

BOLETIM

NA MEDIDA

ANO 5 • NÚMERO 11 • FEVEREIRO 2017

EDITORIAL

ARTIGOS

- A desigualdade do acesso à creche no Brasil considerando características de raça/cor, localização e renda das crianças
- As escolas com localização diferenciada e o perfil dos estudantes (2012 - 2015)
- Concordância e discordância na declaração racial de estudantes: cruzando dados do censo escolar, Saeb e Enem
- As provas das quatro áreas do Enem vistas como uma prova única na ótica de modelos da teoria da resposta ao item uni e multidimensional

CONTRIBUIÇÃO EXTERNA

- O Enem e o impacto de variáveis socioeconômicas e educacionais

INEP

MINISTÉRIO DA
EDUCAÇÃO



BOLETIM

NA MEDIDA

ANO 5 • NÚMERO 11 • FEVEREIRO 2017

COMISSÃO EDITORIAL

Daniel Jaime Capistrano de Oliveira

Lenice Medeiros

Márcio Andrade Monteiro

Robson Quintilio

Valdir Quintana Gomes Júnior

EDITOR-EXECUTIVO

Márcio Andrade Monteiro

REVISÃO/NORMALIZAÇÃO

Amanda Mendes Casal

Clarice Rodrigues da Costa

PROJETO GRÁFICO

Marcos Hartwich

Raphael C. Freitas

DIAGRAMAÇÃO E ARTE-FINAL

Lilian dos Santos Lopes

EDITORIA

Inep/MEC – Instituto Nacional de Estudos e

Pesquisas Educacionais Anísio Teixeira

SIG Quadra 4, Lote 327, Edifício Villa Lobos, Térreo

Brasília-DF – CEP: 70610-908

Fones: (61) 2022-3070, 2022-3077 – editoracao@Inep.gov.br

SUMÁRIO

Editorial.....	3
----------------	---

ARTIGOS

A desigualdade do acesso à creche no Brasil considerando características de raça/cor, localização e renda das crianças Marcelo Lopes de Souza	5
As escolas com localização diferenciada e o perfil dos estudantes (2012-2015) Robson dos Santos	13
Concordância e discordância na declaração racial de estudantes: cruzando dados do censo escolar, Saeb e Enem Adolfo Samuel de Oliveira; Adriano Souza Senkevics; Rosilene Cerri; Pamela Sada Dias Edokawa	21
As provas das quatro áreas do Enem vistas como uma prova única na ótica de modelos da teoria da resposta ao item uni e multidimensional Nara Núbia Vieira	29

CONTRIBUIÇÃO EXTERNA

O Enem e o impacto de variáveis socioeconômicas e educacionais Maurício Urban Kleinke	41
---	----



EDITORIAL

O Plano Nacional de Educação (PNE 2014-2024) apresenta vários desafios para a educação brasileira, principalmente no alcance de metas e na implementação de estratégias que contribuam para que as metas sejam alcançadas ao final do Plano. Nesse contexto, o primeiro boletim desta edição do Na Medida trata da desigualdade de acesso à creche no Brasil, temática para a qual temos a meta do PNE de 50% das crianças de 0 a 3 anos atendidas até 2024. Seu objetivo é analisar se essa expansão do acesso de 2004 para 2014, a partir dos dados da Pesquisa Nacional por Amostra de Domicílios (Pnad), do IBGE, ocorreu com redução das desigualdades, comparando os níveis de acesso segundo alguns fatores: raça/cor da criança, localização do domicílio (urbano/rural) e renda *per capita* domiciliar. Já o segundo boletim aborda uma temática presente em estratégias de várias metas do PNE: políticas educacionais para populações específicas, como populações do campo, indígenas e quilombolas. Assim, o Boletim apresenta, a partir de dados do censo escolar, estatísticas e indicadores para escolas que possuem a chamada “localização diferenciada”, como área de assentamento, terra indígena e área remanescente de quilombos, além de estatísticas e indicadores para seus respectivos estudantes. É importante destacar a importância de considerar as especificidades populacionais se queremos universalizar o acesso e a educação de qualidade no nosso País.

No número 9 do Na Medida, foi publicado um boletim por pesquisadores da Diretoria de Avaliação Básica (Daeb) sobre a forma de coleta do quesito cor/raça nos instrumentos de pesquisa do Inep, mostrando as diferenças nessa coleta.

Neste Boletim, pesquisadores da Daeb retomam o tema com um interessante texto que analisa o grau de concordância e discordância na classificação racial obtido por meio de cruzamento dos dados de cor/raça referentes aos mesmos indivíduos respondentes de questionários diferentes do Inep. Isso mostra o potencial analítico de pesquisas e estudos feitos a partir da junção de bases de dados diferentes, mas referentes ao mesmo objeto de análise, seja esse objeto escola, docente ou estudante. Os resultados mostram a necessidade de reflexão sobre a produção de dados raciais em educação para o aprimoramento da qualidade da informação produzida.

O último boletim de pesquisadores do Inep foi elaborado com base em uma dissertação de mestrado de uma Pesquisadora-Tecnologista da Diretoria de Estatísticas Educacionais (Deed). Esse boletim retoma um tema altamente relevante presente no Na Medida 10, que é a discussão de dimensionalidade das provas do Enem. Apesar de a parte objetiva do Enem ser composta por quatro provas diferentes, esse estudo de dimensionalidade considera a edição de 2012 uma prova única composta por 180 itens (45 de cada prova objetiva). As análises foram realizadas a partir de uma amostra dos alunos concluintes do ensino médio regular seriado que estiveram presentes em todas as quatro provas objetivas do Exame. A principal pergunta a ser respondida no boletim é: considerando uma prova única com 180 itens, há um traço latente dominante?

Para fechar o Boletim 11, similarmente ao número 10 a contribuição externa também vem de um Núcleo de Estudos, gerenciado pela Coordenação-Geral de Instrumentos e Medidas Educacionais (Cgime/Dired), na Universidade Federal de São Carlos, que realizou estudos sobre o Enem, mas apenas para as provas de Matemática e Ciências da Natureza. Nesse estudo, a partir de um modelo de regressão logística aplicado às provas de 2009 a 2012, é estimada a chance de participantes do Enem pertencerem ao decil superior nas provas de Ciências da Natureza e Matemática em função de quatro dimensões: pessoal, escolar, familiar e regional, cujas informações constam do questionário socioeconômico do Exame. Essa opção de classificação quanto a pertencer ou não ao decil superior foi feita devido ao uso dos resultados do Enem para acesso à educação superior, principalmente via Sistema de Seleção Unificada (Sisu) do MEC, considerando que os participantes que pertencem ao decil superior disputarão vagas em importantes universidades em cursos considerados de maior status social, como Medicina, Engenharias e Direito.

Esperamos que a leitura deste Boletim contribua com reflexões importantes para a pesquisa educacional brasileira.

A DESIGUALDADE DO ACESSO À CRECHE NO BRASIL CONSIDERANDO CARACTERÍSTICAS DE RAÇA/COR, LOCALIZAÇÃO E RENDA DAS CRIANÇAS

MARCELO LOPES DE SOUZA*



INTRODUÇÃO

Neste artigo são apresentados alguns resultados de uma pesquisa mais ampla que tem por objetivo analisar os níveis de equidade do acesso à creche no Brasil no contexto da Meta 1 do atual Plano Nacional de Educação (PNE), aprovado em junho de 2014 (Lei nº 13.005/2014).

A Meta 1 do PNE tem por objetivo a ampliação do acesso à creche para um atendimento mínimo de 50% das crianças de até três anos até 2024. Além da ampliação do acesso, outro desafio apresentado pelo PNE diz respeito à redução das desigualdades de acesso. Em primeiro lugar, a superação das desigualdades educacionais é uma das diretrizes do PNE. Já a Estratégia 1.2 tem por intuito, diretamente, a redução da desigualdade educacional segundo a renda das famílias.¹ Além disso, a questão da redução das desigualdades de acesso também deve ser vista no contexto do estabelecimento da própria meta, que não é de universalização, mas de ampliação. Quando a meta é de universalização, todos os estratos populacionais devem, obrigatoriamente, ser atendidos. Como no caso da creche a meta é a ampliação para, no mínimo, 50% de acesso, é importante que se produzam

*Mestre em Economia pela Universidade Federal de Uberlândia (UFU). Pesquisador-Tecnologista em exercício na Diretoria de Estudos Educacionais (Dired) do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). E-mail: marcelo.souza@inep.gov.br

¹ Essa estratégia trata de “garantir que, ao final da vigência deste PNE, seja inferior a 10% (dez por cento) a diferença entre as taxas de frequência à educação infantil das crianças de até três anos oriundas do quinto de renda familiar *per capita* mais elevado e as do quinto de renda familiar *per capita* mais baixo.” (Brasil, 2014).

informações sobre o perfil da inclusão educacional, para que os governos possam avaliar a definição de prioridades de atendimento.

Como pode ser observado no Gráfico 1, no período de 2004 a 2014, década que antecede o atual PNE – que vai de 2014 a 2024 –, registrou-se aumento no percentual de crianças de até três anos que frequentavam escola ou creche. Em 2004, o percentual era de 19,2% e, em 2014, o acesso chegou a 33,3% – uma expansão de 14,1 pontos percentuais (p.p.).² Neste artigo, o objetivo é analisar, para o nível Brasil, se essa expansão do acesso ocorreu com redução das desigualdades, por meio da comparação dos níveis de acesso de alguns estratos populacionais.

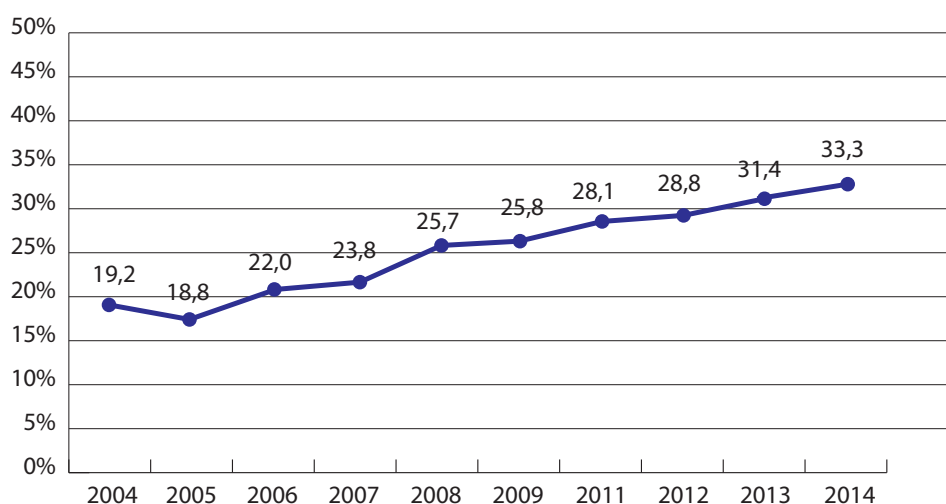


Gráfico 1 Evolução do percentual das crianças de 0 a 3 anos de idade que frequentavam a escola ou creche – Brasil – 2004-2014

Fonte: Elaborado pelo autor com base na Pnad/IBGE.

METODOLOGIA

Para analisar se houve redução da desigualdade, utilizou-se um modelo de regressão logística para comparar as chances de acesso à creche dos estratos populacionais selecionados. As variáveis incluídas na regressão para as crianças de até três anos foram raça/cor, localização (urbano/rural) e renda *per capita* do domicílio. Embora seu resultado em termos de desigualdade não tenha sido tratado neste artigo, a idade das crianças também foi incluída

² Para a análise desse resultado e demais desagregações, consultar o Relatório do Primeiro Ciclo de Monitoramento das Metas do PNE: Biênio (2014-2016). Assim como define o Relatório, a data de referência para o cálculo da idade em anos completos das crianças foi o dia 31 de março, conforme Resolução CNE/CEB nº 6, de 20 de outubro de 2010. Para mais informações, consultar o Relatório.

na regressão logística, a fim de que as comparações levassem em consideração essa importante variável explicativa.³

A base de dados utilizada foi a Pnad dos anos de 2004 e 2014. Os resultados apresentados consideram a estrutura do plano amostral complexo da Pnad.⁴ Foi empregado o módulo *Complex samples* do programa IBM SPSS Statistics Software® e foram utilizadas as seguintes variáveis para uma aproximação do desenho amostral da Pnad: V4617 (STRAT), V4618 (PSU) e V4611 (PESO DO DOMICÍLIO).

RESULTADOS

As estatísticas descritivas das comparações realizadas entre 2004 e 2014 já mostram que, apesar de todos os estratos populacionais analisados terem aumentado o percentual de acesso à creche, a desigualdade se manteve em um patamar muito elevado (Tabela 1).

O acesso à creche depende muito da idade das crianças. Nos dois anos analisados, quanto maior a idade, maior foi o percentual de acesso. Apesar de o acesso ter aumentado para todos os grupos de idade de 2004 para 2014, a diferença entre os percentuais se manteve muito elevada. Embora o acesso à creche tenha aumentado bastante para a população residente na área rural (o aumento foi de 9,8 p.p.), o direito à creche ainda está muito relacionado às famílias residentes na área urbana. O acesso à creche também aumentou para as crianças negras (pretas e pardas), passando de 16,9%, em 2004, para 29,3%, em 2014, totalizando um aumento de 12,4 p.p. Contudo, a diferença entre os percentuais de acesso entre as crianças brancas e negras passou de 4,5 p.p., em 2004, para 8,4 p.p., em 2014. As crianças pertencentes às famílias do quintil mais pobre de renda também tiveram ampliação do acesso quando se compara 2004 com 2014 (um aumento de 9,8 p.p). Contudo, a desigualdade permaneceu muito elevada também para esse tipo de grupo de comparação.

