

## A Evolução do Desempenho Educacional dos Jovens Brasileiros ao Final da Educação Básica: Acompanhamento de Gerações Sucessivas

Reynaldo Fernandes<sup>1</sup>  
Fabiana de Felício<sup>2</sup>  
David Saad<sup>3</sup>

### Introdução

Quando se observa a evolução temporal dos resultados do SAEB ou IDEB, um aspecto chama a atenção: a melhora significativa apresentada a partir de 2005 para os estudantes do final da primeira e, em menor grau, da segunda etapa do ensino fundamental não se verifica para os estudantes do ensino médio. Por exemplo, o aumento na pontuação do SAEB de matemática, entre 2005 e 2019, foi de 45,5 pontos para a primeira etapa do ensino fundamental, 23,5 para a segunda etapa do fundamental e apenas 6,0 pontos para o ensino médio. Ainda assim, o crescimento observado no ensino médio ocorreu todo entre 2017 e 2019: entre 2005 e 2017 a pontuação caiu 1,5 pontos.<sup>4</sup> Para se ter uma ideia da magnitude do crescimento observado para o ensino fundamental, podemos destacar que a pontuação observada ao final da primeira etapa do ensino fundamental em 2019 (227,9) é próxima da pontuação observada para os estudantes ao final da segunda etapa do fundamental em 2005 (239,5). Dada a escala vertical do SAEB (uma única escala para as séries finais das duas etapas do ensino fundamental e do ensino médio), observamos que os alunos do 5º ano em 2019 possuem uma proficiência em matemática que os alunos que finalizaram a etapa em 2001 só iriam obter no 9º ano (em 2005).

Tal cenário tem levado à percepção que o ensino médio seria o “grande gargalo” da educação básica brasileira. Na exposição de motivos (EM 084/21026/MEC) para justificar a reforma do ensino médio (Medida Provisória 746/2016 e Lei 13425/2017), após analisar o desempenho da etapa no SAEB/IDEB, lê-se que: “Resta claro, portanto, que o ensino médio

---

<sup>1</sup> Reynaldo Fernandes é professor titular do FEARP-USP e consultor associado da Metas Sociais.

<sup>2</sup> Fabiana de Felício é sócia-diretora da Metas Sociais.

<sup>3</sup> David Saad é Diretor Presidente do Instituto Natura.

<sup>4</sup> A partir de 2017, o SAEB passou a ser universal para as escolas públicas de ensino médio, com resultados divulgados por escolas. Alguns analistas desconfiam que tal mudança alterou o empenho com que os alunos respondem a prova e, desse modo, esse crescimento na pontuação pode não refletir um aumento do aprendizado. Um fenômeno conhecido na literatura como “inflação de notas”.

brasileiro está em retrocesso, o que justifica uma reforma e uma reorganização ainda este ano, de tal forma que, em 2017, os sistemas estaduais de ensino consigam oferecer um currículo atrativo e convergente com as demandas para um desenvolvimento sustentável”. Vale ressaltar que, por essa interpretação, o quadro de estagnação apresentado não seria limitado ao ensino médio, mas sim à toda educação básica. Afinal, os alunos do ensino médio hoje estariam finalizando a educação básica com o mesmo nível de aprendizado daqueles que finalizaram o ensino médio em 2005, de modo que todos os ganhos de proficiência obtidos no ensino fundamental teriam sido perdidos ao longo do percurso.

Esse tipo de diagnóstico sobre o ensino médio contrasta com um cenário mais positivo apresentado pelos indicadores de cobertura (números e taxas de matrícula e de conclusão). Por exemplo, a taxa líquida de matrícula do ensino médio cresceu 26,1 pontos percentuais entre 2005 e 2019: de 45,3 para 71,4% (Pnad e PnadC, IBGE).<sup>5</sup> Os indicadores apontam que a parcela de uma geração que conclui o ensino médio aumentou significativamente ao longo dos últimos 15 anos. Se uma parcela que antes não concluía o ensino médio (de modo geral, economicamente menos favorecida) agora está concluindo e, ao mesmo tempo, o desempenho médio dos concluintes não está piorando, então, isso poderia ser interpretado como um avanço, ao invés de um retrocesso ou uma estagnação educacional.

Quando se analisa os diagnósticos extremamente negativos sobre a evolução do desempenho das escolas de ensino médio, verificamos que eles, de modo geral, admitem três hipóteses em relação à forma de interpretar a evolução da pontuação do SAEB (no tempo e entre etapas) e, por consequência, do IDEB. A primeira diz que a evolução de 10 pontos na escala do SAEB representa o mesmo ganho de “conhecimento” ou “aprendizado” em qualquer ponto da escala.<sup>6</sup> Assim, se um aluno inicia o ano letivo com 200 pontos no SAEB de matemática e outro inicia com 300, dizemos que eles tiveram a mesma evolução de “conhecimentos” ou “aprendizado” se eles terminarem o ano com pontuações de, digamos, 230 e 330 pontos, respectivamente. Isso mesmo que o primeiro estudante esteja no 9º ano do ensino fundamental e o segundo no 3º ano do ensino médio.

---

<sup>5</sup> A taxa líquida de matrícula é calculada a partir da razão entre o número de matriculados de uma determinada faixa etária (15 a 17, no caso da matrícula líquida do ensino médio) na etapa adequada à idade (o ensino médio, nesse exemplo), e o total de jovens na faixa etária.

<sup>6</sup> Em termos técnicos, a hipótese é que a escala do SAEB é uma escala de intervalos em termos de conhecimentos ou aprendizado. Assim, elevar 10 pontos, entre 250 e 260, seria um ganho de aprendizagem equivalente a variação de 350 para 360.

A segunda hipótese refere-se ao esforço a ser empregado pela escola para elevar a proficiência média de seus alunos em 10 pontos é o mesmo para qualquer série da educação básica em avaliações que estão na mesma escala de proficiência. Assim, se as escolas elevam a pontuação média de matemática de seus estudantes em 20 pontos na segunda etapa do ensino fundamental e em apenas 10 pontos no ensino médio, poderíamos afirmar que, em média, as escolas de ensino médio teriam evoluído menos no ensino de matemática do que as escolas da etapa final do ensino fundamental.

Por fim, a terceira hipótese, a qual o presente artigo lida mais diretamente, refere-se à suposição de inexistência ou a minimização do impacto das mudanças no fluxo escolar (taxas de repetência, promoção e evasão) e cobertura escolar nas pontuações médias do SAEB.

Admitindo a validade dessas hipóteses, o fato de a pontuação do SAEB de matemática na primeira fase do ensino fundamental ter aumentado 45 pontos levaria a expectativa de um aumento similar nas demais etapas do ensino básico. Isso admitindo que a qualidade de ensino não se alterasse nessas outras etapas. Assim, a constatação de que o aumento tem sido bem inferior nas etapas finais seria uma evidência que o ensino estaria piorando o desempenho dos estudantes tanto na segunda fase do ensino fundamental quanto no ensino médio. Entretanto, as hipóteses descritas acima são extremamente fortes, colocando sérias dúvidas quanto à validade dessa interpretação.

Um ponto que vale destacar é que, em escalas verticais com base na TRI, o menor crescimento na pontuação entre séries mais elevadas não é um fenômeno novo e nem exclusivo do SAEB. Por exemplo, a tabela 1 apresenta as pontuações da amostra de normatização do *Comprehensive Test of Basic Skills* (CTBS), modelo U, desenvolvido pela CTB/McGraw-Hill, verticalmente escalado com base na TRI, e utilizando um modelo logístico de três parâmetros - o mesmo modelo utilizado no SAEB. Em ambas as disciplinas, o ganho em pontos cai drasticamente entre as séries mais baixas e as mais elevadas. Além disso, o desvio padrão das pontuações diminui na medida em que a pontuação média sobe. Assim, um aumento de 50 pontos em matemática na 1ª série equivale 0,32 desvio-padrão, enquanto que na 12ª série ele corresponderia a 2,5 desvios-padrão. Então, se a evolução temporal do desempenho das séries fosse avaliada com base no desvio-padrão da série (como usualmente é feito) o resultado seria bem diferente de quando a evolução é avaliada com base na pontuação do teste.

Tabela 1. Pontuação no Comprehensive Test of Basic Skills (CTBS/U) – 1981, Amostra de Normalização

Série	Leitura/Vocabulário			Matemática		
	Pontuação Média	Desvio Padrão	Variação	Pontuação Média	Desvio Padrão	Variação
1	488	85	-	390	158	-
2	579	78	91	576	77	186
3	622	65	43	643	44	67
4	652	60	30	676	35	33
5	678	59	26	699	24	23
6	697	59	19	713	20	14
7	711	57	14	721	23	6
8	724	54	13	728	23	7
9	741	52	17	736	17	8
10	758	52	17	739	16	3
11	768	53	10	741	18	2
12	773	55	5	741	20	0

Fonte: Yen (1986).

A Teoria de Resposta ao Item - TRI modela a proficiência (ou habilidade) como uma variável latente que, dados os parâmetros dos itens, determina a probabilidade de um estudante acertar um determinado item: quanto maior a proficiência, maior a probabilidade. Há uma controvérsia entre psicometristas se a escala de proficiência obtida pode ser interpretada como uma escala de intervalos.<sup>7</sup> Mesmo admitindo que a TRI produzisse uma escala de intervalos, isso não significaria que a medida de proficiência obtida case com a noção intuitiva que as pessoas possuem sobre os termos conhecimentos, habilidades e aprendizados escolares. Assim, interpretar a pontuação do SAEB com base na noção intuitiva que as pessoas têm sobre aprendizado escolar pode ser uma fonte de erros na interpretação da evolução da pontuação do SAEB ao longo do tempo.

As escalas de proficiência da TRI possuem como propriedade que o crescimento de proficiência requerido para aumentar a probabilidade de acerto de um item em um determinado montante é independente da dificuldade do item. Isso significa que o aumento de proficiência requerido para um estudante com baixa proficiência elevar, de 0,1 para 0,9, a probabilidade de responder corretamente um item fácil é o mesmo que o requerido para outro estudante de alta proficiência elevar a probabilidade de acerto, também de 0,1 para 0,9, só que para um item muito difícil.

<sup>7</sup> Como proficiência é uma variável latente, não existe qualquer meio físico capaz de testar essa propriedade. Para uma discussão dessa questão acessível para não psicometristas, ver Ballou (2009).

