

Considerações Finais

Procurei, neste estudo, comparar os desempenhos alcançados pelos alunos brasileiros e portugueses, em Matemática, no Programa Internacional de Avaliação de Estudantes (PISA). Considerei como ponto de partida a concepção de que os resultados de avaliações dessa natureza refletem o que foi apreendido pelos alunos a partir do currículo ensinado. Ou seja, a idéia de que as avaliações de sistemas educacionais podem ser utilizadas como estratégia de apreensão do “currículo aprendido”.

Esses estudos são viáveis, hoje em dia, graças ao desenvolvimento do campo da Avaliação Educacional no Brasil, com reflexo nas modernas técnicas e metodologias aplicadas nesta área. Desde a criação e implementação do SAEB, o país tem investido pesadamente nos sistemas de informações e nas avaliações educacionais, por entender que estes são instrumentos imprescindíveis para subsidiar a formulação de políticas públicas, bem como para estimular a melhoria da qualidade do ensino. Pode-se afirmar que esse campo está consolidado e dá sinais de que as avaliações devem continuar sendo aprimoradas, com vistas a um eficaz monitoramento dos sistemas educacionais, quanto aos níveis de desempenho alcançados pelos alunos e sua evolução, ao longo dos anos. Voltarei a este tema, mais adiante.

Esse esforço pela consolidação do campo e os avanços alcançados, tanto na concepção como na elaboração dos instrumentos utilizados na avaliação educacional em larga escala, convergiram, naturalmente, para a participação do Brasil na avaliação internacional do PISA, juntamente com países que têm forte tradição na área e cujos sistemas educacionais são os mais avançados do mundo.

As informações produzidas a cada ciclo de avaliação reafirmam dois objetivos bastante claros: primeiro, possibilitar uma melhor definição das políticas públicas necessárias, com vistas ao melhoramento da qualidade da educação e à redução das desigualdades ainda existentes no sistema educacional brasileiro; segundo, contribuir para a afirmação de uma cultura avaliativa no país, estimulando o debate público e despertando, na sociedade, maior interesse pela situação da educação oferecida pelo Estado.

Cabe destacar que foi com o foco nesse segundo objetivo que este estudo se desenvolveu. Já que, há alguns anos, as pesquisas que buscam aferir os níveis de desempenho dos alunos brasileiros, particularmente em Matemática, vêm indicando que a escola não tem conseguido cumprir, de modo adequado, sua tarefa de promover as aprendizagens as quais os estudantes almejam, necessitam e têm direito.

Nesse sentido, apresentei os indicadores do aprendizado nacional, extraídos do SAEB, onde se constata o baixo aprendizado dos alunos brasileiros e a distribuição desigual desse aprendizado ao longo do tempo. Mostrei, também, que o descompasso entre o que é proposto pelos currículos e o desempenho real dos alunos se acentua a partir das séries finais do Ensino Fundamental e do Ensino Médio. O acúmulo dos déficits de aprendizagem ao longo dos anos faz com que os alunos da 3ª série do Ensino Médio agreguem ganhos menores em relação aos alunos da 8ª série do Ensino Fundamental do que estes, em comparação com os da 4ª série, considerando-se as médias de proficiências obtidas. Os dados do SAEB 2005 mostram que 70% dos estudantes da 3ª série do Ensino Médio obtiveram pontuação inferior a 300 pontos, valor correspondente à média mínima satisfatória para alunos da 8ª série.

Igualmente, no PISA, os resultados ficaram aquém do desejável. Em 2003, alcançamos 356 pontos na escala do PISA, onde a média era 500 pontos. Esse resultado nos classificou, em média, abaixo do nível 1, na escala de proficiência. Mais da metade dos nossos alunos, 54,4 %, posicionaram-se abaixo do nível 1, na escala global de Matemática. Considerando-se as quatro subáreas da Matemática avaliadas, os percentuais de alunos situados abaixo do nível 1, ou seja, aqueles que não foram capazes de utilizar as capacidades matemáticas requeridas pelas tarefas mais simples da avaliação, também foram bastante elevados: Espaço e Forma (56,4%), Mudança e Relações (61,3%), Incerteza (43,8%) e Quantidade (52,3%), caracterizando-se um desempenho bastante sofrível.

Considerando-se o conceito de letramento adotado no PISA, onde se busca avaliar até que ponto os conhecimentos e as habilidades essenciais para a participação efetiva na sociedade foram adquiridos pelos alunos em idade próxima ao término da escolarização básica obrigatória, esse resultado é, no mínimo, preocupante. Afinal, essas competências e habilidades não são diferentes das que constam nos PCNs e das recomendadas pela literatura na área da Educação

Matemática, há algumas décadas. Mesmo assim, por meio de diferentes inquéritos (SAEB, PISA, Avaliações Estaduais), constata-se que persistem os problemas na aprendizagem da Matemática, o que nos leva a crer que seu ensino, nas escolas, continua priorizando as suas antigas tradições e andando na contramão dos discursos, no âmbito da Educação Matemática. Tais problemas sinalizam para a necessidade de mudança nas práticas de alguns professores e para uma atenção especial na elaboração das propostas curriculares, principalmente no que tange à compartimentalização dos conteúdos e à concepção acerca do saber matemático.

