

Os cavalos também caem: Tratado das inconsistências do IDEB

Dalson Figueiredo ^a

Erinaldo Carmo ^b

Romero Maia ^c

Lucas Silva ^d

Resumo

Este artigo apresenta algumas incoerências identificadas na análise dos resultados do IDEB 2015, referentes ao ensino público. Metodologicamente, o desenho de pesquisa combina estatística descritiva e multivariada. As conclusões indicam que: 1) a alta taxa de aprovação interna infla artificialmente a magnitude do IDEB; 2) não existe correlação, em determinados estados, entre o desempenho do IDEB nas diferentes etapas da educação básica; 3) em alguns estados, não há correspondência entre o IDEB do ensino médio e o ENEM. Em conjunto, essas limitações comprometem a validade da inferência do Índice sobre a educação básica do país e justificam a imposição de dúvidas plausíveis sobre o próprio modelo de avaliação adotado pelo INEP.

Palavras-chave: Avaliação. Educação básica. Ensino público. IDEB.

1 Introdução

O título deste artigo faz referência ao filme de Sydney Pollack, *Os cavalos também se abatem*, um clássico do cinema americano, adaptado do romance de Horace McCoy, que relata o esforço desumano de dançarinos, em meio à Grande Depressão, ao participarem de uma maratona, testando seus limites e resistências em troca de um prêmio¹. Nas competições avaliativas (no filme e na escola), a busca pelo prêmio suprime o prazer da ação avaliada, e o resultado do processo suplanta o próprio processo. A analogia tem o sentido de discutir

¹ No original, *They shoot horses, don't they?* Ver: <https://www.youtube.com/watch?v=hWcc4PptU44>

^a Universidade Federal de Pernambuco. Recife, Pernambuco, Brasil.

^b Universidade Federal de Pernambuco. Recife, Pernambuco, Brasil.

^c Universidade Federal de Minas Gerais. Belo Horizonte, Minas Gerais, Brasil.

^d Universidade Federal de Pernambuco. Recife, Pernambuco, Brasil.

Recebido em: 05 nov. 2016

Aceito em: 11 jan. 2018

como as escolas e redes de ensino criam esforços para o *ranking*, como em uma corrida, onde pontua mais quem chega na frente, e, ao público, é divulgado o placar, como fruto de um espetáculo, e não o processo em si.

As avaliações externas, aplicadas em larga escala com provas padronizadas, são consideradas instrumentos apropriados de monitoramento das políticas públicas educacionais empreendidas pelos governos em nível federal, estadual e municipal, como observam Oliveira e Waldhelm (2016), desde que sejam consistentes e sirvam de referencial à realidade observada. No caso do Índice de Desenvolvimento da Educação Básica (IDEB), desde a sua criação, em 2007, esse instrumento de avaliação de larga escala tem recebido duras críticas ao seu formato e à sua utilização na elaboração e implementação de políticas públicas voltadas à educação básica do país. Essas críticas são, geralmente, reveladoras das inconsistências e incoerências verificadas no instrumento utilizado ou nos resultados por ele apontados.

Antes mesmo das contestações normalmente feitas ao instrumento e ao método, há críticas anteriores dirigidas ao formato de avaliação, baseado em testes padronizados com questões de múltipla escolha e tratamento quantitativo dos dados, o que não permite a real compreensão da realidade escolar por não ter elementos capazes de colher informações qualitativas dos estudantes, dos professores, da comunidade e da vivência escolar. Carmo et al. (2014) também lembram que há críticas à aplicação de uma avaliação em larga escala, no âmbito nacional, por essa pressupor que a base comum de conteúdos esteja em pleno vigor, isto é, que os alunos estão submetidos aos mesmos conteúdos e ao mesmo nível de aprendizagem, a ponto de serem, nas mais distintas realidades, aferidos por um mesmo instrumento.

De um lado, a Lei nº 9.394/96 (BRASIL, 1996) e o Plano Nacional de Educação (PNE) reafirmaram o papel da avaliação externa e tornaram imperativo o processo avaliativo, exigindo a sua universalização, como ressalta Werle (2011), de outro, o resultado pragmático da avaliação em larga escala tem sido o simples ranqueamento de escolas e redes, como afirmam Travitzki, Calero e Boto (2014), pouco contribuindo com a formulação de políticas públicas para a melhoria da educação. Bauer e Sousa (2015) ainda lembram que a construção de indicadores para a avaliação de resultados envolve questões valorativas, que expressam opções políticas e posicionamentos ideológicos. Mas, não nos estenderemos nesse debate, por não ser esse o foco do nosso estudo. Vamos limitar-nos à descrição e à análise de algumas incoerências encontradas nos resultados do IDEB, a partir dos números divulgados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), em setembro de 2016, referentes ao ano letivo de 2015, e apresentar como essas incoerências revelam a fragilidade do Índice.

Para isso, levantamos as principais críticas já feitas ao formato do Índice, em estudos anteriores, e procedemos a inclusão de três novas observações, a partir da exploração dos números oficiais do IDEB 2015, utilizando como metodologia a análise estatística de dados secundários, disponibilizados pelo INEP. Dessa forma, o presente artigo encontra-se organizado da seguinte maneira: na próxima seção abordamos o papel das avaliações padronizadas de desempenho como mecanismo de monitoramento da qualidade da educação; depois examinamos a composição do IDEB, especificamente; em seguida, apresentamos as principais críticas já feitas ao Índice da educação básica em estudos de diferentes pesquisadores da educação; na sequência, expomos as incoerências registradas na análise do IDEB 2015; e, por fim, apresentamos as conclusões deste trabalho de pesquisa.

2 O sentido da avaliação externa

A avaliação não deve ser confundida com a aplicação de provas ou testes, nem ter como objetivo a classificação, recompensa ou punição dos avaliados, mas, sim, segundo André (1996), no sentido da avaliação formativa difundida por Perrenoud, na qual ela deve permitir aos estudantes, professores e gestores a identificação de falhas e lacunas para ajudá-los a corrigi-las. Nessa mesma linha de pensamento, Neves (2010) defende que a avaliação deve ser compreendida como mais um processo de aprendizagem, e, não apenas, um mero instrumento de medida. Assim, só por meio da reflexão sobre o sentido do trabalho avaliativo realizado, é que se pode vislumbrar ações de melhoria e desenvolvimento do ensino.

Ainda de acordo com Neves (2010), a avaliação realizada dificilmente proporcionará oportunidades de desenvolvimento escolar, caso não sejam consideradas as informações sobre a situação atual e, seja o seu significado analisado para se encaminharem, de forma propositiva, ações concretas de melhoria. Essas ações podem ser emergenciais, principalmente quando dizem respeito às escolas públicas, que, por sua natureza social, precisam contemplar os estudantes das camadas socioeconômicas mais vulneráveis. E os estudantes mais pobres não podem ser penalizados na escola, quando essa reproduz as mesmas carências a que o seu público está sujeito fora dela.

