

Do Saeb ao Sinaeb: prolongamentos críticos da avaliação da educação básica

João Luiz Horta Neto

Rogério Diniz Junqueira

Adolfo Samuel de Oliveira

Resumo

A evolução do Sistema de Avaliação da Educação Básica (Saeb) e sua influência no desenvolvimento de outros testes em larga escala no Brasil é analisada com destaque para as repercussões das avaliações nas redes de ensino, a expansão dos testes em larga escala nos estados e municípios e a limitação do uso que se tem feito de seus resultados, enfatizando a competição que se instaura entre as escolas e as redes na busca de melhores desempenhos, sem que isso signifique melhorias da aprendizagem. Também são discutidas as vantagens de se utilizar um sistema de indicadores em vez de um índice sintético, como o Índice de Desenvolvimento da Educação Básica (Ideb), para diagnosticar, monitorar e avaliar a educação básica, tendo em vista o modelo de Sistema Nacional de Avaliação da Educação Básica (Sinaeb) proposto pelo Plano Nacional de Educação (PNE). Por fim, destaca-se a necessidade de esse novo sistema de avaliação valer-se de instrumentos e processos que permitam mensurar de maneira apropriada as habilidades dos estudantes com deficiência, elaborar indicadores sobre as políticas de educação especial e averiguar as condições de acessibilidade e aprendizagem nas realidades escolares.

Palavras-chave: Saeb; avaliação educacional; educação básica; acessibilidade.

Abstract

From Saeb to Sinaeb: needed improvements for the basic education evaluation system

This paper discusses the evolution of Sistema de Avaliação da Educação Básica (SAEB) and its influence on the development of other large-scale testing in Brazil, highlighting the impact of assessments in educational networks, the expansion of large-scale tests in the states and municipalities of the country and the limitation the use that is made of the results, emphasizing the competition between schools and educational networks for better performances, that does not necessarily mean improvements in learning. Further, it discusses the advantages of using a system of indicators, instead of a synthetic index, as the Índice de Desenvolvimento da Educação Básica (Ideb), to diagnose, monitor and evaluate basic education, according to the model proposed by Plano Nacional de Educação (PNE) for Sistema Nacional de Avaliação da Educação Básica (Sinaeb). Finally, it highlights the need that this new evaluation system make use of tools and processes to properly measure the disabilities student's skills, to develop indicators on special education policies and determine the conditions of accessibility and learning in school realities.

Keywords: SAEB; educational evaluation; basic education; accessibility.

Introdução

No mundo inteiro, ao longo nos últimos anos – em especial a partir dos anos 1990 –, muita importância tem sido dada às medidas dos resultados do processo educacional para iluminar as políticas da área. Prova disso é que se amplia o número de países que desenvolvem testes cognitivos para medir o desempenho de seus sistemas educacionais. Da mesma forma, vem aumentando o número de países que têm participado de estudos internacionais comparados, como o Programme for International Student Assessment (Pisa), desenvolvido pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE): em 2000, ano de sua primeira edição, 45 países participaram do teste e, em 2015, foram 76.

Ao mesmo tempo que se instaura um movimento global na direção dos testes cognitivos, crescem também as críticas ao seu uso indiscriminado. Nos Estados Unidos, país em que as políticas de responsabilização mais se desenvolveram, milhares de famílias se organizam em torno do movimento conhecido como *opt-out*, impedindo que seus filhos participem dos testes.

De forma geral, quando os testes se concentram apenas nas medidas do desempenho dos alunos e na divulgação desses resultados, limita-se o escopo da avaliação educacional. Como consequência, em grande parte das vezes, responsabiliza-se a escola e seus profissionais pelo resultado alcançado, como se ambos tivessem capacidade de, isoladamente, modificar a realidade que os cerca.

No Brasil, o governo federal vem desenvolvendo esforços nos últimos anos para ampliar não somente o número de testes como a quantidade de alunos testados. Inicialmente, no final dos anos 1980, foi criado o Sistema de Avaliação da Educação Básica (Saeb). Com o passar dos anos, o governo federal desenvolveu um vasto conjunto de instrumentos para medir o desempenho dos estudantes da educação básica: o Exame Nacional de Certificação de Competências de Jovens e Adultos (Encceja), o Exame Nacional do Ensino Médio (Enem), a Provinha Brasil e a Avaliação Nacional da Alfabetização (ANA).

Inicialmente, o Saeb foi previsto para fornecer informações sobre os sistemas educacionais, e era aplicado de forma amostral. A partir de 2005, ele passou a ser aplicado de forma censitária para as escolas públicas de ensino fundamental. Apesar da ampliação do número de instrumentos e da quantidade de alunos testados, os estudantes com deficiência não contam com instrumentos adequados para medir seu desempenho, dificultando, com isso, a avaliação das políticas voltadas a eles.

A existência de desempenhos calculados por escola permitiu a criação, em 2007, do Índice de Desenvolvimento da Educação Básica (Ideb), e a definição, pelo governo federal, de metas para escolas. Esse índice sintético, envolvendo desempenho no teste e fluxo escolar, passou a representar a qualidade da escola, apesar de não levar em consideração outras importantes dimensões do processo educacional.

O presente texto, tomando por base os aspectos tratados aqui, se propõe a discutir, na primeira seção, a expansão dos testes em larga escala nos estados e municípios e a limitação do uso que se tem feito de seus resultados; na segunda, as vantagens de utilizar um sistema de indicadores, ao invés de um índice sintético, como o Ideb, para diagnosticar, monitorar e avaliar a educação básica, conforme requer o Sistema Nacional de Avaliação da Educação Básica (Sinaeb), instituído pelo Plano Nacional de Educação (PNE); e, na terceira e última, a necessidade de construir, dentro desse Sistema, testes capazes de mensurar de maneira apropriada as habilidades dos estudantes com deficiência, bem como indicadores que iluminem as condições de acessibilidade e aprendizagem a que estão submetidos.

