

A Avaliação de Desempenho na Área de Língua na Argentina: a Utilização de Provas Objetivas(1)

Marta Elena Costa(2)

Apresentação

Neste trabalho, propomo-nos a:

- convidar os leitores a uma reflexão conjunta e, ao mesmo tempo, expor conceitos, critérios e pautas que possibilitem uma melhor compreensão dos processos externos de avaliação da qualidade, isto é, daqueles que não estão integrados ao processo de ensino-aprendizagem;
- incentivar as pessoas que delineiam as avaliações a explorai novos estilos;
- incentivar os professores a uma participação mais lúcida e construtiva no processo de avaliação.

1 Título do original em espanhol: la evaluación del desempeño en el área de lengua: utilización *de pruebas objetivas*. Tradução de Cláudia Schilling revisão técnica da tradução de Maria José do Amaral Ferreira.

2 Professora e pesquisadora da Faculdade de Filosofia e Letras da Universidade de Buenos Aires, tendo lecionado em outras universidades na Argentina e na Venezuela e trabalhado em organismos de planejamento e pesquisa argentinos e internacionais É autora de diversos livros e artigos na área de Educação em geral e na área de Avaliação em particular.

O trabalho trata da avaliação em língua, com exemplos do espanhol, que é a língua nacional da Argentina e a língua materna da maioria de seus habitantes, o que no Brasil seria equivalente à avaliação em língua portuguesa. Ressaltamos, porém, que, apesar de focalizarmos sobretudo a avaliação em língua, alguns dos problemas que colocamos aqui também podem ser aplicados à avaliação em outras áreas da aprendizagem.

Em primeiro lugar, realizaremos uma exposição teórica, que inclui a revisão de diversos autores que tratam da avaliação da qualidade e dos problemas que o tipo de instrumento nela utilizado coloca.

Como este número da Série Idéias refere-se fundamentalmente aos sistemas de avaliação de rendimento escolar no Brasil, e como nos foi solicitado que expuséssemos a experiência argentina, em seguida apresentaremos, como ilustração, provas utilizadas na avaliação da qualidade em nosso país - com uma descrição prévia dos procedimentos nacionais de avaliação, de modo que se possa conhecer o contexto no qual elas foram aplicadas. Ressaltamos contudo, novamente, que os problemas que aqui são colocados ultrapassam o âmbito desta ou de qualquer outra avaliação diagnóstica ou de promoção de um ciclo para outro etc., realizada externamente ao processo de ensino-aprendizagem, em qualquer um de nossos países; o que pretendemos, de fato, é encontrar os pontos críticos da prova e dos instrumentos utilizados para neles focar nossa atenção, a fim de que seja possível transferir para outras experiências uma observação perspicaz que tentaremos exercitar a propósito dos exemplos aqui considerados.

Introdução

Nos mais variados setores tem-se apontado a diminuição da qualidade em educação, e atualmente este fato parece ser uma preocupação tanto nos países centrais como nos de outras regiões.

No campo da língua materna, parece haver uma aceitação generalizada de que, embora hoje em dia uma porcentagem mais elevada de pessoas tenha acesso à educação formal e à leitura, em geral aqueles que lêem, lêem menos e mais superficialmente ou com menor compreensão, em comparação com o modo como antes se lia. A mesma coisa parece suceder com as opiniões correntes e os diagnósticos referentes ao desempenho em escrita nos diferentes níveis - primário, médio e universitário.

Diante disso, parece ter emergido a necessidade de ponderar, com relação a padrões de medida, a qualidade da educação ministrada nos sistemas educacionais e, entre os múltiplos componentes a serem avaliados, têm-se levado em conta principalmente as aprendizagens dos alunos.

A avaliação educacional, que tem décadas de tradição nos países centrais, desde os anos 80 aumentou muito em nível dos sistemas educativos na América Latina, por impulso tanto dos organismos que operam na região - os de educação, ciência e cultura -, como dos organismos que propõem as políticas econômicas.

Esta tendência acentuou-se nos anos 90, chegando a ser incorporada como uma instância institucionalizada nos processos de reformas ou de transformações educacionais

impulsionadas pelo Estado, inclusive nas leis que regulam a educação dos países. Citamos como exemplos os casos da Argentina (Lei Federal da Educação - 1993), Bolívia (Lei da Reforma Educacional - 1994) e Colômbia (Lei Geral de Educação - 1994).

Paralelamente ao auge desse tipo de avaliação, iniciou-se a discussão sobre ele nos setores envolvidos - professores, pais de alunos -, no âmbito político, sindical e em outros, que se somou àquela já existente no âmbito acadêmico sobre o sentido e a finalidade dessas práticas e dos instrumentos utilizados.

Na Argentina, no final de 1993, levou-se a cabo o primeiro processo de avaliação da qualidade, o qual, ao ser divulgado, provocou um intenso debate, acompanhado pelos meios de comunicação. Discutia-se a questão da qualidade, a forma em que deveria ser apreciada e questionava-se quem deveria ser responsabilizado pelos resultados.

Parte das reações provocadas deve ser entendida no contexto de que na Argentina não há exames de massa - vestibulares - para ingressar na Universidade ou no secundário, nem para o encerramento de nenhum ciclo de estudos⁽³⁾. Outro aspecto que provocou polêmica foi que, para o público, a avaliação das aprendizagens adquiridas pelos alunos apareceu separada dos demais componentes da avaliação do sistema educacional, igualando-se, portanto, ao significado total da avaliação.

A Questão da Qualidade da Educação e sua Avaliação

Embora exista uma preocupação generalizada com a qualidade da educação, não há consenso sobre o que se entende por qualidade e menos ainda sobre até que ponto e em que modalidades é possível avaliá-la.

A medição da qualidade em educação surgiu, como já foi dito reiteradamente, como decorrência de um modelo de eficiência econômica. A qualidade dos sistemas educativos e dos processos que neles ocorrem é vista, assim, em termos de eficiência. Estes modelos de avaliação alimentam-se de uma racionalidade própria ao modo de produção industrial.

Sandra Zákia LEAN DE SOUSA (1992) realiza uma cuidadosa revisão das principais teorias de avaliação da aprendizagem desde o início do século nos Estados Unidos e, particularmente, dos autores que tiveram influência no Brasil - influência que, na nossa opinião, estendeu-se a outros países da região -, assinalando como característica comum a eles

“... a tendência a valorizar a dimensão tecnológica da avaliação, enfatizando o caráter cientificista e os métodos e procedimentos operacionais. Essa tendência manifestada nos estudos sobre a avaliação da aprendizagem reflete os princípios da pedagogia tecnicista, provenientes da teoria geral da administração, que direcionaram

³ Em 1984, com a chegada do governo democrático, foram suprimidos os exames de ingresso às universidades estabelecidos pelo governo *de facto* criando-se o ingresso direto. Nos períodos anteriores, nos quais a autonomia universitária estava em vigor, cada Universidade estabelecia seus próprios regimes de admissão. Atualmente, com a vigência da Lei de Educação Superior sancionada em 1995, reafirmou-se essa atribuição das universidades.

a política educacional do Estado brasileiro pós-64. De cunho empresarial-tecnocrático, delineou-se a partir de padrões de racionalidade, eficiência e redução de gastos"(4). A autora ressalta que deve ser elaborada urgentemente uma teoria da avaliação que, baseando-se em princípios e valores comprometidos com a transformação social, permita a superação destas concepções e contribua para que a avaliação se converta em um processo de pesquisa que tenha como ponto de partida e de chegada o processo pedagógico que ocorre na escola(5). Nos mesmos países onde surgiram as posturas que se caracterizavam por valorizar a eficiência, produziu-se nos últimos anos uma maior abertura em relação aos enfoques denominados qualitativos. Margaret LeCOMPTE (1995) celebra o fato de que, aos padrões para a avaliação de programas nos Estados Unidos, tenha se incorporado, em 1993,

"um reconhecimento oficial daquilo que (..) tem ocorrido na educação e nas ciências .sociais ao longo dos últimos cinquenta anos: um distanciamento da exclusiva dependência das normas positivistas para medir a qualidade da pesquisa"(6).

Assim, ante as posturas tecnicistas no tocante à qualidade da educação e a certas práticas correntes na avaliação, surgem posturas alternativas, tanto nos países centrais como na nossa região, algumas delas enunciadas por destacados teóricos da Educação a partir de contextos mais amplos; em outros casos, a crítica concentra-se nos aspectos referentes aos tipos de avaliação, seus parâmetros e instrumentos utilizados.

No caso de autores latino-americanos, existem posições como as de Abraham MAGENDZO e outros, que vinculam a qualidade à cultura da população sujeito da educação. Esta corrente vincula-se a linhas de pensamento e práticas educacionais muito arraigadas na América Latina e em outras regiões do mundo. A partir desta perspectiva propõe-se que, numa educação de qualidade, é imprescindível que as instituições educacionais levem em conta as diferentes culturas e se alimentem delas para poder levar à cabo as aprendizagens pelas quais a escola é responsável. MAGENDZO (1984) afirma que

"existe uma relação direta entre qualidade e cultura, isto é, a educação aumenta a sua qualidade na medida em que for capaz, através do currículo, de valorizar e apropriar-se das formas de pensar, sentir e atuar de uma comunidade para enfrentar e dar significado à sua cotidianidade [...]. Não são os estudos ou as avaliações comparativas de níveis de rendimento nacionais ou internacionais que permitem aquilatar a qualidade da educação"(7).

4 LEAN DE SOUSA, Sandra Zákia. Revisando a teoria da avaliação da aprendizagem. In: PRADO DE SOUSA, Clarilza (Org), avaliação do **rendimento escolar** Campinas: papirus, 1992. p.45.

5 Idem cf p.45/46.

6 LeCOMPTE, Margaret. Un rosnáramo conveniente: diseno de investigación cualitativa y estándares para la evaluación de programas. Revista Electrónica de Investigación y Evaluación Educativa, AIDIPE, e 1, n.1,1995.

7 MAGENDZO, Abraham. Calidad de la educación y su relación con la cultura: síntesis de una investigación en un área indígena de Guatemala. la Educación: Revista Interamericana de Desarrollo Educativo, OEA, Washington, n. 96, p. 32,1984.

De acordo com este autor, podemos afirmar que numerosos estudos transculturais realizados

permitem inferir que muitas das categorias conceituais desenvolvidas para medir a qualidade da educação desconhecem as diferenças cognitivas, perceptuais, de valor e de aprendizagem dos diferentes grupos culturais(8).

A discussão sobre a avaliação da qualidade é relativamente nova com essa denominação, porém está vinculada à questão que foi levantada nas primeiras décadas deste século sobre a legitimidade da utilização de instrumentos padronizados para a avaliação em contextos educacionais.

DAMASCO BARRIGA (1994) analisou recentemente a questão do exame e de sua versão moderna, os testes ou provas objetivas(9). Segundo este autor, o exame deixou de ser um aspecto do método ligado à aprendizagem, como o era na concepção de Comenio, para converter-se em um meio para a promoção, mudando-se assim a prática pedagógica, pois ao se levar em conta apenas a promoção, os esforços de professores e alunos descentram-se dos processos de ensino-aprendizagem. A avaliação foi-se configurando como um âmbito desvinculado do didático; adquiriu uma autonomia como campo de estudo e como prática, baseada apenas em aspectos técnicos, e dissociou-se do campo educacional. Os reducionismos técnicos empobreceram o debate educativo.