Alguém pode considerar que a noção de proficiência implícita nos modelos de TRI não casa com a sua noção intuitiva de aquisição de conhecimentos e habilidades, segundo a qual a aquisição de conhecimentos e habilidade seria maior no segundo caso (e.g. exigiria mais esforço e mais tempo do estudante). Sendo isso verdade, avançar na escala de proficiência iria se tornando mais difícil na medida em que atingimos níveis mais elevados de proficiência. Então, a compressão da escala em sua parte superior seria, ao menos em parte, uma consequência desse processo. E mais, caso ocorresse um aumento de esforço igual para todos os estudantes de todas as séries (ou uma melhora equivalente na qualidade de ensino), seria de se esperar que o aumento de proficiência fosse maior nas séries iniciais (onde a proficiência é menor) do que nas séries mais elevadas. Do mesmo modo, o ganho de proficiência nas séries iniciais, em virtude de um melhor ensino, tenderia a se reduzir na medida em que se vai avançando no sistema, supondo que as condições de ensino tenham se mantido constantes nas séries mais elevadas. Portanto, o fato de a pontuação evoluir menos nas séries mais elevadas não seria uma evidência suficiente para afirmar que os alunos estariam “aprendendo” menos nessas séries.

Esses argumentos põem em xeque as hipóteses 1 e 2, acima apresentadas. No entanto, uma melhora do aprendizado nas séries iniciais deveria elevar a proficiência nas séries mais avançadas, admitindo que a qualidade de ensino não tenha sido alterada para essas últimas. Se as escolas de ensino médio estão recebendo alunos com proficiências mais elevadas e estão entregando, ao final, alunos com o mesmo nível de proficiência do que antes, isso poderia ser interpretado como uma evidência da piora de qualidade dessas escolas. Para muitos, esse é o quadro apresentado pelo SAEB. Uma explicação alternativa seria que alterações de fluxo e cobertura têm feito com que uma parcela da geração que antes não frequentaria todo o ensino médio e, em alguns casos, nem parte dele, agora está concluindo o ensino básico. É esta questão que o presente artigo investiga, além de propor uma forma alternativa de acompanhar a evolução dos resultados da educação básica.

A composição daqueles que fazem a prova em determinada série pode estar mudando ao longo do tempo em decorrência de alterações nos padrões de repetência e evasão (ver Figura 2 na próxima seção). Na presença de mudanças nas variáveis de fluxo escolar, não é tão claro como deveríamos avaliar se o desempenho dos estudantes está melhorando ao longo do tempo, a partir da proficiência média. Podemos pensar em três alternativas: (i) comparar o desempenho de determinada série ao longo do tempo, independentemente da idade dos alunos (como é o caso do SAEB); (ii) comparar o desempenho de gerações sucessivas em determinada idade,

independentemente da série cursada (como é o caso do PISA); e (iii) comparar o desempenho de gerações sucessivas em determinada série, independentemente do ano que o aluno é testado.

Se todos os alunos ingressassem na escola na idade correta e não houvesse repetência nem evasão, essas três comparações seriam idênticas. No entanto, na presença de repetência e evasão, elas podem apresentar resultados muito distintos, e com padrões diversos de evolução, caso essas taxas sofram alteração ao longo do tempo. Por exemplo, uma redução nas taxas de evasão pode levar a um aumento no desempenho entre gerações sucessivas, avaliadas na mesma idade (frequentando ou não a escola), mas a uma redução no desempenho dos estudantes avaliados em determinada série, ao longo do tempo. Isso porque uma parcela de estudantes avaliados em determinada geração (que antes não seriam esperados a atingir a série em questão) agora são observados na série avaliada e esses estudantes possuem desempenho inferior àqueles que, nas condições anteriores, atingiriam a série em consideração.

Se o objetivo é verificar em que medida o ensino vem melhorando ao longo do tempo, o mais adequado seria avaliar o desempenho entre gerações sucessivas. Desse modo, as opções (ii) e (iii) seriam mais apropriadas do que a opção (i). Neste artigo, a evolução do desempenho educacional entre gerações sucessivas será avaliada com base em dois instrumentos: o Programme for International Student Assessment - PISA e o Índice de Inclusão Educacional - IIE (que será apresentado nas seções seguintes).

O PISA avalia estudantes de 15 anos e, em tese, independentemente da série que frequentam. Além da série apropriada para a idade, o PISA considera duas séries anteriores e uma série posterior. Se para a maior parte dos países da OCDE essa restrição não é relevante para os resultados, no Brasil, isso significa deixar de fora da avaliação uma parcela significativa de alunos de uma geração que, portanto, não é considerada no PISA: aquela formada por alunos com mais de três séries de atraso - 7º ano ou mais (até a edição de 2009 eram 2 séries de atraso – 7ª série ou mais, no ensino fundamental de 8 anos) ou que já evadiram a escola. Na primeira edição do PISA (2000) essa parcela era de 40% da geração apropriada. Esse número vem caindo (ver Figura 4 na seção 3) e, portanto, a proficiência da geração tende a ser superestimada, em especial nas primeiras edições, e a evolução do PISA tende a ser subestimada. O artigo propõe uma forma de corrigir esse viés.

Outra questão relacionada ao uso do PISA para acompanhar a evolução dos resultados educacionais da educação básica é que a idade dos participantes da avaliação é, no Brasil, a recomendada para cursar o primeiro ano do ensino médio, ou 9º ano do fundamental (a depender

do mês de nascimento). Assim, o PISA seria mais diretamente comparável ao SAEB da série final do ensino fundamental, ao passo que uma avaliação de jovens na idade recomendada para o final do ensino médio seria a mais recomendada para representar a evolução dos resultados da educação básica.

Para analisar o desempenho, entre gerações sucessivas, ao final do ensino médio propõe-se, então, o Índice de Inclusão Educacional – IIE, desenvolvido pela Metas Sociais a pedido do Instituto Natura. O IIE considera como incluída, do ponto de vista educacional, a parcela de jovens de uma mesma coorte de nascimento (que será denominada geração neste estudo) que termina o ensino médio proficiente (com pontuação em leitura e matemática acima de 300 pontos) e com, no máximo, um ano de atraso em relação ao recomendado.

O artigo é composto de 5 seções, além desta introdução. A seção 2 apresenta uma análise descritiva da evolução do SAEB, IDEB e matrículas do ensino médio. Na seção 3 realiza-se a análise da evolução do PISA, enquanto a evolução do IIE é realizada na seção 4. A seção 5 elabora uma análise de sensibilidade do IIE, que consiste em alterar alguns parâmetros utilizados na construção do indicador (intervalo de nascimento da geração e nota de corte para um jovem ser considerado proficiente). Por fim, a seção 6 conclui o artigo.

## Acompanhando a Evolução da Educação Básica por Meio das Avaliações por Série: o caso da média do SAEB e do IDEB do Ensino Médio

### Período foco das análises

Antes de iniciar a análise dos resultados do SAEB e IDEB como medidas para avaliar a educação básica ao longo do tempo, cabe justificar seleções de períodos de tempo que serão utilizados para tais análises, que muitas vezes não se justificam pela aparente disponibilidade de dados na série histórica. Se, por um lado, o SAEB e o IDEB estão disponíveis para o período de interesse deste estudo (2005 a 2019), por outro, alguns cruzamentos de dados muito relevantes para as análises não são possíveis para todo o intervalo.

Na Tabela 2, apresenta-se um breve descritivo das principais mudanças nas bases de dados utilizadas neste artigo que afetam as análises realizadas e, portanto justificam o fato de se priorizar



o período de 2005 a 2017, porém, nas diferentes abordagens do estudo, é necessário modificar o intervalo. Basicamente, 2005 é o ano em que se inicia a universalização do SAEB (inicialmente para o ensino fundamental) e o cálculo do IDEB. A partir de 2007, a coleta de dados pode aluno no Censo Escolar. Outro fato relevante é que a edição de 2019 do SAEB não coletou a informação de data de nascimento (ou sequer idade) inviabilizando algumas das análises mais relevantes neste ano<sup>8</sup>. Sempre que possível, os anos de 2005 e 2019 também serão apresentados, porém o foco da análise pode variar, a depender da seção e do objetivo.

Tabela 2. Breve descritivo das principais mudanças nas bases de dados empregadas no artigo que podem afetar os períodos apresentados e analisados

Base de Dados	Disponibilidade das informações necessários no período de interesse
SAEB / Inep	Criado em 1992, nas primeiras edições cada estudante fazia uma única prova (língua portuguesa ou matemática). Desde a universalização para a maior parte da rede pública (Prova Brasil), em 2005, cada participante realiza ambas e responde ao questionário de contexto sociodemográfico que contém a informação de idade e sexo, porém ambas foram suprimidas da edição de 2019.
Censo Escolar / Inep	A primeira edição foi em 1995 porém as informações por indivíduo (aluno e professores) só passaram a ser coletadas a partir de 2007.
PISA	Brasil participa desde a 1ª edição da avaliação organizada pela OCDE, no ano 2000. Houve mudanças nos critérios para participação, relativas à data de nascimento, e na data da aplicação da prova, que passou do 2º para o 1º trimestre em 2003.
PNAD e PNADc / IBGE	A PNAD teve suas primeiras coletas na década de 70, passou por muitas mudanças metodológicas ao longo dos anos e não é realizada em anos de Censo Populacional (1980, 1990, 2000 e 2022), porém a maior mudança foi em 2012, quando a PNADc (contínua) passou a ser implementada e veio a substituir a PNAD em 2016. Essa mudança afeta alguns recortes e estatísticas e outros não, o que é possível avaliar comparando o ano de 2015 para o qual as duas estão disponíveis.

## Evolução da Educação Básica a Partir dos Resultados do SAEB

Analisando a evolução da proficiência média do Brasil no SAEB da 3ª série do ensino médio observa-se que os jovens concluintes da educação básica em 2017 obtiveram a mesma média de proficiência que os concluintes de 12 anos antes em matemática, e um crescimento de apenas 11 pontos em língua portuguesa, o mesmo crescimento observado entre 2017 e 2019, biênio este em

<sup>8</sup> Para utilizar os dados de 2019 em análises que exigem a informação de ano de nascimento (ou idade) e necessário fazer alguns exercícios e/ou assumir algumas hipóteses que poderiam colocar em cheque algumas conclusões fundamentais para o artigo, portanto, elas serão utilizadas apenas na construção do IIE 2019, como descrito posteriormente.



que a média de matemática cresceu 8 pontos, muito aquém do progresso observado no final dos anos iniciais e finais do fundamental (ver Figura 1).

