



QUAL É O EFEITO DA DESIGUALDADE SOCIAL NO DESEMPENHO EM CIÊNCIAS DOS ESTUDANTES BRASILEIROS? UMA ANÁLISE DO EXAME NACIONAL DO ENSINO MÉDIO (2012-2019)

How does social inequality affect Brazilian students' performance in science? An inquiry into the Brazilian High School Exam (2012-2019)

Paulo Lima Junior [paulolimajr@unb.br]
Instituto de Física (IF/UnB)
Universidade de Brasília (UnB)
Sala BT 297 – ICC Centro.
Asa Norte. Brasília (DF). CEP: 70919-970.

Jailton Correia Fraga Junior [jailtoncfjr@gmail.com]
Instituto de Física (IF/UnB)
Universidade de Brasília (UnB)

Resumo

Recentemente, o baixo desempenho dos estudantes brasileiros nas avaliações internacionais ganhou os noticiários, levantando a suspeita de que a qualidade da educação científica no país esteja em queda. Neste artigo, inspirados pela sociologia da educação de Pierre Bourdieu, nós investigamos o efeito da desigualdade social no desempenho dos estudantes brasileiros em Ciências da Natureza. Uma amostra representativa de todos os participantes do ENEM entre 2012 e 2019 ($n = 489.167$) foi submetida a uma análise de correspondência múltipla ajustada e inserida em um modelo linear multivariado. Como primeiro resultado, observamos que a origem social é capaz de explicar mais de 35% do desempenho em ciências dos estudantes. Logo, grande parte do sucesso escolar é devido ao privilégio de classe. Por outro lado, frequentar uma escola privilegiada é mais importante que nascer em uma família privilegiada, pois o efeito social da escola mostrou-se quase quatro vezes superior ao efeito social individual. Esses resultados trazem implicações importantes para pensarmos criticamente a educação científica em vista da ideologia do mérito e da distribuição desigual de recursos e oportunidades no sistema educacional.

Palavras-Chave: Desigualdade social; Luta de classes; Bourdieu; ENEM; Ciências da Natureza.

Abstract

Recently, Brazilian students' low performance in international assessments has reached the news, raising the suspicion that the quality of scientific education in the country is declining. In this article, inspired by Pierre Bourdieu's sociology of education, we investigate the effect of social inequality on the scientific performance of Brazilian students. A representative sample of ENEM participants between 2012 and 2019 ($n = 489,167$) was subjected to an adjusted correspondence analysis and inserted in a multivariate linear model. As a first result, we observed that social background explains more than 35% of the students' performance in science. Therefore, school success is largely indebted to class privilege. On the other hand, attending a privileged school is more important than being born in a privileged family, as the school social effect proved to be four times superior to the individual social effect. These results have important implications to critically approach science education regarding the ideology of merit as well as the unequal distribution of resources and opportunities in the education system.

Keywords: Social inequality; Class struggle; Bourdieu; ENEM; Natural Sciences.

INTRODUÇÃO

Desde a década de 1960, a relação entre origem social e destino escolar é considerada um tema fundamental da pesquisa e da política educacional, colocando em xeque a ideia de que o sucesso escolar resulta somente de dom ou mérito individual (Salej, 2005). Ao longo do tempo, os pesquisadores deixaram de se surpreender com a associação entre desigualdades sociais e escolares, dedicando-se a compreender como essas desigualdades se constroem no plano individual (Massi, Muzzeti, & Suficier, 2017). De fato, há muito tempo sabemos que variáveis sociais e escolares tendem a estar associadas, mas essa associação nem sempre apresenta a mesma intensidade e interpretação (OECD, 2010). Movidas por esse espírito, diversas pesquisas têm mapeado os condicionantes sociais do sucesso e do fracasso escolar nas Ciências da Natureza (Silva & Pereira, 2019), contribuindo para a construção de um olhar mais crítico sobre a educação científica.

Por exemplo, ao analisar os relatos e memórias de 136 professores de Ciências, Watanabe e Gurgel (2017) identificaram como o desenvolvimento profissional desses professores é marcado por discriminações de classe e cor que foram sistematicamente reforçadas pelas avaliações escolares. Com respeito ao curso de graduação em Física da Universidade de Brasília (UnB), é possível argumentar que, mesmo não havendo relação significativa entre evasão e desempenho acadêmico, a pressão por excelência estrutura as experiências dos discentes (Lima Junior *et al.*, 2020), prejudicando a integração daqueles que têm origem nas periferias urbanas (Lima Junior, Fraga Junior, Andrade, & Bernardino, 2020). Essas pesquisas ilustram como as avaliações escolares podem contribuir, à escala individual, para a manutenção das desigualdades sociais.

