



Organización
de Estados
Iberoamericanos

Para la Educación,
la Ciencia
y la Cultura

Revista Iberoamericana de Educación

Número 10 - Enero Abril 1996

Evaluación de la Calidad de la Educación

Datos Artículo

Título: Aportes de las Ciencias del Lenguaje para la consideración de la calidad en Educación y su evaluación

Autor: Marta Elena Costa

Aportes de las Ciencias del Lenguaje para la consideración de la calidad en Educación y su evaluación¹

Marta Elena Costa (*)

Ha aumentado en la región la circulación de pruebas objetivas utilizadas en programas de evaluación de la calidad, en exámenes de acreditación o admisión y en diversos estudios diagnósticos. En este trabajo se consideran las concepciones de 'calidad' y de 'evaluación de la calidad' que subyacen en los programas del primer tipo y se analizan algunas posturas alternativas. Desde una perspectiva que integra las ciencias del lenguaje, se plantean -en particular con referencia a las pruebas de lengua y, dentro de ellas, a las de lectura- aportes para la consideración de la diversidad lingüística, cultural y cognoscitiva de las poblaciones y de los sujetos evaluados. Se reconoce la utilidad del empleo de pruebas estandarizadas, del tipo denominado objetivo, cuando se trata de poblaciones numerosas, y por otra parte se analizan desde la perspectiva señalada los aspectos considerados críticos, tanto por los modelos a los que responden como por las limitaciones que de hecho plantean.

79

1. Introducción

La evaluación educacional, de décadas de tradición en los países centrales, ha tenido fuerte incremento en América Latina desde los años 80 por impulso tanto de los organismos de educación, ciencia y cultura

(*) Marta Elena Costa, profesora en Letras por la Universidad de Buenos Aires, es especialista en el campo educacional y en lingüística aplicada. Coordina proyectos en dicha Universidad y en el Ministerio de Cultura y Educación. Fue coordinadora del Área de Investigación de la Dirección de Educación del Adulto, especialista para el Curso Introductorio de la Universidad Nacional Abierta de Venezuela y directora de Planeamiento Educacional de la Universidad de Buenos Aires.

que operan en la región como de los organismos rectores de las políticas económicas o que coadyuvan a ellas.

Esta tendencia se ha acentuado en los 90 hasta el punto de incorporarse como una instancia institucionalizada en los procesos de reformas o de transformaciones educativas impulsadas desde el Estado e incluso en las leyes que regulan la educación de los países. Valgan como ejemplos los casos de Argentina (Ley Federal de Educación, 1993), Bolivia (Ley de Reforma Educativa, 1994) y Colombia (Ley General de Educación, 1994).

Así en Argentina, la Ley Federal de Educación, en su título IX, «De la calidad de la educación y su evaluación», arts. 48 a 50, establece la evaluación permanente del sistema, que incluye la evaluación de los aprendizajes a cargo de la Nación y de las demás jurisdicciones.

Paralelamente al auge de este tipo de evaluaciones se ha instalado la discusión en los sectores involucrados -docentes, padres de alumnos-, en el ámbito político, gremial y otros, que ha venido a sumarse a la ya existente en el ámbito académico, acerca del sentido y finalidad de dichas prácticas y de los instrumentos empleados.

80

Hacia fines de 1993 y comienzos del 94 tomó estado público en la Argentina la evaluación de la calidad llevada a cabo por el Ministerio de Cultura y Educación. Esto suscitó un intenso debate reflejado en los medios, superado sólo por el que anteriormente provocó la transferencia de los servicios educativos a las jurisdicciones y la ley de educación, debate éste último mucho más inserto en la sociedad misma.

Uno de los elementos que se toman en cuenta para la evaluación de los sistemas educativos o de los niveles de enseñanza es la calidad de los aprendizajes de los alumnos y ha sido justamente este aspecto el que provocó el debate mencionado. Al aparecer a través de los medios recortado de los demás componentes de la evaluación, absorbió ante la mirada pública el significado total de la evaluación, lo que tornaba imposible establecer relaciones que permitieran dilucidar las causas del rendimiento, fuera éste aceptable o no.

El campo universitario no es ajeno a la cuestión: se ha visto la necesidad de autoevaluación institucional o de evaluación externa según las posturas adoptadas, fomentado esto por los organismos internacionales, en especial los de universidades². La inquietud ha surgido también desde el propio quehacer académico, sobre todo por parte de docentes e investigadores vinculados a los primeros años de estudios que confrontan las dificultades en el rendimiento de los estudiantes. Se han venido

realizando en diferentes universidades investigaciones que incluyen pruebas diagnósticas de diversos tipos.

Sabemos que de hecho, y por razones fundadas, los saberes que se privilegian al momento de este tipo de evaluaciones suelen ser los de lengua y matemática. Cabe considerar entonces los aportes que desde la perspectiva de las ciencias del lenguaje, y conjuntamente con otras ciencias, podrían hacerse para una conceptualización y también para una práctica evaluativa de calidad.

En este trabajo nos referiremos sobre todo a las pruebas que intentan evaluar dominio en sentido amplio, aptitud verbal, competencias globales, o pruebas de nivel, etc., ya sea con fines de admisión, acreditación o investigación. Así por ejemplo las pruebas que se hacen en algunos países para ingreso al nivel medio o universitario, o las que corresponden a programas de evaluación de la calidad, o las que forman parte de diversos tipos de estudios o investigaciones: todas ellas buscan diagnosticar las competencias lingüísticas de poblaciones numerosas o de muestras de volumen considerable.

2. Concepciones de 'calidad' y de 'evaluación de la calidad'

Bajo este título podrían considerarse multiplicidad de teorizaciones y de prácticas, lo cual está fuera del alcance de este trabajo³. Si bien hay una preocupación generalizada por la calidad de la educación, no hay consenso sobre qué se entiende por calidad y menos aún sobre hasta qué punto y bajo qué modalidades es posible evaluarla. En atención a nuestro propósito, haremos una referencia breve a algunas posturas.