Nesse sentido, entendemos que avaliar significa aferir determinado fato, situação ou circunstância para, através do seu resultado, empreender determinadas ações propositivas no sentido de conhecer, corrigir ou alterar o que se avaliou. Por isso, avaliar uma escola ou uma rede de ensino, seu sistema de funcionamento, a capacidade do seu corpo docente e o desempenho dos seus estudantes pressupõe uma proposta de melhoria ou aperfeiçoamento dessa mesma escola ou rede avaliada. Isso significa que a avaliação não se basta em si. Ela só tem sentido se proporcionar resultados posteriores.

Diferente da avaliação interna, que está focada nos sujeitos avaliados e em situações pontuais, a avaliação externa adquire uma abrangência bem mais ampla, tecendo uma avaliação de todo sistêmico, da estrutura e da conjuntura onde o ensino se desenvolve. Por isso, a avaliação externa não deve ser interpretada como resultado de desempenho do estudante, isoladamente, ou mesmo da escola e da rede de ensino, sem compreender os fatores intra e extraescolares que engendram, de forma direta ou indireta, as ações escolares. A comunidade onde a escola está localizada, o contexto social no qual os estudantes estão inseridos, tudo influencia o resultado da avaliação. Oliveira e Waldhelm (2016) relatam que, até mesmo, a relação de liderança do diretor e o clima organizacional escolar interferem no resultado dos testes dos estudantes.

Por isso, com o intuito de estabelecer uma estrutura de avaliação permanente para identificar, reformular e melhorar a educação básica do país, no início da década de 1990, o INEP criou e implantou o Sistema de Avaliação da Educação Básica (SAEB), que compreende a Avaliação Nacional do Rendimento Escolar, mais conhecida como Prova Brasil, e a Avaliação Nacional da Educação Básica (ANEB). Porém, ambas se limitam à medição do desempenho dos estudantes, sem englobar aspectos intra e extraescolares. As avaliações do SAEB têm como objeto comum o exame em língua portuguesa, com ênfase na leitura e interpretação de textos, e em matemática, com ênfase na resolução de problemas, por meio de provas com itens de múltipla escolha aplicadas aos estudantes do 5º e do 9º anos do ensino fundamental e do 3º ano do ensino médio.

O início dessas avaliações de larga escala, segundo Alavarse, Bravo e Machado (2013), esteve baseado na necessidade de monitorar o funcionamento das unidades escolares e das redes de ensino e fornecer aos seus gestores subsídios para a formulação de políticas públicas educacionais com foco nos resultados. Mas o sentido da avaliação se foi modificando com o tempo e passou a representar, também, a prestação de contas à sociedade e a cobrança por resultados aos professores e gestores escolares. Segundo Bonamino e Franco (1999), naquele momento inicial, houve resistência a esse tipo de avaliação, sendo as primeiras experiências marcadas por contestações, por conta de diferenças de objetivos e papéis no contexto das reformas educacionais implementadas a partir da segunda metade da década de 1990, com fortes influências das demandas do Banco Mundial. De acordo com Sousa e Oliveira (2010), esse modelo avaliativo sofreu, ainda, pressões políticas dos estados e municípios, tendo em vista que o debate da avaliação de sistemas públicos educacionais abrange um amplo leque de questões de natureza técnica, mas também de ordem política e econômica. Isso se justifica pelo fato de a avaliação externa estar diretamente vinculada às políticas públicas educacionais e orientar o repasse de verbas do governo federal aos projetos educacionais das redes públicas de ensino.

3 O Índice da Educação Básica

O IDEB é um indicador de abrangência nacional que permite a análise comparada entre escolas e redes de ensino. No entendimento de Padilha et al. (2012), o Índice serve de parâmetro para uma estimativa básica da situação geral da educação, sendo, ainda, um recurso de grande utilidade empregado como indicador estatístico condutor de políticas públicas, uma ferramenta de acompanhamento das metas do Plano de Desenvolvimento da Educação (PDE), que estabelece que, em 2021, o IDEB do país deve alcançar seis pontos como média nacional, resultado que será divulgado em 2022, com o simbolismo do bicentenário da Independência do Brasil, quando o país deverá atingir o patamar compatível à média dos países desenvolvidos.

Esse instrumento avaliativo emprega duas avaliações distintas: uma interna à escola, aferida pelo fluxo escolar, e outra externa, correspondente à média de desempenho nas provas do SAEB, que corresponde às notas obtidas nas provas de língua portuguesa e matemática. O INEP justifica esse formato de composição do Índice com a alegação de que essa combinação estabelece um equilíbrio à avaliação do IDEB, pois, se o sistema de ensino reprovar seus estudantes para que esses alcancem resultados melhores no SAEB, o fator fluxo será alterado, indicando a necessidade de melhoria do sistema. E se, de maneira oposta, o sistema de ensino aprovar seus estudantes sem qualidade, o resultado das avaliações externas também indicará a necessidade de melhoria do sistema.

O Índice é apresentado de forma simplificada, utilizando uma escala de 0 a 10 e pode ser representado na seguinte equação:

$$\text{IDEB} = (1/T) \times (\text{Nota SAEB})$$

Onde, T representa o tempo médio de conclusão de uma série, tratando-se do fluxo escolar, que observa o rendimento nas avaliações internas. A Nota SAEB corresponde ao resultado na avaliação externa (Prova Brasil para o ensino fundamental e Prova SAEB² para o ensino médio).³ Aqui, o conceito de rendimento é usado para sintetizar a aprovação dos estudantes de uma escola ou sistema de ensino. O indicador de rendimento usado pelo IDEB representa a razão entre o número de anos da etapa avaliada e o número de anos letivos que os estudantes

² Desde 2005 o nome do exame amostral do Sistema de Avaliação da Educação Básica foi mudado para Avaliação Nacional da Educação Básica (ANEB), mas, por sua tradição, o termo SAEB permanece sendo usado até hoje, mesmo nos documentos oficiais do INEP.

³ O INEP previa que a partir da edição de 2017 do IDEB, a avaliação do ensino médio passasse a considerar como avaliação externa o resultado da escola no ENEM, e não mais a Prova SAEB. Ver: <https://undine.org.br/noticia/ministerio-estuda-incluir-enem-no-calculo-do-ideb-para-universalizar-indice>

da escola levam para completar uma série. Isso quer dizer que o cálculo do IDEB dos anos iniciais do ensino fundamental considera a Prova Brasil realizada com os estudantes do 5º ano, e o fluxo escolar dos estudantes do 1º ao 5º ano; o cálculo dos anos finais do ensino fundamental considera a Prova Brasil do 9º ano e o fluxo do 6º ao 9º ano; e o cálculo do ensino médio considera a Prova SAEB do 3º ano, e o fluxo dos três anos que compõem essa etapa de ensino⁴.