A revisão dos trabalhos dos autores mencionados sugere-nos que a busca da qualidade deveria desvincular-se do afã por sua medição, pois este não parece ter sido saudável para a educação por descentrar o olhar e a tarefa.

*Mas, de fato, a avaliação a partir de instâncias externas às instituições onde o ensino é ministrado já está instalada nos sistemas educacionais e legitimada pelas normas em vigor. Seria interessante então orientar o debate neste campo específico tanto para seus fundamentos e para os significados atribuídos às práticas de avaliação nas diferentes representações sociais, como para o reconhecimento de seu valor real em termos daquilo **que se pode e não se pode diagnosticar** através delas, recuperando ao mesmo tempo fundamentos teóricos que favoreçam novas experimentações.*

A nosso ver, além do contexto político de ajuste em que surgiram em muitos casos - e por influência dos organismos que operam na região da América Latina e do Caribe -, é preciso identificar como tais práticas de avaliação podem ser utilizadas positivamente nos países democráticos e analisar os limites da informação que elas fornecem - em estreita vinculação com os tipos de instrumentos utilizados -, além de verificar em que medida é possível melhorá-las e obter uma interpretação mais abrangente dos resultados; em suma, que proveito poder-se-ia obter com essas avaliações e que outras estratégias complementares ou paralelas poderiam ser utilizadas.

8 Idem, p. 33.

9 Cf DÍAZ BARRIGA, Angel. Una polémica en relación al examen. Revista Iberoamericana de Educación, n. 5, p.161-181,1994.

Aspectos Críticos nas Provas de Língua

Seguindo o enquadramento que propusemos, faremos referência, agora, às provas utilizadas para avaliar o desempenho em língua. Centraremos a nossa análise nas provas padronizadas denominadas "objetivas", ou de "resposta estruturada", que são utilizadas não apenas em ações de avaliação da qualidade, mas também em outras situações de avaliação referentes à competência em sentido amplo - provas de nível, provas de aptidão etc. para a aprovação ou admissão em um ciclo de estudos ou para pesquisas diagnósticas. Embora nosso trabalho, aqui, refira-se a testes de língua, alguns indicadores podem ser aplicados às avaliações em outras áreas, na medida em que sejam introduzidos instrumentos nos quais a linguagem intervenha.

Embora possuam múltiplas vantagens no caso de populações numerosas, as provas objetivas apresentam problemas em sua elaboração e aplicação e na interpretação dos seus resultados. Abordaremos alguns dos problemas vinculados tanto às taxonomias (que se procura avaliar) como à natureza das provas.

Em nossos trabalhos anteriores¹⁰, apontamos aspectos críticos desse tipo de provas. De uma perspectiva que integra as contribuições da ciência da linguagem, consideraremos aqui alguns desses aspectos críticos: **o tipo de linguagem** utilizado, **a contextualização**, **a compreensibilidade e as perguntas** formuladas.

A Linguagem

Nas avaliações da língua oficial do país, geralmente toma-se como referência a variedade de língua denominada "padrão"⁽¹¹⁾ e, dentro dela, as modalidades próprias da língua escrita. Deste modo, limita-se ou restringe-se o objeto de avaliação a uma parte da competência lingüística da pessoa examinada. Ficam de fora a linguagem oral e as formas coloquiais da linguagem escrita. Embora estas últimas costumem ser avaliadas através de exercícios de expressão escrita, não se incorporam nem aos textos apresentados nem aos itens de resposta estruturada.

Assim, podem encontrar-se em situação desvantajosa para responder aquelas pessoas que, pela sua origem sócio-econômica, tenham menos contato com a língua escrita, ou cujas culturas de origem se expressem oralmente, e/ou cujo uso da língua está mais distanciado da variedade-padrão, tanto a da escola e dos textos escolares como a dos instrumentos de avaliação.

Uma avaliação que abranja uma gama mais ampla das competências lingüísticas também pode ser interessante para o diagnóstico, na medida em que possibilita o estabelecimento de relações entre a linguagem oral e a escrita ou entre o uso do próprio dialeto e a compreensão da variedade-padrão, todos eles correspondentes à mesma língua.

¹⁰ Ver na bibliografia obras de nossa autoria citadas

¹¹ Quanto ao conceito de variedade "padrão", entendemos que não tem havido suficiente fundamentação científica, e menos ainda na consideração de determinadas formas dialetais como padrão ou não. Outra denominação, a de variantes "prestigiosas", evidencia que neste caso se trata mais de avaliações ou de valorações sociais que lingüísticas.

Quando se trata de avaliações em nível nacional, é necessário utilizar os estudos de dialetologia e consultar especialistas para que possamos incorporar como um dos aspectos a serem avaliados nos ciclos primário e secundário os desempenhos nas variedades próprias. Se isso não for possível, será preciso levar em conta essa limitação na hora de analisar os resultados.

As contribuições da sociolinguística podem ajudar a enfrentar essas questões, que se tornam mais graves em zonas de línguas de contato(12), ou em relação a falantes para os quais a língua avaliada é sua segunda língua. Seria desacertado ignorar as diferenciações internas existentes nas línguas, vinculadas ao âmbito geográfico-espacial, à estratificação sociocultural, ao âmbito estilístico etc., tendendo-se, assim, a uma falsa homogeneidade. Por exemplo, o espanhol, que se estendeu por um enorme espaço geográfico, conservou sua unidade e sua riqueza. Uma forma, entretanto, de garantir esta unidade é favorecer a compreensibilidade entre os habitantes destas e daquelas zonas, destas e daquelas camadas sociais e o acesso de todos eles a diferentes tipos de discursos.

Na realidade, esses pontos que assinalamos são problemas que surgem no currículo de língua e que adquirem ressonância na hora da avaliação; trata-se de questões vinculadas às políticas lingüísticas, sejam estas explícitas ou não.

Vemos, então, que as avaliações diagnósticas não só oferecem informação sobre o desempenho das populações avaliadas, como também evidenciam os usos da linguagem considerados válidos no sistema educacional. Dependendo da forma em que forem elaboradas, elas podem proporcionar um *feedback* a esse sistema, esclarecendo inúmeros aspectos a serem levados em conta por professores e planejadores para que possam resultar em uma melhoria das atitudes lingüísticas; em caso contrário, podem reforçar situações de segmentação social.

Nos últimos anos, tem havido um aumento da consciência de professores e pesquisadores em torno dessa problemática, e foram levadas a cabo valiosas experiências no ensino da língua e na avaliação na sala de aula. Aos currículos oficiais de nossos países muitas vezes foram incorporadas concepções amplas e respeitosas da diversidade lingüística e cultural(13)', mas foi mais difícil introduzi-las na teoria e na prática do tipo de avaliação em questão.

A Contextualização

Freqüentemente são apresentados diferentes tipos de itens nas provas, e cada um dos quais se refere a aspectos pontuais e tem pouca ou nenhuma relação com o resto. Perguntas clássicas deste tipo são aquelas que interrogam sobre o significado ou o sinônimo de uma palavra isolada, ou as que apresentam uma oração na qual é preciso colocar uma palavra que falta. Assim, desta ou de outras formas são apresentados exercícios artificiais que se distanciam dos usos naturais ou habituais da linguagem.

12 Alguns exemplos de línguas em contato na Argentina são, conforme as diferentes zonas: espanhol-português, espanhol-língua indígena, espanhol-língua estrangeira de migrantes de outros continentes, e numerosas combinações possíveis.

13 Na Argentina isto aconteceu com os conteúdos básicos comuns de língua para o nível inicial e para o ensino geral básico.

No caso da inclusão de textos, estes costumam ser demasiadamente resumidos ou, pelas suas próprias características, não facilitam a formulação de perguntas adequadas, ou as perguntas, mesmo que o texto seja adequado, não conseguem se referir a aspectos significativos. Em outros casos, trata-se de textos com certa complexidade sobre os quais se formulam perguntas simples demais.

Em vez de promover os processos da compreensão e da produção, uma prova assim elaborada os. Parece que, com esse tipo de instrumento, se testa mais a habilidade de desagregar os próprios saberes do que o exercício integral da competência lingüística. Estas e outras questões podem ser englobadas naquilo que denominamos **contextualização**.

Courmey CAZDEN (1982) explica que

o exercício de qualquer processo cognitivo - tanto para nós como para as crianças - ocorre em algum contexto: um formato particular de tarefa, condições físicas determinadas, uma certa organização social, regras convencionadas etc.; e as características desse contexto darão apoio ou criarão obstáculos para as tarefas cognitivas nele desempenhadas. A cognição está sempre em um contexto e não pode ser ensinada ou avaliada separadamente de um contexto particular .

Em um trabalho posterior (1986), CAZDEN distingue os contextos internos - os contextos mentais do leitor e do escritor - e os **contextos sociais externos**, nos quais ambos se encontram. Frisa a incidência de ambos os tipos de contexto na leitura e na escrita e assinala a necessidade de incentiva-los para se obter um bom desempenho`.

Consideramos pertinente aplicar a noção de contextualização - e da noção contrária, a de não -, tanto à situação em que a prova é aplicada como aos instrumentos em si.

A Situação de Prova

As pessoas avaliadas devem saber por que são avaliadas, para que serão utilizados os resultados e a partir de que ótica eles serão interpretados. A posteriori, deveriam ter alguma devolutiva sobre a sua própria atuação e os resultados gerais, e deveria haver formas de aproveitamento imediato dessa informação, além dos fins diagnósticos.

Com relação às provas objetivas, lembramos que, em alguns ambientes, não se está suficientemente familiarizado com esse tipo de instrumento; neste caso, poder-se-á exigir previamente - se possível, em uma sessão anterior - um treinamento para sua utilização. Ambos os aspectos podem eliminar insegurança ou qualquer outro tipo de mal-estar na hora da avaliação. Outras recomendações de tipo pedagógico podem criar um bom clima, que

14 CAZDEN, Courtney B. Social context of learning to read. In: MACGINITIE, Walter. (Comp.). *Foundastions of reading*. Readings for ED-B 342, University of Victoria, 1982, p.135. [A tradução é nossa.]

15 Cf CAZDEN, Courtney B. La lendária escrita em contextos escolares .In: FERREIRO, E., GÓMEZ FALACIO, M. (Comps.). *Nueves perspectivas sobre los procesos de lecturas y escritura*. Buenos Aires: Siglo XXI, 1986.

possibilitará um desempenho adequado. Assim, por exemplo, pode-se destinar um tempo inicial para conversar, criando um espaço de contextualização que prepare para a leitura, no qual sejam trabalhados os conhecimentos prévios imprescindíveis para a compreensão do texto. Uma atividade desse tipo deveria estar convenientemente definida para não avançar sobre o próprio conteúdo da prova.