A responsabilização, a expansão dos testes e os limites dos usos de seus resultados

A discussão acerca da responsabilização tem tomado conta das agendas de diversos países. Sahlberg (2012) afirma que existe um fenômeno que está se espalhando pelos sistemas escolares ao redor do mundo, ao qual o autor dá a sugestiva alcunha de GERM, abreviatura de Global Educational Reform Movement. Esse movimento global possui três características: a competição na educação, a transferência de escolas públicas para instituições privadas e a forte responsabilização das escolas. Segundo o autor, esse fenômeno é como uma epidemia que se espalha e infecta os sistemas de ensino, fazendo com que as escolas fiquem doentes, os professores não se sintam bem e as crianças aprendam menos. Afirma, ainda, que o aumento dos testes avaliativos tem causado efeitos danosos e indesejados, como

o crescimento do ensino focado nos testes, o estreitamento do currículo para priorizar a leitura e a matemática e aproximação da pedagogia a uma instrução mecânica (Sahlberg, 2011).

Esse movimento em direção aos testes também se propaga no Brasil e, com ele, cresce a utilização do ranqueamento das escolas – caminho trilhado a partir de três movimentos: identificar as escolas de melhor desempenho, consideradas como exemplos a serem seguidos; divulgar os resultados dos testes, acreditando que as famílias irão mobilizar aquelas com baixo desempenho; e criar competição entre as escolas por desempenhos cada vez melhores. Como consequência, o ranqueamento firma-se como uma necessidade para as políticas e essencial para aprimorar o sistema educacional, levando, inclusive, os governos à utilização ampla dos testes acompanhados de prêmios a escolas, seus professores e seus alunos (Horta Neto, 2013).

O aumento do uso dos testes pode ser observado no Gráfico 1, a partir do número de secretarias de educação estaduais que o utilizam.

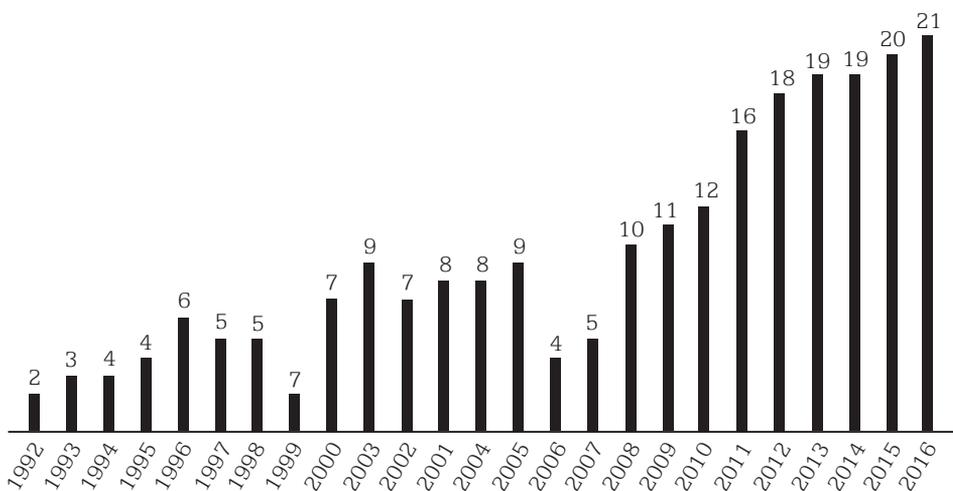


Gráfico 1 – Quantidade de estados realizando testes, por ano de sua aplicação

Fonte: Horta Neto (2014), com atualização dos autores junto às secretarias de educação estaduais.

Observa-se, pelo Gráfico 1, que, principalmente a partir de 2007, ano de criação do Ideb, o número de estados que fazem uso de testes como parte de suas ações educacionais tem aumentado. Esse aumento pode estar baseado na crença de que seu uso, por si só, seja capaz de garantir a qualidade da educação.

O mesmo movimento em direção aos testes é observado nos municípios. Pesquisa realizada pelo Inep em parceria com a Fundação Carlos Chagas, junto aos secretários de educação municipais e respondido por 4.309 deles, indicou que 1.784 municípios (41% dos respondentes) contam com avaliações próprias e que 905 deles (21% dos respondentes), apesar de ainda não possuírem uma, pretendem desenvolvê-la (Bauer *et al.*, 2015). A maior parte dessas avaliações refere-se unicamente a testes aplicados aos alunos, os quais, assim como aconteceu com os estados, também passaram a ser mais utilizados depois de 2007.

Ainda no sentido da responsabilização, tem-se reinterpretado o direito constitucional à educação, um direito social inerente à cidadania e aos direitos humanos, como sendo simplesmente um direito à aprendizagem (Horta Neto, 2013). Dessa forma, tira-se o foco das condições que o Estado e a sociedade devem construir para garantir uma educação de qualidade e centra-se a discussão no trabalho da escola e dos professores para garantir a aprendizagem, medida pelos testes de leitura e matemática. Apequena-se a aprendizagem dos alunos e o direito à educação, como se este direito já estivesse amplamente consolidado no Brasil, faltando apenas responsabilizar os atores da escola por não estarem garantindo o aprendizado dos alunos.

Imaginar que uma boa escola possa ser caracterizada apenas a partir de medidas de alguns de seus resultados é desconhecer a importância que têm os processos utilizados para se chegar a eles. Exemplo disso é que um bom e eficiente sistema educacional, capaz de preparar cidadãos que possam contribuir ativamente para a vida social e que tenham adquirido a habilidade de aprender a aprender ao longo da vida, tem relação com o processo educacional desenvolvido no dia a dia escolar, algo que vai além do resultado nos testes.

A educação escolar envolve a apreensão não só de conhecimentos, mas de valores e atitudes. Todos eles devem estar presentes no currículo, no projeto político pedagógico da escola, no seu relacionamento com a comunidade e com seus diferentes atores. Com relação aos conhecimentos, não somente leitura e matemática contribuem para o desenvolvimento dos estudantes: a educação física, as artes, os estudos sociais, entre outros, permitem agregar outras habilidades também importantes para permitir que os alunos façam sua leitura do mundo. Além disso, as diversas possibilidades de construir uma sociedade mais justa e igualitária são decorrentes do convívio com o outro, com a diferença, com as deficiências e com as infinitas formas de se estabelecer parâmetros éticos que permitam uma interação social baseada no respeito ao outro e às suas ideias e concepções de vida. Se todos esses aspectos são importantes para a formação integral do aluno, fazer inferências sobre a qualidade da educação a partir dos resultados nos testes de leitura e matemática é insuficiente. Mesmo que os testes fossem ampliados para outras áreas, um processo difícil e que exigiria tempo e recursos de monta, o resultado continuaria sendo uma fotografia sem nitidez suficiente para captar toda a realidade educacional.