Em razão de diversos programas federais, o Exame Nacional do Ensino Médio (ENEM) tem funcionado como o principal regulador do acesso à educação pública não compulsória e, por isso, faz sentido que ele seja cuidadosamente monitorado. Sabemos, por exemplo, que as relações entre desempenho no ENEM e origem social dos estudantes são tipicamente mais intensas nas provas de Ciências da Natureza e Matemática (Lima Junior, 2015). Além disso, ao analisar o ENEM nos anos de 2011, 2012 e 2014, Kleinke (2017) observou que a taxa de acerto média nos itens de Física entre estudantes de elite foi duas vezes superior ao desempenho dos estudantes de classe popular. Ao olhar para os tipos de questões, Kleinke concluiu que os itens que contextualizaram o conteúdo de Física, apesar de possuírem baixo índice de acerto, tendem a discriminar menos os estudantes com respeito à sua origem social. Nascimento, Cavalcanti e Ostermann (2018), ao analisar as edições de 2009, 2012 e 2015, ratificaram que os melhores desempenhos no ENEM estão relacionados a um alto volume de capital cultural e econômico dos estudantes e que as questões qualitativas são menos discriminatórias.

Recentemente, o baixo desempenho dos estudantes brasileiros nas avaliações internacionais ganhou os noticiários (Moreno & Oliveira, 2019), levantando a suspeita de que a qualidade da educação científica no país esteja em queda. Por outro lado, sabemos que os *rankings* internacionais produzem uma imagem muito aproximada do que realmente se passa em nosso país, sendo necessário levar em consideração as desigualdades internas de nosso sistema educacional para produzirmos um diagnóstico mais preciso da nossa situação. Enfim, visando expandir o conhecimento já disponível sobre o efeito da desigualdade social no desempenho em ciências da natureza dos estudantes brasileiros, tomamos os dados do ENEM de 2012 a 2019 com o propósito de avaliar o poder preditivo da origem social dos estudantes sobre o desempenho na prova de Ciências da Natureza. Aparentemente, o poder discriminatório do ENEM cresceu de 2006 a 2012 (Lima Junior, 2015). Ou seja, a associação entre origem social e desempenho escolar observada na população brasileira aumentou durante esse período. Nesta análise, nós tomamos uma amostra representativa de todos os participantes do ENEM entre os anos 2012 e 2019 ($n = 489.167$). Com respeito a eles, perguntamos:

- Qual é o efeito da origem social sobre o desempenho na prova de Ciências da Natureza?
- Esse efeito sofre algum tipo de transformação ao longo do período investigado? Qual?

Para responder às questões acima, apresentamos uma análise baseada no método estatístico que se consagrou nos estudos empíricos de Pierre Bourdieu (1984, 1990): a análise de correspondência. Redescoberto recentemente pela pesquisa em educação em Ciências (Claussen & Osborne, 2013), Bourdieu pode ser considerado um autor relevante para pensarmos diversos desafios contemporâneos, tais como a evasão dos cursos de graduação (Fernandes, Ueno Guimarães, Robert, & Passos, 2020; Lima Junior, 2013; Massi & Villani, 2015), o desenvolvimento do interesse por ciência em sala de aula (Anderhag, Hamza, & Wickman, 2015); as aspirações profissionais relacionadas às Ciências de jovens estudantes (Archer, Dawson,

DeWitt, Seakins, & Wong, 2015); a formação e desenvolvimento profissional docente (Genovese, Queiroz, & Castilho, 2015; Mello & Higa, 2018), as estratégias de distinção dos estudantes de Ciências (Almeida & Watanabe, 2019), e, finalmente, o desempenho dos estudantes nas avaliações de Ciências da Natureza (Lima Junior, Ostermann, & Rezende, 2013; Nascimento *et al.*, 2018).