La **medición de la calidad** en educación surgió, como es sabido, en relación con un modelo de eficiencia económica. La «calidad» de los sistemas educativos y de los procesos que tienen lugar en ellos es vista en términos de eficiencia. Los modelos evaluativos se nutren de una racionalidad propia del modo de producción industrial.

Algunas manifestaciones recientes de las visiones economicistas se vinculan con el «nuevo orden mundial». Según esto, los sujetos de la educación son vistos predominantemente desde la mira de su inserción en el mercado laboral, el cual está diferenciado a nivel mundial.

Así de algunas regiones o países podría esperarse, en un extremo de una escala ocupada por un pequeño segmento de su población, una educación «de excelencia», para que se pueda cooperar en la carrera científica y económica y, en el otro extremo, los niveles mínimos para una

eficiencia en los sectores más descalificados de la escala ocupacional, o para la desocupación, subocupación, etc.

Esta caracterización, un tanto simplista quizás, hace referencia sin embargo a la concepción que subyace en las denominadas posturas eficientistas en educación o en algunos intentos de tornar más eficiente un sistema.

En otros casos estas posturas aparecen combinadas con elementos de otras visiones en las que se toman en cuenta aspectos que tienen que ver con concepciones más humanistas, más interesadas en lo pedagógico, en la adquisición de saberes socialmente válidos, etc. Y en países donde se quiere construir una sociedad democrática conviven o entran en conflicto con el principio de la igualdad de oportunidades y con la voluntad política de llevarlo a la práctica.

Frente a las posturas eficientistas, existen posiciones como las de Magendzo y otros que vinculan la calidad a la cultura de la población sujeto de educación. Corriente ésta que entronca con líneas de pensamiento y prácticas educacionales muy arraigadas en América Latina y en otras regiones del mundo.

82

Desde esta perspectiva se plantea que en una educación de calidad resulta imprescindible que las instituciones educativas tomen en cuenta las diferentes culturas y se nutran de ellas para llevar adelante los aprendizajes que son de incumbencia de la escuela.

Magendzo demostró a través de una investigación etnográfica -llevada a cabo en una población indígena de Guatemala, hablante de cakchiquel y español-, que los estilos de enseñanza-aprendizaje de la escuela, al apartarse de los patrones de socialización domésticos de los niños y de otros aspectos de la cultura de la comunidad, conducían al fracaso escolar, a la repitencia y a la deserción.

La tesis fundamental de este autor es *«que existe una relación directa entre calidad y cultura, es decir que la educación incrementa su calidad en la medida en que es capaz, a través del currículo, de valorizar y hacer suya las formas de pensar, sentir y actuar que una comunidad tiene para enfrentar y darle significado a su cotidianidad.(...) No son los estudios o las evaluaciones comparativas con niveles de rendimiento nacionales o internacionales los que permitan aquilatar la calidad de la educación»* (Magendzo, 1984:32).

Siguiendo a este autor puede afirmarse que una cantidad de estudios transculturales que se han llevado a cabo *«permiten inferir que*

muchas de las categorías conceptuales que se han desarrollado para medir la calidad de la educación desconocen las diferencias cognoscitivas, perceptuales, valóricas y de aprendizaje que se presentan en distintos grupos culturales» (Magendzo, 1984:33).

La discusión acerca de la evaluación de la calidad es relativamente nueva bajo esa denominación pero entronca con la cuestión planteada desde las primeras décadas de este siglo, acerca de la legitimidad de la utilización de instrumentos estandarizados para la evaluación en contextos educativos.

Recientemente Díaz Barriga (1994) arroja nueva luz sobre el tema del examen -y su versión moderna, las pruebas objetivas-. Plantea que debido a que la sociedad no puede resolver sus problemas de orden social, económico y psicopedagógico, se observa excesivamente al examen y se espera que a través de él se obtengan conocimientos objetivos sobre el saber de los alumnos.

Enuncia la que denomina «inversiones» -en el sentido foucaultiano del término, en tanto se presentan como relaciones de saber las que son en realidad relaciones de poder-; en torno de ellas realiza un análisis histórico y sincrónico del examen. Según tales inversiones, se ocultan los problemas sociales vinculados a la posibilidad de acceso a la educación y otros a través de la problematización acerca de las características de objetividad, validez y confiabilidad atribuidas al instrumento.

El examen, según Díaz Barriga, ha dejado de ser un aspecto del método ligado al aprendizaje, como lo era en la concepción comeniana, para convertirse en un medio para la acreditación, trastocándose así la práctica pedagógica, al tenerse la mira puesta en la acreditación. Los esfuerzos de docentes y alumnos se descentran de la enseñanza - aprendizaje.

La evaluación se ha configurado como un ámbito desvinculado de lo didáctico; ha adquirido una autonomía como campo de estudio y como práctica, afincada en lo técnico, y se ha dissociado de lo educativo. Los reduccionismos técnicos han empobrecido el debate educativo.

La revisión de los trabajos de los autores mencionados nos sugiere que la búsqueda de la calidad debería desvincularse del afán por su medición, ya que no parece haber sido saludable para la educación el intento por convertir el *quale* en *quantum*, sino que más bien ha descentrado la mirada y la tarea.

En lo que concierne a la reflexión, la cuestión de fondo debería girar en torno de **cuál** educación sería mejor y **cómo** generar transformaciones

en su búsqueda. Asimismo y dado que la evaluación ya está instalada en los sistemas educativos y legitimada por las normativas vigentes, sería de interés orientar el debate en este campo específico tanto hacia sus fundamentos y a los significados que se atribuyen en las distintas representaciones sociales a las prácticas evaluativas, como al reconocimiento de su valor real en términos de **qué pueden** y **qué no pueden** diagnosticar este tipo de pruebas, a la vez que a recuperar fundamentos teóricos que favorezcan nuevas experimentaciones.

Al trabajo de educadores, sociólogos, antropólogos y otros estudiosos habría que sumar los aportes de las ciencias del lenguaje que llevarían al reconocimiento y valorización de la diversidad cultural, lingüística y cognoscitiva, en vinculación con la atribución o no de calidad a la educación o a los logros de aprendizaje.