Outro aspecto a ser considerado é que, por melhor que sejam os instrumentos utilizados para medir o desempenho cognitivo, eles apresentam limitações, pois nem tudo o que deve ser ensinado na escola é passível de ser medido por meio de testes de múltipla escolha, tipo de instrumento normalmente utilizado para aferir esse desempenho. Um exemplo disso é o Cálculo Mental, uma importantíssima habilidade ligada à área de matemática. Essa habilidade – que supera em muito a habilidade de fazer “contas de cabeça” – envolve, *grosso modo*, a habilidade de utilizar variados procedimentos para buscar diversas formas de solucionar um mesmo problema utilizando os diferentes mecanismos de cálculo. Ora, é extremamente difícil desenvolver itens de múltipla escolha que sejam capazes de medir essa habilidade, pois seria necessário acompanhar as muitas possibilidades que um aluno pode vir

a utilizar para solucionar o problema que lhe é proposto. Da mesma forma, ao testar a leitura, uma parte importante do ensino de língua portuguesa, deixa-se de lado, entre outros aspectos, a redação de textos.

Dada a centralidade que os testes vêm ganhando e a crescente preocupação em preparar os alunos para eles, o foco em testar apenas leitura e matemática e a dificuldade em elaborar itens para testar algumas habilidades fazem com que temas importantes deixem de ser tratados em sala de aula. Com isso, o estreitamento do currículo torna-se inevitável.

As considerações apresentadas até aqui não significam que os testes em larga escala, especialmente aqueles aplicados pelo governo federal, não sejam instrumentos importantes para o aprimoramento do processo educacional. Não se trata de satanizar tais testes cognitivos, mas de compreender que eles fornecem medidas limitadas da realidade que se pretende conhecer. O que se argumenta é que centrar todo esforço de coleta de informações para aquilatar a qualidade da educação usando apenas os testes e basear as políticas unicamente nos seus resultados gera muitas distorções. Nesse sentido, seriam necessários outros instrumentos, tais como: a autoavaliação institucional; o apoio constante da gestão das redes às suas escolas, especialmente aquelas que atendem as populações mais vulneráveis; e a promoção de políticas intersetoriais, com foco nos alunos e suas famílias. Essas ações podem fornecer as condições adequadas para que as crianças e os jovens possam desenvolver suas potencialidades.

26

Uma alternativa para os limites impostos pelo Ideb

O objetivo do Ideb é monitorar a regularidade da trajetória e o aprendizado dos alunos, fornecendo resultados, principalmente, para escolas e redes. Do ponto de vista da avaliação da educação básica, com a sua criação, o governo federal fincou as bases para a introdução do conceito de responsabilização e de mobilização social, o que redundou no Plano de Metas do Compromisso Todos pela Educação (Brasil, 2007). Tal política fomentou uma parceria entre a União, os estados, os municípios e a sociedade civil organizada, e estabeleceu metas intermediárias e finais para o Ideb do País, com o objetivo de possibilitar o acompanhamento e a *accountability* da oferta da educação escolar (Oliveira, 2014).

Apesar dos avanços trazidos para o monitoramento e a avaliação da qualidade da educação oferecida, a positividade do impacto que o Ideb e suas metas ensejaram nas políticas públicas educacionais e nos programas governamentais não é consensual, especialmente pelo modo como os conceitos de qualidade da educação e de responsabilização foram colocados em prática por determinadas redes de ensino (Oliveira, 2014). Como discutido anteriormente, iluminar apenas os resultados educacionais deixa de lado outros importantes elementos do processo de escolarização e fazê-lo através de um índice sintético acaba por encobrir as informações que foram por ele resumidas. Nesse sentido, vamos retomar aqui, brevemente, a discussão que contrasta o uso de índices sintéticos a um sistema de indicadores, tendo em vista refletir sobre qual é a maneira mais proveitosa de disponibilizar tais informações.

Scandar Neto, Silva e Jannuzzi (2006) realizam uma importante discussão acerca das vantagens e desvantagens do diagnóstico, monitoramento e avaliação realizados com base em indicadores sintéticos. Segundo os autores, o uso de tais índices, como o Índice de Desenvolvimento Humano (IDH), tem como vantagens, de modo geral, a legitimidade técnica, social e política, devido à explicitação da metodologia utilizada; a visibilidade que proporciona para os assuntos abordados, como se percebe no debate sobre o desenvolvimento humano e qualidade da educação; e seu uso, pelos governos e pela sociedade civil organizada, tanto para o diagnóstico de desigualdades sociais ou escolares quanto para o monitoramento e a avaliação de políticas e programas sociais. Por outro lado, esses índices sintéticos também apresentam desvantagens, quais sejam:

- transmitir uma visão imprecisa e até equivocada do assunto, caso tenham problemas técnicos ou sejam mal interpretados;
- ensejar uma visão simplista do problema e das propostas para sua solução, fragilizando as decisões de gestores governamentais;
- dificultar a explicitação dos julgamentos e das escolhas políticas que nortearam os procedimentos técnicos para a construção do índice, tal como o peso atribuído a determinada dimensão em detrimento de outra;
- substituir o conceito que se pretende mensurar pela própria medida, fazendo com que um retrato simplificado seja considerado capaz de abarcar a complexidade inerente ao conceito ou, o que é pior, à própria realidade enfocada; e, por fim,
- utilizar escalas adimensionais ou abstratas – que vão de 0 a 1 por exemplo –, uma vez que dificultam a compreensão das mudanças detectadas pelos indicadores que os compõem, por não terem uma relação mais próxima com os aspectos da realidade empírica que se deseja apreender.