Diante de diagnósticos semelhantes, cabe destacar que a troca de experiências entre especialistas e técnicos em avaliação e a perspectiva de uma comparação internacional constituem-se no grande diferencial trazido pelos resultados, com a participação no PISA. Essa comparação, no sentido mais próximo que pretendi desenvolver nesta tese, que é o ato de *examinar simultaneamente, a fim de conhecer as semelhanças, as diferenças ou relações* (Aurélio, p.439), visando esclarecer questões teóricas e práticas relativas ao ensino e ao aprendizado da Matemática, desenvolveu-se de duas maneiras diferentes, porém, conectadas. No capítulo 4, comparei os resultados brutos obtidos pelos alunos dos dois países, basicamente em termos de médias e percentuais de acerto, e algumas características dos diferentes sistemas educacionais. No capítulo 5, as comparações entre as diferenças foram feitas sempre controladas pela habilidade, ou seja, comparei a performance, frente ao item, de alunos dos dois países que tinham a mesma habilidade cognitiva em Matemática.

A possibilidade de se compararem grupos diferentes de alunos, desde que submetidos a provas que tenham alguns itens em comum ou, ainda, a comparação entre indivíduos do mesmo grupo que tenham sido submetidos a provas totalmente distintas, é uma das grandes vantagens trazidas pela TRI para a área educacional. Graças às escalas comuns de proficiência, interpretadas em termos do que os alunos conhecem, compreendem e são capazes de fazer, com base nos resultados do seu desempenho, várias questões práticas em educação passam a ser possíveis, dentre elas, o acompanhamento da evolução do desempenho dos alunos e a efetividade ou não do ensino ministrado nas escolas, além da comparação de resultados entre diferentes séries, por disciplinas e de ano para ano.

Os modelos mais tradicionais da TRI pressupõem que os itens apresentem o mesmo funcionamento em diferentes grupos. Uma boa comparação entre

resultados de grupos diferentes de alunos requer, portanto, que os itens que compõem os testes não apresentem DIF, pois do contrário, isso significaria que algum particular grupo de alunos estaria sendo privilegiado em detrimento dos demais, ainda que as proficiências estimadas sejam relativamente robustas em situações nas quais algum DIF está presente. Face ao pressuposto acima mencionado, usualmente busca-se produzir itens de teste que não apresentem DIF, ainda que produzir testes completamente sem DIF seja uma tarefa muito difícil quando as populações são muito diferentes, o que é o caso de alunos de diferentes países. Diante desta situação, abrem-se duas oportunidades de pesquisa: (a) desenvolvimento de modelos de TRI que relaxem o pressuposto acima mencionado; e (b) investigação do sentido educacional do DIF. A presente pesquisa explorou esta segunda oportunidade.

De fato, demonstrei que, não obstante os cuidados que cercam a elaboração e seleção de itens para uma avaliação internacional desse porte, há grande presença de DIF nos itens do PISA, entre o Brasil e Portugal. Através da exposição dos itens públicos, mostrei também que, de um modo geral, todos os itens estão adequados à idade e à escolaridade dos examinandos; ainda, são objetivos e claros acerca da habilidade ou competência que pretendem avaliar e, embora sejam traduzidos para o idioma do país avaliado, o vocabulário e os termos utilizados não se constituem, a priori, num obstáculo à resolução do item. Após essa primeira análise, ficou claro que os fatores causadores desses DIFs não eram tão triviais; ao contrário, são fatores bem mais complexos, que têm a ver com diferenças culturais entre países, refletidas na seleção e ênfases no currículo oficial de Matemática.

Nesse caso, cabe a compreensão de que, por razões dessa natureza, nem sempre itens diagnosticados apresentando DIF (desde que não tenha impacto apreciável sobre a proficiência estimada) são totalmente maléficos ao processo avaliativo e, por isso, devam ser retirados dos testes ou dos cálculos da proficiência. Isto porque, além do fato desse procedimento ser muitas vezes desnecessário e prejudicial à qualidade da estimativa da proficiência do aluno, a análise do DIF, da forma como vimos neste estudo, pode ser uma ferramenta de diagnóstico do sistema educacional bastante útil. A detecção de diferenças nos currículos ou conteúdos que são abordados com ênfases diferenciadas entre Brasil

e Portugal emergiu da análise desses itens que apresentaram funcionamento diferencial.