Por outro lado, deve-se considerar a possibilidade de acrescentar às provas clássicas respondidas de forma individual e com um estrito controle de não-comunicação entre os avaliados alguma exercitação que exija uma atividade cooperativa por grupos. Seria interessante a comparação dos resultados de ambos os tipos de atividades.

CAZDEN assinala que geralmente é difícil as crianças realizarem em um exame individual o que habitualmente fazem de forma grupal, pois diferentemente do leitor adulto, que lê quase sempre individual e silenciosamente, na escola a leitura é uma atividade social. Baseando-se em pesquisas próprias e de outros autores - entre eles um estudo de W LABOV sobre o desempenho de crianças negras na leitura - que ressaltam as comunicações entre pares, ele sugere que estas sejam aproveitadas no ensino da leitura "de modo a que o poder das interações grupais possa ser usado diretamente como contexto de aprendizagem".

Parece-nos que as considerações dessa autora, que se referem principalmente ao ensino e ocasionalmente às avaliações em um marco curricular, poderiam ser aproveitadas com relação ao tipo de provas em questão.

Os Instrumentos

Através de seus materiais e itens, uma prova bem-elaborada deveria possibilitar o exercício da **predição**.

Em sua análise Psicolingüística da leitura, Frank SMITH (1978) apresenta o conceito da predição como base da compreensão. Destaca a importância da informação não visual que está armazenada no cérebro - e da predição. Esta reduz a incerteza, na medida em que permite eliminar antecipadamente certas alternativas dentro de um amplo leque de significados possíveis, possibilitando a compreensão".

De outro enfoque sociolingüístico, HARSTE & BURKE (1986) afirmam que o contexto do escrito e do impresso possibilita a formulação de esquemas antecipatórios. A identificação do contexto permite a detecção de uma gama possível de significados e a delimitação das opções semânticas, isto é, o exercício da predição. Ressaltam também a gênese social da predizibilidade e assinalam que, quando isto não é levado em conta no ensino da leitura e da escrita, altera-se o processo da compreensão".

16 CAZDEN, C., 1982, p.138.

17 CE SMITH, Frank. Understang reading New York: Holt, Rinehart and Winston, 1978.

18 Cf HARSTE, Jerome G, BURKE, Carolyn N. Predicibilidad: un universal en lecto-escritura. In: FERREIRO, E., GÓMEZ PALÁCIO, M. (Comes.). Nu- *Perspectivas...*, op. cit

A nosso ver, o fato de se *levarem em* consideração os conceitos de predizibilidade e de contextualização, e de aplicá-los ao campo da avaliação, implicaria levar em conta questões como:

- as características gráficas, de diagramação, a transcrição correta dos textos e outros aspectos materiais que deveriam confluir para uma perfeita legibilidade;
- as instruções gerais das provas e dos exercícios, de modo que elas resultem claras, concisas e não permitam qualquer dúvida sobre a tarefa a ser realizada;
- o modo de apresentação dos materiais, que deveria levar a uma rápida identificação do gênero. Assim, caso se trate de uma matéria jornalística, a inclusão de uma fotocópia dela contribuirá mais para a contextualização do que a mera transcrição. Em outro tipo de textos será conveniente citar dados do autor, obra ou outros aspectos que ajudem sua localização; tampouco seria bom "recortar" um texto em pequenos fragmentos ou realizar adaptações nas quais fossem suprimidas palavras ou se alterasse a sintaxe, pois nesses casos é possível que se *eliminem "marcas"* da escrita referentes ao significado e à coesão do texto;
- o significado dos materiais utilizados, em função do contexto cultural e do nível de informação;
- a extensão dos textos vinculados ao tipo de atividade requerida; como critério geral, seria preferível selecionar poucos textos que, por sua natureza e extensão, permitissem uma maior quantidade e gama de itens, em vez de muitos textos muito breves;
- algumas "ajudas para a leitura" ou recursos dos avaliadores para que o avaliado se detenha em um parágrafo ou *preste atenção* a algum aspecto em particular -como introduzir entre cada parágrafo um espaço maior que o habitual ou algum tipo de instrução ou subtítulos que não existam no original - podem interferir na predição, na medida em que não estão guiados pela natureza do texto ou pelas estratégias do leitor e sim pelo objetivo do avaliador;
- as perguntas formuladas poderão interferir na compreensão da pessoa avaliada se, em vez de considerarem os processos espontâneos do leitor em interação com o texto e tratarem de ressaltar aspectos significativos, forem formuladas a partir de estereótipos ou práticas de avaliação correntes que não foram revisadas à luz das perspectivas colocadas. Alice OMAGGIO (1993) analisa os tipos de provas utilizadas nos Estados Unidos para a avaliação em língua, *particularmente em* línguas estrangeiras; distingue os *achievement tests* utilizados com relação a um curso determinado e os *proficiency tests*, que são provas gerais de

nível ou de competências, utilizadas, por exemplo, nas universidades, para a aprovação em um idioma estrangeiro'.

Conforme sua descrição, as provas do primeiro tipo limitam-se a aspectos pontuais, a segmentos do currículo de língua, formulando perguntas de respostas estruturadas sobre questões gramaticais ou léxicas muito específicas, a partir de frases isoladas ou orações, tais como colocar a preposição que corresponde ao regime de um verbo dado, completar a tradução de uma oração etc. Os *proficiency tests*, ao contrário, incluem um nível textual, e as perguntas utilizadas são muito mais abrangentes. Embora as perguntas sejam pontuais, são requeridas competências mais globais no tocante a conteúdos, contextos, tipos de discurso etc.

Com a perspectiva da contextualização - descontextualização, OMAGGIO fornece um esquema interessante para analisar itens e para incentivar os professores à elaboração de exercícios adequados:

Seqüência de orações ou frases não relacionadas entre si

Itens convergentes: exige-se uma resposta correta; podem ser itens	De final aberto: muitas respostas possíveis; requer produção divergente,
centrados em aspectos pontuais ou um formato integrativo que inclua estes itens	itens de compreensão global Discurso natural, seqüencial

Fonte: OMAGGIO H., A., *Teaching language in context*, p. 420, ilustração 9.2 (a tradução é da autora).

Cada um dos eixos representa um *continuum*. Os itens analisados ou elaborados serão colocados em um ponto do *continuum* em função de sua maior ou menor proximidade aos extremos. Poucos itens poderão então ser considerados totalmente "contextualizados" ou "descontextualizados".

A autora postula que nas provas utilizadas na escola sejam incorporados mais exercícios do tipo dos *proficiency tests*. Aceita que sejam incluídos itens situados em diferentes pontos do

19 OMAGGIO HADLEY, Alice teaching language in context. Boston, Massachusetts: Heine and Heine, 1993. CE Capítulo 9, "Classroom testing".

eixo horizontal, na medida em que este representa uma gradação do campo lingüístico-comunicacional; em troca, no tocante ao eixo vertical, os itens deveriam situar-se apenas na metade inferior, pois a parte superior fica fora do âmbito da linguagem natural. Assim, os dois quadrantes inferiores constituiriam espaços de inserção das provas que seriam indicadas para a avaliação na escola, para as quais OMAGGIO fornece uma série bastante rica de exercícios para avaliar compreensão oral, compreensão na leitura e produção escrita.

A caracterização realizada pela autora e as recomendações que formula seriam aplicáveis, em nosso contexto latino-americano, a provas diagnósticas, de admissão na universidade etc.

A Compreensibilidade

Em provas para avaliar a compreensão, geralmente são utilizados textos pouco compreensíveis, são formuladas perguntas confusas ou ambíguas ou as respostas, mesmo as consideradas corretas, não estão formuladas claramente.

Os estudos de compreensibilidade referem-se aos fatores de tipo lingüístico e psicolingüístico -vocabulário, sintaxe, organização do texto etc. - que fazem um texto se tornar mais ou menos compreensível. Seria interessante levá-los em conta, em relação tanto aos materiais incluídos nos testes, como aos itens.

Uma parte deste campo de estudo refere-se à complexidade sintática, e descreve as estruturas que são mais compreensíveis em função do desenvolvimento psicolingüístico. No caso de se tratar de uma estrutura pouco compreensível na comunicação oral - na qual, por exemplo, seja difícil identificar o sujeito -, mais difícil ainda será seu processamento na leitura. No outro extremo, uma sintaxe simples demais - baseada, por exemplo, na falta de uso de conectivos e, portanto, no uso de repetições fictícias - tampouco ajuda a compreensão, na medida em que se distancia da linguagem corrente".

Mais do que as avaliações de outro tipo, as avaliações para diagnóstico possibilitam a validação dos instrumentos na etapa de sua elaboração, antes do teste-piloto. Os melhores juízes da compreensibilidade serão os usuários da prova e, se eles puderem expressar-se através de estratégias de validação em pequenos grupos, será possível melhorá-las.

As Perguntas

A questão do conteúdo da pergunta está vinculada à **validade**, isto é, em que medida o conteúdo das perguntas e da prova em seu conjunto refere-se a aspectos significativos e relevantes em relação à competência que se pretende avaliar.

Nas teorias sobre avaliação em língua predomina atualmente a da "competência unitária", isto é, as diversas habilidades da linguagem são vistas como manifestação de uma única

capacidade subjacente, ficando em segundo plano as teorias segundo as quais se devem considerar separadamente diferentes subabilidades (DAVIES, 1992)(21).

Independentemente da postura adotada a respeito, e mesmo nos modelos de avaliação por objetivos(22), parece-nos importante ressaltar que um conhecimento, seja ele limitado ou mais amplo, pode ser avaliado enfatizando-se a informação, o conhecimento de uma noção gramatical, textual etc., ou então a **competência lingüística** vinculada a esta noção. Nas provas correntes encontram-se com maior freqüência os itens do primeiro tipo.

No campo dos estudos sobre a leitura, após décadas de predomínio das teorias que apontavam três níveis "clássicos" de compreensão - a literal, a capacidade de inferência e a leitura crítica -, desde o final dos anos 70 e início dos 80 começaram a predominar as visões da leitura como processo global.

VAN DIJK & KINTSH (1983) propõem um modelo **estratégico** que abarca desde o processamento das unidades de palavras até o das unidades abrangentes de temas ou macroestruturas. Este modelo é chamado de "estratégico" porque, segundo ele, o ouvinte ou leitor exercita um entendimento linear rápido que é hipotético e poderá depois ser retificado durante o decorrer da audição ou leitura. Durante o processamento, os níveis interagem de uma maneira intrincada e complexa(23).

As estratégias de análise utilizadas dependem não apenas das características do texto mas também das do ouvinte/leitor, tais como seu conhecimento do mundo e outros. Assim, o leitor reconstruirá tanto o sentido do texto assinalado pelo escritor, quanto o sentido que é mais relevante para seus próprios objetivos e interesses. .

Em um primeiro passo, através de uma estratégia abrangente que inclui subestratégias decodificação de traços fonológicos/gráficos, identificação de fonemas/letras e de morfemas, dependentes da interpretação semântica e sintática subjacentes -, chega-se à construção de orações relacionadas entre si que constituem um **texto-base**, fundamentado em uma coerência de tipo **local** que ocorre na memória de curto prazo.