³ Outros trabalhos com objetivos similares também incluem essas variáveis nas análises, tais como o Banco Mundial (2011) e Berlingeri e Santos (2014).

⁴ Sobre os aspectos principais referentes aos motivos da incorporação da estrutura do plano amostral complexo da Pnad e sobre quais variáveis devem ser consideradas, consultar Silva, Pessoa e Lila (2002).

Tabela 1 Estatística descritiva do percentual de acesso à creche das crianças de até três anos de idade – Brasil – 2004/2014

Estrato Populacional		2004	2014
Brasil		19,2	33,3
Idade (em anos)	menos de 1 ano	4,1	8,0
	1 ano	9,3	19,7
	2 anos	20,3	40,8
	3 anos	38,9	63,7
Localização	urbano	22,1	36,3
	rural	8,1	17,9
Raça/cor	branco	21,4	37,7
	negro	16,9	29,3
Quintis de Renda	20% mais pobres	12,6	22,4
	20% maior renda	40,4	57,0

Fonte: Elaborada pelo autor com base na Pnad/IBGE.

Os resultados principais da regressão logística são apresentados nos Gráficos 2, 3 e 4. Em cada um desses gráficos, a regressão logística permite comparar as chances de acesso de um grupo em relação ao outro, mantendo constantes as demais características incluídas no modelo⁵. Isso permite que se comparem, por exemplo, as chances de acesso à creche de duas crianças com as mesmas características referentes às variáveis incluídas no modelo, mas cuja localização de domicílio difere.

O Gráfico 2 mostra justamente esse tipo de comparação considerando a diferença de chances existente entre as crianças residentes nas áreas urbanas e as crianças residentes na área rural. Em 2004, o acréscimo de chances de acesso à creche em virtude da residência nas áreas urbanas era de 211,9%. Em 2014, a estimativa diminuiu para 175,0%. Contudo, considerando a magnitude do percentual de 2014 (175,0%), constata-se que o acesso à creche se manteve como uma característica muito relacionada às famílias das áreas urbanas, apesar dessa diferença ter diminuído um pouco.

⁵ O Pseudo R2 de Nagelkerke do modelo adotado foi de 0,25, em 2004, e de 0,35, em 2014. Esse resultado fornece uma medida aproximada do poder explicativo das variáveis incluídas no modelo (idade, raça/cor, localização e renda) para prever as chances de uma criança ter acesso à creche. Assim, em 2004, por exemplo, apenas as quatro variáveis incluídas no modelo eram responsáveis por explicar cerca de 25% das chances de uma criança ter acesso à creche.

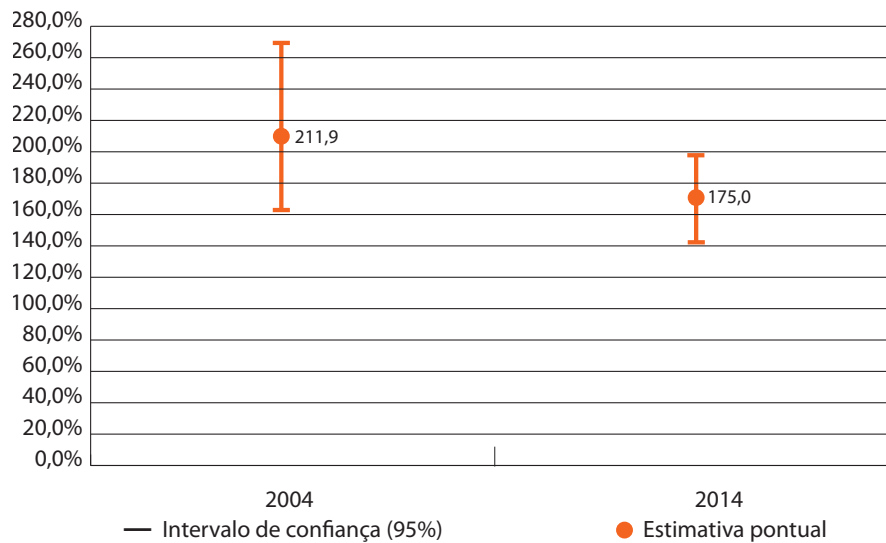


Gráfico 2 Percentual de acréscimo das chances de acesso à creche das crianças de até três anos residentes na área urbana – Brasil – 2004/2014

Fonte: Elaborado pelo autor com base na Pnad/IBGE.

Nota: Estimativas pontuais estatisticamente significantes a 0,1%.

A desigualdade de acesso à creche aumentou quando se considera a raça/cor das crianças (Gráfico 3). A comparação é feita entre as chances das crianças de raça/cor branca e as crianças negras (pretas e pardas). Em 2004, as crianças brancas apresentavam um acréscimo de chances de 10,7%. Em 2014, o acréscimo de chances das crianças brancas foi maior, passando a ser de 39,8%, com intervalo de 95% de confiança, ficando bem acima do intervalo que foi estimado para 2004, conforme pode ser observado no Gráfico 3.

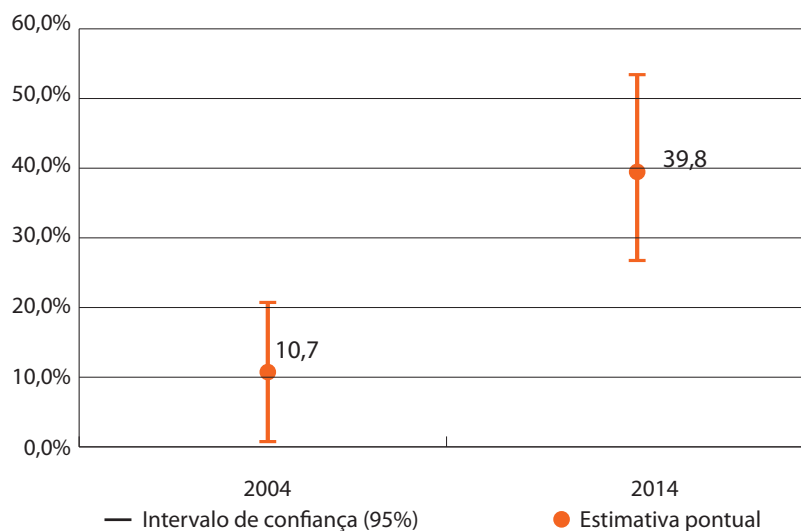


Gráfico 3 Percentual de acréscimo das chances de acesso à creche das crianças de até três anos brancas – Brasil – 2004/2014

Fonte: Elaborado pelo autor com base na Pnad/IBGE.

Nota: Estimativas pontuais estatisticamente significante a 5%, em 2004, e a 0,1%, em 2014.

A desigualdade no acesso à creche se manteve elevada em relação à renda *per capita* (Gráfico 4). Nos dois anos analisados, apesar de ligeira queda em 2014, a estimativa pontual ficou no mesmo nível de acréscimo de 5% de chances a cada variação de R\$ 100,00 na renda domiciliar *per capita*. Apesar de pequena variação, o intervalo de 95% de confiança também ficou no mesmo patamar, 4,1% a 6,7%, em 2004, e 3,8% a 5,7%, em 2014.

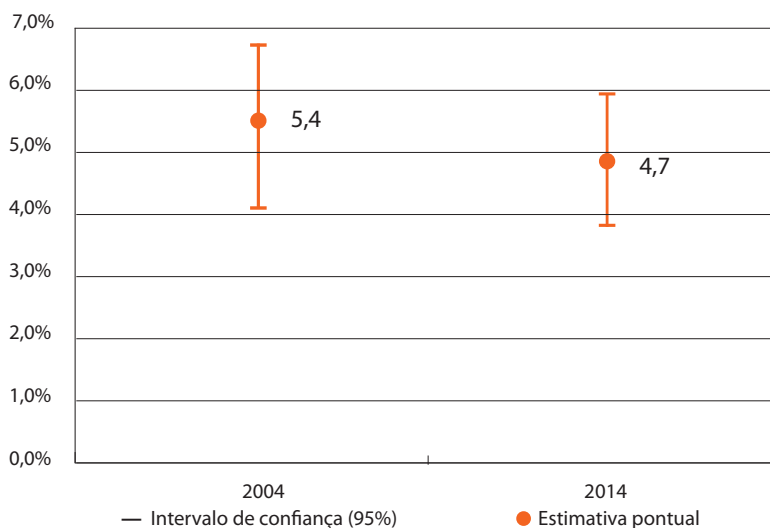


Gráfico 4 Percentual de acréscimo das chances de acesso à creche das crianças de até três anos a cada variação de R\$ 100 na renda domiciliar *per capita** – Brasil – 2004/2014

Fonte: Elaborado pelo autor com base na Pnad/IBGE.

Nota: Estimativas pontuais estatisticamente significantes a 0,1%.

*R\$ 100,00 a preços de 2014.

CONSIDERAÇÕES FINAIS

Apesar do acesso à creche ter aumentado para todos os estratos considerados, a desigualdade de acesso permaneceu em um patamar elevado em 2004 e 2014. Apesar de a diferença ter diminuído um pouco, a chance de acesso à creche continuou sendo muito maior para as crianças das áreas urbanas. A desigualdade de chances de acesso medida pela renda se manteve no mesmo nível. Ademais, verificou-se que a desigualdade entre brancos e negros aumentou – uma má notícia para a busca de equidade de acesso.

REFERÊNCIAS

BANCO MUNDIAL. *Educação infantil: programas para a geração mais importante do Brasil*. São Paulo, 2011.

BERLINGERI, M. M.; SANTOS, D. Projeção da demanda por creche incorporando tendências econômicas e demográficas recentes. In: CAMARANO, A. A. (Org.). *Novo regime demográfico: uma nova relação entre população e desenvolvimento?* Rio de Janeiro: IPEA, 2014. p. 445-468.

BRASIL. Conselho Nacional de Educação (CNE). Câmara de Educação Básica (CEB). Resolução nº 6, de 20 de outubro de 2010. Define Diretrizes Operacionais para a matrícula no Ensino Fundamental e na Educação Infantil. *Diário Oficial da União*, Brasília, DF, 21 out. 2010. Seção 1, p. 17. Disponível em: http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=15542-rceb006-10-pdf-1&Itemid=30192 . Acesso em: 5 fev. 2015.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). *Relatório do 1º Ciclo de monitoramento das metas do PNE: biênio 2014-2016*. Brasília, 2016.

BRASIL. Lei nº 13.005, de 25 de junho de 2014. Aprova o Plano Nacional de Educação- PNE e dá outras providências. *Diário Oficial da União*, Brasília, DF, 16 jun. 2014. Disponível em: http://www.planalto.gov.br/CCIVIL_03/_Ato2011-2014/2014/Lei/L13005.htm.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). *Pesquisa Nacional por Amostra de Domicílios (Pnad): microdados 2004-2014*. Disponível em: <<http://www.ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/Pnad2013/microdados.shtm>>.

SILVA, P. L.; PESSOA, D. G.; LILA, M. F. Análise estatística de dados da PNAD: incorporando a estrutura do plano amostral. *Ciência & Saúde Coletiva*, Rio de Janeiro, v. 7, n. 4, p. 659-670, 2002.

AS ESCOLAS COM LOCALIZAÇÃO DIFERENCIADA E O PERFIL DOS ESTUDANTES (2012-2015)

ROBSON DOS SANTOS*



INTRODUÇÃO

O Plano Nacional de Educação (PNE), em seu artigo 8º, inciso II, estabeleceu que as políticas educacionais devem considerar as especificidades das populações do campo, indígenas e quilombolas. Várias estratégias do PNE se voltam para comunidades específicas, como a 1.10, 2.6, 2.10, 3.7, 5.5, 6.7, 7.26, entre outras, abrangendo a universalização da educação básica, a formação de professores, a produção de materiais didáticos e conteúdos curriculares. Por isso, é relevante apreender quais as características das instituições existentes que buscam responder a tais demandas de contextualização educacional, em específico sob a perspectiva da localização em que ocorre a oferta.

O Censo Escolar da Educação Básica coleta informações importantes sobre a localização da escola e a zona de residência do aluno: pertencimento a área rural ou urbana e a assentamentos de reforma agrária, áreas de preservação ambiental, terras indígenas e quilombolas etc., o que o instrumento denomina “localização diferenciada”. A partir dessas informações, o presente texto busca responder às seguintes indagações: quantos são e qual é o perfil dos estudantes dessas unidades? Para isso, expõe descritivamente a distribuição das escolas com localização diferenciada, a localização de residência dos seus estudantes e apresenta alguns apontamentos sobre o desempenho das unidades.

*Sociólogo. Doutor em Sociologia (Unicamp). Pesquisador-Tecnologista em Informações e Avaliações Educacionais, lotado na Diretoria de Estudos Educacionais (Dired/Inep).

O QUE E QUANTAS SÃO AS ESCOLAS COM “LOCALIZAÇÃO DIFERENCIADA”?