Identifiquei 23 itens apresentando DIF. As hipóteses formuladas para explicar o DIF entre alunos brasileiros e portugueses, surgiram, assim, de uma análise exploratória com apenas 23 itens. As características observadas desses itens traduziram-se em variáveis, que depois foram aplicadas à estrutura do modelo integrado para uma confirmação mais rigorosa e formal das hipóteses levantadas. No entanto, à medida que seleciono uma característica para análise, o conjunto de itens, no universo dos 84 disponíveis, apresentando essa particular característica, fica muito reduzido, comprometendo, muitas vezes, a significância estatística dos resultados. Considerando essas limitações, as conclusões a que cheguei, após os resultados estimados pelo modelo, podem ser expressas nos seguintes termos:

- Para um nível de confiança de 95%, a prova foi estatisticamente significativa mais fácil para os alunos portugueses, em comparação aos alunos brasileiros, e discrimina mais os alunos portugueses;
- Para um nível de confiança de 95%, quanto ao conteúdo da Matemática, itens da subárea “Quantidade” tendem a ser mais fáceis para alunos brasileiros (desde que envolvendo apenas operações algébricas básicas com números inteiros); e itens da subárea “Mudança e Relações” são, aparentemente, mais fáceis para alunos portugueses;
- Quanto ao contexto em que os conhecimentos são aplicados, itens de contexto “Pessoal”, para um nível de confiança de 90%, tendem a ser mais fáceis para alunos brasileiros e itens de contexto “Científico”, no mesmo nível de significância, tendem a ser mais fáceis para alunos portugueses;
- Quanto ao processo da Matemática aplicado na resolução do item, “Conexão” mostrou-se, aparentemente, favorável aos alunos brasileiros, enquanto itens envolvendo a “Reprodução” parecem ser mais fáceis para alunos portugueses;
- Itens cujos enunciados são longos, descrevendo procedimentos a serem seguidos, ou orientações passo a passo de como resolver uma questão, às vezes acompanhados de exemplos, são estatisticamente

significativos, a um nível de 95%, e mais fáceis para os alunos brasileiros;

- Quando os itens envolvem números na representação decimal, eles tendem a ser mais fáceis para os alunos portugueses;
- Itens envolvendo interpretação de gráficos, de variados tipos, tendem a ser mais fáceis para alunos brasileiros.

As análises comparativas realizadas e as conclusões que ora apresento direcionam o nosso olhar para duas constatações. A primeira são os grandes desafios a serem vencidos na educação básica brasileira, particularmente no Ensino Médio. Seu projeto oficial, arrojado por seu caráter prospectivo, ainda não conseguiu sair do papel e superar os privilégios, diminuir as desigualdades e atender plenamente as expectativas de uma economia em desenvolvimento por recursos humanos mais qualificados, como deseja e expressa a Lei de Diretrizes e Bases da Educação Nacional (LDB). O acúmulo dos déficits de aprendizagem ao longo dos anos, conforme mostrado, tem como consequência real o baixo desempenho dos alunos no Ensino Médio, já que estes não consolidam as competências e habilidades adquiridas no Ensino Fundamental e, por isso, não conseguem adquirir as novas competências que imprimem a este nível de escolaridade uma característica própria, uma terminalidade própria e que o distingue como finalização da educação básica.

A segunda constatação a que me referi tem a ver com a gama de estudos que são viabilizados a partir dos dados das avaliações. A preocupação e questionamentos em relação ao Funcionamento Diferencial do Item (DIF), consequência da crescente utilização da TRI na área de avaliação educacional, é apenas um exemplo. E aqui retomo a questão sobre a consolidação do campo da avaliação educacional no Brasil e o aprimoramento dos mecanismos de avaliação e monitoramento da melhoria da qualidade, da equidade e da eficiência do ensino básico brasileiro.

Um bom exemplo desse movimento é que o diagnóstico da educação básica, que até então era feito amostralmente pelo SAEB, foi expandido com o “Prova Brasil”. Isso significa que as inferências estatísticas acerca dos resultados sobre a qualidade dos sistemas educacionais, que só podiam ser feitas ao nível de país e unidades da federação, passam a ser possíveis também para os municípios e para as escolas públicas urbanas de ensino fundamental. Essa possibilidade de

divulgação de resultados individualizados representa um grande avanço, em termos de avaliação (Fernandes, 2008). Além disso, o Inep criou, em 2004, o Índice de Desenvolvimento da Educação Básica (IDEB), cujo cálculo combina taxas de aprovação com resultados da aprendizagem, medidos pelo SAEB e Prova Brasil. Assim, para uma escola obter um bom índice, será necessário que seus alunos progredam em sua trajetória escolar, aprendendo o que se espera que aprendam.

Finalizo, sinalizando que a superação de muitos desafios na área da educação, mesmo aqueles relacionados com o que se aprende em sala de aula, dependerá cada vez mais da apropriação, pelos gestores, pesquisadores e educadores, dos dados gerados por eficientes sistemas de informações.