Em uma situação ideal, como explicam Soares e Xavier (2013), onde não haja reprovação ou abandono, o valor do fluxo seria 1. Quer dizer, um ano letivo seria o tempo que os estudantes levariam para concluir cada série escolar. Entretanto, nas situações reais, há reprovações e abandonos. Por isso, a média do tempo que os estudantes levam para concluir uma série é sempre maior que 1. Dessa forma, como observa Mesquita (2012), os valores de rendimento escolar, registrados em $1/T$, resultam no tempo inverso de conclusão da etapa de ensino e, por isso, serão sempre menores que 1.

Em um exemplo prático, numa escola onde o tempo médio de conclusão das séries que compõem a etapa de ensino avaliada foi de 1,3 ano, e a média obtida na Prova Brasil pela última série dessa etapa foi 6,0, o IDEB será calculado da seguinte forma: $(1/1,3) \times 6,0 = 4,6$.

Ampliando o exemplo, considerando que a taxa de aprovação do primeiro ano do ensino médio de uma escola foi de 0,757, o fluxo será obtido na equação $1/0,757 = 1,32$. Isso significa que a média é de 1,32 ano para que o estudante dessa escola conclua o primeiro ano. Digamos que para os anos seguintes os valores foram, respectivamente, 1,31 e 1,27. A média de anos letivos para completar os três anos do ensino médio é de 1,3. Ou seja, $(1,32 + 1,31 + 1,27) = 3,9 / 3 = 1,3$.

Em cada unidade escolar, ao final do ano letivo, os estudantes matriculados são classificados em uma dessas três categorias: a) aprovados, para os estudantes que preencheram os requisitos mínimos de desempenho e frequência; b) reprovados, para aqueles que não preencheram esses requisitos; e c) abandono, para os que deixaram de frequentar a escola. A taxa de aprovação resulta da razão entre o número de aprovados e a soma dos estudantes das três categorias. Assim, quando a taxa de abandono é alta, tanto a taxa de aprovação quanto a de reprovação podem ser baixas, ao mesmo tempo. Por isso, algumas escolas adotam a estratégia de omitir as evasões, registrando formalmente os estudantes que abandonam a escola como transferidos, com isso, esses nomes deixam de ser computados como matriculados e assim são retirados dessa equação.

⁴ Para mais informações sobre a estimação do IDEB, ver: < http://download.inep.gov.br/educacao_basica/portaal_ideb/o_que_e_o_ideb/Nota_Tecnica_n1_concepcaoIDEB.pdf>.

Como sinaliza Fernandes (2007), quanto maior o abandono e a reprovação, maior será o valor de T, e o Índice resultará em um valor menor. Contudo, nessa equação, como observam Soares e Xavier (2013), o IDEB penaliza as escolas que mais reprovam, atribuindo-lhes valores mais baixos, e beneficia as que mais aprovam. Diante dessa realidade, e sofrendo cobranças de gestores políticos para a produção de resultados, como em uma corrida de cavalos, onde o mais importante é estar na frente, algumas escolas e redes de ensino desenvolveram táticas internas que são disfuncionais à educação, mas acabam por elevar seus resultados no IDEB, como veremos na próxima seção.

4 As principais críticas ao IDEB

Dentre as variadas críticas atribuídas ao IDEB, apresentadas por diferentes autores em diversos trabalhos acadêmicos já publicados até o presente, apontamos agora algumas que são mais difundidas. As críticas mais usuais e que abordaremos aqui são: 1º) há redes de ensino que fazem preparações nas escolas selecionadas para a Prova Brasil, eliminando assim o sentido de aleatoriedade da amostra e a confiabilidade na generalização dos resultados; 2º) a divulgação do Índice adquire um formato de classificação que expõe e cobra resultados de professores e estudantes; 3º) nem todos os estudantes que deveriam realizar as provas participam do exame, levantando a suspeita de que estudantes mais fracos são orientados a não participar do processo; 4º) a proficiência em matemática tem peso diferenciado, o que pode levar algumas escolas a investir no trabalho em sala de aula com a máxima valorização dessa disciplina, em detrimento das demais; 5º) o sentido da nota empregada no IDEB se diferencia do conceito de nota comumente empregado nas escolas; e 6º) o Índice compreende o desempenho dos estudantes, sem considerar as questões extraescolares que interferem na vida desses discentes, como as questões socioeconômicas e o capital cultural. Não nos alongaremos nas explicações, porém as referências informadas podem dar, ao leitor que desejar, o aprofundamento em cada questão.

1º) Uma estratégia adotada por algumas escolas para melhorar seus índices é a de concentrar esforços na preparação dos estudantes para a avaliação externa, seguindo a prática de testes simulados com questões da Prova Brasil de anos anteriores, preparando o estudante para se acostumar com esse tipo de teste de questões de múltipla escolha e preenchimento do cartão-resposta. Outra estratégia é voltada aos professores, com incentivo financeiro, para que esses preparem seus estudantes para o teste. Essas práticas foram apontadas como algo presente em escolas de São Paulo, segundo Alavarse, Bravo e Machado (2013), do Rio de Janeiro, segundo Mesquita (2012), e de Pernambuco⁵. A preparação

⁵ Ver: <https://periodicos.ufpe.br/revistas/cadernoscap/article/view/14978>

diferenciada em determinadas escolas viola o pressuposto da aleatoriedade da amostra, que fundamenta a confiabilidade da generalização estatística. Ou seja, os resultados simplesmente não têm validade para serem referendados à educação como um todo, visto que essa opção de formação para o teste impacta de maneira artificial o Índice.

2º) O objetivo primeiro da avaliação deveria ser o de servir de base para as políticas públicas de melhoria da qualidade da educação, mas, como sinalizam Travitzki, Calero e Boto (2014), a divulgação dos resultados do IDEB adquire uma abordagem centrada no ranqueamento e na exposição das escolas e dos seus professores e estudantes. A divulgação do Índice não considera os esforços locais e pouco colaboram para a construção de uma política de melhoramento das escolas, a partir dos resultados das avaliações, mas apenas torna pública a simples classificação dessas escolas e redes de ensino, mostrando a fragilidade das unidades deficitárias, o que pode produzir nessas instituições uma grave desmobilização dos professores, que não podem ser responsabilizados pela situação posta.