A localização da escola constitui aspecto importante para a realização do direito à educação com equidade. Escolas próximas ao local de residência dos estudantes podem garantir o acesso sem a necessidade de longos deslocamentos, sobretudo nas áreas rurais. Além disso, podem garantir uma associação mais efetiva entre as práticas escolares e o contexto sociocultural. Daí que formas específicas de organização dos processos educativos receberam atenção do Conselho Nacional de Educação (CNE): no caso da educação escolar indígena, a resolução CNE/CEB nº 5, de 22 de junho de 2012, que institui as Diretrizes Curriculares Nacionais para a educação escolar indígena na educação básica; quanto à educação quilombola, a resolução CNE/CEB nº 8, de 20 de novembro de 2012, define as Diretrizes Curriculares Nacionais para a educação escolar quilombola na educação básica; em relação à educação do campo, a resolução CNE/CEB nº 1, de 3 de abril de 2002 estabelece as Diretrizes Operacionais para a educação básica nas escolas do campo, entre outras, que consagram o direito à diversidade das práticas educacionais e visam garantir e regular as formas distintas de organização escolar que decorrem dessas localizações diferenciadas.¹

A partir de 2012, o Censo Escolar da Educação Básica passou a coletar informações sobre sete categorias da variável “localização diferenciada da escola”. Sob essa designação existem instituições e públicos com características socioculturais distintas, como explicitam as definições da variável contidas no Quadro 1.

Quadro 1 Localização diferenciada da escola

(continua)

Código	Descrição
0- Não se aplica	A escola não possui localização diferenciada.
1- Área de assentamento	Área de terra na qual uma população está instalada, destinada à exploração agrícola, obtida por meio do programa de reforma agrária.
2- Terra indígena	Área demarcada como território indígena, de propriedade da União, habitada por um ou mais povos indígenas, utilizada para suas atividades produtivas, imprescindível à preservação dos recursos ambientais necessários para seu bem-estar e essencial para sua reprodução física e cultural, de acordo com seus usos, costumes e tradições.

¹ As referidas resoluções, bem como demais normas relativas à educação básica podem ser consultadas em: <<http://portal.mec.gov.br/par/323-secretarias-112877938/orgaos-vinculados-82187207/12992-diretrizes-para-a-educacao-basica>>.

Quadro 1 Localização diferenciada da escola

(conclusão)

Código	Descrição
3- Área remanescente de quilombos	Área que abriga os grupos étnico-raciais, segundo critérios de autoatribuição, com trajetória histórica própria, dotados de relações territoriais específicas, com presunção de ancestralidade negra relacionada à resistência à opressão histórica sofrida. Essas áreas são demarcadas e reconhecidas pelo Instituto Nacional de Colonização e Reforma Agrária (Incra) e certificadas pela Fundação Palmares.
4- Unidade de uso sustentável	Espaço territorial e seus recursos ambientais, incluindo as águas jurisdicionais, com características naturais relevantes. Está legalmente instituído pelo poder público, com objetivos de conservação e limites definidos, em regime especial de administração, ao qual se aplicam garantias adequadas de proteção, a fim de compatibilizar a conservação da natureza com o uso sustentável de parcela dos seus recursos naturais.
5- Unidade de uso sustentável em terra indígena	Área demarcada como terra indígena e também instituída como uma unidade de uso sustentável.
6- Unidade de uso sustentável em área remanescente de quilombos	Área demarcada como área remanescente de quilombos e também instituída como uma unidade de uso sustentável.

Fonte: Elaborado pelo autor com base no Censo da Educação Básica e no Caderno de Instruções do Censo Escolar da Educação Básica.

O Censo registrava, em 2012, 9.709 escolas com localização diferenciada, ou 5,0% do total de estabelecimentos de educação básica em atividade naquele momento no Brasil, número que se expandiu para 11.080, em 2015, isto é, 5,9% das escolas existentes. Esse acréscimo pode provir da ampliação do número de instituições, mas também do sub-registro existente no início da coleta. Tais escolas se concentravam, sobretudo, em áreas de assentamento da reforma agrária, em terras indígenas e áreas remanescentes de quilombolas (Tabela 1).²

A maior parte das escolas com localização diferenciada tem como endereço o campo. Todavia, o número das que ficam nas áreas urbanas passou por uma pequena ampliação (Tabela 1). Em 2012, 403, ou 4,2% das escolas com localização diferenciada, estavam em áreas urbanas, já, em 2015, o número absoluto se ampliou para 582, de modo que 5,2% se posicionavam em áreas urbanas.

² É importante destacar que, em 2015, o Brasil contava com 2.474 comunidades remanescentes de quilombolas (Fundação Cultural Palmares, 2015), 9.337 áreas de assentamento da reforma agrária (Brasil. Incra, 2015) – em ambos os casos a maioria se localizava no Nordeste – e 588 terras indígenas, principalmente na Amazônia Legal (Brasil. Funai, 2015).

Tabela 1 Distribuição das escolas de educação básica com localização diferenciada, por localização (rural e urbana) – Brasil – 2012/2015

Localização diferenciada	2012		2015	
	Rural	Urbana	Rural	Urbana
Área de assentamento	4.420	0	4.498	0
Terra indígena	2.822	40	2.993	52
Área remanescente de quilombos	1.832	123	2.132	144
Unidade de uso sustentável	193	227	778	369
Unidade de uso sustentável em terra indígena	9	1	26	1
Unidade de uso sustentável em área remanescente de quilombos	30	12	71	16
Total de escolas com localização diferenciada	9306	403	10.498	582
Total de escolas de educação básica	74.112	118.564	64.704	121.737

Fonte: Elaborada pelo autor com base no Censo da Educação Básica.

Considerando as escolas no campo de modo geral, cabe ressaltar que, entre 2012 e 2015, ocorreu uma redução no número de estabelecimentos de educação básica, de 74.112 escolas para 64.704. Já o número de escolas com localização diferenciada no campo se ampliou de 9.306 para 10.498. Em relação à dependência administrativa, entre os anos de 2012 e 2015 a rede municipal era responsável por 80% das escolas com localização diferenciada, enquanto a rede estadual abrigava 18%, não havendo grandes variações entre os dois anos.

Os estados do Amazonas, Pará e Maranhão abrigavam, em 2015, 47,5% das referidas escolas. Já o Distrito Federal, por sua vez, era a única unidade da Federação sem instituições com localização diferenciada (Gráfico 1).



Gráfico 1 Escolas com localização diferenciada, por unidade da Federação – Brasil – 2015

Fonte: Elaborado pelo autor com base no Censo da Educação Básica.

No que se refere aos indicadores de qualidade, uma breve análise do Ideb, em específico dos anos iniciais e finais do ensino fundamental, aponta que era pequeno o número de escolas com localização diferenciada que possuíam o indicador (Tabela 2). Em 2013, 399 escolas com localização diferenciada e oferta de anos iniciais do ensino fundamental apresentavam Ideb: a média geral era de 3,7, com mediana de 3,5 e desvio-padrão de 1,02. Em relação aos anos finais, 390 escolas dispunham de dados e a média do Ideb era de 3,3, com mediana próxima à média e desvio-padrão de 0,74. Para os anos iniciais, 67,6% das escolas detinham um Ideb igual ou inferior a 4. Nos anos finais, esse percentual se ampliava, de modo que 83,9% das escolas com localização diferenciada apresentavam índice igual a ou menor que 4 (Tabela 2). Tais valores precisam ser considerados com ressalvas, pois, para a composição do indicador, é exigido um número mínimo de alunos na série avaliada, condição nem sempre atendida pelas escolas com localização diferenciada.

Tabela 2 Faixa de Ideb para os anos iniciais e finais do ensino fundamental em escolas com localização diferenciada – Brasil – 2013

Ideb	Escolas	
	Ensino fundamental (anos iniciais)	Ensino fundamental (anos finais)
	%	%
De 1,0 a 2,0	2,0	4,1
De 2,1 a 3,0	27,3	32,5
De 3,1 a 4,0	38,3	47,3
De 4,1 a 5,0	19,8	15,4
De 5,1 a 6,0	10,8	0,7
6,1 ou mais	1,8	0,0
Escolas com Ideb	399	390

Fonte: Elaborada pelo autor com base no Censo da Educação Básica e no Índice de Desenvolvimento da Educação Básica (IDEB).

ESTUDANTES EM ESCOLAS COM “LOCALIZAÇÃO DIFERENCIADA”

A oferta de uma educação contextualizada constitui uma das potencialidades das escolas com localização diferenciada. Em 2015, 1.047.209 matrículas na educação básica ocorriam em tais instituições. Dessas, 80,1% eram ocupadas por estudantes residentes no campo e 19,9% nas áreas urbanas. Escolas em áreas de assentamento, em terras indígenas

e em áreas remanescentes de quilombolas registravam o maior percentual de jovens residentes no campo (Tabela 3). Nas escolas em unidades de uso sustentável, os estudantes residentes nas áreas urbanas compunham a maioria das matrículas: 70,8% em 2012 e 62,5% em 2015. Já em relação às escolas em unidades de uso sustentável em área remanescente de quilombos, em 2012, 57,1% dos matriculados residiam nas áreas urbanas e, em 2015, esse percentual caiu para 44,4% (Tabela 3).

Tabela 3 Distribuição das matrículas em escolas com localização diferenciada, por localização de residência do estudante – Brasil – 2012/2015

Localização da escola em que ocorre a matrícula	Local de residência do estudante							
	2012				2015			
	Urbano		Rural		Urbano		Rural	
	Total	%	Total	%	Total	%	Total	%
Área de assentamento	23.232	5,9	369.124	94,1	23.855	6,0	374.843	94,0
Terra indígena	14.324	7,0	188.911	93,0	17.862	7,8	212.289	92,2
Área remanescente de quilombos	42.550	21,0	159.840	79,0	54.031	23,0	180.985	77,0
Unidade de uso sustentável	57.762	70,8	23.861	29,2	106.291	62,5	63.708	37,5
Unidade de uso sustentável em terra indígena	825	32,3	1.727	67,7	1.021	52,2	936	47,8
Unidade de uso sustentável em área remanescente de quilombos	6.050	57,1	4.547	42,9	5.061	44,4	6.327	55,6
Total	144.743	16,2	748.010	83,8	208.121	19,9	839.088	80,1

Fonte: Elaborada pelo autor com base no Censo da Educação Básica.

Em relação à etapa de ensino, a maior parte das matrículas se concentrava, em 2012, no ensino fundamental, cuja representatividade era de 72,2%, enquanto a educação infantil representava 11,9%, o ensino médio, 5,8% e o EJA fundamental, 9,7% das matrículas. Ao se cotejarem esses dados com a situação em 2015, verifica-se que ocorreu uma redução no percentual dos matriculados no ensino fundamental, que passaram a representar 69,6% das matrículas, e uma ampliação no ensino médio e na educação infantil, que atingiram 6,9% e 13,6% das matrículas, respectivamente. A dependência administrativa das matrículas se concentrava nas redes municipais – 73,0% em ambos os períodos –, seguidas pelas redes estaduais, que abrigavam 25% das matrículas, em 2012, e 24,9%, em 2015.

O Censo da Educação Básica possui um elevado número de dados não declarados para o quesito raça/cor – na faixa de 36,0%, em 2012, e de 31,0%, em 2015. Todavia, observar a distribuição dessa variável nas escolas com localização diferenciada pode ajudar a compreender os grupos que as frequentam. O percentual de raça/cor não declarada é menor em escolas em terras indígenas – 15,3%, em 2012, e 11,0%, em 2015. No mesmo ano, por exemplo, o percentual de pretos era maior em áreas remanescentes de quilombolas – 13,0% –, já, nas escolas em áreas de assentamento, o maior percentual era de pardos – 50,8% (Tabela 4).

Tabela 4 Percentual de matrículas em escolas com localização diferenciada, por raça/cor – Brasil – 2012/2015

Localização da escola	Raça/cor											
	Não declarada		Branca		Preta		Parda		Amarela		Indígena	
	2012	2015	2012	2015	2012	2015	2012	2015	2012	2015	2012	2015
Área de assentamento	38,0	34,9	10,9	11,2	2,5	2,3	47,9	50,8	0,4	0,4	0,3	0,4
Terra indígena	15,3	11,0	0,8	0,8	0,3	0,2	7,8	8,4	0,2	0,1	75,6	79,4
Área remanescente de quilombos	36,5	30,7	7,1	8,1	12,9	13,0	42,9	47,7	0,4	0,4	0,3	0,2
Unidade de uso sustentável	35,2	26,0	33,3	31,9	2,8	4,7	28,3	36,9	0,2	0,3	0,1	0,2
Unidade de uso sustentável em terra indígena	30,9	15,5	1,1	19,1	0,3	2,5	10,6	17,2	0,1	0,00	57,1	45,0
Unidade de uso sustentável em área remanescente de quilombos	49,1	22,9	8,1	17,7	8,6	11,8	33,9	46,8	0,2	0,6	0,1	0,2

Fonte: Elaborada pelo autor com base no Censo da Educação Básica.

CONSIDERAÇÕES FINAIS

A comparação entre os anos de 2012 e 2015 revela que ocorreu uma ampliação no número de escolas com localização diferenciada, bem como no total de matrículas em tais instituições, o que pode indicar a ampliação do acesso a uma educação contextualizada ou até

mesmo uma redução no sub-registro de tais instituições pelos instrumentos de coleta. A maior parte dessas escolas se distribuía pelas áreas rurais, espaço em que viviam, em 2015, mais de 80,0% dos estudantes matriculados em tais instituições, sobretudo em escolas localizadas em áreas de assentamento e terras indígenas.

Ao se considerarem apenas as unidades que possuem Ideb, nota-se a predominância de escolas com índice inferior a 4, tanto nos anos iniciais como finais. Resta lembrar que o PNE, em sua meta 7, estabelece como objetivo, para 2015, um Ideb de 5,2 para os anos iniciais do ensino fundamental e de 4,7 para os finais. Caracterizações mais aprofundadas precisam ser feitas, a fim de compreender se a oferta de educação em escolas com localização diferenciada tem sido feita com garantia de qualidade.