La evaluación aparece como un espacio de confluencia de múltiples disciplinas que requiere de interacciones mucho más ricas y complejas que las habituales de «especialistas en evaluación» y «especialistas en contenidos».

3. **La cuestión de la diversidad. Aportes de las Ciencias del Lenguaje**

84

En lo concerniente a la lengua, la consideración de la diversidad implica plantearse cuestiones vinculadas al campo de la sociolingüística y de la dialectología.

Lo primero que aparece en las evaluaciones en la lengua oficial del país, en este caso el español, es que habitualmente se toma como referencia la variedad denominada «estándar»⁴ y, dentro de ella, las modalidades propias de la lengua escrita.

De este modo, queda acotado el objeto de evaluación a una parte de la competencia lingüística de la persona examinada. Quedan fuera el lenguaje oral y las formas coloquiales del lenguaje escrito.

Así, podrían encontrarse en situación desventajosa para responder aquellas personas que por su origen socioeconómico tengan menos contacto con la lengua escrita o cuyas culturas de origen se expresen por la oralidad, y/o cuyo uso del español esté más alejado de la variedad estándar, ya sea de la de la escuela y los textos escolares, o de la de los instrumentos de evaluación.

Una evaluación que abarcara una gama más amplia de las competencias lingüísticas podría asimismo ser de interés en la medida en que

posibilitara establecer relaciones entre el lenguaje oral y el escrito, o entre el uso del propio dialecto del español y la comprensión de la variedad estándar, etc.⁵.

Cuando se trata de evaluaciones de ámbito nacional debería echarse mano a los estudios de dialectología y formular consultas a expertos de modo de incorporar, como uno de los aspectos por evaluar en los ciclos primario o secundario, el desempeño en las variantes propias. De no ser ello posible, habrá que tomar en cuenta esta limitación a la hora de analizar los resultados.

Los aportes de la sociolingüística pueden ayudar a encarar estas cuestiones que se tornan más álgidas en zonas de lenguas en contacto o en hablantes del español como segunda lengua.

Estos puntos que hemos señalado son en realidad problemas que se plantean en el currículum de lengua y que cobran resonancia a la hora de la evaluación. Cuestiones que se vinculan a las políticas lingüísticas, sean éstas explícitas o no.

Vemos entonces que las evaluaciones diagnósticas no sólo brindan información sobre el rendimiento de las poblaciones evaluadas sino que ponen en evidencia qué usos del lenguaje son considerados válidos en el sistema educativo. Según cómo estén encaradas, pueden retroalimentar al mismo arrojando luz sobre innumerables aspectos a ser tenidos en cuenta por docentes y planificadores y que incidan en un mejoramiento de las actitudes lingüísticas o, por el contrario, pueden reforzar situaciones de discriminación, como por ejemplo favorecer indirectamente en la sociedad o en el sistema educativo situaciones de diglosia tomando el término en el sentido amplio en que lo emplean L. Krysin⁶ y otros autores, con referencia a las variedades estándar y no estándar dentro de la misma lengua.

Sería desacertado ignorar las diferenciaciones internas que existen en las lenguas, vinculadas a lo geográfico-espacial, a la estratificación socio-cultural, a lo estilístico, etc., y tender a una falsa homogeneidad.

El español, extendido en un enorme espacio geográfico, ha conservado su unidad y su riqueza. Una forma de afianzar esta unidad es favorecer la comprensibilidad entre los hablantes de unas y otras zonas, de unos y otros estratos sociales, y el acceso de todos ellos a diferentes tipos de discursos.

La cuestión de cuáles variedades se enseñan en la escuela ha sido largamente debatida. En los años 70 y 80 numerosos estudios de

sociolingüística aplicada analizaron cómo operan en el ámbito escolar estadounidense las variedades no estándar y su incidencia en el aprendizaje de la lengua, en particular, de la lectura. Algunas de las investigaciones llamaban la atención sobre el *code switching* -la alternancia de códigos y el pasaje rápido de uno a otro en función de situaciones comunicacionales- (Lamb, 1981).

Sus resultados llevaron a plantear como aconsejable el bidualismo, es decir la enseñanza de la variedad estándar como una alternativa a la propia y sin la erradicación de ésta.

Criterios similares, aunque partiendo quizá de concepciones muy diferentes, han guiado múltiples prácticas educativas en América Latina. Más allá de las posturas adoptadas, ha habido en los últimos años un incremento de la conciencia por parte de docentes e investigadores en torno de estas problemáticas. Más difícil ha resultado introducirlos en la teoría y en la práctica evaluativas.

4. Aspectos críticos en las pruebas de lengua con especial referencia a las de lectura

86

Nos referiremos pues a las pruebas estandarizadas, de tipo objetivo, con énfasis en las de comprensión de lectura. A la vez que poseen múltiples ventajas cuando se trata de poblaciones numerosas, presentan problemas en su elaboración, administración y en el análisis e interpretación de sus resultados. Consideraremos desde una perspectiva que integre los aportes de las ciencias del lenguaje algunos de estos aspectos críticos: **la contextualización, la comprensibilidad y las preguntas**. No abordamos aquí otras cuestiones igualmente importantes como la validación y la interpretación de los resultados.

Aunque el trabajo está referido a las pruebas de lengua, algunos señalamientos pueden resultar aplicables a las evaluaciones en otras áreas en la medida en que se utilicen instrumentos en los que intervenga el lenguaje.

4.1. La contextualización

Frecuentemente se presentan en las pruebas distintos tipos de ítems, cada uno de los cuales, o a veces más de uno, se refieren a aspectos muy puntuales y no guardan relación con el resto. Preguntas clásicas de este tipo son aquellas que interrogan sobre el significado o el sinónimo de una palabra aislada, o las que presentan una oración en la que hay que colocar una palabra que falta. De ésta u otras formas se

plantean ejercitaciones artificiales, que se alejan de los usos naturales o habituales del lenguaje.

