos y análisis cualitativos que, como hemos visto, pueden resultar de tanta o mayor utilidad para orientar una intervención psicopedagógica efectiva.

BIBLIOGRAFÍA

ALONSO TAPIA, J. y CARRIEDO, N.: *Leer, comprender y pensar: desarrollo de estrategias y técnicas de evaluación.* Madrid, Servicio de publicaciones del MEC-CIDE, 1992.

— «Problemas de comprensión lectora: evaluación e intervención», en C. Monereo, I. Solé (coord.), *El asesoramiento psicopedagógico: una perspectiva profesional y constructivista*. Madrid, Alianza, 1996.

ALONSO TAPIA, J. y CORRAL, C.: *Evaluación de la comprensión lectora al término de la ESO Test CL-4*. Madrid, UAM, 1994.

ARTOLA, T.: «El procedimiento *cloze* como medida procesual de comprensión lectora», en *Revista española de Pedagogía,* 180 (1988), pp. 323-334.

BAEZA, M.: *Inventario de métodos de estudio*. Madrid, Miñón, 1979.

BOVAIR, S. y KIEBAS, D. E.: «A guide to propositional analysis for research on technical prose», en B. K. Britton y J. B. Black (comps.), *Understanding expository text.* Hillsdale, LEA, 1985.

CAIN, K. y OAKHILL, J. V.: «Inference making ability and its relation to comprehension failure in young children», en *Reading and Writing*, 11 (1999), pp. 489-503.

CARRIEDO, N. y ALONSO TAPIA, J.: *¿Cómo enseñar a comprender un texto?* Madrid, ICE de la Universidad Autónoma, 1995.

CONDEMARIN, M. y MILICIC, N.: *Test de Cloze. Aplicaciones psicopedagógicas*. Madrid, Visor, 1990.

CUETOS, F., RODRÍGUEZ, B. y RUANO, E. *(PROLEC): Evaluación de los procesos lectores.* Madrid, TEA, 1996.

DUTHIE, J.: «The Web: a powerful tool for teaching and evaluation of the expository essay», en *The history social science teacher,* 21 (1986), pp. 232-236.

FILLMORE, C. J.: «The case for case», en E. Bach, y R. T. Harms (eds.), *Universals in linguistic theory.* Chicago, Holt, Rinehart and Winston, 1968.

GARCÍA VIDAL, J. y GONZÁLEZ MANJÓN, D.: *EVALUA-4*. Madrid, EOS, 1996.

GARNER, R.; WAGONER, S. y SMITH, T.: «Externalizing question answering of poor and good comprehender», en *Reading Research Quaterly,* 18 (1983), pp. 437-439.

GONZÁLEZ, L. y MONTANERO, M.: *Evaluación y entrenamiento de la comprensión lectora mediante textos «desordenados». Dos estudios preliminares en el tercer ciclo de primaria.* Memoria de investigación no publicada. Universidad de Extremadura, 2002.

INCE: *Los resultados escolares. Diagnóstico del sistema educativo.* 1997. Madrid, MEC, 1998.

LÁZARO, A.: *Prueba de comprensión lectora.* Madrid, TEA, 1982.

LEÓN, J. A. y MARTÍN, A.: «El título como recurso didáctico», en *Comunicación, lenguaje y educación*, 19-20 (1993), pp. 159-170.

MEC: *La evaluación psicopedagógica: Modelo, orientaciones, instrumentos*. Madrid, MEC, 1996.

MEYER, B. J. F.: «Prose analysis. puposes, procedures, and problems», en K. Britton y J. B. Black (Eds), *Understanding expository text*. Hillsdale, LEA, 1985.

MONTANERO, M.: *La instrucción de estrategias de comprensión en el ámbito sociolingüístico del segundo ciclo de la ESO. Una propuesta didáctica para enseñar a comprender desde el currículo*. Tesis doctoral no publicada, Universidad de Extremadura, 2000.

MONTANERO, M., BLÁZQUEZ, F. y LEÓN, J. A.: «Enfoques de intervención psicopedagógica para la mejora de las capacidades de comprensión en la Educación Secundaria», en *Infancia y aprendizaje,* 25 (2002), pp. 37-52.

NOVAK, J. D. y GOWIN, D. B.: *Aprendiendo a aprender*. Barcelona, Martínez Roca, 1988.

ORRANTIA, J.; ROSALES, J. y SÁNCHEZ, E.: «La enseñanza de estrategias para identificar y reducir la información importante de un texto. Consecuencias para la construcción del modelo de la situación», en *Infancia y aprendizaje,* 83 (1998), pp. 29-57 .

POZAR, F. F.: *Inventario de hábitos de estudio*. Madrid, TEA, 1972.

PRESSLEY, M. y AFFLERBACH, P.: *Verbal protocols of reading*. Hillsdale, N. J, Erlbaum, 1995.

RAMOS, J. L. y CUETOS, F.: *Evaluación de procesos lectores. PROLEC-SE*. Madrid, TEA, 1999.

ROMÁN, J. M. y GALLEGO, S.: *ACRA*. Madrid, TEA, 1994.

SÁNCHEZ, E.: *Procedimientos para instruir en la comprensión de los textos*. Madrid, CIDE, 1989.

VAN DIJK, T. A. y KINTSCH, W.: *Strategies of discourse comprehension*. New York, Academic Press, 1983.

VIZCARRO, C.: «Development of an inventory to mesure learning strategies», en M. Birembaum y F. Dochy (eds.), *Alternatives in assessment of achivements learning processes and prior knowledge*. Boston, Kluwer Academic Press, 1996.

WEINSTEIN, C.; GOEZTS, E. T. y ALEXANDER, P. A.: *Learning and study strategies, issues in assessment, instruction and evaluation*. New York, Academic Press, 1983.

YUILL, N. y OAKHILL, J. V.: «Understanding of anaphoric relatins in skilled and less-skilled comprehenders», en *British Journal of Psychology,* 79 (1988), pp. 173-186.