REFERÊNCIAS

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). *Censo Escolar da Educação Básica 2015*: caderno de instruções. Brasília, 2015a.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). *Microdados do Censo Escolar (2012-2015)*. Brasília, 2015b. Disponível em: <<http://portal.inep.gov.br/basica-levantamentos-acessar>>. Acesso em: 30 maio 2016.

BRASIL. Lei nº 13.005, de 25 de junho de 2014. Aprova o Plano Nacional de Educação- PNE e dá outras providências. *Diário Oficial da União*, Brasília, DF, 16 jun. 2014. Disponível em: http://www.planalto.gov.br/CCIVIL_03/_Ato2011-2014/2014/Lei/L13005.htm.

FUNDAÇÃO CULTURAL PALMARES. *Quadro geral de comunidades remanescentes de quilombos (CRQs)*. 2015. Disponível em: <<http://www.palmares.gov.br/wp-content/uploads/crq/quadro-geral-por-estado-ate-23-02-2015.pdf>>.

BRASIL. Instituto Nacional de Colonização e Reforma Agrária (Incra). *Incra nos estados: informações gerais sobre os assentamentos da reforma agrária*. 2015. Disponível em: <<http://painel.incra.gov.br/sistemas/index.php>>.

BRASIL. Fundação Nacional do Índio (Funai). *Modalidades de terras indígenas*. 2015. Disponível em: <<http://www.funai.gov.br/index.php/indios-no-brasil/terras-indigenas>>.

CONCORDÂNCIA E DISCORDÂNCIA NA DECLARAÇÃO RACIAL DE ESTUDANTES: CRUZANDO DADOS DO CENSO ESCOLAR, SAEB E ENEM*

ADOLFO SAMUEL DE OLIVEIRA**

ADRIANO SOUZA SENKEVICS***

ROSILENE CERR****

PAMELLA SADA DIAS EDOKAWA*****



INTRODUÇÃO

Há muito se reconhece que o sistema de classificação racial no Brasil é caracterizado pela predominância do fenótipo sobre outros critérios de classificação e, conseqüentemente, por certa fluidez que torna ambígua e complexa a tarefa de designar a população brasileira em grupos de cor ou raça. Entretanto, obter informações acerca do pertencimento racial dos sujeitos, mesmo com as dificuldades que tal ação enseja, se apresenta como uma ferramenta indispensável para a compreensão das desigualdades raciais. Em razão disso, é fundamental colocar em permanente discussão o processo de produção de dados raciais, visando a aprimorá-los.

Quando se trata do método de classificação racial, a literatura da área distingue os conceitos de auto e heterodeclarado, para tratar do pertencimento atribuído pelo próprio

**Doutor em Educação pela Universidade de São Paulo (USP). Pesquisador da Diretoria de Avaliação da Educação Básica do Inep (Daeb/Inep). E-mail: adolfo.oliveira@inep.gov.br

***Mestre em Educação pela Universidade de São Paulo (USP). Pesquisador da Diretoria de Avaliação da Educação Básica do Inep (Daeb/Inep). E-mail: adriano.senkevics@inep.gov.br

****Mestre em Gestão da Avaliação Educacional pela Universidade Federal de Santa Catarina (UFSC). Pesquisadora da Diretoria de Avaliação da Educação Básica do Inep (Daeb/Inep). E-mail: rosilene.cerri@inep.gov.br

*****Bacharel em Estatística pela Universidade de Brasília (UnB). Analista Judiciária – Especialidade Estatística do Supremo Tribunal Federal (STF). E-mail: pamella.edokawa@stf.jus.br

* Este texto é uma versão resumida de trabalho apresentado no 5º Colóquio Luso-Brasileiro de Sociologia da Educação, organizado pela Universidade de São Paulo, em abril de 2016. O texto completo foi submetido à publicação nos anais do evento.

sujeito ou por outrem (Osório, 2003). No tocante às pesquisas do Inep referentes aos estudantes, conforme evidencia o Quadro 1, a autodeclaração é a técnica de coleta empregada nos questionários¹ do Sistema de Avaliação da Educação Básica (Saeb), nos formulários do Exame Nacional do Ensino Médio (Enem) e do Censo Escolar da Educação Básica, quando este diz respeito aos alunos que possuem 16 anos de idade completos ou mais (para os alunos abaixo dessa idade, a coleta é heterodeclarada). Outro ponto importante é diferenciar o método de preenchimento do instrumento de coleta de dados entre preenchimento pelo próprio sujeito (autopreenchido) ou por terceiros (heteropreenchido) – esse quesito diferencia os instrumentos do Saeb e do Enem, preenchidos pelos próprios respondentes, dos formulários do Censo Escolar, preenchidos por membros das equipes escolares ou das secretarias de ensino com base nos dados de matrículas.

Quadro 1 Classificação das pesquisas do Inep segundo o tipo de instrumento, declaração e preenchimento do item cor/raça

Pesquisa	Instrumento	Declaração racial	Preenchimento
Saeb	Questionário	Autodeclarado	Autopreenchido
Enem	Formulário	Autodeclarado	Autopreenchido
Censo Escolar	Formulário	Hetero/autodeclarado	Heteropreenchido

Fonte: Adaptado de Senkevics, Machado e Oliveira (2016).

Em vista das diferenças metodológicas na obtenção dos dados raciais nesses instrumentos, o objetivo deste artigo é analisar o grau de concordância e discordância na classificação racial, obtido por meio de cruzamento dos dados de cor/raça referentes aos mesmos indivíduos, em determinado ano, de distintos instrumentos de coleta. Com isso, espera-se levantar informações sobre o sistema classificatório de cor/raça e subsidiar uma reflexão metodológica sobre a produção de dados raciais em educação, tendo em vista o aprimoramento da qualidade da informação produzida.

METODOLOGIA

Para realizar este estudo, cruzamos as seguintes bases de dados, referentes ao ano de 2013: Saeb e Censo Escolar, no caso dos estudantes matriculados nos 5º e 9º anos do

³ Para fins operacionais, neste texto diferenciamos formulário de questionário da seguinte forma: de modo geral, o primeiro refere-se à coleta de dados por meio de campos que apenas indicam qual informação deve ser preenchida, e o segundo, a partir de perguntas ou de orientações para a resposta dos itens, estabelecendo, assim, um nível maior de interatividade entre o instrumento e o respondente.

ensino fundamental (EF), e Enem, Censo Escolar e Saeb, no caso dos concluintes da 3ª série do ensino médio (EM). Para unir essas bases, foram utilizadas diferentes variáveis de identificação, dentre as quais o código do aluno no Censo Escolar, o Identificador (ID) do estudante no Saeb e o número do CPF. Por se tratar de informações pessoais protegidas, foi solicitada a autorização do Inep no âmbito da Portaria nº 467, de 19 de setembro de 2014, que permite o uso dessas informações para estudos com fins científicos e assevera que os resultados divulgados não permitam identificar os sujeitos presentes em tais bases de dados. Nesse sentido, é importante frisar que os resultados aqui apresentados estão em conformidade com esses preceitos legais e éticos, na medida em que não oferecem risco de quebra do anonimato nem de publicização das informações pessoais.

Em razão das diferentes possibilidades de cruzamento das bases de dados para públicos distintos, foram produzidas quatro bases para o nosso estudo, de acordo com as informações do Quadro 2. Em todas essas bases, utilizaram-se para a análise do item cor/raça apenas os estudantes que obedeciam a dois critérios: primeiramente, não apresentavam dados ausentes (*missing*) para o quesito cor/raça no Saeb e/ou Enem. Em segundo lugar, mantivemos na base de dados apenas os indivíduos que apresentavam uma declaração racial válida no Censo Escolar. Dessa maneira, foi preciso retirar os alunos cujo campo cor/raça no Censo constasse como “não declarada”. Com esses filtros, reduziu-se substancialmente o quantitativo de alunos para a análise.

Quadro 2 Bases de dados produzidas para esta pesquisa a partir de cruzamentos de bases do Inep

Série/ano	Cruzamentos de dados	Quantitativo de alunos	
		Total	Análise de cor/raça
5º ano do EF	Saeb x Censo Escolar	2.303.221	1.322.077 (57%)
9º ano do EF	Saeb x Censo Escolar	2.709.147	1.155.997 (43%)
Concluintes EM	Enem x Censo Escolar	2.128.081	864.277 (41%)
Concluintes EM	Saeb x Enem x Censo Escolar	129.687	35.072 (27%)

Fonte: Elaborado pelos autores com base em dados do Saeb, Censo Escolar e Enem.

Além da variável cor/raça, foram acrescentadas outras duas variáveis – sexo e mês de nascimento –, para fins de controle na comparação dos instrumentos, considerando que a resposta dos alunos a esses dois quesitos deveria ser idêntica a despeito do instrumento analisado.

RESULTADOS

Em uma primeira leitura do grau de concordância e discordância das variáveis analisadas, percebe-se, de acordo com a Tabela 1, uma tendência de as discordâncias reduzirem à medida que a idade do aluno aumenta. Nota-se que a discordância na classificação racial cai de 47,7% para 30,6% entre as séries analisadas. Esses valores, contudo, sempre se apresentaram inferiores para sexo e mês de nascimento.

Tabela 1 Concordância e discordância (%) quanto ao sexo, mês de nascimento e cor/raça entre estudantes dos 5º e 9º anos do ensino fundamental (Saeb x Censo) e concluintes do ensino médio (Enem x Censo) – Brasil – 2013

Variável	Comparação	5º ano EF	9º ano EF	Concluintes EM
Sexo	Discordância	3,4	1,4	0,9
	Concordância	96,6	98,6	99,1
Mês de Nascimento	Discordância	7,4	2,4	0,7
	Concordância	92,6	97,6	99,3
Cor/Raça	Discordância	47,7	40,6	30,6
	Concordância	52,3	59,4	69,4

Fonte: Elaborada pelos autores com base em dados do Saeb, Censo Escolar e Enem.

Com relação aos estudantes do 5º ano do EF, é possível notar uma tendência de branqueamento da classificação racial no Censo Escolar, em comparação ao Saeb. Pela Tabela 2, vê-se que, embora 30,8% dos respondentes tenham se classificado na cor branca no Saeb, esse índice se eleva para 44,0% no Censo. Ademais, percebe-se que 22,9% dos indivíduos estão classificados como brancos nos dois levantamentos. É curioso notar que 13,9% das crianças classificadas como brancas pelos seus pais ou responsáveis no Censo Escolar atribuíram a si mesmas a cor parda no Saeb – por outro lado, apenas 7,5% das crianças que se declararam brancas no Saeb foram classificadas como pardas no Censo. Logo, é possível perceber que é mais provável uma criança autodeclarada parda ser entendida como branca pelos seus pais ou responsáveis do que o contrário.

Tabela 2 Declaração racial (%) dos alunos do 5º ano do EF no Saeb e Censo Escolar – Brasil – 2013

5º ano EF Cor/raça Branca		Censo Escolar					
		Branca	Preta	Parda	Amarela	Indígena	Total
Saeb (Aneb/Anresc)	Branca	22,9	0,3	7,5	0,1	0,0	30,8
	Preta	1,2	2,0	6,1	0,0	0,0	9,3
	Parda	13,9	1,6	27,1	0,2	0,1	42,9
	Amarela	1,0	0,1	1,2	0,1	0,0	2,3
	Indígena	0,6	0,2	1,7	0,0	0,2	2,7
	Não sei	4,3	0,6	6,8	0,1	0,0	11,9
	Total	44,0	4,8	50,2	0,5	0,4	100,0

Fonte: Elaborada pelos autores com base em dados do Saeb e do Censo Escolar.

Tendências similares se repetem se considerarmos a declaração racial de estudantes do 9º ano do EF que realizaram o Saeb (Aneb/Anresc) em 2013, bem como quanto ao pareamento das bases de dados do Enem e do Censo Escolar para os concluintes do EM (resultados não apresentados neste texto).

De agora em diante, analisaremos o nível de concordância entre a declaração racial de um mesmo indivíduo concluinte do EM, no ano de 2013, considerando uma base de dados formada pela junção do Censo Escolar, Saeb (Aneb) e Enem, como se vê na Tabela 3.

Tabela 3 Frequência absoluta (N) e relativa (%) da informação racial dos concluintes do EM no Censo Escolar, Enem e Saeb (Aneb) – Brasil – 2013

		Branca	Preta	Parda	Amarela	Indígena	Não declarada	Não sei	Total
Censo	N	15.439	1.181	17.976	326	150	-	-	35.072
	%	44,0	3,4	51,3	0,9	0,4	-	-	100,0
Enem	N	14.191	3.444	15.822	841	293	481	-	35.072
	%	40,5	9,8	45,1	2,4	0,8	1,4	-	100,0
Saeb	N	12.884	3.480	16.151	1.473	469	-	615	35.072
	%	36,7	9,9	46,1	4,2	1,3	-	1,8	100,0

Fonte: Elaborada pelos autores com base em dados do Saeb, Censo Escolar e Enem.

Embora os alunos sejam os mesmos, há, no Censo Escolar, em relação ao Enem e ao Saeb, um percentual maior de brancos e menor de pretos. O percentual de estudantes classificados como amarelos e indígenas no Censo também é menor que os assim designados no Enem e no Saeb, ao passo que, no que diz respeito aos pardos, essa relação se inverte.

De modo geral, parece existir uma tendência de branqueamento desse conjunto de alunos quando se observa a classificação racial no Censo, cotejado com os outros dois levantamentos.

Para compreender o nível de discordância na declaração racial do mesmo estudante nessa base, analisamos os levantamentos de maneira pareada e, depois, triangulada, tal como se vê na Tabela 4.