Si se incluyen textos, éstos suelen ser demasiado escuetos o por sus características no facilitan la formulación de preguntas adecuadas, o las preguntas, aun cuando el texto sea apto, no logran referirse a aspectos significativos. En otros casos, se trata de textos con cierta complejidad sobre los que se formulan preguntas demasiado simples.

Una prueba así construida, lejos de promover los procesos de la comprensión y la producción, los obstaculizan. En todo caso, si la persona da respuestas acertadas, éstas no demostrarán de por sí un buen desempeño. Parecería que con este tipo de instrumentos se pone más a prueba la habilidad para disgregar los propios saberes que el ejercicio integral de la competencia lingüística.

Estas y otras cuestiones que se vinculan con algunos puntos críticos en la elaboración de la prueba pueden englobarse en lo que se denomina **contextualización**.

Courtney Cazden (ca 1982:135) explica al respecto: «*El ejercicio de cualquier proceso cognitivo -tanto para nosotros como para los niños- tiene lugar en algún contexto: un formato particular de la tarea, condiciones físicas, organización social, reglas convencionales, etc.; y las características de ese contexto aportarán soporte u obstáculos a las tareas cognitivas desempeñadas en él. La cognición está siempre en un contexto y no puede ser enseñada o evaluada en forma separada de un contexto particular*».

En un trabajo posterior distingue los **contextos internos** -los contextos **mentales** del lector y del escritor- y los **contextos sociales externos** en los que ambos se encuentran. Subraya la incidencia de ambos tipos de contexto en la lectoescritura y la necesidad de incentivarlos para un buen desempeño (Cazden, 1986).

Consideramos pertinente aplicar la noción de contextualización -y su contraria, la no contextualización- tanto a la **situación** en que se administra la prueba como a los **instrumentos** en sí.

La situación de prueba

Resulta fundamental que las personas evaluadas conozcan por qué se las evalúa, para qué se van a utilizar los resultados, desde qué óptica van a ser interpretados. A posteriori, debería hacerseles alguna devolución acerca de su propia actuación y de los resultados generales y favorecerse formas de aprovechamiento inmediato de esa información.

Esto cabe mencionarse en particular con relación a investigaciones diagnósticas o a programas de evaluación de la calidad. Ello contribuiría a **dar sentido** a una evaluación que no se vincula con el currículum ni con ningún requisito de acreditación o de admisión. De hecho, la devolución de resultados se practica en algunos casos.

Respecto de las pruebas objetivas, no en todos los ámbitos se está suficientemente familiarizado con este tipo de instrumento, en cuyo caso se requerirá previamente -de ser posible en una sesión anterior- un entrenamiento para su utilización.

Ambos aspectos pueden incidir en la eliminación de inseguridad u otro tipo de malestar a la hora de la evaluación. Otras recomendaciones de tipo pedagógico tenderían a crear un buen clima, favorecedor de un adecuado desempeño.

Por otra parte habría que considerar la posibilidad de que a las pruebas clásicas contestadas en forma individual y con un estricto control de no comunicación entre los evaluados, se añadiera alguna ejercitación que requiriera una actividad cooperativa por grupos. Sería interesante la contrastación de los resultados de ambos tipos de actividades.

Cazden señala que a los niños suele resultarles difícil realizar en un examen individual lo que habitualmente hacen en forma grupal, ya que, a diferencia del lector adulto que lee casi siempre individual y silenciosamente, en la escuela la lectura es una actividad social. Asimismo, apoyándose en investigaciones propias y de otros autores -entre ellos en un estudio de W. Labov sobre el desempeño en lectura de niños negros - que ponen de relieve las comunicaciones entre pares, sugiere que éstas se aprovechen en la enseñanza de la lectura «*de modo que el poder de las interacciones grupales pueda ser usado directamente como contextos de aprendizaje*» (Cazden, ca. 1982:138).

Estimamos que los señalamientos de esta autora, que se refieren principalmente a la enseñanza y ocasionalmente a las evaluaciones en un marco curricular, pueden ser aprovechados en relación con el tipo de pruebas que nos ocupa.

Los instrumentos

Hemos dado ejemplos de tipos de ejercicios no contextualizados. Una prueba bien diseñada debería en cambio, a través de sus materiales e ítems, posibilitar el ejercicio de la **predicción**.

Frank Smith (1978) en su análisis psicolingüístico de la lectura plantea el concepto de la predicción como base de la comprensión⁷.

Harste y Burke (1986) desde un enfoque sociolingüístico sostienen que el contexto de lo escrito y de lo impreso posibilita que se formulen esquemas anticipatorios. La identificación del contexto permite la detección de un rango posible de significados y la delimitación de las opciones semánticas, es decir, el ejercicio de la predicción. Remarcan asimismo la génesis social de la predictibilidad. Señalan que cuando no se toma esto en cuenta en la enseñanza de la lectura y escritura, se altera el proceso de la comprensión.

Tomar en consideración los conceptos de predictibilidad y contextualización y aplicarlos al campo de la evaluación implicaría, a nuestro entender, tomar en cuenta cuestiones como:

- las características gráficas, de diagramación, la transcripción correcta de los textos, la ausencia de errores -al respecto se requiere un trabajo muy cuidadoso de revisión de la versión final impresa de la prueba, a cargo de quienes la han elaborado-, y otros aspectos materiales, los cuales deberían confluir a una perfecta legibilidad;

- las instrucciones generales de las pruebas y las de los ejercicios de modo que resulten claras, concisas y no permitan albergar dudas sobre la tarea asignada;

- el modo de presentación de los materiales, que debería llevar a una rápida identificación del género; así, si se trata de un artículo periodístico, la inclusión de una fotocopia del mismo contribuirá más a la contextualización que la mera transcripción; en otro tipo de textos será conveniente citar datos de autor, obra u otros que ayuden a su ubicación; tampoco sería bueno «recortar» un texto en pequeños fragmentos o suprimir palabras por razones de espacio;

- la significatividad de los materiales que se utilicen, en función del marco cultural y del nivel de información;

- la extensión de los textos, en vinculación con el tipo de actividad que se requiera; como criterio general sería preferible seleccionar pocos textos que por su índole y extensión permitan una mayor cantidad y gama de ítems y no muchos textos muy breves;