A análise de correspondência parte de dados contidos em uma tabela de contingência e reproduz, graficamente, as relações entre esses dados (Greenacre, 2007). Trata-se de uma ferramenta muito útil para a pesquisa sobre desigualdades sociais, pois a possibilidade de gerar uma imagem gráfica da associação entre variáveis categóricas torna intuitiva a interpretação de questionários socioeconômicos. Além disso, ela representa as variáveis qualitativas em uma escala, que pode ser submetida a análises posteriores. Visando contribuir para a popularização da análise de correspondência, neste artigo, nós a empregamos sem supor que o leitor tenha familiaridade com outras ferramentas estatísticas. Enfim, esperamos que nossa análise contribua para um olhar crítico sobre as avaliações escolares em Ciências da Natureza, explicitando a associação entre origem social e oportunidades escolares.

Contribuições de Pierre Bourdieu

A partir dos anos 1970, Bourdieu foi se tornando progressivamente popular no Brasil como o autor do livro “A Reprodução” (Bourdieu & Passeron, 2009), obra que fala diretamente aos que foram iludidos com a promessa de que a garantia de acesso à educação produziria uma sociedade menos desigual e mais baseada no mérito individual. No Brasil, a expansão da educação básica teve início sob a ditadura militar (1964-1986) e foi acompanhada de uma profunda precarização do ensino e proletarização do ofício docente. Segundo dados recentes (Gráfico 1), considerando uma lista numerosa de países investigados, o menor de todos os salários iniciais pagos a professores do ensino secundário é brasileiro.

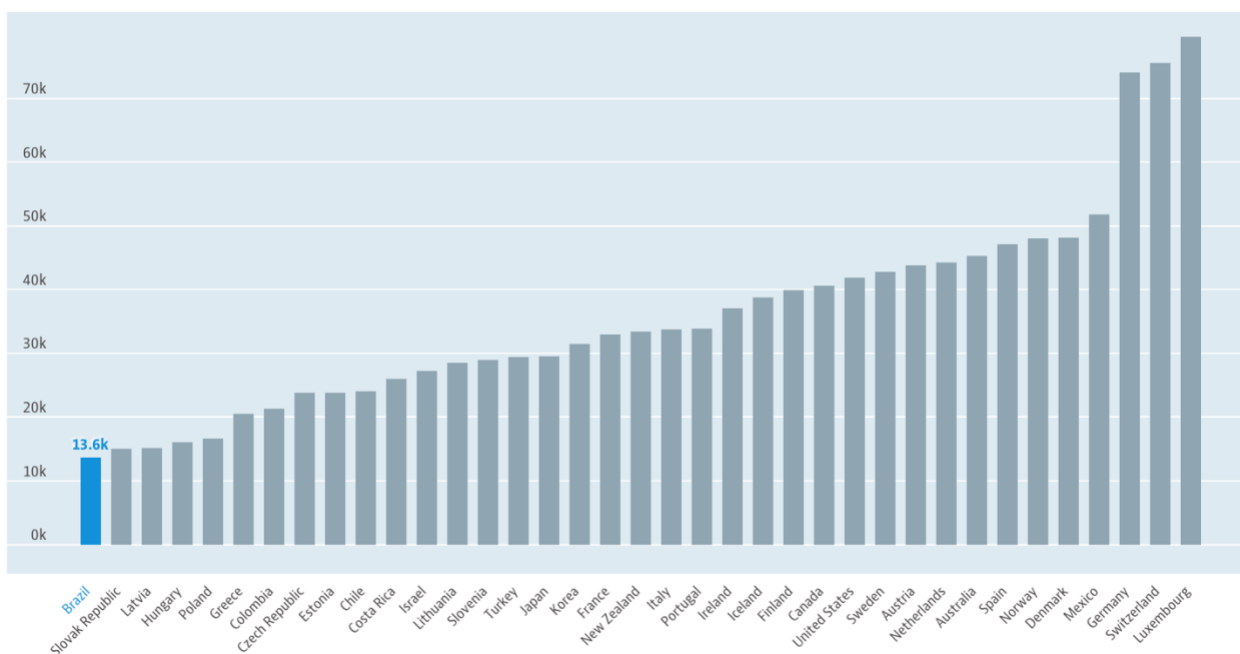


Gráfico 1. Rendimento inicial anual médio pago, em dólares, a professores dos últimos anos da educação básica (*upper secondary*) em diversos países do mundo. **Fonte:** OECD (2020), Teachers' salaries (indicator). doi: 10.1787/f689fb91-en (Acessado em 21 de novembro de 2020).