Tabela 4 Discordância e concordância da informação racial dos concluintes do EM no Censo Escolar, Enem e Saeb – Brasil – 2013

	Censo x Saeb		Censo x Enem		Enem x Saeb		Saeb x Enem x Censo	
	N	%	N	%	N	%	N	%
Discordância	12.460	35,5	11.540	32,9	7.660	21,8	15.161	43,2
Concordância	22.612	64,5	23.532	67,1	27.412	78,2	19.911	56,8
Total	35.072	100,0	35.072	100,0	35.072	100,0	35.072	100,0

Fonte: Elaborada pelos autores com base em dados do Saeb, Censo Escolar e Enem.

Nota-se que o nível de concordância entre as informações coletadas por um instrumento heteropreenchido, que pode ser tanto auto quanto heterodeclarado – no caso do formulário do Censo –, e por outro autopreenchido e autodeclarado – no caso do questionário do Saeb ou do formulário do Enem – é bastante próximo e gira em torno de 66%. Já a concordância entre os dois instrumentos autopreenchidos e autodeclarados é maior, chegando a 78,2%.

A fim de verificar quanto cada um dos levantamentos contribui para a discordância total (43,2%), realizamos a seguinte análise: uma decomposição em discordâncias parciais (aquela em que a informação é divergente em apenas um dos levantamentos) e em discordância total (aquela em que a cor ou raça é diferente em cada um dos três levantamentos), conforme mostra a Tabela 5.

Tabela 5 Decomposição da discordância da informação racial na triangulação do Censo Escolar, Enem e Saeb – Brasil – 2013

Instrumento(s) em que ocorre(m) a discordância	Saeb	Enem	Censo	Saeb, Enem e Censo	Total
N	3.621	2.701	7.501	1.338	15.161
%	10,3	7,7	21,4	3,8	43,2

Fonte: Elaborada pelos autores com base em dados do Saeb, Censo Escolar e Enem.

Podemos observar que a maior fonte de discordância parcial é o Censo (21%), seguido pelo Saeb (10%) e pelo Enem (8%), e que a discordância total é bastante baixa (inferior a 4%). Os resultados da triangulação dos levantamentos, do mesmo modo que os decorrentes das análises relativas ao pareamento desses, apontam que a menor concordância advém do Censo.

CONSIDERAÇÕES FINAIS

Embora a pesquisa esteja em andamento, é possível vislumbrar, a partir das bases analisadas, alguns resultados preliminares: a concordância na declaração racial tende a aumentar à medida que avançam os anos de escolarização e quando os instrumentos usados são autodeclarados e autopreenchidos (Saeb e Enem), e tende a diminuir quando se cruzam os dados coletados por um instrumento autopreenchido (Saeb ou Enem) e um heteropreenchido (Censo), principalmente para os pretos, amarelos e indígenas. De modo geral, as discordâncias nas declarações raciais têm apontado uma tendência de branqueamento dos estudantes no Censo Escolar, o qual pode estar relacionado tanto ao fato de serem, na maioria das vezes, os adultos que classificam racialmente as crianças e jovens (Carvalho, 2005) quanto por ser adotado, nesse recenseamento, um método de heterodeclaração racial (Osório, 2003).

A respeito da produção dos dados raciais, um dos principais desafios que temos à frente é trabalhar pela redução da não declaração racial no Censo Escolar. Ao lado disso, é importante enfatizar que essa coleta solicite, para os alunos matriculados no EM e que tenham pelo menos 16 anos de idade, que o campo cor/raça seja preenchido a partir da declaração do próprio estudante. Em relação às avaliações componentes do Saeb, é preciso conhecer em maior profundidade o processo de coleta dos dados, em função da alta taxa de não preenchimento do questionário pelos alunos do EF e, em especial, do EM.

Essas medidas, em conjunto, poderiam aprimorar a qualidade da coleta desses dados e, talvez, elevar o nível de concordância entre as informações do mesmo indivíduo nos três levantamentos. Não obstante, é preciso frisar que as discordâncias não invalidam os dados produzidos, uma vez que o dado, por ser construído e não meramente “coletado”, inevitavelmente sofre interferência da metodologia empregada em cada levantamento, bem como de outras fontes de variação, tais como as mudanças na percepção de pertencimento racial dos indivíduos ao longo do tempo.

Preende-se, com a melhoria da qualidade dos dados produzidos e, por conseguinte, das informações geradas a partir deles, possibilitar diagnósticos mais finos e subsidiar a formulação de políticas públicas voltadas para a elevação da qualidade da educação básica e o combate às desigualdades escolares e raciais.

REFERÊNCIAS

CARVALHO, M. P. Quem é negro, quem é branco: desempenho escolar e classificação racial de alunos. *Revista Brasileira de Educação*, n. 28, p. 77-95, jan./abr. 2005.

SENKEVICS, A. S.; MACHADO, T. S.; OLIVEIRA, A. S. *A cor ou raça nas estatísticas educacionais: uma análise dos instrumentos de pesquisa do Inep*. Brasília: Inep, 2016. (Texto para Discussão, n. 41).

OSÓRIO, R. G. *O sistema classificatório de “cor e raça” do IBGE*. Rio de Janeiro: Ipea, 2003. (Texto para Discussão, n. 996).

AS PROVAS DAS QUATRO ÁREAS DO ENEM VISTAS COMO UMA PROVA ÚNICA NA ÓTICA DE MODELOS DA TEORIA DA RESPOSTA AO ITEM UNI E MULTIDIMENSIONAL

NARA NÚBIA VIEIRA*



INTRODUÇÃO

A Teoria da Resposta ao Item (TRI) vem sendo amplamente utilizada em avaliações educacionais de larga escala. Internacionalmente, como exemplos há o Programme for International Student Assessment (Pisa), que usa a TRI desde 2000, e o Test of English as a Foreign Language (Toefl), desde 1978. No Brasil, o Sistema de Avaliação da Educação Básica (Saeb), realizado pelo Inep, adotou a TRI em 1995 e hoje a utiliza também no Exame Nacional para Certificação de Competências de Jovens e Adultos (Encceja), na Prova Brasil e no Exame Nacional do Ensino Médio (Enem).

O Enem é de caráter voluntário e “foi criado em 1998 com o objetivo de avaliar o desempenho do estudante ao fim da educação básica, buscando contribuir para a melhoria da qualidade desse nível de escolaridade” (Brasil. Inep, 2015a). No entanto, somente em 2009 que o exame passou a adotar a metodologia da TRI, a qual permite a comparabilidade dos resultados da avaliação inclusive entre diferentes edições do exame.

Hoje, o exame é formado por uma redação e quatro provas objetivas, contendo cada uma 45 questões de múltipla escolha. As cinco provas avaliam as seguintes áreas do conhecimento com seus respectivos componentes curriculares (Brasil. Inep, 2015b):

*Estatística pela UnB e Mestre em Métodos e Gestão em Avaliação pela UFSC. Pesquisadora-Tecnologista da Diretoria de Estatísticas Educacionais do Inep (Deed/Inep). E-mail: nara.vieira@inep.gov.br

Tabela 1 Tabela relacional das áreas do conhecimento e componentes curriculares das provas do Enem

Áreas de Conhecimento	Componentes Curriculares
Ciências Humanas e suas Tecnologias (CH)	História, Geografia, Filosofia e Sociologia
Ciências da Natureza e suas Tecnologias (CN)	Química, Física e Biologia
Linguagens, Códigos e suas Tecnologias (LC) e Redação (RD)	Língua Portuguesa, Literatura, Língua Estrangeira (Inglês ou Espanhol), Artes, Educação Física e Tecnologias da Informação e Comunicação
Matemática e suas Tecnologias (MT)	Matemática

Fonte: Brasil. Inep, 2015b.

O modelo da TRI utilizado hoje no Enem é o unidimensional de três parâmetros. No entanto, como o exame avalia quatro áreas do conhecimento, foram feitos quatro modelos, um para cada área.

Neste contexto, dado que provas multidisciplinares estão sendo estudadas com modelos uni e multidimensionais de TRI (Barbetta et al., 2014; Quaresma, 2014), surgiu o interesse de estudar a dimensionalidade do Enem como prova única composta por 180 itens (45 de cada prova objetiva).

APLICAÇÃO E ANÁLISE DOS MODELOS

Para realizar este trabalho foram utilizados os microdados do Enem 2012. Os cálculos foram realizados a partir de uma amostra dos alunos concluintes do ensino médio regular seriado, que estiveram presentes em todas as quatro provas objetivas do Exame e que receberam o primeiro caderno de provas do dia.

Com base nesta amostra, pretendeu-se aplicar um modelo unidimensional da TRI na prova agregada de 180 itens. Como uma das premissas desse modelo é a existência de um traço latente dominante na prova, realizou-se, primeiramente, uma análise de componentes principais sobre a matriz de correlação tetracórica dos itens com o objetivo de identificar um fator dominante.

O Gráfico 1 mostra um resumo dessa análise, apresentando na ordenada os autovalores da matriz de correlação (os vinte maiores), que podem ser interpretados como variâncias explicadas por cada dimensão considerada. Por esse gráfico, verifica-se que a variância explicada pela primeira dimensão é muito maior que as demais, o que caracteriza uma dimensão dominante.

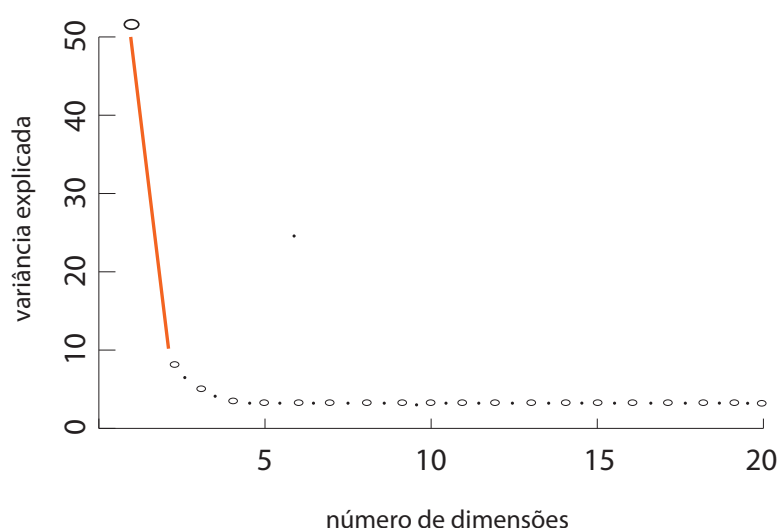


Gráfico 1 Screeplot da variância explicada pelo número de dimensões

Fonte: Elaborado pela autora com base nos microdados do Enem 2012.

Uma vez encontrados indícios de que na prova do Enem há um traço latente dominante, os dados foram ajustados à TRI unidimensional de três parâmetros, gerando uma escala com média zero e desvio-padrão igual a um.

A partir dos resultados, foi possível observar que não houve problemas na calibração dos parâmetros dos itens. A Tabela 2 foi elaborada para comparar a média e o desvio-padrão dos parâmetros ‘a’ (índice de discriminação), ‘b’ (dificuldade do item) e ‘c’ (probabilidade de um item ser respondido corretamente quando ele é considerado difícil para o avaliado – acerto casual) em cada uma das áreas do conhecimento.

Tabela 2 Média e desvio-padrão dos parâmetros ‘a’, ‘b’ e ‘c’

Prova	a		b		c	
	Média	DP	Média	DP	Média	DP
CH	2,08	0,88	0,68	1,16	0,19	0,08
CN	3,03	1,55	0,91	0,63	0,18	0,07
LC	1,84	0,63	0,56	0,73	0,18	0,09
MT	3,23	1,10	0,53	0,62	0,18	0,06

Fonte: Elaborada pela autora com base nos microdados do Enem 2012.

Com base na Tabela 2, é possível concluir que os itens de CN, em média, são os mais difíceis e que seus índices de discriminação são os de maior variabilidade entre todas as áreas, ou seja, há itens com discriminação alta e outros com baixa. Apesar de a área de MT ser a mais

fácil, é a que discrimina mais. A área de CH, em relação à dificuldade do item, é a que possui maior variabilidade, de forma que há itens muito fáceis e itens muito difíceis. A área de LC é a que discrimina menos e a segunda mais fácil.

Para verificar melhor a distribuição dos itens segundo seu índice de dificuldade, o Gráfico 2 foi elaborado para posicionar os itens segundo o valor do parâmetro 'b', considerando apenas aqueles que possuem valores de 'a' maiores que 0,7 e menores que 2,5 – bom índice de discriminação.

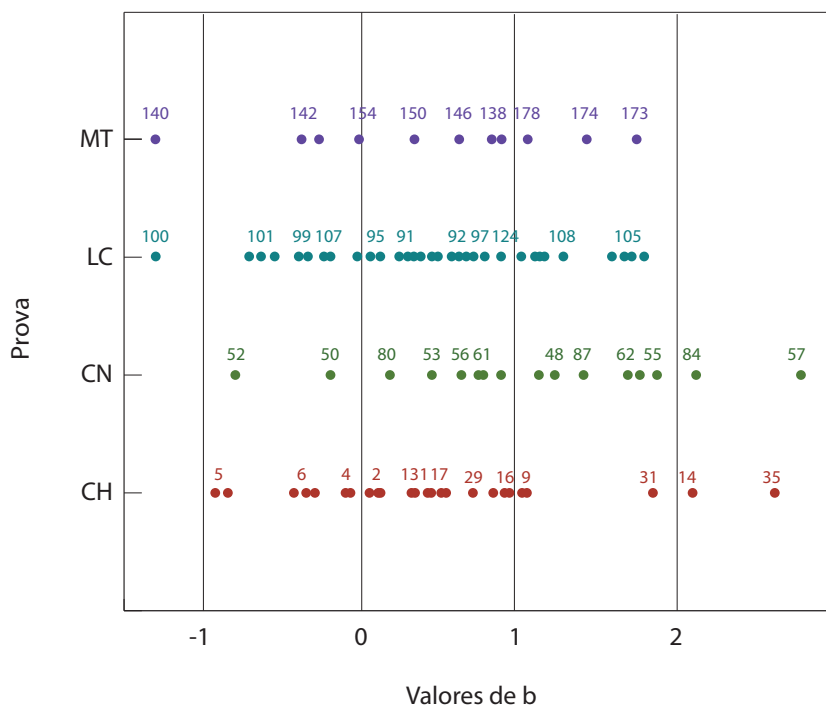


Gráfico 2 Posição dos itens que possuem valores de 'a' maiores que 0,7 e menores que 2,5 segundo valores de 'b', por área de conhecimento

Fonte: Elaborado pela autora com base nos microdados do Enem 2012.