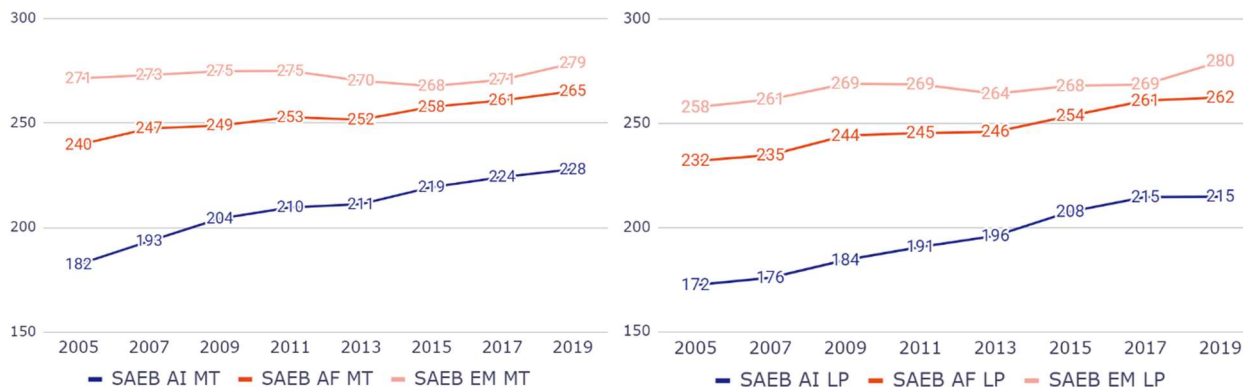


Figura 1. Proficiência média no Saeb, dos estudantes do 5º e 9º ano do EF e da 3ª série do EM, entre 2005 e 2019 – Língua Portuguesa e Matemática

Não é esperada uma elasticidade tão alta para a proficiência do ensino médio quanto é possível de se observar para os anos iniciais. O desempenho dos estudantes tem diversos determinantes e as melhorias ocorridas nos últimos anos da educação básica tendem a surtir efeitos pequenos diante de 10 ou mais anos de contexto educacional, em que têm grande peso o background familiar, o contexto social do território e as experiências escolares anteriores.

Entretanto, ainda que se considere uma defasagem de 4 anos entre as etapas de ensino (prazo em que uma geração deveria passar pelas séries avaliadas pelo SAEB), para a observação de reflexos dos avanços, compreender a estagnação do SAEB do ensino médio entre 2005 e 2017 fica ainda mais desafiador. Seguindo esse raciocínio, analisando os resultados de matemática, as crianças que frequentavam o 5º ano em 2009 obtiveram 22 pontos a mais na média do SAEB, relativamente àqueles que frequentavam em 2005. Em 2013, quando seria esperado que a maior parte daquela geração chegaria ao 9º ano, o aumento foi de 11 pontos, relativamente aos que concluíram em 2005 e apenas 3 pontos em relação aos que passaram por essa mesma série 4 anos antes. Em 2017, quando a geração que estava no 5º ano em 2009 deveria estar terminando o ensino médio, nenhum avanço foi captado pelo SAEB da 3ª série do ensino médio (seja na comparação com os concluintes de 2005, seja com de 2013).

A questão colocada neste estudo refere-se ao fato de a média do SAEB ser um indicador que representa a proficiência dos jovens que estão naquela série, naquele ano. Porém, ainda que

cada média seja uma boa representação do universo de jovens que frequentam aquela série naquele mesmo ano de aplicação da prova, o universo representado mudou ao longo desses 14 anos.

A principal evidência dessa mudança de composição está ilustrada na Figura 2, por meio da Taxa de Matrícula Líquida, que é calculada a partir da razão entre o número de matriculados no ensino médio com a idade adequada à etapa – 15 a 17 anos, e o total de jovens nessa mesma faixa de idade (incluindo os atrasados e os que estão fora da escola). Trata-se, portanto, de um indicador populacional, e sua evolução reflete as mudanças ocorridas no fluxo escolar no ensino médio, mas principalmente nas etapas anteriores, ou seja, a redução da repetência e da evasão no território de abrangência resulta em uma maior parcela dos jovens de 15 a 17 anos no ensino médio.

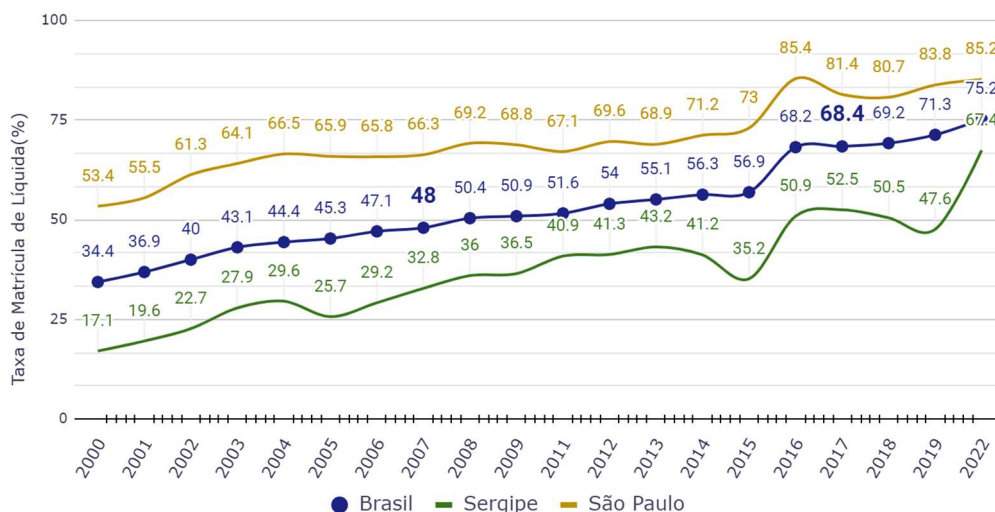


Figura 2. Evolução da Taxa de Matrícula Líquida para o Brasil, Sergipe e São Paulo, entre os anos 2000 e 2022 (Fonte: PNAD e PNADc / IBGE)

Ainda na Figura 2 estão apresentadas as séries históricas do indicador para o Brasil, Sergipe e São Paulo, a fim de ilustrar que a tendência do indicador é comum a todas as unidades da federação, entre os que apresentam os níveis mais baixos e mais altos nesse critério. A melhora desse indicador é expressiva, segundo os dados do IBGE. Se em 2007, 5 de cada 10 jovens de 15 a 17 anos estavam no ensino médio, em 2017 já eram quase 7 (aumento de 42,5%).

Um dos reflexos dessa mudança no SAEB está apresentado na Figura 3, Nela nota-se que a composição daqueles que fazem a prova do SAEB na 3ª série do ensino médio mudou significativamente entre 2005 e 2017. Por exemplo, a proporção de alunos com 19 anos ou mais saltou de 27,36% para 47,57% entre o início e o final do período. Tal mudança de composição pode ter impactos importantes na evolução das pontuações no SAEB.

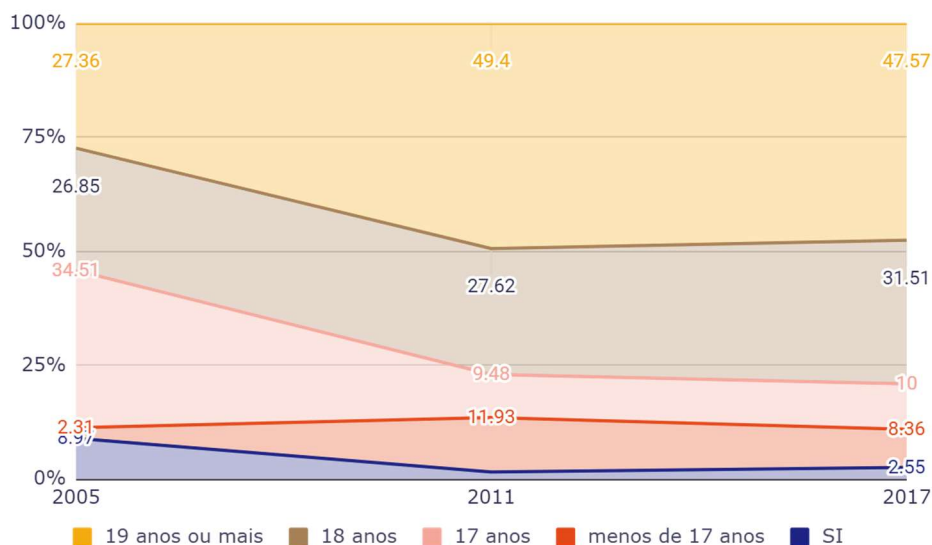


Figura 3. Distribuição dos Estudantes da 3ª Série do EM Segundo Alinhamento Idade-Série no Saeb entre 2005 e 2017

Se a evasão ou a defasagem idade-série não fossem fatores correlacionados ao desempenho escolar, provavelmente a melhora do fluxo escolar ocorrida neste período não levaria a ressalvas quanto ao acompanhamento da evolução da qualidade da educação básica por meio da proficiência média da avaliação por série escolar. Entretanto, há evidências de que essa hipótese não seja razoável. A redução da evasão nas etapas iniciais da escolaridade, aumenta a participação da parcela de jovens com atraso escolar na série avaliada, e como estes têm, em média, uma proficiência mais baixa, reduziriam a média geral se tudo o mais ficasse constante.