Não surpreende observar que o desempenho médio dos nossos estudantes nas avaliações de Ciências da Natureza siga de perto a precarização estrutural do nosso sistema de ensino. Segundo os dados mais atuais (Gráfico 2), a Indonésia é o único país conhecido em que o desempenho médio dos estudantes em Ciências da Natureza fica aquém do brasileiro. Como se não bastasse, o sistema educacional brasileiro não é somente precário, mas internamente heterogêneo, desigual. Essa desigualdade pode ser claramente percebida nos *rankings* nacionais que posicionam as instituições privadas de Ensino Médio à frente das escolas públicas (Evolucional, 2020).

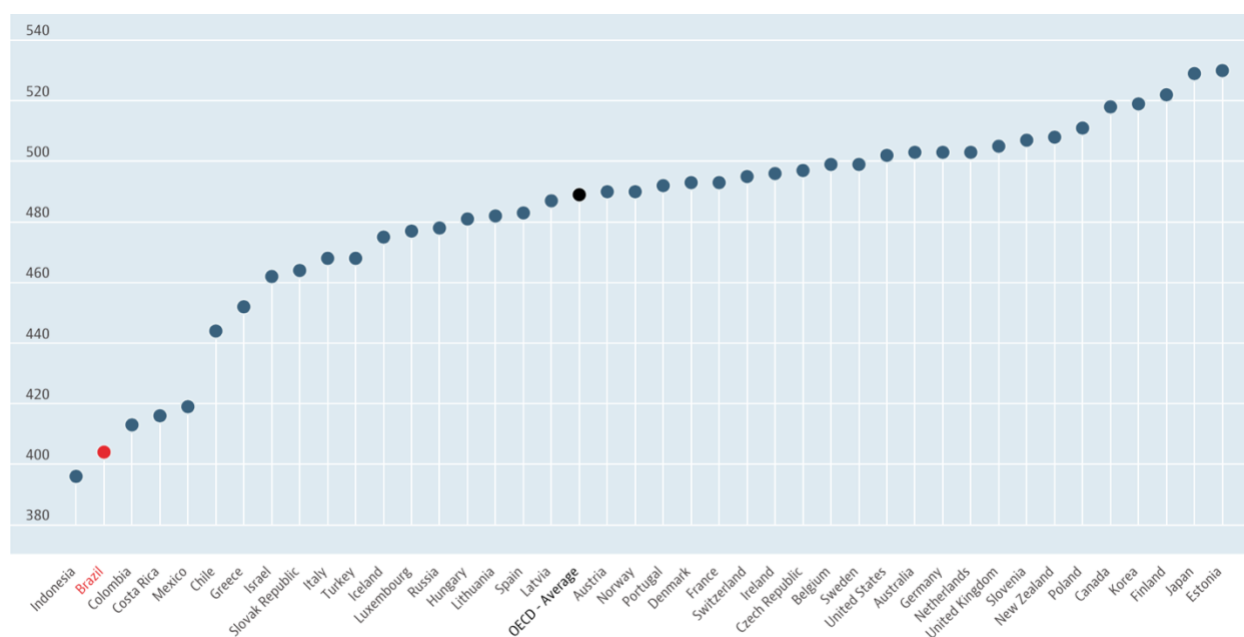


Gráfico 2. Desempenho médio dos estudantes em Ciências da Natureza em diversos países do mundo.
Fonte: OECD (2020), Science performance (PISA) (indicator). doi: 10.1787/91952204-en (Acessado em 21 de novembro de 2020).