Observa-se pelo Gráfico 2 que a prova de MT possui poucos itens com valor de $0,7 < a < 2,5$. Grande parte dos itens dessa prova que não apareceram no gráfico possui $a > 2,5$. Os itens que possuem um bom índice de discriminação estão em sua maioria distribuídos entre $-0,5 < b < 1,5$, o que indica que, para o avaliado responder corretamente a todos esses itens, é necessário que sua proficiência esteja 1,5 desvios-padrão acima da média dos avaliados.

A prova de LC é a que possui mais itens com $0,7 < a < 2,5$ – bom índice de discriminação – e esses itens estão bem distribuídos entre $-1 < b < 2$. Ou seja, uma pessoa que tenha proficiência igual a -1, por exemplo, teria uma alta probabilidade de acertar o item 100 e de errar os demais. Enquanto uma com proficiência igual a 2 teria uma alta probabilidade de acertar todos os itens.

Já na prova de CN, os itens estão mais concentrados em $0 < b < 2$, indicando que há itens mais difíceis, de modo que, para o avaliado respondê-los corretamente, sua proficiência deveria ser dois desvios-padrão acima da média.

Por fim, na prova de CH os itens estão mais concentrados para $b < 1$ com alguns pontos discrepantes posicionados com valores mais altos de 'b'. Isto é, em geral, uma pessoa com proficiência igual a um desvio-padrão acima da média teria baixa probabilidade de acertar os itens 31, 14 e 35, e alta para responder corretamente aos demais itens.

Já em relação ao parâmetro 'c', não há diferenças expressivas entre os valores das quatro áreas do conhecimento.

Em seguida, buscou-se verificar a correlação entre a média aritmética das proficiências das quatro provas estimadas pelo Inep com a proficiência calculada, considerando a prova única com 180 itens, a fim de observar se a primeira medida seria bem representada pela segunda.

O Gráfico 3 apresenta o diagrama de dispersão entre as duas medidas de proficiências.

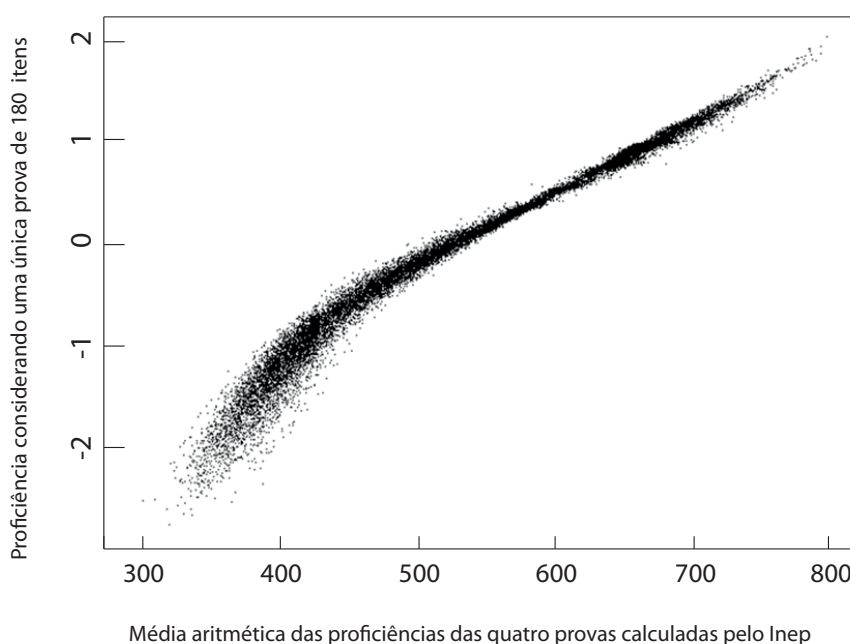


Gráfico 3 Diagrama de dispersão entre a proficiência, considerando uma prova única, e a média das quatro proficiências estimadas pelo Inep

Fonte: Elaborado pela autora com base nos microdados do Enem 2012.

De acordo com o Gráfico 3 e dado que o coeficiente de correlação de Pearson dessas medidas foi igual a 0,98, observa-se que há uma alta correlação entre essas variáveis. No entanto, pode-se perceber que, para as proficiências mais baixas, a associação entre as duas proficiências está mais dispersa.

Por esse motivo, fez-se necessário observar a quantidade de informação que o modelo deste estudo tem para cada nível de proficiência. O Gráfico 4 mostra a curva de informação do teste que indica as regiões em que o teste funciona melhor. Verifica-se que a curva é mais alta com θ entre 0 e 2, ou seja, o teste discrimina melhor os avaliados entre a média e a média mais dois desvios-padrão.

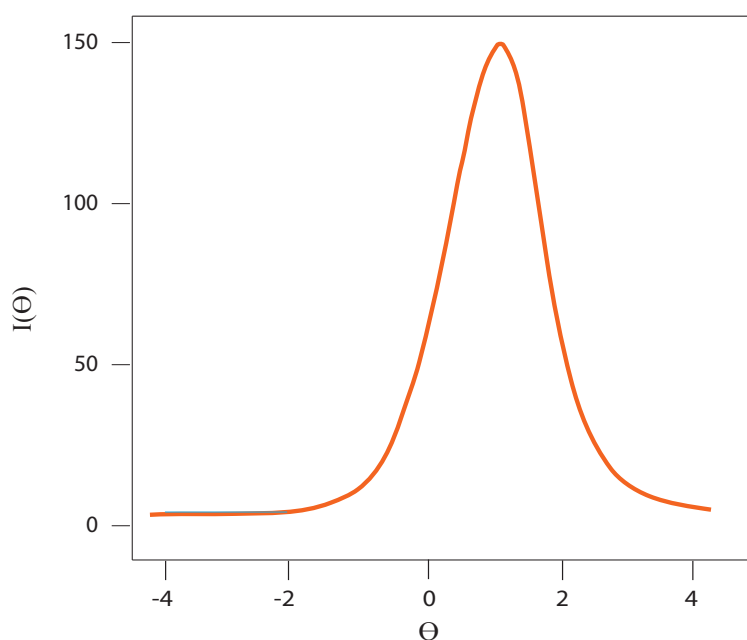


Gráfico 4 Curva de informação da prova agregada das quatro áreas: CH, CN, LC e MT

Fonte: Elaborado pela autora com base nos microdados do Enem 2012.

Este dado vai ao encontro do fato de os pontos do Gráfico 3 estarem mais dispersos para as pessoas de proficiência mais baixa e estarem mais próximos a partir do ponto de $\theta=0,5$.

Reckase (2009) menciona que os modelos unidimensionais têm vantagens por possuírem um formato matemático mais simples, por permitirem numerosas aplicações de exemplos e por existirem evidências de que ele é robusto à violação de suposições. Além disso, apesar de normalmente as avaliações exigirem que o estudante possua mais de uma proficiência para responder corretamente a um item, os modelos unidimensionais têm provado que, sob certas condições, representam bem sua proficiência.

No entanto, há uma necessidade de utilização de modelos de TRI mais complexos para aumentar sua capacidade de reflexão das interações entre estudantes e itens de forma mais acurada. Uma forma de aumentá-la é utilizar modelos de TRI que descrevam a interação de vetores de traços latentes com as características dos itens em vez de assumir um parâmetro de traço latente único. Como estes modelos possuem mais de um parâmetro para cada aluno, eles são chamados de teoria da resposta ao item multidimensional – TRIM.

Diversos modelos de TRIM foram aplicados à prova agregada de 180 itens. Entre estes, o modelo de TRI de quatro dimensões foi ajustado aos dados, determinando previamente que os itens de cada área pertencem a uma dimensão distinta, com o objetivo de verificar se as diferentes áreas do conhecimento são avaliadas como componentes distintos no Enem. No entanto, o modelo bidimensional no qual não foram delimitados os itens de cada uma das dimensões apresentou melhores resultados.

A Tabela 3 mostra a carga fatorial de todos os itens do modelo bidimensional. As cargas em células coloridas são aquelas em que a correlação do item com o fator correspondente é de moderada a alta (acima de 0,4 ou abaixo de -0,4).

Tabela 3 Carga fatorial dos itens do Enem para o modelo bidimensional

(continua)

Item	F1	F2	Item	F1	F2	Item	F1	F2	Item	F1	F2
1	0,14	0,69	46	0,38	0,51	91	0,08	0,69	136	0,70	0,25
2	0,07	0,75	47	0,56	0,36	92	0,30	0,54	137	0,65	0,29
3	0,06	0,13	48	0,37	0,47	93	0,01	0,87	138	0,60	0,16
4	0,16	0,68	49	0,50	0,45	94	0,08	0,70	139	0,70	0,22
5	0,04	0,61	50	0,40	0,28	95	0,17	0,41	140	0,62	0,09
6	0,06	0,70	51	0,30	0,65	96	0,24	0,58	141	0,60	0,29
7	0,12	0,78	52	0,06	0,45	97	0,04	0,64	142	0,64	0,17
8	0,14	0,73	53	0,33	0,52	98	0,03	0,28	143	0,52	0,33
9	0,23	0,63	54	0,65	0,24	99	0,19	0,62	144	0,72	0,20
10	0,06	0,32	55	0,68	0,17	100	0,06	0,65	145	0,73	0,23
11	0,25	0,69	56	0,35	0,43	101	-0,07	0,69	146	0,44	0,18
12	0,21	0,71	57	0,18	0,22	102	0,03	0,79	147	0,76	0,21
13	-0,01	0,68	58	0,57	0,35	103	-0,04	0,86	148	0,62	0,31
14	0,24	0,42	59	0,59	0,40	104	0,12	0,69	149	0,82	0,16
15	-0,03	0,79	60	0,72	0,28	105	0,06	0,61	150	0,51	-0,02
16	0,12	0,62	61	0,57	0,26	106	0,01	0,69	151	0,72	0,24
17	0,17	0,52	62	0,22	0,63	107	0,13	0,69	152	0,64	0,26
18	0,07	0,78	63	0,42	0,49	108	0,12	0,67	153	0,63	0,34
19	0,32	0,62	64	0,52	0,45	109	-0,06	0,45	154	0,73	0,08
20	0,33	0,59	65	0,37	0,60	110	-0,02	0,66	155	0,71	0,23
21	0,24	0,70	66	0,51	0,42	111	0,06	0,69	156	0,74	0,23
22	0,05	0,78	67	0,80	0,18	112	0,10	0,64	157	0,71	0,16
23	0,19	0,02	68	0,26	0,48	113	0,09	0,65	158	0,76	0,20
24	0,42	0,50	69	0,47	0,44	114	-0,06	0,69	159	0,62	0,30
25	0,08	0,38	70	0,45	0,39	115	0,09	0,79	160	0,90	-0,03
26	0,30	0,61	71	0,38	0,53	116	0,10	0,45	161	0,66	0,27
27	0,06	0,73	72	0,68	0,28	117	0,01	0,89	162	0,49	-0,04
28	0,09	0,74	73	0,89	0,01	118	-0,04	0,89	163	0,69	0,25

Tabela 3 Carga fatorial dos itens do Enem para o modelo bidimensional

(conclusão)

Item	F1	F2	Item	F1	F2	Item	F1	F2	Item	F1	F2
29	0,31	0,48	74	0,65	0,22	119	0,00	0,79	164	0,69	0,24
30	0,08	0,77	75	0,27	0,66	120	-0,02	0,83	165	0,79	0,22
31	0,08	0,77	76	0,45	0,49	121	-0,06	0,86	166	0,71	0,12
32	0,17	0,69	77	0,57	0,21	122	0,00	0,61	167	0,87	0,02
33	0,20	0,64	78	0,62	0,29	123	0,05	0,71	168	0,75	0,22
34	0,11	0,17	79	0,50	0,45	124	-0,02	0,87	169	0,75	0,19
35	-0,08	0,72	80	0,20	0,62	125	-0,06	0,51	170	0,68	0,17
36	0,35	0,60	81	0,39	0,54	126	-0,04	0,88	171	0,84	0,05
37	0,17	0,48	82	0,57	0,41	127	0,26	0,66	172	0,74	0,23
38	0,16	0,48	83	0,44	0,48	128	-0,01	0,82	173	0,54	0,33
39	0,14	0,75	84	0,32	0,21	129	0,03	0,53	174	0,66	0,14
40	0,03	0,82	85	0,21	0,51	130	0,08	0,85	175	0,55	0,36
41	-0,01	0,63	86	0,49	0,50	131	0,08	0,59	176	0,69	0,28
42	0,00	0,62	87	0,28	0,58	132	0,20	0,67	177	0,77	0,22
43	-0,04	0,63	88	0,49	0,44	133	0,04	0,76	178	0,66	0,19
44	0,02	0,77	89	0,42	0,52	134	-0,02	0,58	179	0,67	0,32
45	0,18	0,76	90	0,63	0,37	135	-0,01	0,56	180	0,70	0,29

Fonte: Elaborada pela autora com base nos microdados do Enem 2012.