Transpondo algumas dessas críticas ao Ideb das escolas, é possível afirmar que esse índice sintético, dependendo do seu uso, pode transmitir visões imprecisas a respeito dos ganhos dos alunos em conhecimento, pois, quando não se considera nas análises o nível socioeconômico, pode-se atribuir o alto desempenho ao trabalho pedagógico, quando, na verdade, este se deve mais ao público privilegiado que a escola atende; ou, ao contrário, considerar o trabalho pedagógico da escola como ineficaz, com base no seu baixo desempenho no Ideb, quando, na realidade, o valor que a escola agrega, em termos de conhecimento, é alto, pois, mesmo atendendo um público de nível socioeconômico baixo, cria um ambiente propício para sua aprendizagem, porém, como o ponto de partida é precário, não atinge o desempenho considerado adequado (Almeida; Dalben; Freitas, 2013).

Do ponto de vista técnico, o Ideb apresenta várias limitações, como apontam Soares e Xavier (2013), ao fazerem uma análise estatística e pedagógica desse índice. Dentre elas, destacamos o peso maior que matemática acaba tendo no computo final do índice e, o que consideramos mais importante, o encobrimento da desigualdade intraescolar ao utilizar, no seu desenho metodológico, a média das proficiências em leitura e resolução de problemas em vez da distribuição dos alunos pelos níveis de

proficiência da escala. Tal fato faz com que o aumento da média possa se dar por causa do alto desempenho de poucos alunos, o qual compensaria o baixo desempenho de muitos, colocando, assim, esse grave problema em segundo plano.

Ademais, a escala do Ideb é adimensional, logo, não dá uma ideia mais substantiva do que está sendo mensurado, impossibilitando, por exemplo, ter clareza do que significa a elevação de 1 ponto, num intervalo que, teoricamente, vai de 0 a 10. Também não permite vislumbrar se esse aumento exige muito ou pouco esforço, em termos de trabalho pedagógico e de aprendizagem dos alunos. Além disso, não possibilita saber se esse crescimento está ligado à melhoria da proficiência em leitura, matemática ou da taxa aprovação, por exemplo. Como as taxas de aprovação juntamente com as de reprovação e de abandono totalizam 100%, informar somente a primeira não indica se o percentual de não aprovados se deve à reprovação, ao abandono ou a ambos, o que acaba por dificultar o direcionamento das ações dos gestores – em termos administrativos e pedagógicos –, pois, são problemas distintos, que exigiriam ações também distintas.

Com base no exposto, consideramos que índices sintéticos, quando tomados como orientadores de ações governamentais, requerem muito cuidado na interpretação das informações que disponibilizam. O Ideb, como comentamos, por não controlar as diversas fontes do aprendizado discente e por focalizar as proficiências em apenas duas áreas do conhecimento, deixando de lado outras dimensões do ato educativo, pode estimular um tratamento inadequado do problema, bem como subsidiar de maneira frágil as tomadas de decisões dos gestores.

Não obstante, os apontamentos críticos feitos a esse índice sintético não inviabilizam outros usos produtivos, tais como fornecer um retrato mais geral do desempenho da rede, bem como sinalizar, à primeira vista, as escolas que precisariam de um maior apoio da administração da rede. Porém, parece-nos que as limitações relativas ao emprego do Ideb para o trabalho de diagnóstico, monitoramento e avaliação da educação básica, especialmente do ponto de vista da gestão governamental, podem ser superadas por um sistema de indicadores, que, ao tratar das várias facetas das dimensões educacionais – inclusive as que foram sintetizadas pelo referido índice –, conseguiria iluminar, de maneira mais ampla e específica, os aspectos da realidade educacional que precisam ser conhecidos e aprimorados.

Diante desse quadro, é oportuno pensar em alternativas que valorizem as potencialidades das informações produzidas pelo Inep e minimizem o seu uso menos profícuo para as finalidades do ato educativo. É o que foi proposto pelo Plano Nacional de Educação (PNE).

Além de consagrar o Sistema Nacional de Avaliação da Educação Básica (Sinaeb) no corpo da Lei e de constituir-lo, conforme o Art. 11, em “fonte de informação para a avaliação da qualidade da educação básica e para a orientação das políticas públicas desse nível de ensino”, o atual PNE introduziu uma modificação de extrema relevância para o debate educativo atual no que diz respeito às informações produzidas pelas avaliações em larga escala: alçou os indicadores institucionais ao mesmo patamar dos indicadores de rendimento. Essa guinada não é pouca e nos estimula a pensar um novo uso e um maior aproveitamento de tais informações.

Nesse sentido, o Inep vem construindo uma proposta que, ao invés de monitorar e avaliar a educação a partir de índice sintético, como o Ideb, tem como um de seus componentes um sistema de indicadores capaz de fornecer aos gestores públicos e à sociedade civil informações mais diversificadas e profícuas, que contemplem várias dimensões do processo educativo e possam retroalimentar as ações administrativas e pedagógicas desenvolvidas pelos governos subnacionais e suas secretarias de educação, bem como pelo próprio governo federal e pelo Ministério da Educação (MEC).

A proposta que vem sendo construída aponta que o Sinaeb é mais que um sistema de indicadores, possuindo outros componentes estruturantes, tais como: a pactuação com os entes federados para sua governança, coordenação e implementação; a construção de instrumentos de coleta de dados flexíveis, capazes de investigar novas problemáticas e de produzir dados para avaliar políticas e programas educacionais; e a utilização de outras bases de dados, tais como registros administrativos ou pesquisas do Instituto Brasileiro de Geografia e Estatística (IBGE), tanto para compor seus indicadores quanto para efetuar o monitoramento e as avaliações que estão no seu escopo.

Todavia, para essas vantagens ficarem mais claras, será preciso conceituar o que é um sistema de indicadores. Tiana Ferrer (1997) o define como um conjunto organizado e coerente de medidas, capaz de refletir adequadamente tanto uma realidade complexa quanto as relações existentes entre suas partes, apoiando-se em alguma teoria ou modelo do processo educacional. Para o autor, esse é o ponto crítico do sistema de indicadores, pois, apesar de superar a mera justaposição de informações, as teorias decorrentes das ciências sociais que o fundamentam ainda são frágeis no que diz respeito à integração da organização, do funcionamento e dos resultados do processo educacional.