Na Figura 3 está apresentada a forma mais simples de exemplificar a relação entre a defasagem idade-série e a proficiência medida pelo SAEB 3ª série do EM – matemática. Nela vê-se a média do SAEB por faixa de idade, e é possível notar as diferenças consistentes ao longo do tempo entre elas. Os jovens com maior atraso têm, em média, proficiências mais baixas o que, juntamente à melhora no fluxo escolar nas últimas décadas, corrobora a leitura de que a mudança de composição do universo amostrado pelo SAEB, torna pouco recomendável o acompanhamento da evolução da qualidade da oferta de educação básica, por meio da média observada na avaliação.<sup>9</sup>

<sup>9</sup> Ainda que o desejável seja que os sistemas educacionais ofereçam uma forma de compensação em termos de ensino, que dê oportunidade aos estudantes com maior dificuldade de aprendizagem, de forma que todos as crianças e jovens

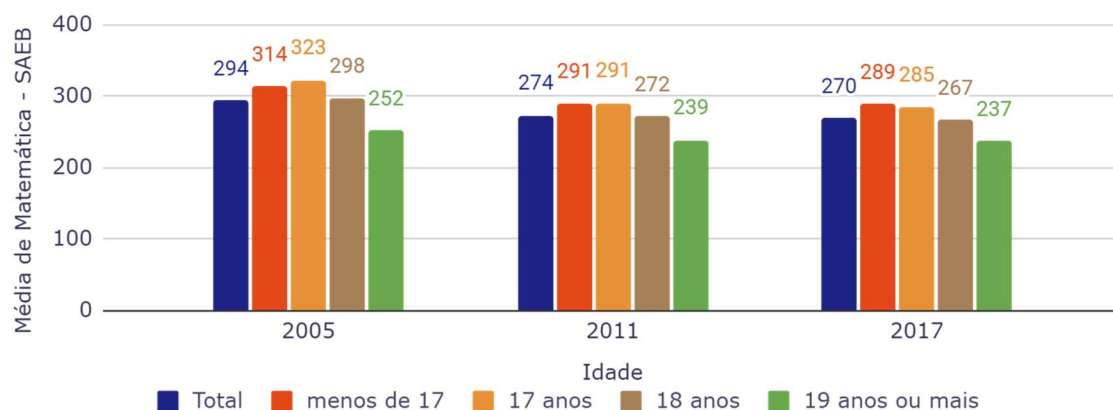


Figura 4. Média do Saeb por dos estudantes da 3a série do EM, segundo proficiência no Saeb entre 2005 e 2019

Não é possível observar, entretanto, as mudanças de composição decorrentes do aumento da taxa de promoção. Mas tal qual a redução da evasão, a consequência é a mudança de composição do universo avaliado: os jovens chegam antes – com menos repetências e portanto, menor idade, ao final do ensino médio. Sob a hipótese de que os jovens que passaram a ser promovido teriam proficiência média menor que a parcela que era promovida anteriormente, a média da avaliação na série final da educação básica subestimaria o progresso alcançado.

Cabe reforçar que não se trata aqui de um questionamento à qualidade da avaliação do SAEB como uma medida de desempenho dos estudantes, objetivo este que ela cumpre, empregando a metodologia que é referência para as principais avaliações internacionais. O que está sendo questionado é o uso de suas médias de forma isolada para avaliar a evolução da educação básica ao longo do tempo, que leva à conclusão, sem ressalvas, de estagnação ou até retrocesso da qualidade da educação no ensino médio durante a última década.

### Evolução da Educação Básica a Partir do IDEB

O Índice de Desenvolvimento da Educação Básica – IDEB, é calculado para as 2 etapas do ensino fundamental e para o ensino médio, a partir dos dados de proficiência média do SAEB e da estimativa de tempo médio necessário para conclusão da etapa em questão (derivada da taxa média de aprovação). Sendo assim, o IDEB mostra-se como um indicador mais recomendado para o acompanhamento da evolução da educação ao longo do tempo, relativamente ao SAEB.

---

pudessem atingir médias similares, independentemente de suas características e da defasagem idade-série, a realidade é que grande parte dessa desigualdade não é compensada, conforme observado nos dados apresentados.

Na Figura 5 apresenta-se a evolução do IDEB para as 3 etapas entre 2005 e 2019. Nela observa-se que no período de 2005 a 2017, foco da análise nesta seção, o índice cresceu 0,4 ponto, 12%, crescimento maior que do SAEB no mesmo período (que foi zero), porém ainda muito menor que a elevação de 52,6% para os anos iniciais, no mesmo período, e 34,3% dos anos finais, porém um pouco menos distante. No período completo, de 2005 a 2019, os crescimentos foram de 55,3% para os anos iniciais, 40% para os anos finais e 23,5% para o ensino médio.

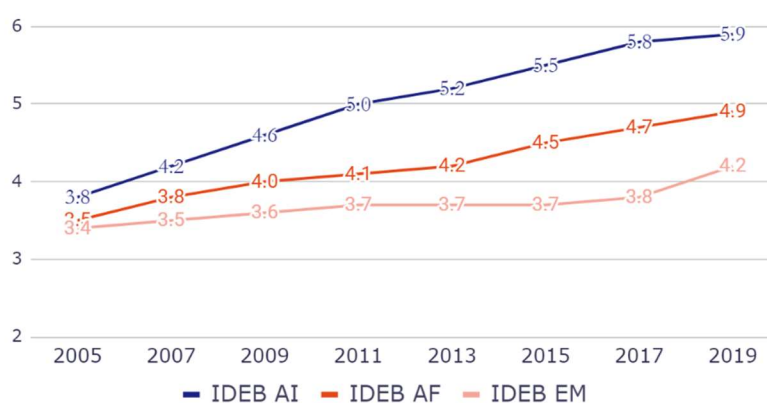


Figura 5. Evolução do IDEB entre 2005 e 2019 para os anos iniciais e finais do ensino fundamental e para o ensino médio (Fonte: INEP)

O melhor desempenho observado por meio da trajetória do IDEB do ensino médio relativa à trajetória do SAEB da 3ª série da mesma etapa, é resultado do fato de o IDEB considerar em parte a questão do fluxo escolar. Como mencionado anteriormente, o período de 2005 a 2017 foi marcado por redução da evasão e aumento da promoção no ensino médio e nas etapas anteriores. O IDEB capta parte dessa mudança, porém apenas uma parcela da mudança ocorrida no ensino médio, já que as taxas de aprovação contempladas no cálculo são desta etapa apenas. Desse modo, o índice capta uma parte da mudança de composição, mas ainda não reflete toda a evolução do desempenho da educação básica que se busca aqui.

## Acompanhando a Evolução da Educação Básica por Meio das Avaliações por Geração: o caso do PISA

Nesta seção, o objetivo é analisar os prós e contras de utilizar o PISA para acompanhar a evolução da educação básica, ou seja, um exame que avalia não uma série escolar, mas uma coorte de nascimento. À primeira vista, a mudança de composição, no molde da observada no SAEB, não seria essa questão para este tipo de avaliação, permitindo uma melhor comparabilidade dos resultados ao longo do tempo. Entretanto, questões de implementação e características típicas do Brasil, entretanto, exigem algumas ressalvas que serão discutidas na sequência.

O Brasil participou de todas as edições do PISA desde seu lançamento no ano de 2000. Os resultados mostram que o país teve um crescimento significativo de desempenho entre 2000 e 2012 e, a partir de então, uma ligeira queda é observada. Trabalhos que analisaram o período de crescimento do Brasil no exame investigaram em que medida as mudanças na data de realização da prova contribuíram para esse crescimento (Klein, 2011 e Carnoy et al., 2015). Eles concluem que tais mudanças explicam apenas uma parcela do crescimento nas pontuações observadas. Portanto, teria ocorrido uma melhora “genuína” de aprendizado entre gerações sucessivas no período considerado. Carnoy et al. (2015) chamam a atenção para o fato que o crescimento observado no PISA entre 2000 e 2006 não ser observado no SAEB dos anos finais do ensino fundamental em período similar.

As ressalvas que esta seção destaca quanto à análise da evolução da educação brasileira por meio do resultado do PISA são relativas às mudanças ocorridas no período de realização das provas, na mudança na delimitação da coorte de nascimento avaliada (alteradas em decorrência das mudanças de período de aplicação), e na diferença entre a coorte de nascimento completa e o universo amostrado pelo PISA no Brasil. Alguns estudos já investiram em isolar a primeira mudança, e esses resultados também serão discutidos a seguir.

Em Klein (2001), o autor analisa a questão da mudança da data da avaliação, propõe métodos de correção e conclui que a antecipação da realização da aplicação do PISA contribuiu para **superestimar o crescimento** da pontuação do país no exame. Entretanto, apresenta-se neste artigo evidências de que os métodos de correção propostos pelo autor podem levar à uma subestimação desse crescimento.

Além disso, o fato de uma parcela significativa de alunos da geração avaliada cursar séries anteriores a 7ª série/ 7º ano, e por tanto não ser amostrada conforme critério do PISA, pode levar

a uma **subestimação importante da evolução** da pontuação do Brasil no exame, na medida que a proporção de atrasados vem reduzindo ao longo dos anos. Assim como apresentado na seção anterior, trata-se de uma melhora do fluxo escolar que afeta a composição do grupo avaliado, o que em tese não deveria ocorrer para uma avaliação por geração, mas que ocorre devido à restrição de tomar como universo uma aproximação de uma coorte nascimento – apenas os que frequentam a escola a partir da 7ª série/7º ano.

Na Figura 6 apresenta-se a evolução da parcela da geração que não foi considerada para a amostragem do PISA. Essa parcela era de 25% em 2003, foi sendo reduzida até atingir 5% em 2018. Admitindo-se que, caso realizassem o exame, a parcela da geração de 15 anos não amostrada no PISA teria um desempenho inferior à mediada da amostra do PISA, a redução dessa parcela na amostragem tenderia a subestimar o progresso do Brasil no PISA ao longo do tempo. Isso ocorre porque, ao amostrar apenas os que frequentam a escola a partir da 7ª série/ 7º ano, estaria havendo uma superestimação do desempenho do Brasil no PISA. Na medida que a parcela da geração não amostrada se reduz, o viés de superestimação também se reduz e a evolução do desempenho da geração corrigida seria melhor do que a observada pelos dados oficialmente divulgados.

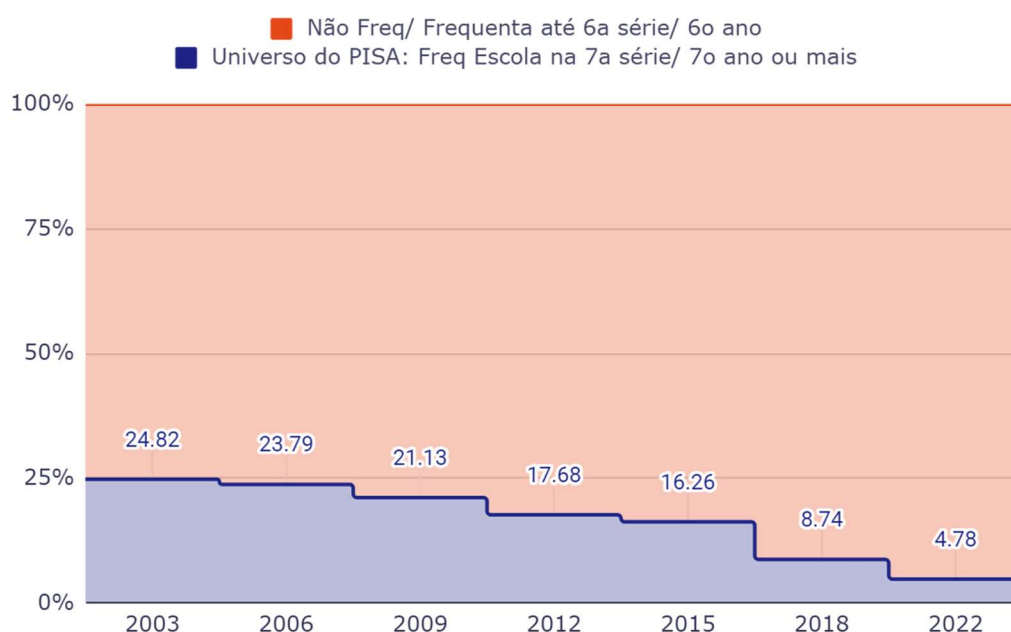


Figura 6. Jovens de 15 anos amostrados no PISA (nascidos no período recomendado e frequentando a escola a partir 7ª séries até 2009 e do 7º ano do ensino fundamental a partir de 2012)



Considerando o viés provocado pela proporção de alunos da geração não amostrada no PISA (aqueles com mais de dois anos de atraso e os fora da escola), a metodologia de correção adotada é aquela proposta por Fernandes e Natenzon (2003). Eles admitem que todos os estudantes atrasados da geração possuem desempenho inferior ao desempenho do estudante mediano (ou qualquer outro percentil de interesse) da geração, enquanto todos os adiantados possuem desempenho superior. Para nossos propósitos, vamos admitir que não há estudantes com mais de dois anos de adiantamento.