Dessa forma, podemos verificar que, para a prova de CH, à exceção do item 23, as cargas fatoriais de maior peso estão no segundo fator, F2. Quanto à prova de CN, observa-se que há cargas fatoriais altas tanto no F1 como no F2. Além disso, verifica-se que há diversos itens em que a carga fatorial é em torno de 0,4 nos dois fatores. Sobre os itens de LC, observa-se que as maiores cargas fatoriais estão predominantemente no F2. Por fim, todos os itens da prova de MT estão moderada ou altamente correlacionados ao primeiro fator, F1.

Em geral, observou-se pelo modelo bidimensional que as provas de CH e de LC pertencem a uma mesma dimensão, enquanto a de MT pertence à outra. Já a prova de CN não possui os itens correlacionados a um mesmo fator, ou seja, possui itens com carga fatorial alta tanto em F1 como em F2. Sendo assim, infere-se que uma dimensão contém itens que exigem mais raciocínio lógico e outra, itens que estão mais relacionados à compreensão de textos.

A partir desse entendimento e do entendimento de que a prova de CN engloba itens de componentes curriculares distintos, como Química, Física e Biologia, buscou-se identificar os itens de cada um desses componentes para verificar se estes são representados por um fator.

A Tabela 4 destaca as cargas fatoriais dos itens da prova de CN por componente curricular. A marcação das cargas é a mesma usada na tabela anterior.

Tabela 4 Carga fatorial dos itens da prova de CN do Enem, por componente curricular

Item	F1	F2	Item	F1	F2	Item	F1	F2
Biologia			Física			Química		
Item	F1	F2	Item	F1	F2	Item	F1	F2
48	0,37	0,47	47	0,56	0,36	46	0,38	0,51
51	0,30	0,65	50	0,40	0,28	49	0,50	0,45
52	0,06	0,45	54	0,65	0,24	53	0,33	0,52
56	0,35	0,43	55	0,68	0,17	58	0,57	0,35
57	0,18	0,22	60	0,72	0,28	59	0,59	0,40
62	0,22	0,63	61	0,57	0,26	66	0,51	0,42
63	0,42	0,49	64	0,52	0,45	69	0,47	0,44
65	0,37	0,60	67	0,80	0,18	70	0,45	0,39
68	0,26	0,48	71	0,38	0,53	76	0,45	0,49
75	0,27	0,66	72	0,68	0,28	79	0,50	0,45
80	0,20	0,62	73	0,89	0,01	82	0,57	0,41
81	0,39	0,54	74	0,65	0,22	84	0,32	0,21
85	0,21	0,51	77	0,57	0,21	90	0,63	0,37
86	0,49	0,50	78	0,62	0,29			
87	0,28	0,58	83	0,44	0,48			
89	0,42	0,52	88	0,49	0,44			

Fonte: Elaborada pela autora com base nos microdados do Enem 2012.

Pela Tabela 4, observa-se que todas as questões de biologia têm a maior carga fatorial em F2 (fator mais relacionado à proficiência de leitura e compreensão de textos), no entanto, alguns itens possuem uma carga moderada em F1 também. Os itens de física, em sua maioria, possuem uma alta carga em F1 (fator mais relacionado ao raciocínio lógico). Já em relação aos itens de química, há alguns com alta carga em F1, outros em F2 e outros com cargas moderadas nos dois fatores.

Em seguida, os itens de todas as áreas foram posicionados segundo os dois traços latentes (raciocínio lógico e leitura e compreensão de textos). O Gráfico 5 apresenta a distribuição dos mesmos itens apresentados no Gráfico 2.

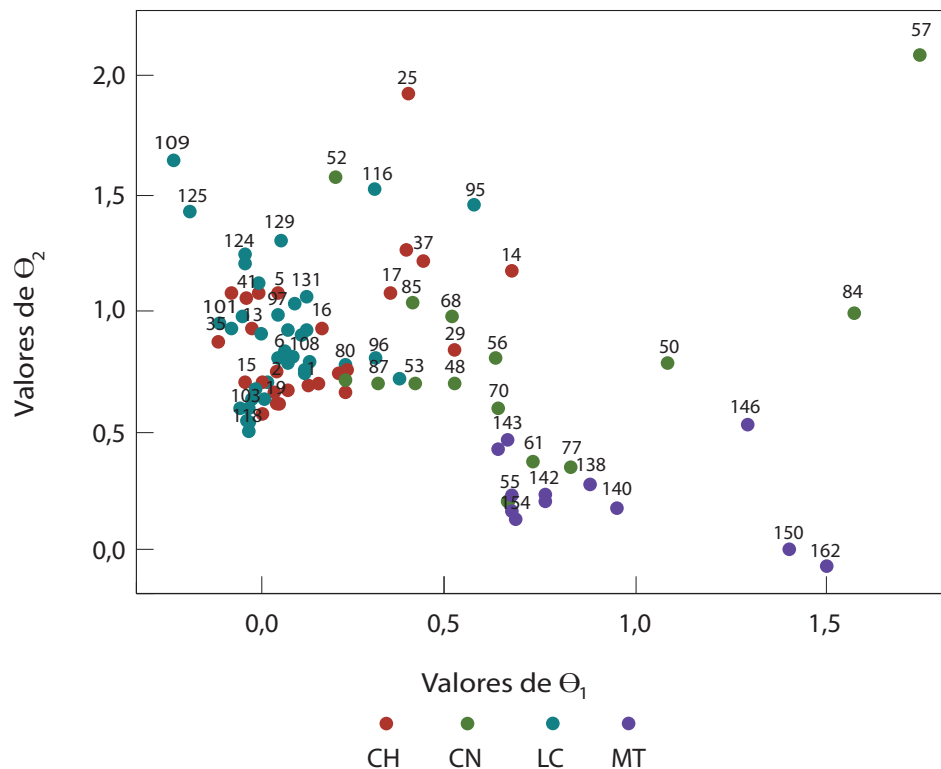


Gráfico 5 Posição de alguns itens – segundo valores de θ_1 e θ_2 – por área de conhecimento do item

Fonte: Elaborado pela autora com base nos microdados do Enem 2012.

Observa-se pelo Gráfico 5 que os itens das provas de CH e LC, para serem respondidos corretamente, exigem do aluno uma proficiência maior em leitura e interpretação de textos e menor em raciocínio lógico, ao passo que os itens das provas de CN e MT estabelecem o contrário – uma proficiência alta em raciocínio lógico e baixa em leitura e interpretação de textos.

Além disso, verifica-se que a maioria dos itens de CH e LC estão próximos uns dos outros, entre os quais se observa que os de LC exigem um pouco mais de compreensão de texto e os de CH exigem um pouco mais de raciocínio lógico.

CONCLUSÃO

De forma geral, observou-se que pelo modelo unidimensional todas as áreas do conhecimento podem ser agregadas em termos matemáticos e gerar uma medida unidimensional de proficiência. Isso porque, ao ajustar o modelo, a análise de componentes principais sobre a matriz de correlação tetracórica dos itens apresentou que o maior autovalor

é muito superior aos demais, demonstrando uma dimensão dominante, o que ratifica a hipótese de que a prova agregada das quatro áreas pode ser considerada essencialmente unidimensional. Ademais, foi identificado que a correlação entre as proficiências geradas pelo modelo estudado e a média das proficiências oficiais divulgadas pelo Inep é bastante alta, indicando que o traço latente único do avaliado é uma boa aproximação para esta outra medida. No entanto, o modelo unidimensional poderia apresentar uma desvantagem para as instituições de educação superior que queiram atribuir pesos diferentes para as provas em que elas julgam que o avaliado deve ter maior domínio para ingressar em um determinado curso.

Em seguida, foi gerado um modelo tetradimensional em que as dimensões eram compostas pelos itens de cada uma das provas do Enem. Todavia, ao se compararem os resultados com o bidimensional, aferiu-se que o modelo bidimensional estava mais bem ajustado aos dados. Nesse sentido, ao se considerarem quatro proficiências, pode-se concluir que essas não são totalmente determinadas pelos itens de cada área.

Por fim, verificou-se que por meio do modelo bidimensional é possível identificar uma perceptível diferenciação entre uma dimensão com itens que estão relacionados à leitura e interpretação de textos – em geral os de CH e LC – e outra ligada ao raciocínio lógico – via de regra, os de MT. Já os itens de CN oscilam entre esses dois fatores. Contudo, ao se considerar cada um dos componentes curriculares da prova de CN (biologia, física e química), foi possível identificar alguns padrões. Os itens de biologia normalmente têm a maior carga fatorial na dimensão de leitura e interpretação de textos. Os de física, na de raciocínio lógico, enquanto, no componente curricular química, há itens com alta carga em raciocínio lógico, outros em leitura e interpretação de textos e outros com cargas moderadas nos dois fatores.

REFERÊNCIAS

ANDRADE, D. F.; TAVARES, H. R.; VALLE, R. C. *Teoria da resposta ao item: conceitos e aplicações*. São Paulo: ABE, 2000.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Microdados para download: microdados Enem 2012. Disponível em: < <http://portal.inep.gov.br/basica-levantamentos-acessar>>. Acesso em: 26 set. 2015.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). *Sobre o Enem*. 2015a. Disponível em: <<http://portal.inep.gov.br/web/enem/sobre-o-enem>>. Acesso em: 26 set. 2015.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Edital nº 6, de 15 de maio de 2015. Exame Nacional do Ensino Médio- Enem 2015. Diário Oficial da União, Brasília, DF, 18 maio 2015b. Seção 3, p. 59.

BARBETTA, P. A. et al. Aplicação da teoria da resposta ao item uni e multidimensional. *Estudos em Avaliação Educacional*, São Paulo, v. 25, n. 57, p. 280–302, jan./abr. 2014.

QUARESMA, E. S. A. *Modelagem para construção de escalas avaliativas e classificatórias em exames seletivos usando a teoria da resposta ao item uni e multidimensional*. 2014. 187f. Tese (Doutorado) – Universidade de São Paulo, Piracicaba, 2014.

RECKASE, M. D. *Multidimensional item response theory*. New York: Springer, 2009.

O ENEM E O IMPACTO DE VARIÁVEIS SOCIOECONÔMICAS E EDUCACIONAIS

MAURÍCIO URBAN KLEINKE*



INTRODUÇÃO

Desde sua criação como avaliação em larga escala em 1998, o Enem vem sendo utilizado em processos seletivos de acesso ao ensino superior. De forma parcial até 2008 e como instrumento preponderante de 2009 até o presente, as universidades federais valem-se do desempenho nesse Exame para selecionar seus candidatos por meio do sistema MEC-Sisu.

Embora cada edição do Enem forneça considerável volume de dados, sua utilização para avaliar sistemas de ensino esbarra em dificuldades, pois o Exame é voluntário e não representa uma amostra estatisticamente representativa dos concluintes do ensino médio.

Tendo em perspectiva essas questões, analisamos os grupos socioeducacionais que apresentaram melhor desempenho nas provas de Ciências da Natureza e Matemática de 2009 até 2012, período em que teve início o chamado Novo Enem. Tal escolha deveu-se ao fato de que alunos com bom desempenho, em uma aproximação empírica, apresentam maiores chances de ingressar em cursos universitários socialmente valorizados.

Como indicador desse grupo, foi considerado o decil superior – estimando-se as notas brutas – nas provas de Ciências da Natureza e Matemática. Essa escolha baseia-se na hipótese de que os 10% com melhor desempenho no Exame disputarão as vagas dos cursos de alto *status* social, como Medicina, Engenharias e Direito. Em nossa proposta de trabalho, caracterizamos

*Doutor em Ciências pela Universidade Estadual de Campinas (Unicamp) e professor do Instituto de Física Gleb Wataghin (Unicamp). E-mail: kleinke@ifi.unicamp.br

o decil superior em comparação com a totalidade dos candidatos, identificando as diferenças entre as características sociais e educacionais dos dois grupos.

CAPITAL CULTURAL E EXAMES DE LARGA ESCALA

A história relativa à realização de exames de larga escala no Brasil é recente. Apesar disso, o Enem é um dos maiores exemplares do mundo, sendo superado apenas pelo GaoKao, da China. Contudo, existem poucos estudos sobre o impacto ou o efeito de variáveis extraclasse – e para além das trajetórias escolares – sobre o desempenho dos estudantes. Considerando os vínculos entre o Enem e o acesso ao ensino superior hoje presentes, compreender as relações entre o desempenho e as variáveis socioeconômicas constitui-se em importante fonte de informações com vistas a sua utilização na definição de políticas públicas.

Para compreender como os ambientes familiar e escolar afetam os estudantes, buscamos apoio no pensamento do sociólogo francês Pierre Bourdieu, que, na década de 1960, elaborou uma teoria sobre a relação entre desigualdades de desempenho escolar e classes sociais, denominada *capital cultural*.

Indo além da associação entre desempenho e aptidões intrínsecas dos estudantes, Bourdieu (1998) estabelece relações entre a herança cultural transmitida pela família e as taxas de êxito de diferentes grupos sociais. Em geral, qualidades tidas como “inatas” estão associadas a todo um investimento familiar em bens culturais. Em sua proposição, o autor associa o conceito de capital cultural a uma “hipótese indispensável para dar conta das desigualdades de desempenho escolar de crianças provenientes de diferentes classes sociais, relacionando o ‘sucesso escolar’ [...] à distribuição de capital cultural entre as classes” (Bourdieu, 1998).