Como a necessidade de conjugar outros indicadores com o Ideb foi apontada por diversos pesquisadores brasileiros da área educacional (Alves; Soares, 2013), o Inep, antes mesmo da proclamação do atual PNE, e com o intuito de complementar o retrato que esse índice e os resultados das suas avaliações da educação básica forneciam, começou a desenvolver um conjunto de novos indicadores (referentes ao nível socioeconômico das escolas, à complexidade da gestão escolar e à formação, regularidade e esforço dos docentes) e passou a utilizá-los, como indicadores de contexto, na divulgação dos resultados de suas avaliações externas. Também passou a divulgar, em um painel *online*, além dessas, outras informações e indicadores produzidos pelo Instituto, tais como taxas de aprovação, reprovação, abandono e distorção idade-série.

Apesar desses avanços, as informações disponibilizadas, do ponto de vista da gestão governamental, ainda são insuficientes, pois não basta divulgar resultados e contextualizá-los, sem oferecer um modelo teórico que permita compreendê-los de maneira integrada, que contemple outras dimensões da educação e que aponte os nós críticos do processo que devem ser conhecidos, corrigidos ou aperfeiçoados. Por esta razão, sugere-se que é preferível disponibilizar um sistema de indicadores que, além de abarcar as principais dimensões educacionais, diferencie as informações

referentes ao contexto, aos recursos, aos processos e aos resultados, pois, mesmo apresentando limitações, ele é capaz de oferecer um retrato mais amplo e coerente da realidade enfocada, ao articular os aspectos mensurados e, assim, potencializar sua capacidade interpretativa.

A título de exemplo e com fins didáticos, se desmembrarmos os indicadores que, direta ou indiretamente, estão contemplados pelo Ideb, notaremos existirem cinco (proficiência em leitura e em matemática, bem como taxas de aprovação, reprovação e abandono), que podem ser classificados como indicadores de resultado. Ao gestor, interessa conhecer também o contexto em que o processo de ensino e aprendizagem ocorre; daí a importância de utilizar um indicador de nível socioeconômico para caracterizar o público atendido. Faz-se necessário, além disso, apreender os recursos disponíveis, por meio de um indicador que aponte, só para ilustrar, o percentual de docentes com nível superior. Como o emprego desses recursos pode ser apreendido por indicadores de processo, é importante saber se os alunos estão tendo aulas com professores com a formação adequada para as disciplinas que lecionam. Mesmo sem entrar nas complexas relações que podemos estabelecer entre esses indicadores e na dificuldade de classificá-los dentro dessas quatro categorias, o modelo sistêmico adotado pode facilitar a compreensão do gestor sobre os pontos críticos da atividade educacional e que merecem atenção.

Assim sendo, consideramos que, para fortalecer o papel de diagnóstico, monitoramento e avaliação da educação básica, a construção de um sistema de indicadores, em contraposição a um índice sintético, deve ser um dos componentes do Sinaeb.

Testes para medir as habilidades ou produzir a deficiência?

A adoção de políticas públicas oficialmente orientadas pelos pressupostos da educação inclusiva, a partir dos anos 1990, coloca os testes em larga escala diante de uma contradição: como falar em avaliar a qualidade da educação e continuar excluindo desse processo os estudantes com deficiência (Schuelka, 2013), deixando de analisar as realidades educacionais de um expressivo contingente de alunos que são objeto de diferentes experiências educacionais e, não raro, enfrentam situações de vulnerabilidade social? Se a educação é um direito de todos, quem lida com testes em larga escala deve enfrentar o desafio de superar a lógica que presidiu a elaboração de instrumentos pensados para uma maioria sem deficiência e que, no melhor dos casos, tende a considerar apenas marginalmente as necessidades das pessoas com deficiência.

No âmbito dos testes em larga escala, há uma crescente preocupação com as necessidades específicas dos estudantes com deficiência (Katsiyannis *et al.*, 2007; Albus; Lazarus; Thurlow, 2015; Lazarus; Heritage, 2016). Nos Estados Unidos, eles foram reconhecidos como legítimos participantes dos testes federais e estaduais (Harr-Robins *et al.*, 2013). Admite-se que suas necessidades precisam ser consideradas nas avaliações educacionais em geral e, com maior razão, compreende-

se que as realidades educacionais vividas pelas pessoas com deficiência e as políticas educacionais a elas dirigidas precisam ser analisadas. Não por acaso, considera-se que os desafios relativos à garantia de uma educação inclusiva não podem ficar de fora do desenvolvimento de avaliações que vão além do desempenho dos estudantes submetidos aos testes, considerando-se outras dimensões da qualidade de maneira mais ampla, profunda e articulada.

Isso implica reconhecer a idoneidade da participação das pessoas com deficiência nos processos avaliativos e a legitimidade de se desenvolver instrumentos e processos que contemplem as suas necessidades educacionais específicas (Thurlow; Quenemoen; Lazarus, 2011). Ser devidamente considerado nas avaliações é um direito das pessoas com deficiência, assim como é um dever do Estado assegurar a elas os meios necessários para garantir-lhes uma participação com equidade, segurança e autonomia. Ao lado disso, como já destacado anteriormente, surge a necessidade de se elaborar indicadores mais sofisticados voltados não apenas a melhor contextualizar os seus desempenhos nos testes, mas, sobretudo, a considerar várias outras dimensões de qualidade que digam respeito aos desafios inerentes à garantia de uma educação efetivamente inclusiva e de qualidade. Especificamente quanto aos testes, além das barreiras comumente situadas na sua aplicação, é preciso identificar e eliminar aquelas geradas no processo de elaboração de itens.