Além desta representação na memória de curto prazo, constitui-se também uma representação do fato ou da situação de que trata o texto, do seu tema, denominada **macroestrutura**.

"... entender um texto significa que a pessoa é capaz de construir um modelo mental para ele. E, inversamente, na produção do texto, o modelo é o ponto de partida de todo processamento: a pessoa sabe algo sobre um fato e este conhecimento está representado em seu modelo do fato; este modelo servirá de base, por exemplo, para contar uma história sobre o fato(24).

Cf. DAVIES, Allan. Language testing and Evaluation In: CRASE, W E.]Caplan. Introduction to applied linguistics. reading, Massachusetts, Addison-Wesley Publ. Comx,1992.

"Nesses modelos habitualmente são construídos itens referentes a objetivos específicos - isolados dos terminais -; através deles pretende-se avaliar a consecução destes últimos.

"Cf VAN DIJK, T. A., KINTSH, I. W Strategies of discourse comprehension. New York: Academic Press, 1983.

"VAN DIJK, T. A. Dela gramática del texto al análisis crítico del discurso. *Boletín de Estudios Lingüísticos Argentino*; n. b.p.2G, maio 1995.

Por último, os autores assinalam que muitos tipos de discurso, por exemplo a narração ou a argumentação, possuem uma estrutura convencional denominada **superestrutura**. As provas de leitura comumente utilizadas geralmente estão mais vinculadas a modelos anteriores a este que acabamos de mencionar, os quais estabeleciam hierarquicamente níveis vinculados a habilidades e subhabilidades, da decodificação à compreensão, que ocorreriam de forma sucessiva e progressiva no indivíduo. Conseqüentemente, um esquema habitual nos exercícios referentes a um texto é fazer perguntas sobre os detalhes dos conteúdos dos parágrafos, acrescentando-se no final alguma pergunta mais global. Mesmo quando as respostas estiverem corretas, em muitos casos poderíamos perguntar-nos se realmente o texto foi compreendido.

Uma das práticas mais comuns nas provas é constituída pelas perguntas sobre **inferências**; isso se deve ao papel-chave que a maioria dos autores atribui à inferência no processo da leitura. No entanto, freqüentemente as perguntas desse tipo referem-se a aspectos quase totalmente explícitos; nas respostas esperadas constroem-se enunciados que são facilmente dedutíveis do enunciado dado no texto, e, portanto, não se requer do leitor nenhuma atividade heurística ou de formulação de hipóteses.

Esse tipo de pergunta poderia ter sua origem na exigência de se responder a uma taxonomia escolhida para a avaliação, pois de fato muitas das inferências possíveis não acrescentam nada relevante à compreensão do texto.

Enquanto, por um lado, a inferência intervém permanentemente na compreensão e não só nos pontos escolhidos para a elaboração de perguntas que procuram avaliar esta capacidade, ao mesmo tempo geralmente esperam-se respostas óbvias, que não são as únicas possíveis, pois o exercício da inferência pode basear-se em diferentes indícios presentes em um texto, levando a conclusões diversas.

Quanto a outros tipos de perguntas de compreensão, é preciso ver quais os conteúdos considerados mais importantes. Embora possa haver consenso sobre alguns aspectos considerados chaves, também existe um campo muito subjetivo, pois diferentes bons leitores poderiam ser impressionados preponderantemente por diferentes aspectos do texto. A validação por pareceres de especialistas ou por diversos leitores pode ajudar nesse sentido.

Ao mesmo tempo seria útil introduzir em uma prova de leitura um exercício como o seguinte: "se você quisesse saber se uma pessoa compreendeu esse texto, que perguntas faria para comprová-lo?". A informação obtida seria interessante em termos de diagnóstico e contribuiria para diminuir a unidirecionalidade própria desse tipo de instrumento.

Uma forma de aumentar a "objetividade" das provas seria a construção de uma "intersubjetividade" que reduzisse o grau de subjetividade das perguntas formuladas e das respostas esperadas. A ampliação da quantidade e dos tipos de pessoas que participam da elaboração das provas, a incorporação de metodologias que incluam a participação de diferentes tipos de juízes - os professores dos cursos onde se aplicam as provas podem se transformar em informantes-chave -, e a atribuição de um papel mais ativo aos avaliados, tanto nas etapas de construção dos instrumentos como no momento da avaliação, poderiam contribuir para isso.

Com relação a outra questão importante, a maneira como os itens são construídos, habitualmente se faz uso da tecnologia "avaliativa", cujo desenvolvimento aumentou consideravelmente. Existe uma grande quantidade de prescrições sobre a redação correta de um item, sobre os aspectos formais que influem na escolha de uma resposta, assim como sobre as formas de validação-análise da correspondência item-objetivo e, depois da aplicação da prova-piloto, análise da consistência interna, do grau de discriminação dos itens, do grau de facilidade etc.

Tudo isso é necessário, porém não é suficiente. De fato, este conhecimento acumulado nem sempre tem sido utilizado de forma apropriada, e mesmo, às vezes, contribuiu para a consolidação de usos rotineiros e poucos criativos, na medida em que se prestou demasiada atenção a essas questões formais, deixando de lado outros aspectos vinculados à natureza da pergunta.

Exemplos de usos rotineiros ou de pouco cuidado quanto ao conteúdo do item podem ser dados pela realização de perguntas referentes a aspectos pouco relevantes; ou por perguntas que podem ser respondidas independentemente da leitura do texto ou - mesmo quando a resposta correta se referir ao texto - por itens nos quais as alternativas incorretas, em vez de se construírem com elementos tomados do texto, propõem conteúdos alheios a ele. É importante que todas as respostas tenham um certo grau de verossimilhança - isto é, que resultem plausíveis - para incentivar o leitor a exercitar seu discernimento e para não o afastar do processo de leitura que estava realizando.

Outra dificuldade em relação aos textos que apresentam um certo grau de complexidade ou trechos ambíguos é que se tende - através das perguntas formuladas ou das respostas consideradas corretas - à eliminação da ambigüidade ou da complexidade.

Embora seja certo que as provas objetivas envolvem uma quantidade de restrições, é possível ampliar e enriquecer este contexto, como foi demonstrado por inúmeras experiências.

Queremos ressaltar que nossa experiência na elaboração de provas ou na revisão de provas elaboradas por outras pessoas nos leva a reconhecer a importância da observação dos próprios processos pelos quais passa o avaliador durante a leitura dos materiais selecionados, como uma das fontes às quais ele pode apelar para construir os itens, em vez de recorrer a modelos preestabelecidos.

Um último comentário refere-se às perguntas abertas ou de desenvolvimento. Muitas vezes se requer um exercício escrito para verificar a compreensão em leitura ou complementa-se um item de resposta estruturada com um espaço em branco, no qual se pede, por exemplo, uma fundamentação da resposta. Nesse caso, a compreensão é avaliada por meio da produção e o material obviamente é mais rico.

Os parâmetros construídos para a correção desses exercícios deveriam ser suficientemente amplos para considerar os modos de abordar a leitura das diferentes pessoas, assim como suas estratégias de escrita, que nem sempre coincidem com as mais valorizadas em nossa tradição escolar e acadêmica, mais vinculadas ao pensamento lógico e aos processos de tipo racional.

Para concluir esse percurso conceitual sobre o que denominamos "aspectos críticos" das provas, consideramos que se deve trabalhar para gerar marcos amplos que permitam o desenvolvimento de uma gama rica e variada de estratégias de avaliação.

A seguir, passamos a expor a experiência de avaliação da qualidade desenvolvida na Argentina.

A Avaliação da Qualidade da Educação na Argentina O Sistema Nacional de Avaliação da Qualidade(25)

A Lei Federal de Educação estabelece a avaliação permanente do sistema educacional(26). Assinala, em seu artigo 49:

“A avaliação da qualidade no sistema educacional verificará a adequação do conteúdos curriculares dos diferentes ciclos, níveis e regime especiais às necessidades sociais e aos requerimento educativos da comunidade, assim como o nível de aprendizagem dor alunos/as e a qualidade da formação docente”.

Embora tivessem sido realizadas ações de avaliação em nível de sistema antes da sanção dessa Lei, a partir de então foi implantado um Sistema Nacional de Avaliação e foi realizada a primeira avaliação nacional. Para viabilizá-la, o Ministério de Cultura e Educação da Nação (MCyE) levou a cabo diversas ações tendentes à obtenção de acordos básicos com as 24 jurisdições nas quais se ministra educação, correspondentes à Prefeitura da Cidade de Buenos Aires e às 23 províncias. Os acordos referiam-se a conteúdos a avaliar, instrumentos de avaliação, processamento e comunicação dos resultados e programação da avaliação.

Em novembro de 1993 realizou-se essa primeira avaliação, que abrangeu a 7^a. série do nível primário e a 5^a. série do nível médio: colegial, comercial e escola técnica(27). Escolheu-se uma amostra de alunos aleatória e estratificada - por jurisdição, por zona rural e urbana e por modalidade.

Em todo o país foram entregues simultaneamente questionários a 9.831 famílias, 1.056 diretores e 1.957 professores de Língua e Matemática; 38.876 alunos foram avaliados através de provas de Língua e Matemática, bem como também tiveram de responder a um questionário de opinião. Os questionários fornecidos solicitavam informações complementares, que contribuíram para explicar os resultados das provas.

Nos anos seguintes -1994 e 1995 -, a avaliação abrangeu os mesmos níveis e modalidades e versou sobre as mesmas disciplinas, porém aumentou consideravelmente a quantidade de sujeitos avaliados. Em 1995, também foram acrescentados outros cursos e disciplinas. Em 1994, em duas províncias - San Juan e Tierra del Fuego -, a avaliação teve caráter de censo, ou seja, foram avaliados todos os alunos da 7^a. série da escola primária e da 5^a. série da escola média. Em outras seis províncias foi feita uma avaliação experimental com alunos

25 Para a descrição dos processos nacionais de avaliação e a referência às provas de língua e aos seus resultados, baseamo-nos em fontes do Ministério da Cultura e Educação. Não contamos com informação proveniente das escolas sobre a experiência de avaliação.

26 lei Federal de Educação (Lei n° 24.195). Título IX, "Dela calidad de la educación y su evaluación", artigos 48 a 50.

27 Atualmente está sendo implementada na Argentina a transformação da estrutura de ensino primário (7 anos, obrigatório) e médio (5 anos) em educação geral básica (9 anos, obrigatória) e polimodal (3 anos).

adultos de cursos de nível médio, os quais tiveram de se submeter às mesmas provas que os alunos dos cursos de nível médio comuns(28).

Em 1995, ampliou-se o universo da avaliação, acrescentando-se alunos da 3ª. série do nível primário e da 2ª. série do nível médio. Os alunos da 7ª. série, por sua vez, foram avaliados em ciências sociais e em ciências naturais, além de efetuarem as provas de Língua e Matemática.