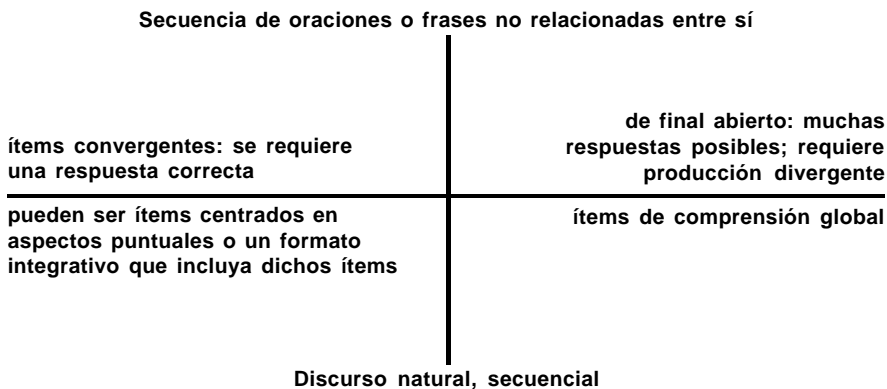
- algunas «ayudas para la lectura» o recursos de los evaluadores tendientes a que el evaluado se detenga en un párrafo o repare en algún aspecto en particular -como por ejemplo introducir entre párrafo y párrafo más espacios que los habituales o algún tipo de instrucción o subtítulos que no existen en el original-, los cuales pueden más bien interferir en la predicción en la medida en que no estén guiados por la índole del texto o por las estrategias del lector, sino por el objetivo del evaluador;

• las preguntas que se formulen -y éste es un punto que retomaremos luego- podrán interferir en la comprensión de la persona evaluada, o al menos no serán suficientemente válidas para su verificación, si en vez de considerar los procesos espontáneos del lector en interacción con el texto y tratar de recalcar en aspectos significativos de ellos, se formulan en cambio a partir de estereotipos o prácticas evaluativas corrientes que no se han revisado a la luz de las perspectivas planteadas.

Alice Omaggio (1993) analiza los tipos de pruebas que se utilizan en EE.UU. para la evaluación en lengua -se refiere en particular a las lenguas extranjeras-, distingue los *achievement tests* que se utilizan en relación con un curso determinado, de los *proficiency tests*, que son pruebas generales de nivel de competencias que se emplean por ejemplo en las universidades para la aprobación de un idioma extranjero, independientemente de que se hayan aprobado los cursos de ese idioma⁸.

Según su descripción, las pruebas del primer tipo se ciñen a aspectos puntuales, a segmentos del currículum de lengua, y formulan preguntas de respuesta estructurada sobre cuestiones gramaticales o léxicas muy específicas, a partir de frases aisladas u oraciones, tales como colocar la preposición que corresponde según el régimen de un verbo dado, completar la traducción de una oración, etc. Los *proficiency tests*, por el contrario, incluyen el nivel textual y las preguntas son mucho más abarcativas; si bien se interroga sobre cuestiones concretas, se requieren también competencias más globales referidas a contenidos, contexto, tipo de discurso, etc.

Omaggio, a la vez que postula que en el currículum se utilicen más las pruebas del segundo tipo, provee desde la perspectiva de la contextualización-descontextualización un esquema interesante para analizar los ítems y para incentivar a los docentes a la construcción de ejercicios adecuados:



Omaggio H.,A, 1993:429 (La traducción es nuestra).

Cada uno de los ejes representa un *continuum*; los ítems que se analicen o se elaboren se ubicarán en un punto del *continuum* en función de su mayor o menor proximidad a los extremos. Pocos ítems podrán entonces ser considerados como totalmente «contextualizados» o «descontextualizados».

Las pruebas escolares podrán incluir ítems ubicados en diferentes puntos del eje horizontal en la medida en que éste representa una gradación del campo lingüístico-comunicacional. En cambio, en lo que respecta al eje vertical, deberían ubicarse sólo en la mitad inferior, ya que la parte superior queda fuera del ámbito del lenguaje natural. Así los dos cuadrantes inferiores constituirían el espacio de inserción de las pruebas que denomina «híbridas» en tanto se utilizan para la evaluación de los aprendizajes vinculados al currículum pero están orientadas a la demostración de competencias más globales (*proficiency oriented*).

Omaggio provee luego una serie bastante rica de ejercitaciones de este tipo para evaluar comprensión oral, comprensión en la lectura y producción escrita.

Nótese que la caracterización que realiza Omaggio y las recomendaciones que formula serían aplicables en nuestro contexto latinoamericano no sólo a la evaluación de los aprendizajes curriculares sino a programas de evaluación diagnóstica o a pruebas de nivel.

4.2. La comprensibilidad

Suele suceder que en pruebas para evaluar la comprensión se utilicen textos poco comprensibles, se formulen preguntas confusas o ambiguas o las respuestas, aun las consideradas correctas, no estén claramente formuladas.

Los estudios de comprensibilidad (*readability*)⁹ se refieren a los factores de tipo lingüístico y psicolingüístico -vocabulario, sintaxis, organización del texto, etc.- que inciden en que un texto resulte más o menos comprensible. Sería de interés tomarlos en cuenta tanto para los materiales que se incluyan en las pruebas como para los ítems.

Una parte de este campo de estudio se refiere a la complejidad sintáctica y describe cuáles estructuras son más comprensibles en función del desarrollo psicolingüístico. Si se trata de una estructura poco comprensible en la comunicación oral -en la cual, por ejemplo, resulte difícil la identificación del sujeto-, más difícil aún será su procesamiento en la lectura. En el otro extremo, una sintaxis demasiado simple -basada por ejemplo en el no uso de conectores y por ende en el uso de

repeticiones ficticias- tampoco ayuda a la comprensión en la medida en que se aleja del lenguaje corriente (Dawkins, 1977).