No terreno das políticas públicas, é importante observar que a Convenção Internacional sobre os Direitos das Pessoas com Deficiência, aprovada pela Organização das Nações Unidas (ONU) em 2007 e ratificada pelo Brasil no ano seguinte (Brasil, 2008), lançou novas bases jurídicas para se conceber as deficiências e assegurar o direito à acessibilidade em todas as áreas, inclusive na educação. A rigor, o tratado dá continuidade e amplia o que foi definido na Classificação Internacional de Funcionalidade, Incapacidade e Saúde (CIF), aprovada pela Organização Mundial da Saúde (OMS) em 2001, que conferiu um caráter sociológico e político à deficiência (Diniz, Medeiros, Squinca, 2007). A partir desses marcos legais, a deficiência passou a ser juridicamente definida como resultante da interação entre pessoas com impedimentos (físicos, sensoriais, cognitivos, entre outros) e as barreiras (sociais, culturais, curriculares, comunicacionais, metodológicas, atitudinais, estruturais, ambientais) que lhes impedem de participar plena e autonomamente da vida social em igualdade de oportunidades.

Resultam disso uma ampliação da noção de deficiência e o gradativo abandono de tomá-la como sinônimo de lesão, desvio anatômico que mereceria cuidados médicos, e que tornaria as pessoas com deficiência alvo de intervenção clínica e reabilitação (Diniz; Santos, 2010; Mello; Nuernberg, 2012). A tônica, até então prevalentemente assimilacionista e assistencialista, é posta em crise, e a sua superação implica entender a deficiência como um produto da discriminação centrada na ideologia da normalidade e refletir sobre a necessidade de se redefinir pressupostos e a lógica das políticas. Em sintonia com essa nova compreensão, a noção de acessibilidade assume centralidade, passando a ser reconhecida como princípio e direito, cuja garantia é indispensável, tanto para o exercício de demais direitos quanto para assegurar a justiça social e a inviolabilidade da dignidade humana.

A concepção de deficiência que emerge da Convenção da ONU leva a convir que nada em relação à acessibilidade possa ser considerado definitivo. É necessário interrogar-se inclusive sobre medidas exitosas que pareçam não requerer novas providências. No terreno dos testes em larga escala, há que se investigar continuamente se barreiras ordinárias seguem presentes e se novas se interpuseram, da elaboração à aplicação dos testes. É igualmente indispensável avaliar as políticas de educação especial e averiguar o estado da arte da acessibilidade nas realidades escolares. Quanto ao conteúdo, na elaboração dos testes, ainda se desconsidera que estudantes com deficiência podem ter direito a adequações curriculares.

Caberia investigar os limites e as possibilidades de continuar a investir na melhoria do processo de adaptação de testes pensados exclusivamente para pessoas sem deficiência. Ao mesmo tempo, é preciso também perguntar se, ao invés disso, não seria mais apropriado construir instrumentos a partir de matrizes mais abertas e plurais, que também considerassem as necessidades e os direitos educacionais das pessoas com deficiência (e de outros grupos) e as particularidades de seus percursos escolares e currículos. Além disso, no terreno da aplicação, é preciso monitorar, avaliar e aprimorar a oferta de recursos e auxílios, bem como investigar outras metodologias de uso de tempo adicional. Há, ainda, que se pesquisar a elaboração de provas em diferentes formatos, dimensões, durações e com ritos de aplicação diversificados, segundo as deficiências e seus distintos graus. Os critérios de correção dos testes, sobretudo os escritos, e os cálculos das médias também precisariam ser objeto de análise permanente.

Como se pode notar, os direitos das pessoas com deficiência e os desafios para assegurar-lhes acessibilidade nos testes em larga escala exigem discussões que extrapolariam os limites deste artigo. Aqui, atemo-nos a chamar a atenção para alguns aspectos relacionados às necessidades das pessoas com deficiências sensoriais no tocante aos testes, esperando contribuir para ampliar a compreensão acerca da legitimidade do aprimoramento da acessibilidade neles.

No Brasil, nos testes em larga escala, é recorrente uma oferta de recursos de acessibilidade arquitetônica e mobiliária a estudantes com deficiência física restrita a rampas, salas de fácil acesso, banheiros adaptados e mesas para cadeira de rodas. Há casos em que se providenciam auxílio para a leitura ou a transcrição de respostas. É infrequente a preocupação com a qualidade dos recursos físicos ou profissionais ofertados. Raramente o uso de tecnologias assistivas é considerado, qualquer que seja a deficiência. Quando se trata de estudantes cegos ou com baixa visão, ofertam-se provas em Braille ou em macrotipos (ampliadas), eventualmente com itens adaptados e, não raro, sem se discutir a qualidade dessa produção, o teor de seus itens e suas eventuais adaptações. No caso de estudantes surdos ou com deficiência auditiva, crê-se comumente que basta providenciar um intérprete de Língua Brasileira de Sinais (Libras) ou apoio de leitura labial. Também é pequena a preocupação em relação à qualidade desse atendimento, e a necessidade de se adaptar itens para estudantes surdos ou deficientes auditivos é amplamente ignorada.

No caso das transcrições para o Braille, não se costuma disponibilizar gráficos, mapas ou figuras em formato tátil. Não raro, os itens têm suas imagens excluídas e

substituídas por descrições. Assim, as pessoas que atuam no auxílio à leitura perdem o acesso às imagens dos itens, embora as descrições possam ser indispensáveis para o seu trabalho. As respostas escritas são ditadas pelo estudante, que deve informar como grafar cada palavra. Para as pessoas com baixa visão, geralmente se oferecem provas ampliadas – que tendem a ser provas comuns fotocopiadas ou impressas em tamanho maior, onde as imagens são mantidas, sem descrição e sem qualquer intervenção adaptativa que garanta, por exemplo, o devido contraste e o acesso aos detalhes. Quando, mesmo com uso de lentes, não consegue ler, distinguir as imagens ou escrever as respostas, o estudante com baixa visão pode recorrer ao auxílio para leitura ou transcrição. Porém, tal como ocorre no caso do atendimento ao estudante cego, essa equipe não costuma possuir formação específica suficiente. Entre as várias providências possíveis, para os dois grupos de estudantes, a oferta de computadores com programas leitores de tela e outras tecnologias assistivas poderia oferecer maior autonomia ao aluno e contribuiria para reduzir os efeitos de adaptações insatisfatórias ou de equipes sem devida qualificação.