Ao contrário do capital financeiro, que se consolida como medida concreta do grau de riqueza de um indivíduo, o conceito de capital cultural tem caráter mais abstrato. O capital financeiro está relacionado à materialidade, como obras de arte, recursos para cinema, teatro, música, ciência, tecnologias modernas etc. Entretanto, se não houver a apropriação cultural do que esses bens materiais têm a oferecer, não ocorre a transmissão de valores que podem impactar na interiorização dos conhecimentos escolares. É nesse ponto que a família desempenha um papel de destaque, pois é no ambiente em que o sujeito vive que essas trocas ocorrem. A escola, por sua vez, constitui-se em um meio no qual o capital cultural é valorizado, principalmente pelos professores, tornando-se um recurso de poder para alguns alunos (Bourdieu, 1998).

De acordo com Bourdieu (1998), o nível de instrução dos membros da família permite apenas situar o nível cultural destes. Não informa quais conteúdos nem as vias pelas quais eles são transferidos. Ressalte-se que, no caso brasileiro, a noção de capital cultural está fortemente atrelada ao capital financeiro.

AVALIAÇÃO DA EQUIDADE

Para avaliar a equidade do processo seletivo, selecionamos um conjunto de perguntas do questionário socioeconômico do Enem, reagrupando-as em quatro dimensões: pessoal, escolar, familiar e regional. A composição dessas dimensões é o que buscamos detalhar a seguir:

Dimensão Pessoal – engloba duas variáveis: o *sexo* (masculino ou feminino) e a *faixa etária*. A variável faixa etária apresenta quatro categorias: 1) Treineiros – formada por candidatos até 16 anos; 2) Em fase – entre 17 e 19 anos; 3) Jovens – com idade entre 20 e 25 anos; e 4) Adultos – acima de 25 anos.

Dimensão Escolar – as variáveis são *o ensino médio e o ensino fundamental*, sendo ambas associadas a três categorias de rede escolar em que estudou: pública, privada e mista.

Dimensão Familiar – formada por três variáveis: *renda familiar, escolaridade do pai e escolaridade da mãe*. As categorias de renda familiar são: 1) Sem renda; 2) Renda baixa (até um salário mínimo – SM); 3) Renda média baixa (de um a dois SM); 4) Renda média (entre dois e cinco SM); 5) Renda média alta (de cinco a dez SM); e 6) Renda alta (acima de dez SM). O nível de escolaridade do pai e da mãe são importantes indicadores de *status* econômico e social, que podem ser associados à renda familiar e ao capital cultural. Ambas as variáveis são subdivididas em categorias de nível de escolaridade: 1) Não estudou; 2) Primeira etapa do ensino fundamental; 3) Ensino fundamental; 4) Ensino médio; 5) Ensino superior; e 6) Pós-graduação.

Dimensão Regional – é composta por uma única variável abrangendo como categoria *as regiões geográficas brasileiras*: Centro-oeste; Nordeste; Norte; Sudeste; e Sul.

Uma análise da distribuição da frequência relativa nas categorias para os candidatos nas dimensões apresentadas acima, bem como a aplicação de um modelo estatístico baseado em regressão logística serão apresentadas a seguir.

FREQUÊNCIAS RELATIVAS POR VARIÁVEL

As frequências relativas foram medidas em cada uma das quatro aplicações das provas de Ciências da Natureza e de Matemática. As médias das oito frequências relativas obtidas para cada variável estão apresentadas na Tabela 1. As análises dos efeitos das dimensões no acesso ao decil superior serão apresentadas na sequência.

Tabela 1 Média das frequências relativas das variáveis para as provas de Ciências da Natureza e Matemática nas quatro edições do Enem analisadas

(continua)

Dimensão	Variáveis e categorias	Frequências Relativas	
		Grupo completo	Decil superior
Pessoal	Sexo		
	Masculino	40,0	58,3
	Feminino	60,0	41,8
	Faixa etária		
	Treineiros	5,6	7,4
	Em fase	44,5	57,0
	Jovens	26,0	23,2
Adultos	23,9	12,4	
Escolar	Ensino fundamental		
	Público	72,1	33,0
	Misto	12,4	17,8
	Privado	14,9	48,9
	Ensino médio		
	Público	76,6	38,2
	Privado	16,7	54,5
Familiar	Escolaridade da mãe		
	Não estudou	5,0	0,9
	1º Ciclo	25,4	8,9
	Fundamental	23,7	13,2
	Médio	28,2	33,0
	Superior	10,1	27,4
	Pós-graduação	4,9	15,5
	Escolaridade do pai		
	Não estudou	6,5	3,5
	1º Ciclo	29,5	1,3
	Fundamental	21,9	11,6
	Médio	23,9	14,5
	Superior	7,9	32,0
	Pós-graduação	2,9	25,0
	Renda familiar		
	Sem Rendimento	1,5	0,6
Baixa (até 1 SM)	20,7	4,2	
Média Baixa (entre 1 e 2 SM)	42,9	21,9	
Média (entre 2 e 5 SM)	21,9	27,8	
Média Alta (entre 5 e 10 SM)	8,2	24,1	
Alta (acima de 10 SM)	4,2	19,8	

Tabela 1 Média das frequências relativas das variáveis para as provas de Ciências da Natureza e Matemática nas quatro edições do Enem analisadas

(conclusão)

Dimensão	Variáveis e categorias	Frequências Relativas	
		Grupo completo	Decil superior
Região Federação	Região		
	Centro-Oeste	8,5	6,7
	Nordeste	30,7	21,0
	Norte	9,2	4,3
	Sudeste	38,2	53,4
	Sul	13,4	14,7

Fonte: Elaborada pelo autor com base nos microdados do Enem.

Na dimensão pessoal, vemos que as mulheres representam 60% dos participantes, porém no decil superior sua presença cai para 42%. Isso indica que, apesar do número de mulheres participantes no Enem ser maior do que o de homens, o desempenho feminino nas áreas de Ciências da Natureza e Matemática ainda é inferior ao masculino. Já com relação às faixas etárias, observamos que os candidatos em fase ampliam sua participação no decil superior, passando de 45% para 57%.

Os egressos da rede pública correspondem a 76% dos candidatos que prestaram o Enem no período analisado, porém no decil superior encontram-se apenas 38% de alunos formados na escola pública. Essa queda à metade da participação no decil superior repete-se com os candidatos que cursaram o ensino fundamental público. Tais dados são bastante expressivos e reforçam a necessidade de discutir a qualidade da escola pública.

Na dimensão familiar, tanto a escolaridade da mãe quanto a escolaridade do pai apresentam comportamento similar na transição para o decil superior: filhos de pais com escolaridade abaixo do ensino médio têm sua participação fortemente reduzida nos 10% com melhor desempenho.

Ao analisar candidatos cujos pais concluíram o ensino médio, verificamos diferenças de comportamentos entre pai e mãe: filhos de mães que possuem ensino médio completo têm um acréscimo na participação no decil superior; no caso dos pais há uma queda na transição. Esse resultado pode estar relacionado ao fato de o processo educacional familiar ainda ser predominantemente responsabilidade das mães, implicando maior interação destas – e de seu conhecimento – com os filhos. Esses dados reforçam a escolha da escolaridade da mãe como um indicador de capital cultural dos candidatos.

Quando consideramos a faixa de renda alta, a participação dos candidatos no decil superior é multiplicada por cinco, passando de 4% para 20%, indicando que estes apresentam melhor formação geral para realizar o Enem.

No caso das regiões da Federação, observa-se que Norte e Nordeste diminuem sua participação na transição do âmbito geral da prova para o decil superior, caindo para a metade e para 2/3, respectivamente.

MODELO DE REGRESSÃO LOGÍSTICA

A influência das diferenças socioeconômicas sobre as chances de um candidato estar no decil superior será estimada por meio de um modelo de regressão logística. Esses modelos permitem comparar o efeito das variáveis socioeconômicas nas distintas categorias elencadas nas dimensões sugeridas, medindo a variação proporcional do acesso ao decil superior.

O nosso modelo propõe estimar o efeito das categorias socioeconômicas sobre a chance de os candidatos acessarem o decil superior das notas de Ciências da Natureza e Matemática. Para calcular essa chance, é necessário definir um perfil de referência, que no nosso caso é formado por um candidato do sexo masculino; que estudou sempre na rede privada; concluiu em fase sua formação; pertence a família de renda média; habita a região Sudeste; e cujos pais concluíram o ensino médio. Esse candidato apresenta uma chance igual a um de atingir o decil superior. Os demais perfis podem apresentar chance maior ou menor que o perfil de referência.

Os resultados médios para os quatro anos do modelo de regressão logística estão apresentados na Tabela 2.

Tabela 2 Modelo de regressão logística para sucesso ao acessar o decil superior nas provas de Ciências da Natureza e Matemática

(continua)

Dimensão	Variáveis e categorias	Medidas de Chance	
		Ciências da Natureza	Matemática
Pessoal	Sexo		
	Masculino	1,000	1,000
	Feminino	0,517	0,421
	Faixa etária		
	Treineiros	0,551	0,687
	Em fase	1,000	1,000
	Jovens	1,230	1,180
	Adultos	1,082	1,036
Escolar	Ensino fundamental		
	Público	0,522	0,475
	Misto	–	–
	Privado	1,000	1,000
	Ensino médio		
	Público	0,429	0,510
	Misto	0,546	0,568
	Privado	1,000	1,000

Tabela 2 Modelo de regressão logística para sucesso ao acessar o decil superior nas provas de Ciências da Natureza e Matemática

(conclusão)

Dimensão	Variáveis e categorias	Medidas de Chance	
		Ciências da Natureza	Matemática
Familiar	Escolaridade da mãe		
	Não estudou	0,454	0,453
	1º Ciclo	0,677	0,696
	Fundamental	0,678	0,773
	Médio	1,000	1,000
	Superior	1,259	1,228
	Pós-graduação	1,357	1,294
	Escolaridade do pai		
	Não estudou	0,601	0,616
	1º Ciclo	0,760	0,777
	Fundamental	0,807	0,823
	Médio	1,000	1,000
	Superior	1,307	1,286
	Pós-graduação	1,442	1,415
	Renda familiar		
	Sem Rendimento	0,598	0,572
	Baixa (até 1 SM)	0,368	0,364
	Média Baixa (entre 1 e 2 SM)	0,577	0,575
Média (entre 2 e 5 SM)	1,000	1,000	
Média Alta (entre 5 e 10 SM)	1,460	1,458	
Alta (acima de 10 SM)	1,910	1,979	
Região da Federação	Região		
	Centro-Oeste	0,658	0,497
	Nordeste	0,607	0,578
	Norte	0,514	0,525
	Sudeste	1,000	1,000
	Sul	0,992	0,998

Fonte: Elaborada pelo autor com base nos microdados do Enem
 Todos os resultados são significativos com $p < 0,001$

Os parâmetros do modelo apresentados na Tabela 2 estão associados à chance de sucesso das diferentes categorias. Segundo essa tabela, a chance de uma pessoa do sexo feminino estar no decil superior em Matemática e em Ciências da Natureza é menor do que a chance de uma pessoa do sexo masculino.

Embora não esteja no escopo dessa discussão explorar as várias hipóteses para o número de mulheres que atingem o decil superior ser menor do que o de homens, trata-se de uma realidade que necessita ser modificada, devendo ser considerada tanto por professores da educação básica quanto por elaboradores dos itens.

Os jovens apresentam maior chance de atingirem o decil superior que os estudantes em fase, podendo esse fato estar associado à preparação para cursos de alta demanda e/ou às diversas realizações do Exame como consequência de mudanças de opção quanto à carreira universitária.

Um aluno que tenha cursado toda a educação básica na rede pública apresenta uma chance de estar no decil superior igual a um quarto da chance de um aluno que tenha realizado o mesmo percurso na educação privada.

Associando-se os efeitos de renda à dimensão escolar, verificamos que um candidato de renda média, que cursou a educação básica na escola pública, quando comparado a um candidato de renda alta, que cursou a educação básica na escola privada, apresenta uma chance oito vezes menor de estar no decil superior.

Com relação às regiões geográficas, o Sudeste não se diferencia do Sul, ambos apresentando coeficientes de regressão muito próximos. Por outro lado, para a região Norte, a chance de acessar o decil superior cai à metade, cabendo aqui questionar se a tentativa de “universalização” do Enem não deveria ser revista em prol de maior equilíbrio entre as regiões.

CONCLUSÕES

Estes resultados causam impacto na medida em que a maior parte do contingente em idade escolar está matriculado em escolas públicas. Há uma aparente passividade social em relação ao desempenho dos egressos da escola pública, que se deve, em parte, ao fato de o acesso ao ensino superior ter se beneficiado das leis de cotas e outras formas de ações afirmativas. Sem desmerecer ou entrar no mérito dessas questões, fato é que podem estar mascarando o grave problema da qualidade da educação básica oferecida à maior parte da população.

Não é a prova do Enem em si que está sendo excludente, são as desigualdades de renda, muito marcantes no Brasil, que se manifestam no desempenho. Em vista disso, o Enem, enquanto instrumento de avaliação neutro, está efetivamente privilegiando candidatos com maior capital cultural e financeiro.

Mais do que possibilitar o acesso ao ensino superior para os oriundos da escola pública, é necessário garantir uma formação acadêmica e cultural de qualidade a essa população, propiciando sua participação de forma ativa e criativa na sociedade.

REFERÊNCIAS

BOURDIEU, P.; PASSERON, J. C. *Escritos de Educação*. Petrópolis: Vozes, 1998.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep).
Microdados para download: microdados Enem. Disponível em: < <http://portal.inep.gov.br/basica-levantamentos-acessar>>.