As Provas de Língua

O que se procurou avaliar

A taxonomia utilizada, pelo menos no tocante à descrição dos conteúdos a avaliar, corresponde a uma visão atualizada no campo dos estudos sobre a língua. Consideraram-se aspectos vinculados a: a) compreensão leitora; b) noções e regras gramaticais; e c) produção escrita (ver no Anexo "Conteúdos da prova" correspondente à 7ª. série de nível primário, 1993). A inclusão de aspectos como estrutura e superestrutura do texto e adequação da escrita à situação comunicativa etc. significa uma superação com relação a certas taxonomias aplicadas tradicionalmente e que é possível reconhecer em algumas provas que atualmente circulam em nossa região.

No entanto, deve-se analisar se o ordenamento classificatório dos conteúdos de compreensão leitora corresponde aos processos da leitura tal como eles ocorrem na mente do leitor, e se o próprio enunciado de uma tabela de especificações não leva a uma segmentação artificial. Também é preciso verificar se, efetivamente, em todos os casos avalia-se o que se pretende avaliar e se predominam os itens vinculados à competência, ou os vinculados a conhecimentos sobre questões gramaticais, sobre a estrutura do texto etc.

É preciso assinalar também que, em termos de pontuação, cada item das questões de resposta estruturada vale a mesma coisa, quer busque avaliar um conhecimento pontual, quer procure verificar competências complexas e mais abrangentes.

A utilização dos resultados

Os resultados foram divulgados e, em particular, comunicados à comunidade educativa, e utilizados com fins diagnósticos. A identificação de determinados fatores - estratégias didático-pedagógicas da escola e outros fatores institucionais, assim como os perfis socioeconômicos, a partir das informações obtidas com os questionários, permitiu associá-los ao desempenho nas provas e contribuir para explicar suas causas.

A partir dos resultados da primeira avaliação e das seguintes foram elaborados materiais destinados às escolas, nos quais se analisam a produção escrita e os itens que apresentaram mais dificuldades, propondo-se estratégias aos professores para trabalhar em torno desses conteúdos na sala de aula.

28 A educação de adultos inclui-se nos "regimes especiais" aos quais deve atingir a avaliação da qualidade segundo o mencionado artigo da Lei.

De outro lado, os resultados serviram de *feedback* para a equipe de especialistas do MCyE: o comportamento dos itens e as observações registradas pelos pesquisadores em uma folha planejada para esse fim no momento da aplicação das provas produziram informações importantes, que foram levadas em conta para a análise destas e para a elaboração dos instrumentos que se seguiram. A esse respeito, nos materiais mencionados registram-se alguns comentários dos especialistas que fazem parte do *feedback*.

O formato das provas. O processo de construção.

As provas utilizadas nas operações de 1993 e 1994 foram tornadas públicas após a sua aplicação e a comunicação dos resultados. No seguinte quadro, pode ser visualizado o formato das provas de Língua.

Nível/Grau ou Ano	Formato	
	1993	1994
Primário: 7ª série	PARTE A Texto 1 (itens 1 a 14) Texto 2 (itens 15 a 25) PARTE B (itens 28 a 40) EXPRESSÃO ESCRITA Exercício	PARTE A Texto 1 (itens 1 a 13) Texto 2 (itens 14 a 24) PARTE B (itens 25 a 40) EXPRESSÃO ESCRITA Exercício
Médio: 5ª série	PARTE A Texto 1 (itens 1 a 17) Texto 2 (itens 8 a 15) Texto 3 (itens 16 a 18) PARTE B (itens 19 a 31) EXPRESSÃO ESCRITA Exercício	PARTE A Texto 1 (itens 1 a 5) Texto 2 (itens 6 a 16) PARTE B (itens 28 a 40) EXPRESSÃO ESCRITA Exercício

Na etapa de elaboração das provas, foram selecionados inicialmente itens de três fontes:

a) material elaborado por equipes de especialistas; b) instrumentos utilizados e validados em experiências nacionais e estrangeiras²⁹; e c) textos nacionais e estrangeiras⁽²⁹⁾. Para as sucessivas

²⁹ Chile, Unguai, México, Costa Rica e Estados Unidos.

validações que levaram à elaboração dos instrumentos finais, contou-se com a participação de peritos das jurisdições nacionais, na qualidade de juízes.

O caminho seguido para chegar à versão definitiva das provas foi árduo e trabalhoso, mas gostaríamos de efetuar uma análise mais abrangente delas, abordando-as com um outro olhar. Na análise que realizaremos, prescindiremos de aspectos estatísticos - que ultrapassem a simples observação da porcentagem de respostas das alternativas de cada item e dos resultados gerais - e das prescrições formais para a redação dos itens (extensão etc.), para o balanço dos itens dentro de uma prova; preferimos deter-nos nos aspectos vinculados ao pensamento e à linguagem, a partir de uma perspectiva pedagógica e das ciências da linguagem, enquanto realizamos algumas apreciações com base em nossa própria competência lingüística e na observação de nossos processos de leitura.

Análise de uma amostra do material de prova

As provas utilizadas em 1993 e em 1994 foram divulgadas após a sua aplicação. Nossa análise centra-se na prova para a última série, a 7ª do nível primário, aplicada em 1993 a um grupo de alunos da Capital Federal e de 23 províncias, no qual estavam representadas escolas urbanas -públicas e particulares - e rurais públicas. A prova transcrita em anexo é formada pela parte A, que inclui dois textos com seus respectivos itens, e pela parte B, com exercícios sobre aspectos morfosintáticos, ortografia, questões de coerência etc. referentes a um texto prévio; o último item requer a produção de um texto escrito (com relação aos resultados, ver em anexo, no quadro 1, a distribuição do grau de dificuldade dos itens por zona urbana e rural).

Nossa análise refere-se aos itens de resposta estruturada, relativos à leitura e/ou a questões gramaticais, de ortografia etc. Não se aborda o exercício de escrita incluído na prova.

A metodologia de trabalho foi a seguinte: realizou-se uma primeira leitura, na qual se tentou ocupar o lugar do leitor para reconstruir as possíveis dificuldades; a seguir, os itens foram analisados detalhadamente; depois, especialistas do MCyE verificaram os resultados da prova e analisaram os itens que apresentaram maior dificuldade. Em muitos casos descobriu-se que as perguntas nas quais detectamos dificuldades tinham sido acertadas por menos de 50% dos alunos; em compensação, outras perguntas em relação às quais tínhamos levantado objeções foram respondidas corretamente pela maioria dos estudantes.

Isto talvez permita que nos aventuremos a afirmar (sem nenhum suporte estatístico) que certas falhas dos itens induzem a uma resposta incorreta, enquanto outras não interferem na resposta e não impedem que o item seja respondido satisfatoriamente nas validações de tipo estatístico; no entanto, elas estariam atentando contra a validade do item, pois a resposta dada não serve para demonstrar efetivamente a competência que se pretende avaliar.

Por último, percebemos que outras perguntas que, a nosso ver, estão bem construídas, tiveram uma frequência de respostas corretas abaixo da média. Nestes casos, consideramos que o item é válido para diagnosticar o domínio ou não daquela competência específica; caberia verificar apenas se seu nível de dificuldade é adequado para o grau ou ano que os alunos estão cursando.

A partir dos itens desta prova (PRIM. 93) que comentamos, oferecemos alguns exemplos de outras provas - média ou secundária 93, primária 94 -, identificadas respectivamente como

MM. 93, PRIM. 94, as quais, por razões de espaço, não transcrevemos aqui. Estamos conscientes de que o fato de citar itens fora do contexto da prova à qual pertencem pode prejudicar uma apreciação mais completa deles; no entanto, esperamos que se entenda o que queremos transmitir em cada caso.

Começamos então pela parte A, com a análise dos itens vinculados ao primeiro texto que denominaremos Faroleiro.

O item 1 refere-se à informação explícita no texto (ver Anexo); é o único item da prova que, de acordo com os resultados, é qualificado de "muito fácil", tanto nas áreas urbanas como nas rurais. O item 2 também se refere à informação explícita, mas a um aspecto muito pontual: o significado da expressão "e tantos" (ver Anexo); foi bem respondido por 75% dos alunos. Com relação ao peso atribuído a cada item na avaliação, pode-se notar que esses itens têm o mesmo valor que outros que se referem a aprendizagens muito mais complexas.

Em nossa opinião, o item 3 está bem construído (ver Anexo); entretanto, foi respondido acertadamente apenas por 42% dos alunos.

O item 4 visa ao reconhecimento da superestrutura do texto em questão, em um aspecto que corresponde à introdução. Ao examinar as alternativas (ver Anexo), verificamos que a opção correta, a D, é de natureza diferente das outras três, já que a **apresentação do personagem, do lugar, do tempo** e do problema são categorias aplicáveis a qualquer texto narrativo; em compensação, as outras respostas referem-se concretamente a elementos deste texto - "farol", "homem" etc.

A opção pela alternativa correta implicaria então conhecer conceitualmente o que é uma introdução, mas não necessariamente discernir qual é a introdução desse texto. Para garantir isto, a resposta correta deveria ter sido construída, assim como as demais, referindo-se concretamente ao texto. Outra forma possível de construir o item teria sido explicar, na raiz, que o texto possui uma introdução e depois perguntar qual é ela. Em suma, se se quiser **avaliar a competência em leitura** referente a um aspecto da superestrutura que corresponde ao gênero narrativo, o item não está muito bem elaborado.

Com relação aos resultados, 43% dos alunos optaram pela resposta correta; e 33% pela alternativa A, que é parcialmente correta e, de qualquer forma, permite demonstrar certo grau de competência para identificar a introdução.

O item 5 visa ao reconhecimento do tema (macroestrutura), perguntando sobre o título que melhor se ajusta ao conteúdo do texto. A resposta correta, A, *Uma noite na vida do faroleiro*, foi escolhida por 53%; no entanto, temos dúvidas de que "*uma noite na vida de...*" realmente faça parte do título adequado.

Segundo nossa percepção na leitura, embora se trate de uma narração de um episódio ou de um acontecimento, o personagem adquire tanto peso no texto que outro título possível **bem poderia ser** "O faroleiro" ou "O velho faroleiro". O segundo e o terceiro parágrafos referem-se a questões que vão além daquela noite e servem para introduzir o personagem. Por outro lado, no início se faz referência à tormenta anunciada para essa noite, mas o relato se conclui quando o farol é aceso - certamente no primeiro momento de escuridão - e com a imagem do faroleiro saboreando seu cachimbo. A noite, com a tormenta incluída,

fica fora do relato. Talvez tudo isso tenha influenciado na escolha da alternativa D, *O velho faroleiro de Mar del Plata*, por 26,38% dos alunos, assim como também pode ter influenciado nessa escolha a menção de Mar del Piara, que não figura no texto mas responde a uma necessidade de contextualização do leitor: se os alunos alguma vez viram um farol ou tiveram alguma referência concreta a respeito de um deles, é muito provável que este tenha sido o dessa cidade balneária, a mais importante do país.

Em outros itens nos quais se pergunta sobre o título, por exemplo em SEC. 93, item 9, não encontramos objeção à resposta dada como correta; esta foi selecionada por 70% dos participantes. Esse item não será transcrito porque sua compreensão requereria também a transcrição do texto correspondente.