Vale notar que, enquanto os ouvintes realizam as provas redigidas em sua primeira língua, o português, as pessoas surdas ou com deficiência auditiva apenas raramente têm a oportunidade de fazê-las em Libras. A desigualdade de oportunidades já se coloca desde o início, a partir de uma barreira linguística. O caso se agrava quando se observam os limites estruturais da atuação dos intérpretes. Mesmo que, idealmente, fosse possível garantir, em cada local de aplicação, a presença de intérpretes devidamente certificados, experientes, com treinamento específico para o teste, fluente nas áreas de conhecimento ali avaliadas, autorizados a atuar não apenas na mediação da comunicação entre ouvintes e surdos e nem apenas traduzindo termos ou expressões, mas sim todo o teste, graves barreiras ainda estariam presentes e produzindo efeitos. Nessas condições ideais, como um intérprete traduziria bem uma prova sem ter tido antes acesso a ela e nem tempo para refletir, pesquisar e encontrar soluções exigidas pela tradução? A isonomia, princípio central de toda avaliação, ficará comprometida enquanto não se equacionar devidamente também este problema. Para isso, a solução mais ágil e equitativa parece ser disponibilizar, individualmente, videoprovas em Libras em meio digital, cuidadosamente produzidas, assegurando as devidas condições de acessibilidade e a mesma tradução (e qualidade) a todos os surdos. Nos casos das provas de Língua Portuguesa, elas podem ser elaboradas considerando as especificidades da formação escolar e dos processos de aquisição do português como segunda língua por parte dos estudantes surdos (Quadros; Schmiedt, 2006; Pereira, 2014).

Como se vê, a concepção de acessibilidade em um exame não pode ficar restrita a aspectos pontuais relativos à adaptação de itens e à aplicação da prova, pois envolve todas as suas etapas. Para garanti-la, todo o processo deve ser objeto de constante análise e passível de reformulação. A outra opção seria a de manter uma espécie de corrida de obstáculos que, ao afetar de maneira iníqua os desempenhos das pessoas com deficiência, reitera preconceitos, ensaja a constituição discursiva do corpo deficiente como anormal, inferior e incapaz. Em suma, a inexistência ou a insuficiência de processos avaliativos devidamente acessíveis

implica o aumento das probabilidades de os estudantes com deficiência apresentarem desempenho médio inferior ao dos demais, alimentando-se ulteriormente os preconceitos contra eles e incrementando-se o nível de vulnerabilidade e os riscos de sofrerem novas discriminações. Esse ciclo só pode ser rompido por meio da garantia da acessibilidade aliada à promoção do reconhecimento da diversidade e da diferença.

Considerações finais

No texto, apontamos como o Saeb, ao concentrar suas medidas no desempenho dos estudantes nos testes, apequenou a discussão sobre o processo educacional, ao considerar apenas uma de suas dimensões. Quando o Governo Federal desenvolveu o Ideb e o transformou no indicador de qualidade da educação, deixou de lado as várias outras dimensões que exercem influência sobre a educação escolar.

Uma alternativa a isso foi dada pelo PNE, quando este apontou para a necessidade de criação do Sinaeb, composto por indicadores institucionais e de rendimento, o que nos motivou a defender – tendo em vista facilitar o uso das informações disponibilizadas pelo Inep e aprimorar os processos avaliativos de redes e escolas de educação básica – a utilização de um sistema de indicadores, em vez de índices sintéticos, como o Ideb. Contudo, tal posicionamento não significa o abandono desses índices, nem deixar de lado outros instrumentos, apenas procura trazer para o primeiro plano sinalizadores mais específicos, que guardem uma relação mais direta e significativa com a realidade que procura mensurar, a fim de ampliar e precisar as informações relativas ao processo de escolarização utilizadas por gestores governamentais, profissionais da educação e sociedade em geral, para o acompanhamento e a (re)orientação de políticas e programas educacionais.

Ao lado disso, o fato de o PNE apresentar metas ambiciosas no campo das políticas inclusivas e de dispor sobre a criação de indicadores que permitam avaliar a educação inclusiva coloca o Sinaeb diante de um novo desafio: garantir a acessibilidade nos testes para além de adaptações pontuais e da oferta de alguns recursos de acessibilidade. A acessibilidade nos testes deve implicar todo o processo. Ademais, há a necessidade de se superar uma concepção anacrônica segundo a qual os estudantes com deficiência não precisariam ser considerados no âmbito dos processos de avaliação da educação. Tal anacronismo colide com a adoção de um modelo educacional inclusivo e com o conceito de direito à educação de qualidade para todos. A acessibilidade é um princípio e um direito a ser assegurado em todos os espaços e situações da vida. Nas avaliações, ela requer uma articulação entre políticas e pedagogias alicerçadas na cultura dos direitos humanos e o empenho contínuo para identificar e eliminar barreiras, desmantelar privilégios e fender regimes de verdades, de tal maneira a poder operar como elemento coadjuvante na promoção de equidade, equalização e justiça social. O direito à educação parece depender também disso.