3º) Muitos estudantes, para os quais os testes são enviados, não realizam os exames. A Prova Brasil é encaminhada para as escolas selecionadas, dirigida a cada estudante matriculado, usando para isso a base do Censo Escolar. A portaria do Ministério da Educação sobre o cálculo do IDEB diz que o indicador é calculado desde que mais da metade dos estudantes da escola tenha comparecido. Soares e Xavier (2013) já registraram a ausência de 14% dos estudantes dos anos iniciais e 22% dos anos finais do ensino fundamental na Prova Brasil. São quantitativos expressivos que podem comprometer os resultados, e segundo esses autores, se forem correspondentes aos estudantes mais fracos, fazem com que os valores do IDEB das escolas com muitas ausências sejam superestimados. Isso porque algumas escolas, como estratégia, escolhem seus melhores estudantes para realizar a prova e, com isso, elevar o valor do IDEB. Sobre essa questão, Padilha et al. (2012) acrescentam que isso é um incentivo para que algumas escolas invistam mais nos estudantes mais competitivos, induzindo os “mais fracos” a não participarem do teste, que não tem tanto rigor, seriedade e fiscalização em sua realização.

4º) A proficiência em matemática influencia o resultado da avaliação mais que a proficiência em língua portuguesa. Isso porque o resultado da prova de cálculo e resolução de problemas é expressa em valores maiores que o resultado da prova de leitura e interpretação. Como consequência, a média usada na construção do IDEB é mais influenciada pela nota da prova de matemática. Na análise de Soares e Xavier (2013), o algoritmo de padronização não leva em conta esse efeito e o indicador de aprendizagem expressa, de forma equivocada, que os estudantes são

melhores em matemática do que em língua portuguesa. Essa representação não corresponde à realidade. Nas provas do ENEM 2015, por exemplo, a média dos estudantes das escolas públicas foi 439 em matemática e 479 em linguagem. Para Padilha et al. (2012) e Soares e Xavier (2013), esse equívoco metodológico pode ser aproveitado por algumas escolas como uma forma para melhorar o IDEB, procedendo um esforço maior no ensino de matemática, apenas, reduzindo o tempo e o esforço de outros conhecimentos básicos.

5º) A nota do IDEB não se associa ao conceito de nota usado na escola. De acordo com Mesquita (2012), mesmo sendo o IDEB expresso em valores simples, variando de 0 a 10, ele não se relaciona ao termo usado nas escolas, nem incorpora o conceito de nota empregado nas avaliações escolares. Enquanto a nota 6,0 é desejável no Índice, essa mesma nota é considerada baixa nas avaliações escolares. Essa falta de associação entre as notas dificulta a compreensão da avaliação da escola no IDEB, por esse instrumento não dialogar com o conceito popularizado na educação, mesmo utilizando a mesma escala.

6º) O IDEB se limita às avaliações interna e externa, sem considerar questões endógenas e exógenas que influenciam o resultado do desempenho dos estudantes, como lembra Avancini (2008). Sendo o IDEB um instrumento que identifica se os avaliados estão assimilando as informações fornecidas pelo sistema educacional, precisaria ponderar sobre os aspectos que têm impacto direto e indireto sobre a educação, como a valorização docente, a estrutura física e a localização da escola. As diferenças socioeconômicas e culturais também deveriam influir na análise explicativa da diferença de desempenho das escolas. Padilha et al. (2012) e Almeida, Dalben e Freitas (2013) orientam que a elaboração do IDEB precisa levar em conta o contexto em que a escola realiza o seu trabalho, pois omitir o nível econômico e o capital cultural da comunidade atendida pela escola na constituição do Índice é desconsiderar o verdadeiro trabalho desenvolvido nesse ambiente de ensino.

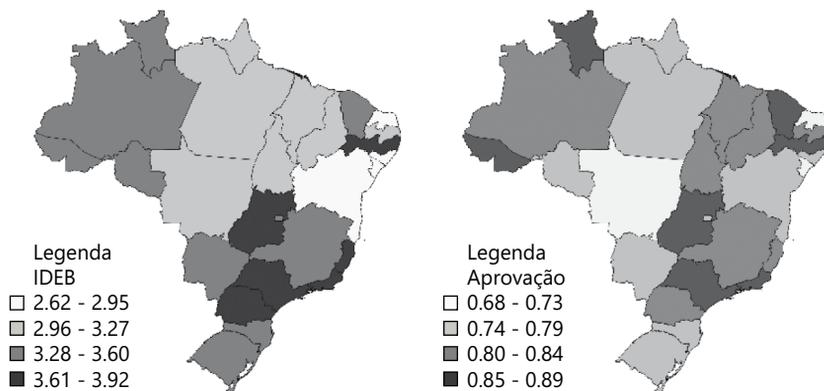
Além desses problemas acima relacionados, também merece destaque a observação de Padilha et al. (2012), de que o Índice não permite qualquer conclusão sobre a cobertura educacional das redes de ensino. Um município pode alcançar uma nota elevada no IDEB investindo na qualidade da educação em poucas escolas e para poucos alunos, atendendo apenas a uma parcela da população, excluindo os mais vulneráveis. Entendemos que diversos outros trabalhos também foram publicados, abordando outras falhas identificadas nas avaliações externas, mas não trazemos essas obras e autores aqui por limitações temporais e espaciais. E, ainda, há outras críticas, originadas a partir de nossa análise sobre os resultados do IDEB 2015, que serão apresentadas na seção seguinte.

5 Incoerências verificadas no índice

Outras questões referentes às inconsistências do IDEB serão apresentadas nessa parte do texto, essas, resultantes da análise dos dados divulgados no IDEB 2015. Aqui notificamos que a alta taxa de aprovação interna das escolas avaliadas infla artificialmente a magnitude do IDEB (Relação IDEB x Aprovação interna). Também registramos que, em alguns estados, não existe correlação no desempenho no IDEB entre os diferentes níveis de ensino (Relação IDEB x IDEB). E, ainda, identificamos que não existe correspondência entre o IDEB do ensino médio e o ENEM (Relação IDEB x ENEM). Isso mostra o quanto o Índice é incoerente em sua ação avaliativa, não apresentando consistência quando comparado às aprovações nas avaliações internas, nem quando relacionado às próprias avaliações do IDEB nos três níveis de ensino, tampouco, quando associado a outro instrumento avaliativo, como veremos a seguir.