Então, definindo  $C$  como o percentil da distribuição de proficiência daqueles considerados pela amostra do PISA que é correspondente à proficiência mediana de todos os estudantes da geração (considerados e não considerados para a amostra do PISA), ele pode ser obtido conforme (1), onde  $p$  é a proporção da geração considerada para a amostra do exame. Com base nessa equação podemos obter uma estimativa da proficiência mediana dos alunos da geração, desde que obtenhamos estimativas para  $C$  e  $p$ . A estimativa de  $C$  é obtida diretamente dos alunos que realizaram o exame, enquanto a estimativa de  $p$  é obtida com base na PNAD. Evidentemente, esse procedimento só pode ser adotado caso a proporção da geração considerada para a amostra do PISA seja superior a 50%.<sup>10</sup>

$$C = \frac{0,5 - (1-p)}{p} \times 100 \quad (1)$$

Fernandes e Natenzon (2003) adotam esse procedimento para corrigir o SAEB dos anos iniciais do ensino fundamental e consideram que os alunos da geração com qualquer atraso possuem desempenho inferior à mediana. Da mesma forma, consideram que todos os alunos adiantados possuem desempenho superior ao mediano. Essas hipóteses têm sido contestadas. Por exemplo, Alves (2007) analisa essa questão com base no PISA-2000. Ela mostra que dentre os alunos com um ano de defasagem, 15% apresentam desempenho acima do percentil 60 da distribuição de desempenho dos alunos na série “ideal”. Para os alunos adiantados em uma série,

---

<sup>10</sup> Como apresentado na Figura 4, em 2003 a parcela da coorte de 15 anos amostrada pelo PISA é próxima de 69% e sobe para mais de 91% em 2018.

41% estão abaixo do percentil 60, indicando que os pressupostos assumidos por Fernandes e Natenzon (2003) podem ser excessivamente fortes.

Vale destacar que as hipóteses aqui adotadas são bastante mais brandas do que as adotadas por Fernandes e Natenzon (2003). Considera-se que os alunos com três ou mais anos de atraso possuem desempenho inferior ao estudante mediano, ao invés de estudantes com qualquer atraso. Assim, as hipóteses adotadas neste artigo são bastantes razoáveis, incapazes de produzir distorções importantes nas análises realizadas.<sup>11</sup>

Em relação às mudanças de data da realização da prova, Klein (2011) chama a atenção para o fato que a determinação da geração avaliada no PISA é dependente da data em que o exame é realizado. A geração a ser avaliada é composta por aqueles com idade entre 15 anos e 3 meses e 16 anos e 2 meses na data do início do teste. No Brasil, a data de realização do PISA teve duas antecipações entre 2000 e 2009 e tais antecipações teriam aumentado a idade média dos examinados. Por sua vez, essa elevação da idade, teria aumentado a série que, em média, os examinados se encontravam. Segundo Klein (2011), o aumento da série do aluno médio, em decorrência da antecipação do exame, explicaria parte significativa do aumento do desempenho do Brasil no PISA, que é observado no período.

A Tabela 3 apresenta o mês de realização das várias edições do PISA e a determinação da geração educacional correspondente. O critério adotado para definir a geração de referencia a partir de 2009 eleva a idade média dos participantes em torno de 6 meses, em comparação ao critério adotado em 2000. Isso aumenta a proporção de estudantes nas séries mais elevadas e, desse modo, contribui para a elevação do desempenho no exame.

---

<sup>11</sup> Analisando dados do SAEB, para os quais é possível observar jovens de uma mesma coorte de nascimento em duas séries diferentes, observa-se que em 2011, 94% dos que completavam 17 no mesmo ano e fizeram o Saeb 9º ano tiveram proficiência em matemática abaixo da mediana daqueles nascidos no mesmo ano que fizeram o Saeb 3ª série do EM, essa parcela é de 95% no caso de língua portuguesa. Em 2017, esse percentual foi de 89% em matemática e 90% em língua portuguesa. Esse elevado percentual reforça a hipótese de que a parcela da geração que tem mais de 2 anos de atraso escolar teria desempenho menor que a mediana, e pode-se dizer, ainda, que provavelmente trata-se uma subestimação do percentual total da coorte que não é amostrada no PISA, pois os dados utilizados referem-se a parcela defasagem idade série menor (2 anos) – os não contemplados no PISA são os que têm 3 anos ou mais de defasagem e os fora da escola (os que não frequentam mais escola representam uma parcela que varia de 14% a 4% entre 2005 e 2017).

Tabela 3. PISA no Brasil: Mês da Realização e Geração de Referência

Edição	Mês da Realização	Geração de Referência	
		15 anos completos em:	Nascidos entre:
2000	Outubro	30/06/2000	01/07/1984 e 30/06/1985
2003	Agosto	30/04/2003	01/05/1987 e 30/04/1988
2006	Agosto	30/04/2006	01/05/1990 e 30/04/1991
2009	Maio	31/12/2008	01/01/1993 e 31/12/1993
2012	Maio	31/12/2011	01/01/1996 e 31/12/1996
2015	Maio	31/12/2014	01/01/1999 e 31/12/1999
2018	Maio	31/12/2017	01/01/20002 e 31/12/2002

Fonte: INEP (2020)

Klein (2011) propõe dois procedimentos para corrigir a distorção causada pela mudança na data de realização do exame. O primeiro consiste em computar para cada ano a média de pontuação por série (da 7ª série ou 8º ano do fundamental ao 2º ano do ensino médio) e obter a média ponderada para cada ano utilizando a distribuição dos estudantes entre as séries de um ano fixo (e.g. a proporção dos alunos em cada série correspondente ao ano de 2000). O segundo considera apenas os estudantes cujo mês de nascimento seria selecionado em qualquer um dos três critérios adotados entre 2000 e 2009. A saber, os nascidos entre julho e dezembro.

Vale destacar que o primeiro procedimento admite que toda mudança na distribuição dos alunos entre as séries é decorrente da mudança na data de realização do exame. O próprio estudo de Klein (2011), entretanto, mostra que esse não é o caso: a proporção de alunos nas séries mais elevadas aumentou, entre 2000 e 2009, para estudantes com o mesmo mês de nascimento. Ao considerar que toda mudança na distribuição dos alunos nas séries é decorrente da mudança da data do exame, esse procedimento impõe um viés em sentido contrário: subestima o crescimento do desempenho ao longo do tempo.

Outro problema da antecipação da data do exame, não considerado por Klein (2011), é que os estudantes, em 2009, fizeram a prova com 4 meses a menos de aula (já descontada as férias de julho) do que os estudantes que fizeram a prova em 2000. Isso, por si só, tenderia a reduzir o desempenho dos alunos. Então, a antecipação na data de realização do PISA produz dois impactos, em sentidos opostos, sobre o desempenho observado. O primeiro, ao elevar a idade média dos participantes, aumenta a proporção de estudantes nas séries mais elevadas e, por esse motivo, tende a aumentar o desempenho observado. O segundo, ao diminuir os meses de aula do ano letivo no momento do exame, tende a reduzir o desempenho dos estudantes. É difícil dizer qual dos dois

impactos é mais forte. Caso sejam de intensidade similares, eles se cancelariam e nenhum procedimento de correção seria necessário.

Nesta seção, analisa-se a evolução do desempenho do PISA sem correção para a data do exame ou para a o intervalo de nascimento avaliados e um segundo, corrigindo o viés de participação. Para isso foram calculados o percentual da geração de 15 anos fora da escola ou frequentando a escola em séries anteriores à 7ª série (2003 e 2006) ou 7º ano (2012, 2018 e 2022) em cada ano de PISA, a partir da PNADc, e verificada qual seria a mediana se esse percentual tivesse proficiência inferior à mediana observada. Na Figura 7 apresentam-se os resultados para o estudante mediano da geração, com e sem correção para o viés de participação.

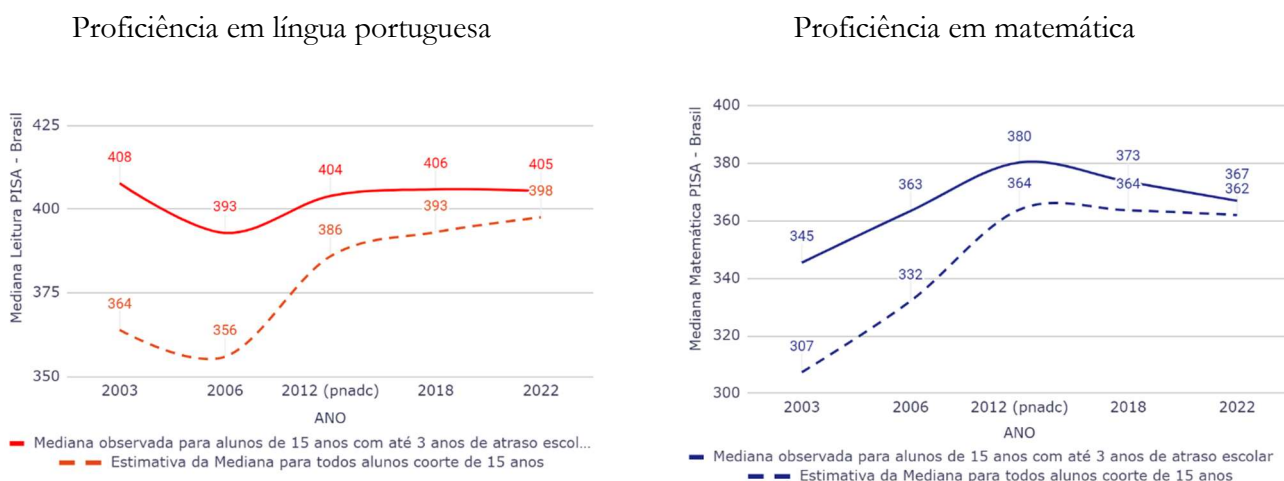


Figura 7. Mediana observada no PISA, entre os estudantes de 15 anos, frequentando a partir do 7º ano do fundamental, 2006, 2012 e 2018, e estimativa da mediana caso todos os jovens de 15 anos participassem do exame

A linha sólida mostra a evolução do desempenho mediano daqueles considerados para a amostra do exame, enquanto a linha tracejada considera a correção para o desempenho mediano da geração. A correção para o viés de participação traz duas novidades: uma ruim e outra boa. A ruim é que o desempenho do Brasil no PISA é pior do que o divulgado pela OCDE. A boa é que o crescimento no desempenho, se observada toda a geração de 15 anos, teria sido melhor do que o oficialmente apresentado e estaria cada vez mais próxima da observada.