Referências bibliográficas

- ALBUS, D.; LAZARUS, S.; THURLOW, M. 2012-13 *Publicly reported assessment results for students with disabilities and ELLs with disabilities*. Minneapolis: National Center on Educational Outcomes/ University of Minnesota, 2015. (Technical Report, 70).
- ALMEIDA, L. C.; DALBEN, A.; FREITAS, L. C. de. O Ideb: limites e ilusões de uma política educacional. *Educação e Sociedade*, Campinas, v. 34, n. 125, p. 1153-1174, dez. 2013.
- ALVES, M. T. G.; SOARES, J. F. Contexto escolar e indicadores educacionais: condições desiguais para a efetivação de uma política de avaliação educacional. *Educação e Pesquisa*, São Paulo, v. 39, n. 1, p. 177-194, mar. 2013.
- BAUER, A.; PIMENTA, C.; HORTA NETO, J. L.; SOUZA, S. Z. Avaliação em larga escala em municípios brasileiros: o que dizem os números? *Estudos em Avaliação Educacional*, São Paulo, v. 26, n. 62, p. 326-352, maio/ago. 2015.
- BRASIL. Decreto Legislativo n. 186, de 9 de julho de 2008. Aprova o texto da Convenção sobre os Direitos das Pessoas com Deficiência e de seu Protocolo Facultativo, assinados em Nova Iorque, em 30 de março de 2007. *Diário Oficial da União*, Brasília-DF, 10 jul. 2008. Seção 1, n. 131, p. 1.
- BRASIL. Decreto Legislativo nº 6.094, de 24 de abril de 2007. Dispõe sobre a implementação do Plano de Metas Compromisso Todos pela Educação. *Diário Oficial da União*, Brasília, DF, 25 de abr. 2007.
- BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). *Relatório do Sistema Nacional de Avaliação da Educação Básica*. Brasília, DF: Inep, 1992.
- BRASIL. Lei nº 13.005, de 25 de junho de 2014. *Diário Oficial da União*, 26 jun. 2014. Seção 1 (ed. Extra, p. 1-7. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l13005.htm>.
- DINIZ, D.; MEDEIROS, M.; SQUINCA, F. Reflexões sobre a versão em português da Classificação Internacional de Funcionalidade, Incapacidade e Saúde. *Cadernos de Saúde Pública*, Rio de Janeiro, v. 23, n. 10, p. 2507-2510, out. 2007.
- DINIZ, D.; SANTOS, W. (Org.). *Deficiência e discriminação*. Brasília: Letras Livres; Ed. UnB, 2010.
- HARR-ROBINS, J.; SONG, M.; HURLBURT, S.; PRUCE, C.; DANIELSON, L.; GARET, M. *The inclusion of students with disabilities in school accountability systems: an update*. Washington-DC: National Center for Education Evaluation and Regional Assistance, 2013.
- HORTA NETO, J. L. *As avaliações externas e seus efeitos sobre as políticas educacionais: uma análise comparada entre a União e os estados de Minas Gerais*

e São Paulo. 2013. Tese (Doutorado em Política Social) – Instituto de Ciências Humanas, Universidade de Brasília, Brasília, DF, 2013.

HORTA NETO, J. L. Avaliações educacionais e seus reflexos em ações federais e na mídia eletrônica. *Estudos em Avaliação Educacional*, v. 25, n. 59, p. 172-201, set./dez. 2014.

KATSIYANNIS, A.; ZHANG, D.; RYAN, J. B.; JONES, J. High-stakes testing and students with disabilities: challenges and promises. *Journal of Disability Policy Studies*, v. 18, n. 3, p. 160-167, Jan. 2007.

LAZARUS, S.; HERITAGE, M. *Lessons learned about assessment from inclusion of students with disabilities in college and career ready assessments*. Minneapolis: National Center on Educational Outcomes; Washington-DC: Arabella, 2016.

MELLO, A.G.; NUERNBERG, A. H. Gênero e deficiência: interseções e perspectivas. *Estudos Feministas*, Florianópolis, v. 20, n. 3, p. 635-655, set./dez. 2012. Disponível em: <<http://www.scielo.br/pdf/ref/v20n3/03.pdf>>.

OLIVEIRA, A. S. de. *Progressão continuada e outros dispositivos escolares: êxito e fracasso escolar nos anos iniciais do ensino fundamental*. 2014. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade de São Paulo, São Paulo, 2014.

36

PEREIRA, M. C. C. O ensino de português como segunda língua para surdos: princípios teóricos e metodológicos. *Educar em Revista*, Curitiba, Ed. Esp. n. 2, p. 143-157, 2014.

QUADROS, R. M.; SCHMIEDT, M. L. P. *Ideias para ensinar português para alunos surdos*. Brasília: MEC, 2006. Disponível em: <http://portal.mec.gov.br/seesp/arquivos/pdf/port_surdos.pdf>.

SAHLBERG, P. *Finnish lessons: what can the world learn from educational change in Finland?* New York: Teacher College Press, 2011.

SAHLBERG, P. How GERM is infecting schools around the world? 2012. *PasiSahlberg.com*. Disponível em: <<http://pasisahlberg.com/text-test/>>.

SCANDAR NETO, J.; SILVA, P. L. N.; JANNUZZI, P. M. Do que precisam os gestores de programas sociais? In: TELES, J. L.; SIGNORI, C. T. (Org.). *Educação na diversidade: como indicar as diferenças?* Brasília: Secretaria de Educação Continuada, Alfabetização e Diversidade, MEC, 2006. (Educação para Todos, 25).

SCHUELKA, M. J. Excluding students with disabilities from the culture of achievement: the case of the TIMSS, PIRLS, and PISA. *Journal of Education Policy*, v. 28, n. 2, p. 216-230, 2013.

SOARES, J. F.; XAVIER, F. P. Pressupostos educacionais e estatísticos do Ideb. *Educação & Sociedade*, Campinas, v. 34, n. 124, p. 903-923, set. 2013.

THURLOW, M.; QUENEMOEN, R.; LAZARUS, S. *Meeting the needs of special education students: recommendations for the race-to-the-top consortia and states*. Minneapolis: National Center on Educational Outcomes; Washington-DC: Arabella, 2011.

TIANA FERRER, A. Qué son y qué pretenden: indicadores educativos, política educativa. *Cuadernos de Pedagogía*, España, n. 256, p. 50-53, mar. 1997.
Disponível em: <<http://www.cubaeduca.cu/medias/pdf/1355.pdf>>.

João Luiz Horta Neto, doutor em Política Social e mestre em Educação pela Universidade de Brasília (UnB), é pesquisador do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Possui diversos capítulos de livros e artigos publicados sobre a avaliação educacional.

jlhorta@gmail.com

Rogério Diniz Junqueira, graduado em Comunicação pela Universidade de Brasília (UnB), especialista em Cooperação Internacional e Políticas de Desenvolvimento pela Universidade de Bolonha e doutor em Sociologia pelas Universidades de Milão e Macerata, é pesquisador do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) e membro da Editoria Científica da Revista Brasileira de Estudos Pedagógicos.

rogerio.junqueira@inep.gov.br

Adolfo Samuel de Oliveira, bacharel e licenciado em Ciências Sociais e mestre e doutor em Educação pela Universidade de São (USP), é pesquisador do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Atuou como gestor de sistema de ensino municipal e professor de sociologia no ensino médio.

adolfo.oliveira@inep.gov.br

Recebido em 2 de maio de 2016

Aprovado em 20 de maio de 2016