Desde otra perspectiva, la de la teoría e investigación lingüísticas, Alfredo Hurtado llevó a cabo un estudio sobre las estructuras tardías en niños de escuelas primarias de México: «(...) en el proceso de adquisición [del lenguaje] aparecen estructuras que sólo se interpretan, se repiten o se producen **después de otras**, en el orden de adquisición. A estas estructuras se las llama tardías con respecto a las anteriores» (Hurtado, 1984:23).

Hurtado plantea la utilidad del conocimiento de las estructuras tardías en español para formularse luego los interrogantes de si es conveniente «*aprender a leer una estructura antes de interpretarla, repetirla y producirla espontáneamente en el habla*» y «*escribir una estructura antes de interpretarla, repetirla y producirla espontáneamente en el habla*» (Hurtado, 1984:16).

Más allá de las consideraciones teóricas y de las investigaciones que pueden ir en ayuda de quienes elaboren las pruebas, las evaluaciones con fines de diagnóstico posibilitan más que otras validar los instrumentos en la etapa de su elaboración, antes de la prueba piloto. Los mejores jueces sobre la comprensibilidad serán los usuarios de la prueba y, si se los deja expresarse a través de estrategias de validación en pequeños grupos, será posible mejorarlas.

92

4.3. *Las preguntas*

Sobre qué se interroga

Este punto se vincula con la cuestión de la **validez**, es decir en qué medida el contenido de las preguntas y de la prueba en su conjunto se refiere a aspectos significativos y pertinentes respecto de la competencia que se pretende evaluar.

En las teorías sobre evaluación en lengua predomina actualmente la de la «competencia unitaria», es decir, que las diversas habilidades del lenguaje son vistas como manifestación de una única capacidad subyacente, frente a las teorías según las cuales hay que considerar independientemente distintas subhabilidades (Davies,1992).

Cualquiera sea la postura que se adopte, aun en los modelos de evaluación por objetivos, un conocimiento acotado puede evaluarse poniendo énfasis ya sea en la información, en el conocimiento de una noción gramatical o en la **competencia** lingüística vinculada a esa noción. En las pruebas corrientes abundan más los ítems del primer tipo.

En el campo de los estudios sobre la lectura, luego de décadas de predominio de las teorías que señalaban los tres niveles «clásicos» de comprensión: la literal, la capacidad de inferencia y la lectura crítica, desde fines de los 70 y comienzos de los 80 empezaron a predominar las visiones de la lectura como proceso global.

Maria y MacGinitie (1982) examinan los modelos sobre la lectura; señalan que tanto en el de Rumelhart como en el de Kintsch, el lector es visto comprometido simultáneamente en un procesamiento de muchos niveles en forma paralela. Interactúan en la comprensión del texto los procesos de tipo «de abajo hacia arriba» (*bottom-up*) y «de arriba hacia abajo» (*top-down*), términos que hacen referencia respectivamente a la identificación de palabras, el acceso a su significado, el análisis sintáctico, y el objetivo del lector al leer, su conocimiento del mundo y los esquemas (*schemata*) que estructuran el texto.

Así, en ambos autores, la distinción entre decodificación y comprensión, que en los modelos anteriores respondía a dos niveles jerárquicos diferentes, resulta desdibujada.

Van Dijk, T.A. y Kintsch, W. (1983) plantean un modelo **estratégico** que da razón desde el procesamiento de las unidades de palabras hasta el de las unidades abarcativas de temas o macroestructuras. Lo denominan «estratégico» porque el oyente o lector ejercita un entendimiento lineal rápido, que es hipotético y podrá rectificarse luego, durante el transcurso de la audición o lectura. Durante el procesamiento, los niveles interactúan de una manera intrincada y compleja.

Las estrategias de análisis que se utilizan dependen no sólo de las características del texto sino de las del oyente/lector, tales como su conocimiento del mundo y otros. Así el lector reconstruirá no sólo el sentido del texto señalado por el escritor, sino el sentido que es más relevante para sus propios intereses y objetivos.

En un primer paso, a través de una estrategia abarcativa que incluye subestrategias -decodificación de rasgos fonológicos/gráficos, identificación de fonemas/letras y de morfemas, dependientes de la interpretación semántica y sintáctica subyacente-, se llega a la construcción de proposiciones relacionadas entre sí que constituyen un **texto base** anclado en una coherencia de tipo **local**, que tiene lugar en la memoria de corto término.

Además de esta representación en la memoria episódica, se constituye también una representación del hecho o situación de que trata el texto, de su tema, que se denomina **macroestructura**. «Entender un texto significa que la gente es capaz de construir un modelo mental para

el texto. E inversamente, en la producción del texto, el modelo es el punto de partida para todo procesamiento: la gente sabe algo acerca de un hecho, y este conocimiento está representado en su modelo del hecho, y este modelo servirá como la base, por ejemplo, para contar una historia acerca del hecho» (van Dijk, 1995:26).

Por último los autores señalan que muchos tipos de discurso poseen una estructura convencional -así por ejemplo la narración o la argumentación-, que se denomina *superestructura*.

Las pruebas de lectura suelen estar más vinculadas a modelos anteriores a los analizados, que establecían jerárquicamente niveles, vinculados a habilidades y subhabilidades, desde la decodificación hasta la comprensión, que tendrían lugar en forma sucesiva y progresiva en el individuo. En las pruebas de lectura, en los ejercicios con referencia a un texto, el esquema más habitual es hacer preguntas sobre detalles de los contenidos de los párrafos y al final alguna pregunta más global. Aun cuando se den respuestas acertadas, uno podría en muchos casos preguntarse si realmente se comprendió el texto.

Una de las prácticas comunes en las pruebas la constituyen las preguntas sobre **inferencias**; ello es debido al papel clave que la mayoría de los autores atribuye a la inferencia en el proceso de la lectura. Sin embargo, es muy frecuente que las preguntas de ese tipo se refieran a aspectos casi totalmente explícitos; en las respuestas esperadas se construyen enunciados muy fácilmente deducibles a partir del enunciado dado en el texto, que por lo tanto no requieren del lector ninguna actividad heurística o de formulación de hipótesis.