## Acompanhando a Evolução da Educação Básica Utilizando o Índice de Inclusão Educacional - IIE

O Índice de Inclusão Educacional (IIE) é um indicador elaborado pela Metas Sociais a pedido do Instituto Natura e consiste na proporção de indivíduos de uma geração que finaliza o ensino médio com até um ano de atraso, e é proficiente em matemática e língua portuguesa. Deste modo, permite avaliar a evolução do desempenho educacional, ao final do ensino básico, entre gerações sucessivas. A dificuldade para realização dessa análise é que parte dos dados necessários para a operacionalização do IIE não estão disponíveis e, assim, precisamos de algum procedimento de imputação. Então, antes de apresentar a evolução do IIE, será apresentada a metodologia desenvolvida para viabilizar seu cálculo.

### Metodologia

Admitindo que os dados contidos na Tabela 4 estivessem disponíveis, o cálculo do IIE seria direto:  $IIE = A_{12} + A_{22} + A_{32}$ , onde o termo  $A_{ij}$  significa a proporção de indivíduos da geração que terminou o ensino médio na condição de atraso  $i$  (adiantado, em linha, 1 ano de atraso, 2 anos de atraso e 3 anos de atraso ou mais) e na condição de proficiência  $j$  (não proficiente e proficiente). Entretanto, não existe nenhuma fonte de dados que nos permita obter essas informações diretamente.

É importante destacar que os dados da Tabela 4 podem ser extraídos do cruzamento dos dados das Tabelas 5 e 6:  $B_{ij} \times C_i = A_{ij}$ . Então, o trabalho consiste em encontrar uma forma de preencher essas tabelas.

Tabela 4. Distribuição da Geração nos Estados Educacionais ao Final do Ensino Médio

	Não Proficiente	Proficiente	Total
<b>Adiantados</b>	$A_{11}$	$A_{12}$	$T_{1(.)}$
<b>Em Linha</b>	$A_{21}$	$A_{22}$	$T_{2(.)}$
<b>Atrasados 1</b>	$A_{31}$	$A_{32}$	$T_{3(.)}$
<b>Atrasados 2</b>	$A_{41}$	$A_{42}$	$T_{4(.)}$
<b>Atrasados 3+</b>	$A_{51}$	$A_{52}$	$T_{5(.)}$
<b>Total</b>	$T_{(.)1}$	$T_{(.)2}$	1

Tabela 5. Distribuição dos Indivíduos da Geração em Determinada Condição de Atraso por Nível de Proficiência ao Final do Ensino Médio

	Não Proficiente	Proficiente	Total
<b>Adiantados</b>	$B_{11}$	$B_{12}$	1
<b>Em Linha</b>	$B_{21}$	$B_{22}$	1
<b>Atrasados 1</b>	$B_{31}$	$B_{32}$	1
<b>Atrasados 2</b>	$B_{41}$	$B_{42}$	1
<b>Atrasados 3+</b>	$B_{51}$	$B_{52}$	1

Tabela 6. Distribuição da Geração Segundo a Condição de Atraso ao Final do Ensino Médio

Alinhamento Idade Série	Quantidade ou proporção de jovens
<b>Adiantados</b>	$C_1$
<b>Em Linha</b>	$C_2$
<b>Atrasados 1</b>	$C_3$
<b>Atrasados 2</b>	$C_4$
<b>Atrasados 3+</b>	$C_5$
<b>TOTAL</b>	1

Para se obter os dados da Tabela 5, foi utilizado o SAEB, e foram considerados proficientes aqueles com pontuações acima de 300 em matemática e leitura. Considerando a geração que deveria finalizar o ensino médio no ano  $t$ , utilizou-se o SAEB do ano  $t$ . Note que o SAEB do ano  $t$  só permite obter as informações dos que se encontram “em linha”: os adiantados e os atrasados não fazem o SAEB no ano que seria o adequado para os membros da geração de interesse finalizar o ensino médio. Uma vez que o SAEB só é realizado a cada dois anos, os adiantados e aqueles com 1 ano de atraso não fizeram o SAEB. Então, para o preenchimento da Tabela 5, considerou-se, como uma aproximação, os dados das gerações vizinhas, obtidos no SAEB do ano  $t$ . Para aqueles com dois anos de atraso seria possível utilizar a edição seguinte do SAEB (realizada em  $t+2$ ). Por

simplicidade, no entanto, todos os dados da tabela foram todos preenchidos com base na SAEB do ano  $t$ , utilizando as gerações vizinhas como uma aproximação.

Em relação à Tabela 5, dois aspectos devem ser destacados. O primeiro é que a realização do SAEB se dá antes que a situação de aprovação e reprovação esteja definida. Assim, alguns daqueles representados no SAEB podem não concluir a etapa. Como a reprovação deve, provavelmente, ser mais elevada para os de baixa proficiência, alguém pode alegar que o procedimento adotado subestima a proporção de pessoas da geração no nível proficiente. Por outro lado, cerca de 25% dos escalados para fazer o SAEB no ensino médio não realizam o exame. Caso os que faltam à prova estejam super-representados entre os de baixa proficiência (muitos já deixaram de frequentar a escola porque já não possuem chances de aprovação), isso geraria uma distorção que vai no sentido oposto à anterior. É difícil dizer qual o efeito líquido dessas duas distorções. Para efeitos do presente estudo nenhum procedimento foi adotado para lidar com esses potenciais problemas.

O segundo aspecto a ser destacado diz respeito ao SAEB-2019. Em 2019 o INEP não coletou no questionário de contexto sociodemográfico informações que permitissem identificar o ano de nascimento dos estudantes, impossibilitando a separação dos estudantes quanto ao alinhamento idade-série - adiantados, em linha e atrasados. Assim, o SAEB 2019 só permite obter uma estimativa para a última linha da Tabela 5. Para lidar com isso, o procedimento adotado foi, em primeiro lugar, distribuir o total de cada coluna da Tabela 5, em 2019, de acordo com a distribuição obtida para 2017 (mantendo o total da coluna de 2019 que é observado). Em segundo lugar, cada linha da tabela deve somar 1, mas mantendo as proporções obtidas pela primeira transformação. Esses são os dados considerados nas linhas da Tabela 5, para o ano de 2019.

Para os dados da Tabela 6 é possível utilizar ao menos dois procedimentos distintos. O primeiro seria utilizar os dados do Censo Escolar, o qual, por não ser amostral, permite que o IIE seja estimado com precisão para diferentes subgrupos da geração (e.g. área geográfica, gênero, raça/cor, educação dos pais etc.). Como o Censo Escolar só possui informação daqueles que estão na escola, teríamos que utilizar a PNAD para estimar a proporção de indivíduos da geração que se encontram fora da escola. O problema em utilizar o Censo Escolar por aluno (necessário para o cálculo do IIE) é que ele só está disponível a partir de 2007, impedindo a verificação da evolução do IIE em um período de crescimento significativo do PISA: 2000-2009. O segundo procedimento,



aqui adotado, seria utilizar apenas a PNAD.<sup>12</sup> Então, para a geração que deveria terminar a educação básica no ano  $t$ , calcula-se a proporção dos alunos da geração que já terminou o ensino médio na PNAD do ano  $t$ . Essa seria a estimativa para a primeira linha da Tabela 6: os que terminaram a educação básica adiantados. Na PNAD de  $t+1$  calcula-se novamente a proporção dos membros da geração que terminaram o ensino médio. Essa seria a estimativa da proporção daqueles que finalizaram o ensino médio em linha ou adiantados. Desse total é subtraída a estimativa dos que terminaram o ensino médio adiantado e se obtém a estimativa da proporção que terminou a educação básica em linha. E assim sucessivamente, até preencher a Tabela 6.

Para definição de uma geração, foram considerados aqueles que nascem entre 1 de janeiro e 31 de dezembro de um determinado ano. Então, são considerados em linha, por exemplo, aqueles que completam 17 anos no ano que cursam o último ano do ensino médio.

## Resultados

A Figura 8 mostra a evolução, entre 2003 e 2017, para uma versão adaptada do IIE, criada para permitir a comparação da evolução deste tipo de índice com os já discutidos nas seções anteriores – SAEB, IDEB e PISA. Na ausência de Censo Escolar por aluno no período anterior a 2007, foi utilizada a PNAD para a informação de alinhamento idade-série, ao invés do Censo Educacional.<sup>13</sup> Como o SAEB, até 2005, era amostral, em algumas edições mais antigas os alunos faziam apenas uma das provas - matemática ou leitura, de modo que não é possível computar a proporção de alunos proficientes em ambas disciplinas. A saída foi considerar apenas uma das disciplinas e foi escolhida matemática por ter a melhor comparabilidade com o PISA.

Seguindo a proposta original do IIE, foi considerado como proficiente o que aluno que obteve 300 pontos ou mais em matemática. Os IIE calculado para o Brasil, Estados e Distrito Federal, entre 2013 e 2019 está apresentado no apêndice.<sup>14</sup>

---

<sup>12</sup> Com possível complementação do censo populacional para os anos que não há PNAD.

<sup>13</sup> O censo escolar por alunos (necessário para o preenchimento da tabela 5) só está disponível a partir de 2007. Como, para acompanhamento dos adiantados, é necessário o censo do ano anterior ao ideal para a geração de referência cursar o 3º ano do ensino médio, só poderíamos calcular o IIE com base no censo a partir de 2009, ano que o SAEB não está disponível. Portanto só poderíamos construir a série do indicador a partir de 2011.