No item 7 pergunta-se sobre *a expressão que substitui a palavra farol no texto*. Esse item destina-se a avaliar o reconhecimento da coesão, neste caso através de paráfrase. A resposta correta, C, *ciclope imóvel*, foi escolhida por 38% dos alunos. A dificuldade parece estar no texto, na palavra *ciclope*. Provavelmente o **significado da palavra seja majoritariamente desconhecido** pelos avaliados, perdendo-se assim o valor da metáfora, a comparação do farol ao gigante mitológico que possui apenas um olho. Por outro lado, em nossas latitudes a metáfora parece alheia ao âmbito marinho, ou ao mundo imaginário do faroleiro e até dissonante em um texto onde o vocabulário utilizado é familiar, com exceção de algum termo especializado, como *espelhos parabólicos* que, mesmo sem se conhecer seu significado, poderia ser entendido como algum aparelho ou dispositivo próprio do farol.

No item 9 pergunta-se sobre *as razões que NÃO explicam por que o velho marinheiro trabalha como faroleiro*. Tenta-se avaliar a capacidade de estabelecer relações de causalidade. A resposta majoritária, com 39,88%, foi a alternativa C, *porque ele gosta de ficar perto do mar*. Outra porcentagem de alunos optou pelas alternativas A, *porque sente saudades de sua vida de marinheiro* (18,35%) e D, *porque gosta de ajudar aos marinheiros* (12,32%). Nossa interpretação é que esses três tipos de alunos suprimiram o NÃO e responderam **sobre as razões que explicam** por que ele trabalha como faroleiro. Por último, 27,62% dos participantes, que escolheram a alternativa correta B, *porque ele gosta das alturas*, demonstraram identificar relações de causa e efeito, e, além disso, foram capazes de estabelecer uma relação negativa, o que pressupõe uma operação lógica de maior nível de abstração.

Em suma, a pergunta apresenta uma complexidade maior que a simples relação causa-efeito; todos os alunos que responderam demonstraram, em nossa opinião, que podiam identificar esta relação na leitura do texto, mas quase três quartas partes dos alunos erraram na resposta ao serem interrogados sobre qual não era a causa.

Além deste item, as perguntas que contêm uma negação na raiz costumam suscitar dificuldades na compreensão da instrução, se implicam converter à forma negativa o que no texto lido é apresentado em forma afirmativa. Observando os itens desta e de outras provas com esse tipo de raiz, encontramos o seguinte.

Dois itens anteriores, o 6 e o 8, também incluem uma negação na raiz.

O item 6 possui o mesmo esquema que o 9 quanto à negação; nele se pergunta *qual dos seguintes elementos nomeados no texto não faz parte do farol*. Através dele, tenta-se avaliar reconhecimento de informação explícita. Embora a pergunta seja aparentemente simples,

no texto não se explicitam em nenhum momento como tais "os elementos do farol". O raio de luz, inclusive, não pode ser equiparado a uma lâmpada de óleo enquanto "elemento". A escolha da resposta correta, B, velas acesas, requer uma interpretação correta da oração **o raio de luz equivalente a milhões de velas acesas surgiu do nada e dirigiu-se para o mar** - 42,66% dos alunos optaram pela resposta correta; 26,58% escolheram a opção D, espelhos parabólicos. Talvez aqui a dificuldade tenha sido provocada pelo tipo de texto e pela falta de informação prévia sobre o tema-mesmo aqueles que sabem o que é um farol provavelmente não estão informados sobre como ele é por dentro -; a isto acrescenta-se a forma negativa presente no enunciado.

O **item 8** (ver Anexo) é bastante mais simples que o anterior com relação ao seu conteúdo. As respostas certas chegaram a 47,80%.

Deve-se assinalar que, na prova da escola primária de 1994, com referência às leituras apresentadas, não foram incluídos itens com negação na raiz; talvez isto tenha ocorrido porque podem ter sido detectadas nas provas do ano anterior as dificuldades que mencionamos.

Voltando à prova que estamos analisando, PRIM. 93, no item 28 pergunta-se qual é o parágrafo que não tem relação com os outros (ver Anexo). Este item não se refere a um texto, mas a relação se estabelece entre as alternativas, sendo que cada uma delas corresponde a uma oração ou a um parágrafo; deve-se buscar a coerência no interior do item e não é preciso realizar uma operação de transformação para a forma negativa. Portanto, neste item o NÃO da raiz não nos suscita objeção, pois não acreditamos que contenha uma dificuldade que influa na compreensão da pergunta. A resposta correta foi escolhida por apenas 40,78%: neste caso, parece-nos adequado vincular o resultado ao pouco domínio dos alunos sobre a detecção de coerência ou não-coerência.

Um item formalmente semelhante a este é o 25 da prova PRIM. 94, que tampouco apresenta dificuldades em sua construção.

Continuando com a observação dos itens que apresentam a forma negativa do verbo em seu enunciado, encontramos em FED. 93 os itens 5 e 6; por razões de espaço não vamos transcrever o texto ao qual se referem. Transcreveremos, porém, ambos os itens, começando pelo 5:

5 -Qual dos seguintes conceitos não está presente no texto?

A- A história não volta atrás

B- A história serve para compreender de onde viemos e para onde vamos

C- A história é a memória daquilo que nossos antepassados fizeram

D- A história não é uma ciência

Para 66% dos que escolheram a resposta correta, não houve obstáculos na compreensão da pergunta. Isto mostraria que na 5ª série do nível médio os alunos já estão em condições de realizar a transformação daquilo que o texto diz de forma explícita e reconhecer informação textual implícita; isso é o que se tenta verificar através deste item. No entanto, 21% dos que escolheram a opção A parecem não ter podido processar a dupla negação - da raiz e da alternativa, cujos verbos respectivos estão na forma negativa

Deve-se esclarecer que neste caso o tipo de pergunta resulta mais funcional que no item 9

da prova do primário, pois se se perguntasse sobre os conceitos presentes no texto, a resposta correta se aproximaria mais do explícito, e aqui tenta-se avaliar aspectos implícitos.

A seguir, transcreve-se o item 6 da X=. 93:

6 -Qual das seguintes afirmações não contradiz a colocação do texto .obre a cultura humana?

- A- A cultura romana incorpora a cultura grega e a difunde pelo mundo
- B- A independência americana significou uma ruptura total com a cultura espanhola
- C- A vontade de Napoleão foi .suficiente para transformar a história da Europa moderna
- D- A Guerra do Golfo entende-se exclusivamente a partir do fundamentalismo muçulmano

A resposta A foi escolhida por quase 40% dos alunos. Os outros alunos que responderam dividiram-se de forma bastante equilibrada entre as demais alternativas, sendo que todas elas têm em comum o fato de contradizerem o que o texto coloca. Uma interpretação possível dos resultados é que não se compreendeu bem o texto e/ou não se soube aplicar seus conceitos às propostas das alternativas. Outra interpretação vincula-se à dificuldade da frase não contradiz onde o verbo, pelo seu significado, implica uma oposição ("dizer o contrário"), acrescentando-se a isso a forma negativa.

O item é relevante para aquilo que se pretende avaliar- estabelecer relações de semelhança-oposição -mas talvez fosse mais compreensível se se tivesse perguntado pelas afirmações que contradizem a proposta do texto, o que naturalmente teria exigido a construção de alternativas diferentes.

Retomando a análise dos itens de PRIM. 93, observa-se que, às vezes, na formulação de um item, a pergunta resulta mais difícil do que a resposta adequada. Este é o caso dos itens 10 e 11, através dos quais se tenta avaliar a identificação ou o reconhecimento de co-referência, outro aspecto relacionado à coesão do texto. Neste caso, trata-se de pronomes possessivo e pessoal, respectivamente. Quanto ao grau de dificuldade, ambos os itens resultaram difíceis na zona urbana e muito difíceis na zona rural.

O item 10 refere-se a um parágrafo do texto "caminhou os metro que reparavam sua cara do farol" e pergunta a que se refere a palavra sublinhada. Apenas 22,63 % dos avaliados optaram pela alternativa correta C a homem. Acima dessa cifra as outras alternativas, A ao farol e B a metro atraíram 28,66% e 29,95% dos alunos, respectivamente. Nossa hipótese é a de que a pergunta não foi compreendida, pois consideramos muito improvável que crianças da 7ª. série não tenham entendido que a casa era do faroleiro, embora sua substituição por homem na alternativa correta acrescente alguma dificuldade.

Situação semelhante ocorre no item 11, no qual se transcreve um parágrafo do qual apresentamos apenas um fragmento `... seu novo trabalho lhe permitia cheirar a água salgada e sentir o vento e a chuva no rosto...'; pergunta-se novamente a que se refere a palavra sublinhada. A opção majoritária -45,12% - foi pela alternativa B, vento e chuva, enquanto só 27,37% escolheram a resposta correta D, a faroleiro.

Se se perguntasse simplesmente de quem era a cara e a quem o seu novo *trabalho permitia* cheirar a água e a chuva, havia uma opção majoritária pelas respostas corretas. Talvez as respostas fossem óbvias, porém refletiriam mais fielmente que as anteriores a compreensão real do texto. De qualquer forma, poder-se-iam buscar outras situações de pronomes nas quais a dificuldade surgisse do texto e não das perguntas. Em suma,, em nossa opinião esses itens não avaliam acertadamente este aspecto da compreensão em leitura, mas é provável que avaliem a conceitualização sobre pronome em termos de "fazer referência a".

No item 13 pergunta-se qual das palavra da seguinte oração é verbo. As alternativas são A, cima; B, se; C, pôr; D, ordem. A alternativa C, apresentada como correta, não o é, pois embora do ponto de vista das classes de palavras se seja um pronome e pôr um verbo, neste caso trata-se de um verbo quase reflexivo onde o se carece de valor pronominal. Em uma análise sintática, se pôs seria considerado núcleo do predicado verbal e essa deveria ser a resposta correta.

Considerando então que não se trata do verbo pôr mas do verbo pôr-se, mencionado no dicionário, resulta inadequado separar se depôs. Observando os resultados, verificamos que 55% optaram pela alternativa considerada correta. Quanto aos alunos que optaram pela alternativa B, que constituem 25%, eles talvez tenham sido influenciados de alguma maneira por essa separação artificial - e equivocadamente escolheram o outro segmento do verbo se pôr.

O item 14, que se refere a outra questão morfossintática, pergunta sobre o sujeito de uma oração (ver Anexo), e foi respondido acertadamente por cerca de 60%.

Depois desse item se introduz o segundo texto que, de forma abreviada, denominaremos Golfinhos bicudos. Pelas suas características, ele corresponde ao gênero matéria jornalística. Desconhecemos se o texto foi elaborado especificamente para a prova ou se se trata efetivamente de uma matéria de jornal - nesse caso seria conveniente fotocopiar o original pois, como se explicou anteriormente, o formato e as características gráficas ajudam a identificação do gênero e oferecem uma melhor contextualização para a leitura. O texto em si apresenta algumas dificuldades para a compreensão, o que nos leva a conjecturar que poderia tratar-se de uma adaptação de um texto preexistente no qual foram efetuadas mudanças ou supressões.