5.1 IDEB x Aprovação interna

A nota do IDEB, como já mencionado anteriormente, é composta pela avaliação interna (fluxo escolar) e pela avaliação externa (Prova Brasil). Quando comparamos as duas avaliações, para cada um dos estados, percebemos o quanto elas não se relacionam. Observando o desempenho das redes estaduais de ensino, verificamos que, em algumas unidades da federação (AC, CE, GO, PE e RR), a avaliação interna apresenta níveis extraordinários de aprovação, entretanto esse conhecimento proporcionado aos estudantes, capaz de produzir elevado nível de aprovação interna, não se reflete nas avaliações externas. A Figura 1 apresenta como essa discrepância é visível no ensino médio, entre a avaliação do IDEB e as avaliações internas.

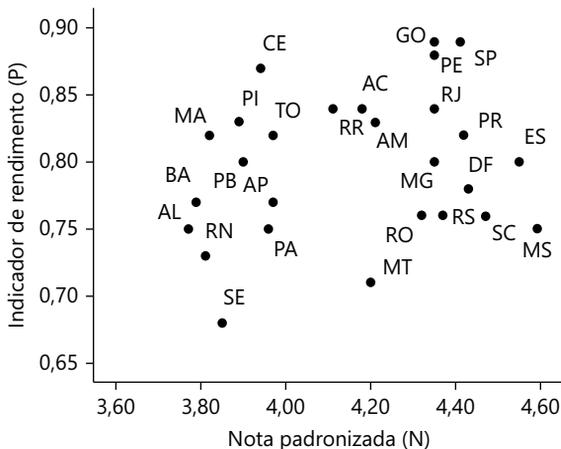


Fonte: INEP, 2016. Elaboração dos autores.

Figura 1. Desempenhos 2015 (IDEB e Aprovação interna).

As legendas mostram que alguns estados apresentam elevado desempenho na avaliação interna (fluxo escolar), representando baixas taxas de retenção e evasão, e não apresentam, paralelamente, o mesmo desempenho na avaliação externa (Prova Brasil). Constata-se, aqui, que essa estrutura do Índice, baseada em duas avaliações distintas (uma interna, feita na própria escola pelos professores de cada disciplina do curso, e outra externa, com testes padronizados em língua portuguesa e matemática), não estabelece o equilíbrio do processo, como pretende o INEP, pois nessa lógica de compensação, a escola pontua mais se aprovar mais internamente, independente do resultado na avaliação externa. Por outro lado, não faz sentido juntar variáveis para medir o mesmo fenômeno se elas não estão correlacionadas. Por isso, não se registra uma coerência entre os resultados do IDEB e as avaliações internas das escolas. Porém, ao inflarem a aprovação interna, essas escolas estrategicamente elevam suas colocações no IDEB. O Gráfico 1 ilustra a relação, no ensino médio, entre o indicador de rendimento (taxa de aprovação) e o desempenho médio padronizado de conhecimento.

Teoricamente, altos níveis de aprovação deveriam estar associados a elevados níveis de aprendizagem. Isso porque o estudante é aprovado quando internaliza, pelo menos, um mínimo necessário de conteúdo programático. A correlação entre a nota padronizada do IDEB (N) e o indicador de rendimento (P) mostra-se pequena ($r = 0,227$) e não significativa ($p\text{-valor} = 0,255$; $n = 27$). Ao se considerar os dados agregados por unidade da federação, a análise estatística indica que não é possível rejeitar a hipótese nula de independência entre as variáveis. Em termos



Fonte: INEP, 2016. Elaboração dos autores.

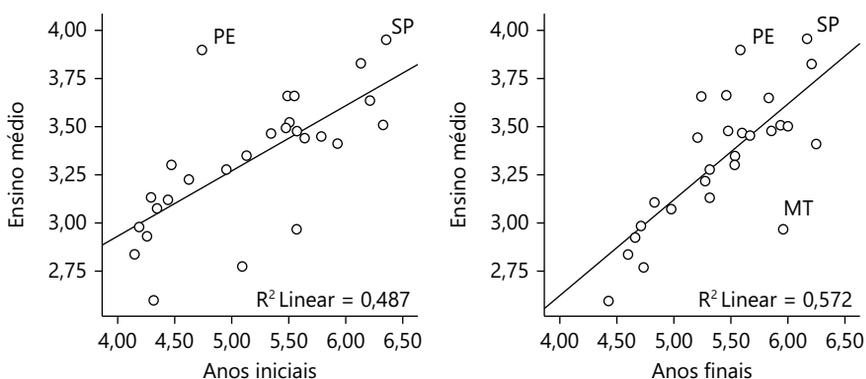
Gráfico 1. Correlação entre nota e rendimento (2015).

menos técnicos, isso significa que não existe relação linear entre o desempenho escolar dos estudantes e a taxa de aprovação, ou seja, um estado com alto grau de aprovação não apresenta, necessariamente, elevados níveis de desempenho. Isso pode representar uma falha na elaboração dos indicadores que formam o Índice, visto que há uma discrepância, em alguns casos, entre alto rendimento na avaliação interna e baixo desempenho na avaliação externa, ou vice-versa.

5.2 IDEB x IDEB

O IDEB avalia três etapas da educação básica. E, comumente, os estudantes das escolas públicas seguem uma trajetória escolar na rede pública. Não é tão frequente, em nosso país, a migração da rede privada para a pública, com algumas exceções. Portanto, é do mesmo grupo de estudantes que os resultados do IDEB falam, em suas três fases de estudos. Entretanto, ao considerarmos a associação do IDEB, em seus próprios componentes, verificamos outra incoerência. Enquanto a maioria dos estados seguem uma progressão linear, outras unidades da federação fogem à regra, com destaque para Pernambuco, onde o desempenho apresentado no ensino médio destoa do desempenho registrado nas etapas anteriores. O Gráfico 2 representa o padrão de associação entre as notas do IDEB por nível de ensino.

Aqui verificamos uma forte incoerência entre o desempenho de alguns estados no IDEB entre os diferentes níveis de ensino. Vejamos: a linha diagonal, no gráfico da esquerda, ilustra o valor esperado do IDEB no ensino médio, dada a nota do IDEB nos Anos iniciais. Algumas unidades da federação se distanciam desse eixo, sendo Pernambuco o estado com o maior resíduo padronizado positivo ($Z = 2,11$), tendo uma média de IDEB do ensino médio desproporcionalmente mais alta do



Fonte: INEP, 2016. Elaboração dos autores.

Gráfico 2. Correlação entre notas do IDEB por nível de ensino.

que o esperado. Similarmente, o gráfico da direita ilustra a nota esperada do IDEB no ensino médio, dada a nota no IDEB dos Anos finais. Dois estados (PE e MT) são os que mais se distanciam da linha, sendo o primeiro desses o que registra, outra vez, o maior resíduo padronizado positivo ($Z = 2,85$). Ou seja, Pernambuco apresenta uma média do IDEB no ensino médio desproporcionalmente mais elevada que a média do IDEB nos Anos finais. A Tabela 1 sumariza a correlação entre os valores do IDEB por etapa de ensino da educação básica.