Este tipo de preguntas podría estar originado en la exigencia de responder a una taxonomía elegida para la evaluación, ya que de hecho muchas de las inferencias posibles no agregan nada relevante para la comprensión del texto.

Mientras por una parte la inferencia interviene permanentemente en la comprensión y no sólo en los puntos que se eligen para construir preguntas *ad hoc*, a la vez respecto de éstas suelen esperarse respuestas obvias, que no son las únicas posibles, ya que el ejercicio de la inferencia puede anclarse en diferentes indicios presentes en un texto que conduzcan a conclusiones diversas.

En cuanto a otros tipos de preguntas de comprensión, hay que ver a qué contenidos se les da mayor importancia. Si bien podría haber consenso sobre algunos aspectos considerados claves, también hay un campo muy subjetivo, ya que diferentes buenos lectores podrían ser

impresionados preponderadamente por distintos aspectos del texto. La validación por juicios de expertos o por diversos lectores puede ayudar en este sentido.

A la vez sería útil introducir en una prueba de lectura una ejercitación que planteara por ejemplo: «si usted quisiera saber si una persona comprendió este texto ¿qué preguntas le haría para comprobarlo?». La información obtenida sería de interés en términos de diagnóstico y contribuiría a disminuir la unidireccionalidad propia de ese tipo de pruebas.

Una forma de acercarse a la característica de «objetividad» de las pruebas sería la construcción de una «intersubjetividad» que redujera el grado de subjetivismo en cuanto a las preguntas que se formulen y a las respuestas esperadas. La ampliación de la cantidad e índole de personas que intervengan en la elaboración de las pruebas, la incorporación de metodologías que incluyan la participación de diferentes tipos de jueces y la asignación de un rol más activo a los evaluados, tanto en las etapas de construcción de los instrumentos como en el momento de la evaluación, podrían contribuir a ello.

Cómo se construyen los ítems

En este punto es imprescindible -aunque no suficiente- apelar a la tecnología evaluativa, muy desarrollada dentro de las ciencias de la educación y de las ciencias sociales. Existen cantidad de prescripciones acerca de la buena redacción de un ítem, de los aspectos formales que inciden en la elección de una respuesta, así como formas de validación: análisis de pertinencia ítem-objetivo, y, *a posteriori*, de su administración, análisis de consistencia interna, de grado de discriminación de los ítems, etc.

Sin embargo, este conocimiento acumulado no siempre se ha utilizado en forma apropiada sino que ha conducido a veces a la consolidación de usos rutinarios y poco creativos. Por otra parte las pruebas objetivas conllevan una cantidad de restricciones. No obstante, como innumerables experiencias lo demuestran, es posible ampliar y enriquecer ese marco.

En una experiencia, expuesta en 1982, de pruebas administradas a lo largo de dos años a un total de 37.300 estudiantes que ingresaban a la Universidad Nacional Abierta de Venezuela se introdujeron diversas innovaciones en el diseño de las pruebas y en la estructuración de los ítems (Costa, 1984).

Uno de los aspectos que observábamos entonces es que a veces en las pruebas de lectura se hacen preguntas que se pueden contestar

independientemente del texto. En otros casos, si bien la respuesta correcta está referida al texto, las alternativas falsas plantean contenidos ajenos a él.

En las pruebas que presentábamos, todas las alternativas - la correcta y las falsas - se construían con palabras o ideas del texto, utilizando para ello transformaciones sintácticas, o alguna alteración en los significados, etc.

Queremos señalar que nuestra experiencia en el diseño de pruebas, o en la revisión de pruebas elaboradas por otras personas, nos lleva a reconocer la importancia de la observación de los propios procesos durante la lectura de los materiales seleccionados, como una de las fuentes a las que puede apelar el evaluador para construir los ítems, en vez de recurrir a los modelos preestablecidos.

Un último comentario se refiere a las preguntas abiertas o de desarrollo. Muchas veces se requiere un ejercicio escrito para verificar la comprensión en lectura. En ese caso se evalúa la comprensión a través de la producción y el material es obviamente más rico.

Los parámetros que se construyan para la corrección de esos ejercicios deberían ser suficientemente amplios como para considerar los modos de aproximación de las distintas personas a la lectura y sus estrategias de escritura, que no siempre coinciden con los más valorizados en nuestra tradición escolar y académica, vinculados más bien al pensamiento lógico y a los procesos de tipo racional.

Para concluir, creemos que hay que trabajar para generar marcos amplios que permitan el desarrollo de una gama rica y variada de estrategias de evaluación.

Conclusiones

- Las evaluaciones que se realizan con el objeto de medir la calidad- u otras afines- ponen en evidencia los criterios y prácticas existentes en el sistema educativo. Sin embargo, múltiples prácticas innovadoras no se han visto convenientemente representadas en la evaluación.

- Existen importantes desarrollos en evaluación de la lengua- aquí y en otros países- que han incorporado criterios de las ciencias del lenguaje. Hay asimismo aspectos de la teoría lingüística, de la sociolingüística y psicolingüística, del campo interdisciplinario de los estudios sobre **lectura** (*reading*), donde confluyen la psicología cognitiva, la lingüística del texto, etc., que pueden ser aplicables a la evaluación.

- Las pruebas objetivas crean un marco constrictivo para el evaluador, sin embargo es posible superar algunas de sus limitaciones. Si queremos que las poblaciones evaluadas puedan desplegar sus saberes hay que desechar algunos estereotipos y construir instrumentos válidos y confiables en los que se formulen preguntas inteligentes, cuyas respuestas esperadas contemplen la diversidad lingüística, cultural y cognoscitiva.

Notas

(1) Una primera versión de este trabajo fue presentada en las Jornadas «De la teoría lingüística a la enseñanza del español», en la Facultad de Filosofía y Letras de la Universidad de Buenos Aires, junio de 1993.

(2) Para una apreciación de la vigencia del asunto en lo que respecta a universidades, cf. *Évaluation et Qualité. Evaluation and Quality*. CRE-action. Revue trimestrielle de la Conférence permanente de recteurs, présidents et vice-chanceliers. N° 96, 1991/4, con referencia al contexto europeo, y en lo que concierne a nuestra región cf. Puiggros, A. y Krotsch, C. (comp.) *Universidad y evaluación. Estado del debate*. Buenos Aires, Rep. Arg. IDEAS_AIQUE, 1994.