A primeira pergunta vinculada a esse texto, o **item 15, não suscita** nenhuma objeção. Através dele tem-se a intenção de avaliar relações de causalidade; foi respondido corretamente por 53% dos alunos. Isto permitiria reforçar a hipótese já exposta de que a baixa porcentagem de rendimento no item 9 não se deve principalmente ao pouco domínio da detecção de relações causa-efeito, mas a problemas na formulação do item.

Os itens 17 e 18 avaliam a mesma coisa; o 17 foi difícil de responder, enquanto o 18 foi considerado difícil na zona urbana e muito difícil na zona rural.

Em nossa observação dos itens referentes ao texto encontramos algumas dificuldades. O **item 17 afirma** que os cientistas pensaram que estavam diante de uma nova espécie porque... A resposta considerada correta, B, os *exemplares apresentavam* características similares, não é correta se considerarmos o conteúdo do texto e em particular o 42, 5º e 62 parágrafos. Esta alternativa explica as razões pelas quais os exemplares pertenciam à **mesma** espécie, mas não que esta

fosse **uma nova espécie**. Para a resposta ser correta faltam os traços característicos desses exemplares que marcam o contraste ou a diferença com as espécies já conhecidas.

Neste item, 30,77% escolheram a alternativa C, 26,62% optaram pela A e apenas 26,27% pela alternativa supostamente correta. Dado que as duas alternativas mais escolhidas tampouco têm elementos que possam ser relacionados às verdadeiras razões dos cientistas, é possível que os alunos não tenham sabido reconhecer no texto quais foram essas razões, mas o fato de que a resposta considerada correta não o é, ou não o é totalmente, invalida o item.

A raiz do **item 18** é a seguinte: *provavelmente esta espécie demorou para ser descoberta porque os exemplares...* A resposta correta C, *aparecem muito pouco*, foi escolhida apenas por 31,15%, enquanto a alternativa D, *tem nadadeira dorsais e peitorais pequenas*, atraiu 40,42%. Podemos interpretar que aqui houve uma falha de compreensão de um texto que é bastante mais complexo que o primeiro, e que os alunos optaram por escolher uma frase literal tomada do texto em vez da resposta correta, na qual com uma frase muito curta com o verbo *se mostra* - cujo significado talvez não esteja muito claro para eles - realiza-se uma substituição das causas explícitas *"deve-se provavelmente à rua notável discricção: vivem em alto-mar e costumam realizar poucas piruetas aéreas, diferentemente de outros golfinho"*.

O item 19, no qual se pergunta a que se refere a expressão sublinhada em *o pesquisador peruano Juan Carlos Reger informou que examinaram dei exemplares da nova espécie*, apresenta dificuldades similares às dos itens 10 e 11, nos quais se avalia a coesão, embora neste caso não se trate de pronomes.

No **item 19**, 47% dos alunos escolheram a resposta correta A, *Meaoplodon peruvianus*. Seguindo o texto, pareceria mais adequado perguntar qual é a nova espécie.

No item 23, pergunta-se por que se fala no título do texto de *golfinhos bicudos*; tenta-se avaliar o reconhecimento de informação explícita (ver Anexo). Só 32% responderam acertadamente, enquanto 34,11 % escolheram a alternativa C. Do ponto de vista da avaliação, e dado que a resposta correta é D, *por nenhuma das razões anteriores*, fica difícil interpretar o item porque não se sabe o que se quis perguntar. Se se queria verificar por que o título é *Golfinhos bicudos*, isto estaria relacionado à identificação do tema; no contexto da prova, em termos de previsibilidade, espera-se algo semelhante com relação ao item 5, que não está de acordo com o conteúdo destas alternativas. Entretanto, se o que se queria perguntar era por que os golfinhos são chamados de bicudos, informação explícita no último parágrafo do texto, deveria ser suprimida, na pergunta, a referência ao título.

Com relação aos resultados, evidencia-se que aqueles que responderam acertadamente não caíram nos erros de interpretação que pressupõem as alternativas não-corretas, mas isso não garante que tenham compreendido bem a pergunta.

Em nossa opinião, a resposta "nenhum dos anteriores" ou semelhantes, que às vezes são utilizadas nas provas como alternativa diante da dificuldade de construir outras alternativas adequadas, não é válida como resposta correta nas provas de língua. Por outro lado, itens desse tipo podem distanciar o leitor do processo de leitura que está realizando, pois não contêm informação relevante sobre o texto.

Outro exemplo, embora menos difícil, é o item 13 do MED. 93, transcrito a seguir:

13. Qual é o sujeito da seguinte oração?

"Há dezesseis anos instalou-se no cume de uma colina"

A- Há dezesseis anos

B- No cume de uma colina

C- No cume

D- Nenhum dos anteriores

Trata-se de uma pergunta simples, bastante mais fácil que a do item 23 PRIM. 93, que foi respondida corretamente por 78% dos alunos. Eles demonstraram que não confundiam o sujeito da oração com o conteúdo de nenhuma das alternativas A, B e C, porém não demonstraram de forma confiável que sabiam qual era esse sujeito tácito que o texto permite identificar.

O item 24 de PRIM. 93 pergunta claramente sobre o tema do texto, e tanto as perguntas corretas como as não-corretas estão formuladas adequadamente. Foi bem respondido por 71,60% e foi catalogado como fácil.

Em compensação, o item 25 apresenta algumas dificuldades e merece especial atenção, pois através dele tenta-se medir uma capacidade global vinculada à compreensão: pergunta-se qual é o melhor resumo do texto (ver o item completo no Anexo). A alternativa correta que transcrevemos a seguir foi escolhida por apenas 20,36% dos alunos.

Há quase 20 anos descobriu-se uma nova espécie de golfinho, o menor de sua espécie, com um bico pronunciado. Provavelmente não tenha sido descoberto antes porque vivem em águas profundas e mostram-se muito pouco.

Examinando a alternativa, observa-se que o fragmento "descobriu-se uma nova espécie de golfinho: o menor de sua espécie..." produz confusão, pois a palavra espécie no primeiro caso alude a *Mesoplodon peruvianus* e, no segundo, parece referir-se aos golfinhos em geral ou ao gênero dos golfinhos bicudos aos quais pertence a espécie descoberta. Além disso, o texto não se refere claramente à classificação dos golfinhos bicudos e do *Mesoplodon peruvianus* como gênero, espécie etc., e este duplo uso da palavra espécie não é válido nem ajuda a compreensão.

Por outro lado, em nossa opinião, o resumo apresentado como correto é um tanto incompleto porque, por exemplo, não se refere à "notícia": o fato importante, segundo esta alternativa, sucedeu há vinte anos e suprime-se do relato o processo posterior, que permitiu confirmar aquele achado e oferecer a informação à imprensa; também desaparece a figura do pesquisador peruano.

É muito provável que os alunos da 7ª. série tenham dificuldade para elaborar resumos - o que se detecta até em alunos que estão iniciando a carreira universitária -; no entanto, aspectos como os assinalados podem ter levado a uma escolha inadequada. Dos avaliados, 45,64%, que optaram pela alternativa incorreta A, escolheram um resumo com referências concretas do texto, relacionadas em forma coerente e cuja primeira oração tem um verbo em modo pessoal e um sujeito, o pesquisador peruano.

A avaliação da capacidade de detectar as idéias principais de um texto sem distorcê-las, atribuindo-lhes uma organização adequada, prestar-se-ia para formular uma pergunta aberta baseada em certas pautas, na qual se exigisse a produção de um resumo escrito.

No tocante à sua medição através de itens de resposta estruturada, poderiam ser elaboradas várias alternativas (inclusive mais de quatro), pois, de fato, cada leitor faria um resumo diferente; entre elas poderia haver mais de uma resposta válida e eventualmente poder-se-ia optar por uma ou mais de uma. A existência de exercícios desse tipo evidenciaria a aceitação de certo grau de divergência válido na compreensão e na produção, no tocante à avaliação de uma competência de compreensão global que é fundamental na leitura, diferentemente de outros itens nos quais são avaliadas questões pontuais.

No **item 26**, em relação a uma oração do texto em que se mencionam os mamíferos marinhos, pergunta-se que classe de palavra é marinhos. Cerca de 33% dos alunos escolheram a resposta correta B, adjetivo, enquanto a maioria, 55,22%, optou pela A, substantivo. O item está bem colocado, pois a palavra é considerada pela função que desempenha no texto lido. No entanto, um aspecto da diagramação - a oração transcrita fica fora do quadro do item pode ter influenciado neste caso, e por isso as pessoas podem ter optado por considerar a palavra marinhos em si, que tanto pode ser substantivo como adjetivo. Também cabe interpretar que os alunos não foram capazes de reconhecer o comportamento da palavra na oração.

Na **parte B** da prova há itens que interrogam sobre aspectos muito pontuais: ortografia, pontuação, separação de sílabas etc., e outros vinculados a questões mais complexas, como coerência e adequação ao gênero.

Dos itens que correspondem à ortografia-37, 38, 39-, os dois primeiros permitiram detectar deficiências dos alunos nesse aspecto: com efeito, as respostas certas para os pontos 37 e 38 não passaram de 24% e 35%, respectivamente. Um pouco acima quanto a resultados - com um pouco mais de 41% de respostas corretas - esteve o **item 32**, através do qual se avalia a separação de palavras em sílabas. Um desempenho bastante melhor ocorreu em concordância e correlação verbal (itens 35 e 36) e sinais de pontuação (**item 33**), com um índice de respostas corretas de 35% ou mais.

Vários desses aspectos da competência em Língua podem ser avaliados com maior contextualização através do exercício de expressão escrita (**item 41**), fazendo com que talvez alguns itens da parte B se tornem desnecessários.

Gostaríamos de comentar particularmente o **item 31**, no qual se apresenta uma família de palavras, navegador, navícula, navegável, naval e pergunta-se qual é a palavra primitiva. A alternativa correta C, nave, foi escolhida por apenas 29,10% dos alunos, enquanto a alternativa A, navegante, atraiu 35,60%. Isto pode ser explicado pelo desconhecimento do conceito de palavra primitiva por parte dos alunos. No entanto, talvez sua competência lingüística lhes permita reconhecer esta palavra.

Se um exercício apresentasse uma série de palavras, inclusive a primitiva, e se perguntasse qual delas é a base sobre a qual as outras se formaram, talvez os alunos pudessem responder acertadamente. Outra maneira possível seria explicar de forma simples, na raiz do item, os conceitos de palavra "primitiva" e "derivada" e depois formular a pergunta

sobre um conjunto determinado de palavras. Em ambos os casos estaria sendo avaliada a competência lingüística.