Como esperado, existe uma correlação positiva entre a nota do IDEB nos níveis de ensino. Em particular, verificamos uma correlação forte entre os Anos iniciais e os Anos finais ($r = 0,815$; p -valor $< 0,0001$). Similarmente, a nota do IDEB, nos Anos iniciais, também está positivamente associada ao desempenho do ensino médio ($r = 0,698$; p -valor $< 0,0001$). E naturalmente, quanto maior a nota nos Anos finais, maior o IDEB do ensino médio ($r = 0,757$; p -valor $< 0,0001$). Isso quer dizer que, em média, quando uma rede tem alta performance em um nível de ensino, deve-se observar comportamento semelhante no outro nível.

Para garantir resultados ainda mais robustos, calculamos a distância de Mahalanobis,⁶ tendo como variáveis de interesse as notas do IDEB em cada nível (Anos iniciais, Anos finais e Ensino médio). A meta é identificar casos extremos na combinação dos valores das variáveis explicativas, o que a Estatística denomina de *outlier* multivariado. Colocando de forma mais simples, estamos interessados

Tabela 1. Correlação do IDEB entre níveis de ensino.

R p-valor	Anos Iniciais	Anos Finais	Ensino Médio
Anos Iniciais	1	0,815 0,0001	0,698 0,0001
Anos Finais		1	0,757 0,0001
Ensino Médio			1

Os resultados são significativos ao nível de 0,1%.

Fonte: INEP, 2016. Elaboração dos autores.

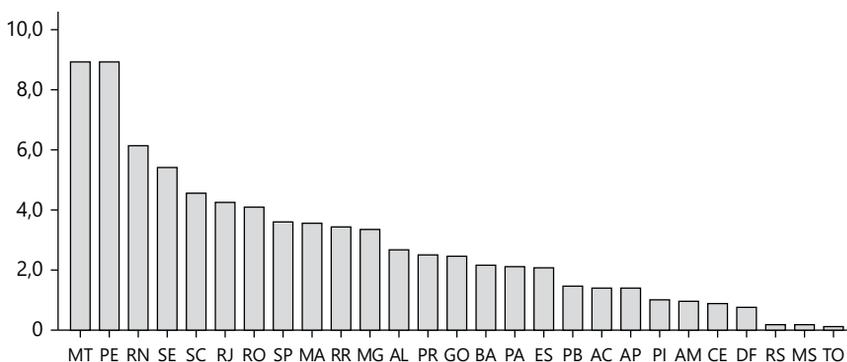
⁶ Algebricamente, essa medida pode ser representada da seguinte maneira: $D^2 = (x - m)^T C^{-1} (x - m)$. Onde x representa o vetor de uma determinada variável, m a matriz com a média de todas as variáveis, C^{-1} a matriz de covariância inversa entre x e m . O termo T indica que a matriz diferença $(x - m)$ é transposta. Em termos práticos, essa medida é amplamente utilizada em análise de conglomerados para classificar observações a partir do nível de similaridade entre os casos. A distância de Mahalanobis também considera a correlação entre as variáveis, o que elimina eventuais problemas de escala. Quanto maior a distância, maior é a diferença entre um determinado caso e o centro do *cluster*. Logo, menos provável é o seu pertencimento a um grupo ou distribuição.

na plausibilidade de combinação da nota do IDEB para os três níveis de ensino examinados. Em particular, qual é a probabilidade de um estado apresentar um desempenho mediano no ensino fundamental e ser excelente no ensino médio? O Gráfico 3 ilustra a variação dessa medida por unidade da federação.

Comparativamente, os estados do Mato Grosso e de Pernambuco aparecem como os casos mais destoantes da amostra. Em outras palavras, são observações incomuns ao se considerar simultaneamente os resultados dessas unidades da federação nos três diferentes níveis de ensino avaliados pelo IDEB. Esses resultados evidenciam como a prática de inflar as taxas de aprovação interna das escolas surte efeito sobre o resultado do IDEB, algo que não se projeta na avaliação do ENEM, como veremos no tópico seguinte.

5.3 IDEB x ENEM

Agora, vamos verificar a correlação entre os resultados de duas avaliações externas distintas, referentes ao mesmo grupo de estudantes.⁷ Vamos comparar os resultados do IDEB 2015 e do ENEM 2015⁸. Os dados oficiais demonstram que não há uma forte correspondência nesse caso. A visualização da Figura 2 permite a comparação do posicionamento de cada unidade da federação em ambas as avaliações: IDEB e ENEM, respectivamente.

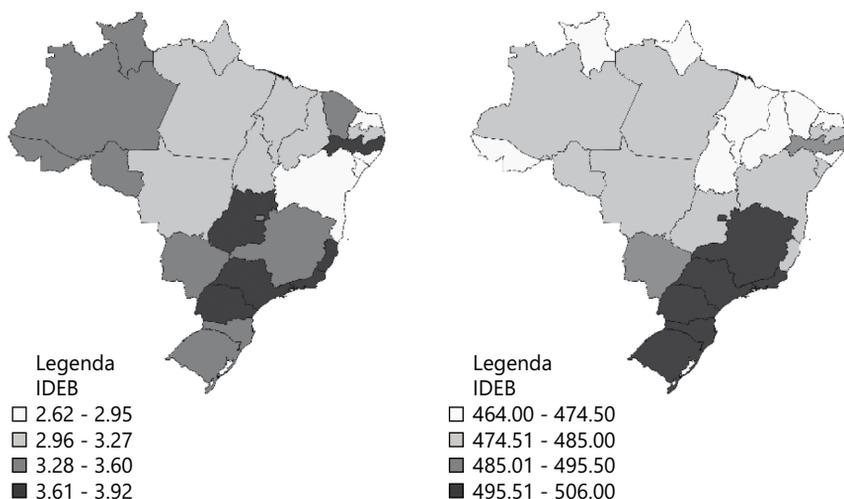


Fonte: INEP, 2016. Elaboração dos autores.

Gráfico 3. Distância de Mahalanobis por estado.

⁷ Nessa parte do estudo, tomamos como referência o IDEB do ensino médio, por ser essa uma etapa do ensino básico que participa de outra avaliação externa, o ENEM, permitindo que os resultados das duas avaliações possam ser comparados.