(3) Cf. en Edwards, Verónica, su análisis en torno del concepto de calidad y las relaciones que establece con las concepciones curriculares.

(4) En cuanto al concepto de variedad «estándar», entendemos que no ha habido suficiente fundamentación científica. Menos aún en la consideración de determinadas formas dialectales como estándar o no. Otra denominación, la de variantes «prestigiosas» pone al descubierto que se trata más de valoraciones sociales que lingüísticas.

(5) En una investigación en la que participamos, llevada a cabo por la Dirección de Educación del Adulto como parte de un estudio regional de la OREALC, se administraron pruebas objetivas de lengua y matemática a 300 adultos con escolaridad básica incompleta. Con 30 de ellos se realizó una entrevista etnográfica; el análisis lingüístico del material grabado permitió contrastar el resultado de las pruebas de lectura y escritura con el desempeño oral. (*Analfabetismo funcional*. Ministerio de Cultura y Educación-OREALC. Informe final en preparación).

(6) Mencionado por de Granda, Germán, en el seminario de doctorado «El contacto lingüístico», en la Facultad de Filosofía y Letras, UBA, nov. 1993.

(7) En su muy valiosa y conocida obra *Understanding Reading*, destaca la importancia de la información no visual -la que está en el cerebro- y de la predicción. La predicción posibilita la reducción de la incertidumbre, a través de la eliminación anticipada de ciertas alternativas y posibilita la comprensión.

(8) Cf. en especial el capítulo 9 «Classroom Testing».

(9) *Readability* se ha traducido al castellano como *lecturabilidad*, seguramente con el objeto de distinguirla de la palabra *legibilidad*, que se refiere a los aspectos materiales de lo escrito. Otros autores usan *legibilidad* en el sentido amplio que tiene el término *readability* en inglés .

(10) Cf. en Dubois (1984), el análisis que realiza sobre el papel de la inferencia en los distintos modelos explicativos de los procesos de la lectura.

Referencias bibliográficas

- BERNSTEIN, Basil «Social Class, Language and Socialization» en *Power and Ideology in Education*, Karabel and Halsey Eds., New York, Oxford University Press, 1977.
- CAZDEN, Courtney B., «Social Context of Learning to Read» en Walter MacGinitie (comp.), *Foundations of Reading*. Readings for ED-B 342. University of Victoria. s/f
- CAZDEN, Courtney B., «La lengua escrita en contextos escolares» en Ferreiro, E. y Gómez Palacio, M. (comp.) *Nuevas perspectivas sobre los procesos de lectura y escritura*. Buenos Aires, Siglo Veintiuno, 1986.
- COSTA, Marta. «Pruebas objetivas para evaluar la comprensión en lectura» en *Lectura y vida*, IRA, año 5, Nº 1, marzo 1984.
- DAVIES, Allan «Language Testing and Evaluation» en Crabe, W. and Kaplan *Introduction to Applied Linguistics*. Reading, Massachusetts, Addison-Wesley Publ.Comp, 1992.
- DAWKINS, John, *Syntax and Readability*. Newar, Delaware, IRA, 1977.
- DÍAZ BARRIGA, Ángel «Una polémica en relación al examen». En *Revista Iberoamericana de Educación*, (Nº 5), 1994, ps. 161-181.
- DIJK, Teun A. van «De la gramática del texto al análisis crítico del discurso» en *Boletín de Estudios Lingüísticos Argentinos*. año 2, (No.6), mayo 1995, ps. 20-40.
- DIJK, T.A. van and KINTSCH, W. *Strategies of Discourse Comprehension*. Academic Press, New York, 1983.
- DUBOIS, María Eugenia «Algunos interrogantes sobre comprensión en la lectura» en *Lectura y Vida*. IRA, Año V, Nº 4, dic. 1984.
- EDWARDS R., Verónica, *El concepto de calidad de la educación*. Santiago, Chile, UNESCO-OREALC, 1991.
- GÓMEZ PALACIO, M. y otros, *Indicadores de la comprensión lectora*. Washington, OEA, 1993.
- HARSTE, Jerome C. y BURKE, Carolyn N., «Predictibilidad: un universal en lecto-escritura» en Ferreiro, E. y Gómez Palacio, M. (comp.), *Nuevas perspectivas sobre los procesos de lectura y escritura*. Buenos Aires, Siglo Veintiuno, 1986.
- HURTADO, Alfredo y colab., *Estructuras tardías en el lenguaje infantil*. México, Dirección General de Educación Especial -SEP-OEA, 1984, tomo I.
- LAMB, Pose, «Dialects and Reading» en Robert Shaffer Ed, *Applied Linguistics and Reading*. Newar, Delaware, IRA, 1981.
- MAGENDZO, Abraham, «Calidad de la educación y su relación con la cultura: síntesis de una investigación en un área indígena de Guatemala» en *La Educación*. Revista

Interamericana de Desarrollo Educativo. OEA, Washington, año XXVIII, N° 96, ps. 32, 47.

MARIA, Katherine and MacGINITIE, W.H. «Reading Comprehension Disabilities: Knowledge Structures and Non-Accommodating Text Processing Strategies» in *Annals of Dyslexia*, vol 32,1982.

OMAGGIO HADLEY, Alice, *Teaching Language in Context*. Boston, Massachusetts, Heinle and Heinle Publ., 1993.

PRADO de SOUZA (org.), *Avaliação do Rendimento Escolar*. Campinas, Papirus. 1992.

RAMA, Germán W, *Qué Aprenden y Quiénes Aprenden en las Escuelas de Uruguay*. Montevideo, CEPAL, 1992.

SMITH, Frank, *Understanding Reading*. New York, Holt, Rinehart and Winston, 1978.

Una Opinión. Caracas, vol 4, N° 2, abril 1982 (número dedicado a la evaluación).