Resumindo os tipos de dificuldade encontrados no material analisado, podemos assinalar o seguinte:

- nem sempre se avalia o que se pretende avaliar, seja por uma construção inadequada do item, seja porque a resposta esperada não se vincula àquilo que a tabela de especificações estipula para esse item;
 - às vezes avalia-se mais um conceito gramatical ou um conhecimento relativo à organização do texto que a competência lingüística ou a competência em leitura;
 - alguns aspectos referentes à apresentação dos textos - formato, falta de referência à fonte - não ajudam a contextualização;
 - embora os textos escolhidos estejam aptos a fornecer uma série de itens que abrangem um amplo leque de aspectos, o primeiro texto pode ficar "descontextualizado" para os alunos que desconheçam o que é um farol e um faroleiro, enquanto o segundo texto resulta pouco claro em alguns pontos;
 - alguns trechos dos textos oferecem dificuldade e isso ou é transferido para a pergunta ou se tenta evitar, mas de alguma maneira fica evidente nos itens;
 - o modo de formulação de algumas perguntas implica a escolha de uma resposta incorreta;
 - em pelo menos cinco itens a resposta apresentada como correta ou não o é ou não o é totalmente, ou não é a única válida com relação às demais alternativas apresentadas.
- Voltando ao que já mencionamos, os avaliadores aproveitaram os resultados das provas e das observações registradas pelos pesquisadores encarregados da aplicação para o seu próprio *feedback*. Provavelmente - embora isso fique fora de nossa análise - as provas sucessivas tenham melhorado quanto aos textos selecionados e à elaboração dos itens.

No entanto, à guisa de um último comentário sobre a validade do instrumento em seu conjunto, é possível colocar algumas questões. A primeira delas é saber se, do ponto de vista lingüístico dos processos da leitura, é válido propor uma taxonomia cujo correlato, os exercícios apresentados na prova, de algum modo desagrega os saberes e talvez até produza interferência na própria competência.

A segunda questão é como interpretar os resultados dos indivíduos avaliados, tanto daqueles que tiveram uma média aceitável ou um nível elevado de respostas corretas, como daqueles que obtiveram uma pontuação considerada deficiente ou abaixo da aceitável.

Nossa hipótese é que os alunos do primeiro grupo demonstraram uma competência global em leitura e em língua. Puderam superar os obstáculos de certas perguntas e, mesmo no

caso de terem escolhido uma opção não totalmente correta segundo a nossa análise, optaram muito provavelmente pela melhor das alternativas apresentadas.

Quanto aos alunos do segundo grupo, resta saber se sua real compreensão dos textos ou seu conhecimento gramatical, ortográfico etc. é superior àquilo que a natureza do instrumento lhes permitiu expressar. Talvez com exercícios de outro tipo tivessem obtido uma pontuação superior; provavelmente estas crianças pertençam às camadas sociais mais baixas ou a ambientes com menores oportunidades de acesso à leitura - como claramente pode ver-se no desempenho da zona rural -, e estejam habituadas a um tipo de pensamento menos analítico que o exigido nessas provas. Estes alunos merecem que se levem em conta seu contexto cultural, seu estilo de compreensão e de resposta e seu ritmo de aprendizagem na avaliação.

De fato, muitas das perguntas formuladas na prova são um tanto artificiais e não são as que habitualmente seriam feitas - fora de uma situação de avaliação formal - para comprovar a compreensão de um texto.

A questão seria como aproveitar as vantagens dos instrumentos objetivos para avaliar populações numerosas, tratando de reduzir ao máximo suas desvantagens que se vinculam às limitações que apontamos. É preciso experimentar novos desenhos que permitam que as perguntas formuladas estejam relacionadas aos processos reais que ocorrem na mente do leitor no tocante a um texto determinado, que deve ser respeitado como tal, e que ampliem a divergência quanto à incorporação de mais de uma resposta possível - nos casos em que isto couber -, ou através de perguntas abertas.

Talvez fosse necessário tentar aplicar outras estratégias complementares, mais vinculadas à atividade da sala de aula, sujeitas a uma verificação externa. Resultam dignas de menção as reflexões realizadas pelos especialistas do MCyE em um documento no qual se analisa a produção em língua e se estabelecem relações com as respostas aos itens objetivos da prova que analisamos. Como sugestões para futuros instrumentos, eles propõem que

... seria conveniente Melhorar o instrumento (itens objetivos), ampliando o campo das operações realizadas pelo aluno e avançando rumo a estratégias mais completas e mais próximas da produção e do uso da língua. Este aspecto é particularmente difícil e envolve um grande desafio, porque pressupõe uma conciliação - que nunca será totalmente bem-sucedida - entre uma abordagem construtivista da didática da língua e um instrumento de avaliação comportamental que "segmenta" o processo de aprendizagem (30).

Do ponto de vista dos avaliadores, a construção das provas é um processo árduo, complexo, requer "ciência" e tem ao mesmo tempo um "componente artesanal" que nem sempre tem sido devidamente considerado. Por outro lado, a avaliação aparece como um espaço de confluência de numerosas disciplinas, exigindo interações muito mais ricas e complexas que as habituais, por parte de "especialistas em avaliação" e "especialistas em conteúdos".

30 Ministério de Cultura e Educação, Secretaria de Programação e Avaliação Educativa, Sistema Nacional de Avaliação. Primeira operação nacional, 1993. Informe de la producción em em água, Buenos Aires, março de 1994.

Em geral, os países latino-americanos têm - além das diferenças entre países - uma escassa cultura avaliativa em comparação com os países centrais; isso tem suas desvantagens mas também algumas vantagens. Enquanto podemos aprender de teorias e práticas de outras regiões, também temos a possibilidade de inovar a partir de nossas raízes e tentar transferir para o campo da avaliação nossos próprios desenvolvimentos, tanto pedagógicos como das diversas disciplinas.

No contexto do Mercosul e tendo-se avançado na revalidação de estudos, certificados e diplomas, seria de sumo interesse que houvesse uma compatibilização mínima dos currículos e um intercâmbio fluído sobre práticas inovadoras em avaliação.

Conclusões e Perspectivas

Em torno da avaliação da qualidade e da atribuição ou não de qualidade às conquistas da aprendizagem, estão sendo gerados em nossa região, e em outras, marcos conceituais mais de acordo com as necessidades sociais e educacionais de nossas populações. Nesta tarefa de reflexão e de construção do conhecimento, podem tornar-se significativas as contribuições das ciências da linguagem, que levam ao reconhecimento e à valorização da diversidade cultural lingüística e cognitiva. Critérios e conceitos da teoria lingüística, da sociolingüística e da psicolingüística, assim como do campo interdisciplinar do estudo sobre leitura, podem ser aplicados à avaliação.

As avaliações realizadas para medir a qualidade - ou outras afins - põem em evidência as concepções e práticas existentes nos sistemas educativos. No entanto, numerosos desenvolvimentos inovadores no ensino da língua não foram convenientemente representados nelas.

Em avaliações de alcance nacional surgem alguns desafios: como construir teoricamente os traços de uma competência que leve em conta os processos lingüísticos de leitura e de escrita dos sujeitos avaliados e que, partindo do local e do regional, abranja o amplo espectro de um país?

Um segundo desafio tem a ver com as estratégias de avaliação e os instrumentos escolhidos. As provas objetivas criam um marco restritivo para o avaliador, mas é possível superar algumas das suas limitações. A valorização da diversidade terá de refletir-se nos textos escolhidos, nas perguntas realizadas, nas respostas esperadas.

Se quisermos que as populações avaliadas possam demonstrar seus saberes, devemos ampliar seu grau de participação, deixando de lado alguns estereótipos e construindo instrumentos válidos e confiáveis que formulem perguntas inteligentes, cujas respostas esperadas levem em conta a diversidade lingüística, cultural e cognitiva.

Bibliografia"

Livros

COSTA, Marta Elena. Pruebas objetivas para evaluar la comprensión en lectura. *Lecturay Vida*, IRA, n. 1, mar. 1984.

_____. Aportes de la ciência del lenguaje para la consideración de la calidad en educación y su evaluación. *Revista Iberoamericana de Educación*, n. 10, p. 79-99, jan./abr. 1996.

EDWARDS R., Verónica. El concepto *de calidad de* la educación. Santiago, Chile: UNESCO OREALC, 1991.

PRADO DE SOUSA, C. (Org.). *Avaliação do rendimento escolar*. Campinas: Papirus, 1992.

RAMA, Germán W. *Quê aprendeu y quiénes aprenden en lis escuelas de Uruguay*. Montevideo: CEPAL, 1992.

Documentos

DOUBLIER, Alejandro. Proceso de construccitín de instrumentos. informe técnico n° 2. Buenos Aires: MCyE, 1993. (Primeira operação nacional 1993).

GLEJBERMAN, David. *Directo de la muestra*: informe técnico n4 1. Buenos Aires: MCyE, 1993. (Primeira operação nacional 1993).

REPÚBLICA ARGENTINA. Ministério de Cultura e Educação. Conselho Federal de Cultura e Educação. *Contenidos básicos comunes para la edivración general básica*. Buenos Aires: MCyE, 1995.

_____. Ministério de Cultura e Educação. Secretaria de Programação e Avaliação Educativa. Sistema Nacional de Avaliação Educativa. Primeira operação nacional 1993. *Instruaivo para el enquestador*. Buenos Aires: MCyE, 1993.

_____. Ministério de Cultura e Educação. Secretaria de Programação e Avaliação Educativa. Sistema Nacional de Avaliação Educativa. Primeira operação nacional 1993. *Resultados nacionales: primer informe*. Buenos Aires: MCyE, 1993.

_____. Ministério de Cultura e Educação. Secretaria de Programação e Avaliação Educativa. Sistema Nacional de Avaliação Educativa. Primeira operação nacional 1993. *Ancifisis de los itens de mayor dicultad: lengua nivel primario*. Buenos Aires: MCyE, 1993.

" Omitem-se as obus e documentos citados no rodapé do artigo.

REPÚBLICA ARGENTINA. Ministério de Cultura e Educação. Secretaria de Programação e Avaliação Educativa. *Recomendaciones metodológica para la enseñanza: lengua nivel primario*. Buenos Aires: MCyE, [s/d].

_____. Ministério de Cultura e Educação. Secretaria de Programação e Avaliação Educativa. *Recomendaciones metodológicas para la enseñanza: lengua nivel medio*. Buenos Aires: MCyE, [s/d].

_____. Ministério de Cultura e Educação. Secretaria de Programação e Avaliação Educativa. *Programa nacional de lecturay producción escrita*. Buenos Aires: MCyE, [s/d].

_____. Ministério de Cultura e Educação. Secretaria de Programação e Avaliação Educativa. *Operativo nacional de evaluación, 1994: primer informe*. Buenos Aires: MCyE, 1994.

_____. Ministério de Cultura e Educação. Secretaria de Programação e Avaliação Educativa. *Operativo nacional de evaluación, 1995*. Buenos Aires: MCyE, 1995.

Instrumentos

REPÚBLICA ARGENTINA. Ministério de Cultura e Educação. Secretaria de Programação e Avaliação Educativa. Sistema Nacional de Avaliação Educativa. *Prueba lengua primaria*. Buenos Aires: MCyE, 1993.

_____. Ministério de Cultura e Educação. Secretaria de Programação e Avaliação Educativa. Sistema Nacional de Avaliação Educativa. *Prueba lingua media*. Buenos Aires: MCyE, 1993.

_____. Ministério de Cultura e Educação. Secretaria de Programação e Avaliação Educativa. Sistema Nacional de Avaliação Educativa. *Prueba de Lingua nivel primario*. Buenos Aires: MCyE, 1994.