⁸ Para acessar os resultados do ENEM 2015, ver: http://download.inep.gov.br/educacao_basica/enem/enem_por_escola/2015/Sistema_Enem_2015.xlsm

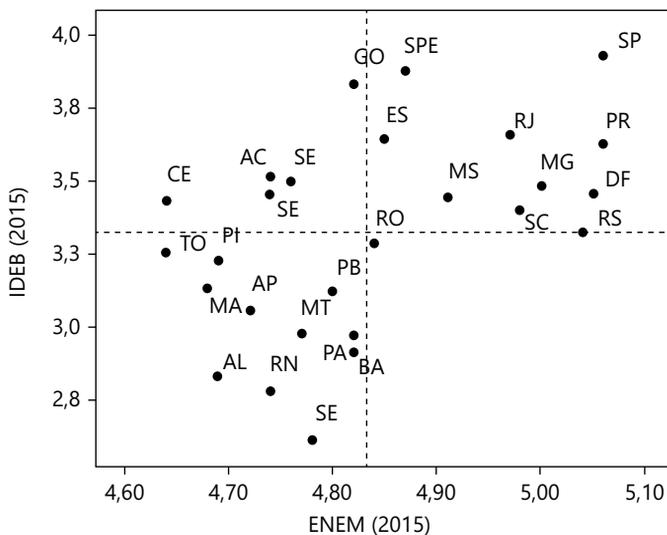


Fonte: INEP, 2016. Elaboração dos autores.

Figura 2. Desempenhos 2015 (IDEB e ENEM).

Em alguns casos, o nível da avaliação verificado no IDEB não corresponde ao verdadeiro nível da educação do estado. Percebe-se que algumas unidades da federação (AC, CE, ES, GO, PE e RR) apresentam bons resultados no IDEB que não se repetem no ENEM. Isso significa que o aparente desempenho no IDEB de algumas redes de ensino estaduais, verificado entre as escolas públicas de ensino médio, não reflete o resultado do ENEM, mesmo sendo ambas as avaliações correspondentes ao mesmo público. O Gráfico 4 ilustra a relação entre as avaliações do ENEM e do IDEB.

As linhas pontilhadas representam as médias dos dois mecanismos de avaliação. Nota-se que há uma correlação positiva, moderada e estatisticamente significativa, entre o rendimento no ENEM e a nota do IDEB para o ano de 2015 ($r = 0,482$; $p\text{-valor} = 0,011$; $n = 27$). Ou seja, em média, quanto maior o IDEB, maior a média no ENEM. Porém, certos estados fogem a essa regra, a exemplo de Sergipe, com baixíssimo desempenho no IDEB e próximo da média nacional no ENEM; Ceará e Tocantins, ao contrário, tiveram baixos desempenhos no ENEM e foram medianos no IDEB; e Pernambuco e Goiás, que despontam no IDEB, em comparação aos demais estados, porém, foram medianos no ENEM. São Paulo e Alagoas, em posições extremas, apresentam uma relação coerente entre ambos os instrumentos avaliativos, assim como os demais estados localizados próximos a essa linha diagonal (AP, MS, MT, PB, RJ e RO).



Fonte: INEP, 2016. Elaboração dos autores.

Gráfico 4. Correlação entre ENEM e IDEB (2015).

O ato de avaliar não se faz no abstrato, e, sim, em relação a um referencial, como pondera Santos (2016). Mas, no caso de alguns estados, com resultados discrepantes entre as distintas avaliações, a falta de referência da nota do IDEB torna a avaliação sem serventia para ser empregada como parâmetro na identificação de falhas e lacunas a serem corrigidas, como pontua André (1996), no sentido de uma avaliação formativa, e não produz reflexões sobre o trabalho realizado para que se possa, por meio de ações propositivas, apontar ações de melhoria do ensino, como afirma Neves (2010).

6 Conclusão

Este artigo teve como objetivo central apontar e discutir algumas incoerências identificadas na análise dos resultados do IDEB 2015, referentes ao ensino público no país. Para isso, foi realizada uma pesquisa combinando estatística descritiva e multivariada. Os resultados mostram que a elevada taxa de aprovação das escolas infla artificialmente a magnitude do IDEB. Identificamos, também, que há correlação entre o desempenho aferido pelo IDEB nas três etapas da educação básica, mas, em alguns casos, a exemplo do estado de Pernambuco, o desempenho no ensino médio destoa bastante do desempenho no ensino fundamental. E verificamos, ainda, que há uma correlação moderada entre o IDEB do ensino médio e o ENEM, porém, em algumas situações específicas, não foi verificada

essa correspondência, mesmo sendo essas duas avaliações referentes ao mesmo universo avaliado. Isso evidencia que tal correlação não se mostra mais alta em razão das estratégias pouco convencionais praticadas em alguns estados para mascarar o real nível de conhecimento do seu alunado.

Registramos que esses problemas aqui apontados, acrescidos aos demais já tratados por diferentes autores em diversos trabalhos anteriores, comprometem a credibilidade do Índice e levantam dúvidas plausíveis sobre o seu formato de estruturação, aferição e divulgação dos resultados. Por isso entendemos que o uso do IDEB, como medida única de verificação da qualidade das escolas e das redes de ensino, faz com que gestores escolares busquem elevar ao máximo os resultados nesse Índice, e isso tem sido feito, em alguns casos, de forma pouco adequada pedagogicamente, podendo resultar, como alertam Soares e Xavier (2013), em um sistema educacional disfuncional.

Temos aqui uma situação arriscada onde o instrumento de avaliação pode tornar-se mais valorizado que o próprio objeto (ou sujeito) avaliado, aproximando-se de uma maratona ou corrida de cavalos, onde os corredores são sacrificados e toda mobilização, dos participantes e da plateia, está voltada para a corrida, acima de tudo, até mesmo acima dos próprios corredores. E como lembra Campbell (1976), quando um indicador quantitativo é utilizado para a tomada de decisão, ele fica mais sujeito às manipulações e, com isso, sua própria existência compromete os processos que ele mesmo deveria monitorar.

Horses also fall: treaty of IDEB's inconsistencies

Abstract

This article analyzes the inconsistencies of the results of the IDEB 2015 related to public education. Methodologically, we use descriptive and multivariate statistics. The findings indicate that: 1) high pass rate artificially inflates the IDEB's magnitude; 2) there is no correspondence between the performance of the IDEB at different levels of education; 3) there is no correlation between the position in the IDEB and in the ENEM; Together, these methodological limitations compromise the validity of the published inference and justify the existence of doubts about the assessment model adopted by INEP.