<sup>14</sup> Também é possível calcular um IIE adaptado que considerasse a proficiência em leitura, mas não foi realizado para este estudo.

A mesma metodologia poderia ser aplicada para outros cortes de proficiência, o que pode contribuir para a análise da evolução da educação a partir da observação de outros pontos da distribuição de proficiências. Além do corte de 300 do IIE, outros dois foram calculados e são apresentados aqui por terem apresentado evoluções diferentes no IIE: um corte mais alto - 375, que foi denominado Índice de Sucesso Educacional - ISE, e outro mais baixo - 250, que seria o Índice de Pré-Inclusão Educacional - IPIE.<sup>15</sup>

O IIE-Matemática apresenta um crescimento de 21% (3,3 pontos) entre 2003 e 2017. Essa mesma variação ocorre para o período de 2005 a 2017, intervalo no qual não se observa nenhum avanço na média do SAEB de matemática e tem-se crescimento de 12% no IDEB. A não ser pela estranha queda observada em 2015, esse crescimento parece consistente, ainda que modesto.<sup>16</sup>

Já o Índice de Sucesso Educacional apresenta uma queda entre 2003 e 2007 e, após 2007, mostra uma estabilidade (com exceção a queda em 2015). Por fim, o Índice de Pré-Inclusão Educacional apresenta um aumento substancial, de 20,1% (10,2 pontos entre 2003 e 2017). De modo geral, a Figura 8 mostra uma melhora do desempenho educacional entre gerações sucessivas e que tal melhora ocorre principalmente na parte inferior da distribuição de desempenho, o que seria compatível com uma redução de desigualdade educacional. Esse resultado é distinto de quando olhamos a evolução do SAEB ou IDEB.

---

<sup>15</sup> Em 2013 os indicadores são calculados usando tanto a PNAD como a PNAD contínua, a fim de verificar se a mudança metodológica afetaria o acompanhamento ao longo do tempo. Como é possível verificar no gráfico, a mudança do uso da PNAD para a PNAD contínua parece não causar distorção.

<sup>16</sup> Vale destacar que essa queda no IIE também é observada para o Índice de Sucesso Educacional, mas não para o Índice de Pré-Inclusão Educacional. Isso poderia sugerir a existência de um “efeito teto” na prova do SAEB-2015. A saber, a falta de itens capazes de identificar os estudantes com proficiências mais elevadas. Outra possibilidade seria alguma deficiência na amostra da rede privada, que representa a maior parte dos casos de desempenho elevado.

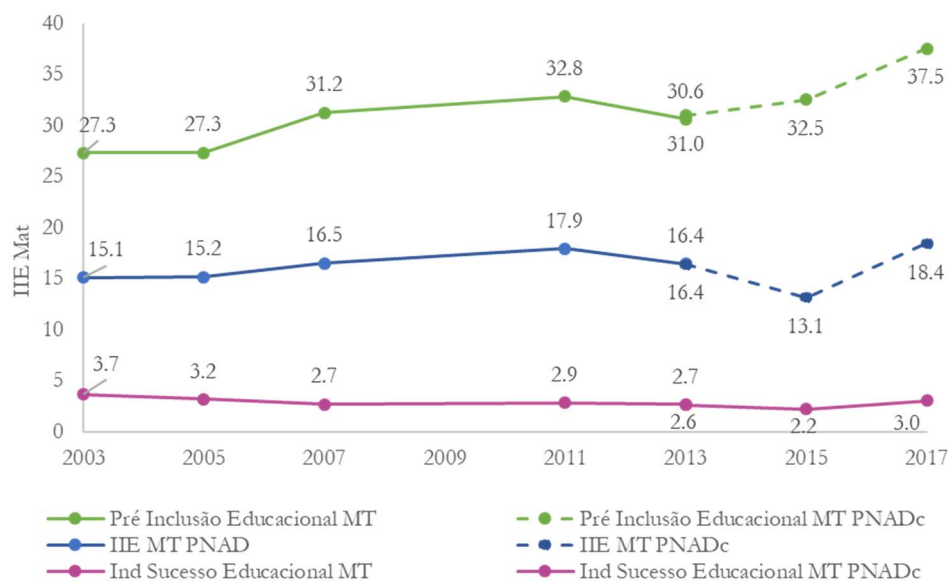


Figura 8. Adaptação do Índice de Inclusão Educacional para Construção da Série Histórica: % de pessoas de uma coorte de nascimento que conclui o ensino médio até os 18 anos acima dos níveis de proficiência em matemática indicados em cada índice: 375, para o Índice de Sucesso Educacional; 300, para o IIE-Matemática; e 250 para o Índice de Pré-Inclusão Educacional

Os indicadores discutidos acima são novamente apresentados na Figura 9, mas agora calculados a partir da metodologia original, com base no Censo Escolar e considerando tanto a proficiência em matemática quanto em leitura. Com esperado, os níveis dos indicadores são menores, pois agora é necessário ser proficiente em leitura e matemática (conforme os cortes definidos anteriormente) para compor o índice. Em termos de evolução, no entanto, as conclusões não são muito diferentes. A inclusão do ano de 2019 mostra um crescimento bastante acentuado do indicador, refletindo o crescimento observado no SAEB. Vale destacar que, como discutido acima, o SAEB 2019 omite a informação de ano de nascimento dos estudantes e, portanto, foi necessário utilizar a distribuição das colunas da Tabela 5 de 2017, conforme já discutido na seção de metodologia do IIE.

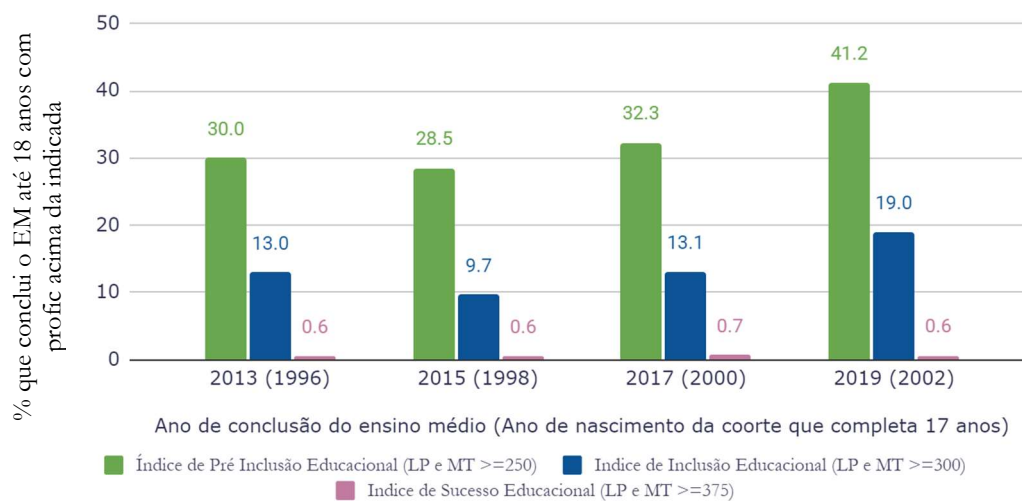


Figura 9. Proporção de pessoas de uma coorte de nascimento que conclui o ensino médio até os 18 anos acima dos níveis de proficiência indicados em cada índice – Índices de Pré-Inclusão, de Inclusão Educacional e de Sucesso Educacional

### Análise de Sensibilidade do IIE

A fim de verificar a sensibilidade do IIE às definições dos parâmetros para sua construção, foram realizados exercícios variando o intervalo de definição da coorte de nascimento e de corte para determinação do status de proficiente. Vale lembrar que, na definição inicial, o intervalo da coorte é coincidente com o ano calendário, 1º de janeiro a 31 de dezembro, e o corte de proficiência é 300 para matemática e língua portuguesa.

Na Figura 10 apresenta-se a evolução do IIE para duas diferentes definições de geração educacional, além da original: os nascidos entre abril do ano  $t$  e março do ano  $t+1$ ; e os nascidos entre julho do ano  $t$  e junho do ano  $t+1$ . Para as duas novas definições de geração temos que, para qualquer ponto no tempo, os estudantes são, em média, mais velhos. Então, não surpreende que o desempenho do indicador é maior para esses novos grupos: encontram-se, em média, em uma série mais elevada. Entretanto, a mudança na definição de geração parece não afetar muito o padrão da evolução temporal do indicador.

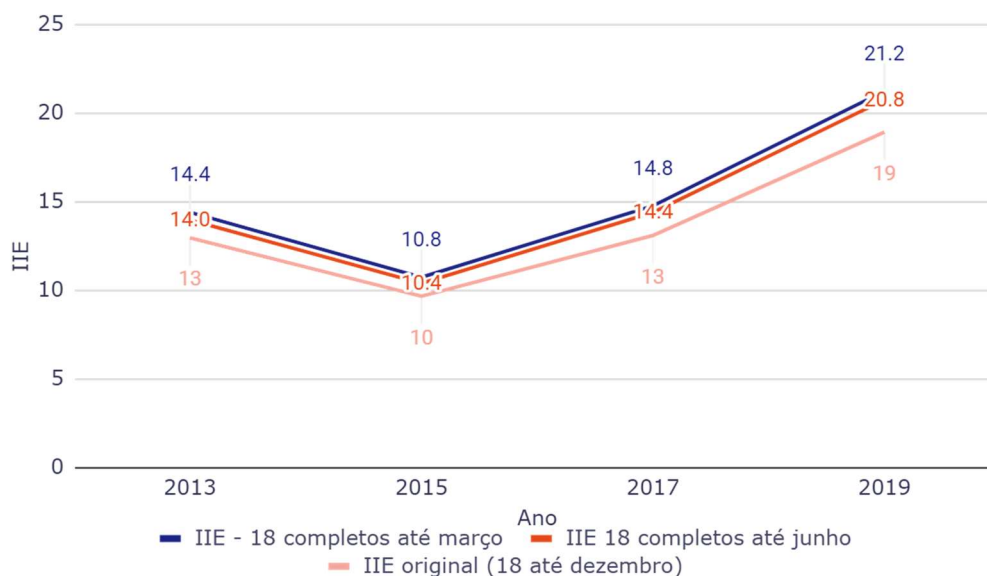


Figura 10. Análise de Sensibilidade da Evolução do IIE à sua Construção - Variação da Idade Adequada para Ingresso no Ensino Fundamental

O mesmo não pode ser dito em relação à mudança da nota de corte para considerar o estudante proficiente, como já visto na seção anterior. Evidentemente, quanto menor a nota de corte, maior o valor do indicador. Mas não é apenas o nível do indicador que é afetado pela mudança na nota de corte. Como nota-se na Figura 11, o padrão de evolução temporal também é afetado. O corte de 250 é onde o indicador apresenta a maior evolução positiva. Isso sugere que além do IIE, seria interessante monitorar indicadores com outras notas de corte. Na seção anterior, foram considerados o Índice de Pré-Inclusão Educacional (corte de 250 pontos) e o Índice de sucesso Educacional (corte de 375 pontos), a fim de oferecer uma análise mais completa da evolução da educação básica nos diferentes nível de desempenho, já que como demonstrado, elas não necessariamente seguem o mesmo padrão.