Keywords: *Evaluation. Basic education. Public education. IDEB.*

Los caballos también se caen: tratado de las inconsistencias del IDEB

Resumen

Este artículo presenta algunas incoherencias identificadas en el análisis de los resultados del IDEB 2015, referentes a la enseñanza pública. Metodológicamente, el diseño de investigación combina estadística descriptiva y multivariada. Las conclusiones indican que: 1) la alta tasa de aprobación interna infla artificialmente la magnitud del IDEB; 2) no existe correlación, en determinados estados, entre el desempeño del IDEB en las diferentes etapas de la educación básica; 3) en algunos estados, no hay correspondencia entre el IDEB de la enseñanza media y el ENEM. En conjunto, esas limitaciones comprometen la validez de la inferencia del Índice sobre la educación básica del país y justifican la imposición de dudas plausibles sobre el propio modelo de evaluación adoptado por el INEP.

Palabras clave: *Evaluación. Educación básica. Enseñanza pública. IDEB.*

Referências

ALAVARSE, O.; BRAVO, M. H.; MACHADO, C. Avaliações externas e qualidade na educação básica: articulações e tendências. *Estudos em Avaliação Educacional*, São Paulo, v. 24, n. 54, P. 12-31, jan./abr. 2013.

ALMEIDA, L. C.; DALBEN, A.; FREITAS, L. C. O IDEB: limites e ilusões de uma política educacional. *Educação & Sociedade*, v. 34, n. 125, p. 1153-74, out./dez. 2013.
<https://doi.org/10.1590/S0101-73302013000400008>

ANDRÉ, M. E. D. A. Avaliação escolar: além da meritocracia e do fracasso. *Cadernos de Pesquisa*, n. 99, p. 16-20, nov. 1996.

AVANCINI, M. Afogados em números. *Revista Educação*, set. 2008.
Disponível em: <<http://revistaeducacao.uol.com.br/textos/136/artigo234432-1.asp>>. Acesso em: 20 ago. 2016.

BAUER, A.; SOUSA, S. Indicadores para avaliação de programas educacionais: desafios metodológicos. *Ensaio: Avaliação e Políticas Públicas em Educação*, v. 23, n. 86, p. 259-84, jan./fev. 2015.
<https://doi.org/10.1590/S0104-40362015000100010>

BONAMINO, A.; FRANCO, C. Avaliação e política educacional: o processo de institucionalização do SAEB. *Cadernos de Pesquisa*, n. 108, p. 101-32, nov. 1999. <https://doi.org/10.1590/S0100-15741999000300005>

BRASIL. Lei Nº 9.394, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. *Diário Oficial da União*, 23 dez. 1996.

CAMPBELL, D. T. Assessing the impact of planned social change. New Hampshire: The Public Affairs Center, Dartmouth College, 1976.
Disponível em: <<https://www.globalhivmeinfo.org/CapacityBuilding/Occasional%20Papers/08%20Assessing%20the%20Impact%20of%20Planned%20Social%20Change.pdf>>. Acesso em: 29 maio 2013.

CARMO, E. et al. Políticas públicas de democratização do acesso ao ensino superior e estrutura básica de formação no ensino médio regular. *Revista Brasileira de Estudos Pedagógicos*, v. 95, n. 240, p. 304-27, mai./ago. 2014.
<https://doi.org/10.1590/S2176-66812014000200004>

FERNANDES, R. Índice de Desenvolvimento da Educação Básica. Brasília, DF: INEP, 2007.

- MESQUITA, S. Os resultados do IDEB no cotidiano escolar. *Ensaio: Avaliação e Políticas Públicas em Educação*, v. 20, n. 76, p. 587-606, jul./set. 2012. <https://doi.org/10.1590/S0104-40362012000300009>
- NEVES, E. F. Estudo de uma escala de autoavaliação da prática docente: contributos para o desenvolvimento profissional. Faculdade de Psicologia e Ciências da Educação, Universidade de Coimbra, 2010. Disponível em: <<https://estudogeral.sib.uc.pt/handle/10316/15594>>. Acesso em: 22 ago. 2016.
- OLIVEIRA, A. C. P.; WALDHELM, A. P. S. Liderança do diretor, clima escolar e desempenho dos estudantes: qual a relação? *Ensaio: Avaliação e Políticas Públicas em Educação*, v. 24, n. 93, p. 824-44, out./dez. 2016. <https://doi.org/10.1590/S0104-40362016000400003>
- PADILHA, F. et al. As regularidades e exceções no desempenho no IDEB dos municípios. *Estudos em Avaliação Educacional*, v. 23, n. 51, p. 58-81, jan./abr. 2012. <https://doi.org/10.18222/ea235120121948>
- SANTOS, L. A articulação entre a avaliação somativa e a formativa, na prática pedagógica: uma impossibilidade ou um desafio? *Ensaio: Avaliação e Políticas Públicas em Educação*, v. 24, n. 92, p. 637-69, jun./set. 2016. <https://doi.org/10.1590/S0104-40362016000300006>
- SOARES, J. F.; XAVIER, F. P. Pressupostos eoucaionais e estatísticos do IDEB. *Educação & Sociedade*, v. 34, n. 124, p. 903-23, jul./set. 2013. <https://doi.org/10.1590/S0101-7330201300030001>
- SOUSA, S. M.; OLIVEIRA, R. Sistemas estaduais de avaliação: uso dos resultados, implicações e tendências. *Cadernos de Pesquisa*, São Paulo, v. 40, n. 141, dez. 2010. <https://doi.org/10.1590/S0100-15742010000300007>
- TRAVITZKI, R.; CALERO, J.; BOTO, C. ¿Qué información proporciona el Examen Nacional de Enseñanza Media (Enem) a la sociedade brasileña? *Revista Cepal*, n. 113, p. 163-81, ago. 2014. Disponível em: <<http://repositorio.cepal.org/handle/11362/36964>>. Acesso em: 12/ out. 2016.
- WERLE, F. Políticas de avaliação em larga escala na educação básica: do controle de resultados à intervenção nos processos de operacionalização do ensino. *Ensaio: Avaliação e Políticas Públicas em Educação*, v. 19, n. 73, p. 769-792, out. 2011. <https://doi.org/10.1590/S0104-40362011000500003>



Informações dos autores

Dalson Figueiredo: Professor do Programa de Pós-Graduação de Ciência Política da Universidade Federal de Pernambuco. Doutor em Ciência Política. Contato: dalsonbritto@yahoo.com.br

Erinaldo Carmo: Professor do Colégio de Aplicação da UFPE e Coordenador da Pós-Graduação em Políticas Públicas da UFPE. Doutor em Ciência Política. Contato: erinaldo.fcarmo@ufpe.br

Romero Maia: Analista de Planejamento e Estatísticas do IBGE. Especialista em Indicadores Sociais pela UFMG. Contato: romeromaia@gmail.com

Lucas Silva: Graduado em Ciência Política pela UFPE. Contato: lucas.eosilva@ufpe.br