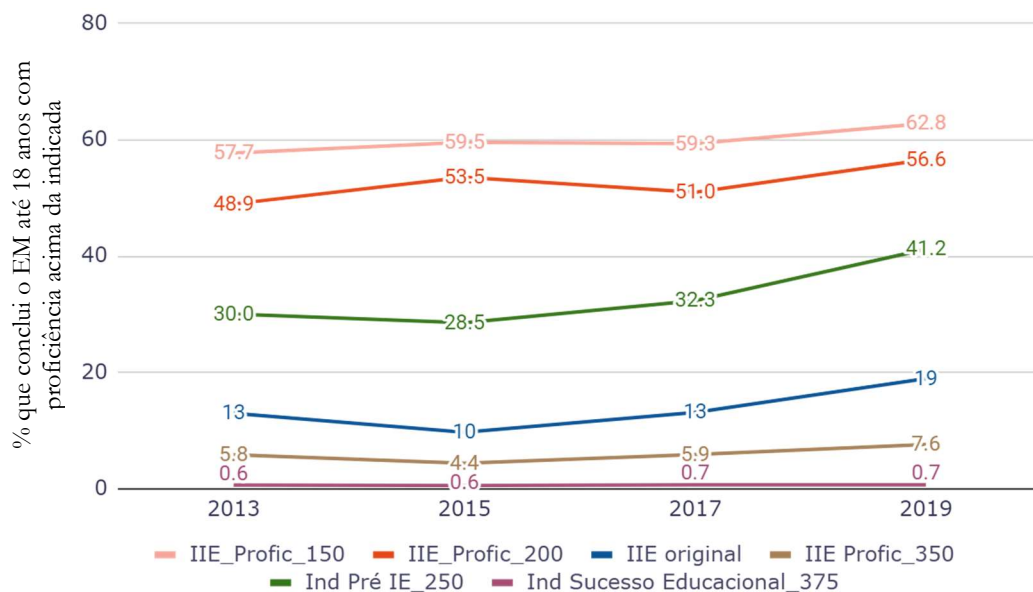


Figura 11. Análise de Sensibilidade da Evolução do IIE à sua Construção - Variação do Corte para Nível de Proficiência

## Considerações Finais

Este artigo avalia as possibilidades de indicadores disponíveis para acompanhar a evolução do desempenho na educação básica e propõem uma nova forma de fazê-lo, por meio da construção de um índice que compara o desempenho escolar entre gerações sucessivas. Vimos que indicadores como o SAEB e IDEB são fortemente afetados por efeitos de composição quando os indicadores de fluxo (taxas de repetência, promoção e evasão) mudam ao longo do tempo. O problema é que o SAEB e, conseqüentemente, o IDEB consideram a evolução do desempenho de uma determinada série entre anos consecutivos e a composição geracional de uma determinada série pode variar significativamente entre dois anos não muito distante.

Argumentamos no artigo que uma medida mais apropriada seria considerar o desempenho dos membros de uma determinada geração ao final da educação básica. Uma avaliação que considera a geração e poderia ser utilizada seria o exame do PISA. Entretanto, o PISA apresenta dois problemas. O primeiro é que ele considera indivíduos com 15 anos de idade e, assim, não abrange todo o ensino básico. O segundo é que, no Brasil, o PISA está sujeito a efeitos de composição. Isso porque, na idade de 15 anos, uma parcela significativa da geração encontra-se

com mais de dois anos de atraso ou já evadiu a escola e, assim, não é considerada para a amostragem do exame.

O artigo considera, para avaliar o desempenho educacional no ensino básico entre gerações sucessivas, o Índice de Inclusão Educacional (IIE). O IIE mede a parcela de jovens de uma mesma coorte de nascimento que termina o ensino médio proficiente (com pontuação em leitura e matemática acima de 300 pontos) e com, no máximo, um ano de atraso em relação ao recomendado. Avaliou-se também dois outros pontos de corte: 375 pontos (Índice de Sucesso Educacional) e 250 pontos (Índice de Pré-Inclusão Educacional). De modo geral, a análise aponta para uma melhora do desempenho educacional entre gerações sucessivas e que tal melhora ocorre principalmente na parte inferior da distribuição de desempenho. O que seria compatível com uma redução de desigualdade educacional. Esse resultado é distinto de quando olhamos a evolução do SAEB ou IDEB.



## Apêndice – Índice de Inclusão Educacional para Brasil, Estados e Distrito Federal

Na Tabela 7 estão apresentados o IIE, que é calculado para o ensino médio e também uma adaptação dele que é calculado para os anos iniciais e anos finais do ensino fundamental, empregando o mesmo conceito de conclusão com até um ano de atraso, e adequando as notas de corte. Para considerar o desempenho como proficiente, foi utilizado, então, 250 para os anos iniciais e 275 para os anos finais, mantendo o corte de 300 para o ensino médio, sempre o mesmo corte para matemática e leitura.

Tabela 7. Índice de Inclusão Educacional – IIE para Anos Iniciais (AI) e Anos Finais (AF) do Ensino Fundamental e para o Ensino Médio (EM) – 2013 a 2019

Território	2013			2015			2017			2019		
	AI	AF	EM	AI	AF	EM	AI	AF	EM	AI	AF	EM
<b>BRASIL</b>	37.9	14.9	12.8	43.2	16.9	9.6	49.8	21.3	13.6	52.4	24.7	19
<b>RO</b>	32.1	9.6	8.4	38.1	11.7	6.3	47.5	19.4	9.9	46.1	21.6	12.1
<b>AC</b>	29.1	9.9	7.2	35.7	9.8	4.4	45.2	14.9	7.5	47.7	17.6	9.2
<b>AM</b>	25	9.1	4.2	31.1	11	6.4	37.3	14.1	6.1	39.2	16.2	7.7
<b>RR</b>	26.6	7.6	5.9	31.6	8	4.9	39.2	11	5.7	36.7	12.8	8.3
<b>PA</b>	14.8	5.1	4.9	19.6	5.9	4.6	24.8	7.8	5.9	28.8	10.7	8.4
<b>AP</b>	13.1	5	5.6	18.3	5.6	2.8	22.1	7.3	5	28.6	8.9	6.5
<b>TO</b>	28.3	11.2	7	32	11.3	4.8	41.9	18.4	9.9	44.3	19.9	11.8
<b>MA</b>	13.7	5.1	3.8	18.8	6.6	2.6	24	7.5	5.8	29.1	11	8.9
<b>PI</b>	21.2	10.2	9.6	27.3	11.4	6.2	36.6	16	10.7	44.3	21.7	15.2
<b>CE</b>	36.7	14.5	10.7	46.7	17.9	7.8	53.9	24.5	13.5	58.2	31.5	21.2
<b>RN</b>	23.8	10.6	9.5	28.3	10.8	5.8	34.1	13.4	8.3	35.4	16.7	15.5
<b>PB</b>	24	8.4	10.1	28.9	9.8	6.3	35.5	13	9.7	40.7	17.7	15.9
<b>PE</b>	26.1	10	11.1	30.6	11.4	9.8	37.2	15.1	13.2	42.6	20.1	19.3
<b>AL</b>	17.9	5.4	5.1	24.6	7.4	3.4	34.9	12.8	7	41.7	17.9	11.9
<b>SE</b>	20.3	7.8	6.8	25.7	9.7	6.4	32.5	11.8	11	34.6	15.2	13.1
<b>BA</b>	19.6	6.7	5.9	25.3	8.8	4.8	32.8	10.7	7	34.8	13.8	7.9
<b>MG</b>	53	22.3	15.9	57.2	22.9	10.7	62.3	25.9	16.2	62	28.9	22
<b>ES</b>	40.6	16.5	15.2	46.9	18.1	12.7	53	23.1	18.6	56	26.9	24.6
<b>RJ</b>	37.5	15.7	14.6	41.5	17.3	10.8	48.2	22.5	15.3	48.8	24.3	18.1
<b>SP</b>	52	21.1	18.1	58.3	24.7	13.2	65.6	30.1	17.6	67.5	32.4	27
<b>PR</b>	51.2	17.9	13.7	56.2	19.4	12.2	64	26.9	16.5	65.3	30.8	24.7
<b>SC</b>	52.2	18.2	17.5	58.7	25.9	14.1	63.1	29.6	15.9	64.1	32.3	17.4
<b>RS</b>	44.1	16.2	16.6	47.7	14.9	11	52	23.3	14.9	55.9	27	17.6
<b>MS</b>	36.3	15.1	12	42	18.3	9.5	47.5	20.9	12.5	48	21.9	16.8
<b>MT</b>	33.9	12.2	9.1	42.3	14.3	7.2	47.4	19.2	9.6	47.3	21.4	13.4
<b>GO</b>	43.2	18.3	13.6	46.8	21.3	9.6	53.5	26.7	17.4	56.4	29.8	25.8
<b>DF</b>	50.6	19.4	18.7	52.5	20	15.2	59.9	26.1	19.3	63.2	30	26

## Referências

- Alves, F. (2007). Qualidade da educação fundamental: integrando desempenho e fluxo escolar. *Ensaio: Avaliação e Políticas Públicas em Educação*, 15(57), p. 525-542.
- Ballou, D. (2009). Test scaling and value-added measurement. *Education Finance and Policy*, 4(4), p. 351–383.
- Carnoy, M. et al. (2015). A educação brasileira está melhorando? Evidências do PISA e do SAEB. *Cadernos de Pesquisa*, 45(157), p. 450-485.
- Fernandes, R. e Natenzon, P. E. (2003). A evolução recente do rendimento escolar das crianças brasileiras: uma reavaliação dos dados do Saeb. *Estudos em Avaliação Educacional*, 28, p. 3-21.
- INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (2020). *Relatório: Brasil no PISA 2018*.
- Klein, R. (2011). Uma re-análise dos resultados do Pisa: problemas de comparabilidade. *Ensaio: Avaliação e Políticas Públicas em Educação*, 19(73) p. 1-20.
- Yen, W. (1986). The choice of scale for educational measurement: an IRT perspective. *Journal of Educational Measurement*, 23(4), p. 299-325.