Antes que essas comparações nacionais e internacionais fossem tão comuns quanto são hoje em dia, Bourdieu teve o mérito de formular uma explicação teoricamente consistente e empiricamente fundamentada para a associação entre origem social e percurso escolar (Lima Junior, Pinheiro, & Ostermann, 2012; Nogueira & Nogueira, 2009). Segundo ele (Bourdieu, 2007), as desigualdades escolares não decorrem imediatamente da desigualdade econômica, mas das relações entre dois sistemas de relações: (1) o sistema das relações de classe (que não pode ser reduzido à distribuição desigual da renda), e (2) o sistema de ensino. Bourdieu (1984) teve, também, o mérito de explicitar a maneira como diferentes origens sociais correspondem a diferentes estratégias familiares de investimento escolar. Fazendo isso, ele permite perceber que a consagração escolar dos cidadãos culturalmente privilegiados não é automática, mas depende da maneira como suas famílias se mobilizam ativamente em vista de sua classificação e desclassificação (Bourdieu, 1984). Essa mobilização ativa, por outro lado, não implica ter consciência do quanto as lutas sociais estruturam nossas ações. Pelo contrário, o princípio gerador das nossas práticas (chamado *habitus*) encontra-se incorporado até mesmo às nossas ações mais instintivas, fazendo com que muitas delas pareçam resultado de um cálculo racional. Isso é o que se observa, por exemplo, em algumas famílias de classe média que investem na escola porque têm a educação dos filhos como um valor fundamental em suas vidas (ou seja, não por reconhecerem que o sucesso escolar da prole é condição indispensável para preservar a posição familiar na estrutura das relações de classe). Frequentemente, são essas famílias que mais experimentam uma sensação de nobreza superior ao se deparar com outras, que não se dedicam tão obstinadamente a posicionar seus filhos nos melhores cursos e instituições de ensino. Dessa maneira, faz sentido afirmar que o *habitus* é a necessidade tornada virtude (Bourdieu, 2013): muitas ações e percepções que, em última análise, representam luta, disputa e manutenção de privilégios são vividas como se partissem das virtudes individuais.

De maneira muito sintética, as análises de Bourdieu apontam que, em cada formação social, os sistemas educacionais são menos revolucionários do que se imaginava, e mais comprometidos com a manutenção das desigualdades sociais. Essa perspectiva, compartilhada por diversos analistas emergentes na década de 1970, permite localizar a obra de Bourdieu como uma *teoria crítica* do pensamento educacional (Silva, 2014). Contudo, mergulhando um pouco mais profundamente no autor, é fácil perceber que a crítica social da educação é um tema relativamente estreito no conjunto de sua obra. Bourdieu estava mais voltado a questões fundamentais da teoria social, tal como a superação do subjetivismo e do objetivismo nas análises da ação individual (Bourdieu, 1990), e a construção de um aparato teórico capaz de explicar tanto as reproduções quanto as transformações do campo social (Bourdieu, 2004).

A análise de correspondência não é somente a técnica estatística mais empregada por Bourdieu, mas ela funciona como uma chave importante para entendermos a obra desse autor, que, há alguns anos, foi

redescoberto pela pesquisa em educação em Ciências (Claussen & Osborne, 2013). A despeito do estilo de escrita muito sofisticado, Bourdieu jamais postulou ideias fora de uma sólida base empírica. De fato, as aproximações entre as formulações básicas que encontramos nesse autor e a análise de correspondência são substanciais. Conforme veremos ao longo deste artigo, a grande conquista da análise de correspondência é a possibilidade de mapear padrões de associação entre variáveis categóricas (nominais ou ordinais). São exemplos de variáveis categóricas: faixa de renda, escolaridade e ocupação dos pais, padrões de consumo, descritores de hábitos culturais, alimentares, religiosos, posicionamento político... Todas essas informações são ativamente empregadas pelos indivíduos de cada formação social para se distinguir uns dos outros, delimitando grupos sociais com os quais podem estabelecer relações de estranhamento e pertencimento. Assim, os mapas da análise de correspondência permitem agrupar os indivíduos e suas qualidades por similaridade. Qualidades similares (e.g., renda de 2 a 3 salários mínimos e escolaridade média) são representadas como pontos próximos no mapa. Enquanto isso, qualidades dissimilares (e.g., baixa renda e alta escolaridade) são afastadas. Assim, faz todo o sentido que o espaço social a que se refere Bourdieu (2007) seja imaginado como um espaço multidimensional em que cada dimensão representa um tipo diferente de desigualdade. O espaço social é, portanto, o espaço das distinções.

Outro destaque relevante que podemos fazer sobre a análise de correspondência diz respeito à maneira recíproca como ela permite representar relações causais (Klüger, 2018). No quadro teórico bourdieusiano, a estrutura social, determinada pela distribuição desigual de capital (econômico, cultural, social), orienta flexivelmente as ações individuais (Bourdieu, 1990). Porém, as formas de capital incluem componentes que são elas mesmas práticas, capazes de distinguir os agentes, por exemplo, em função de sua proximidade com relação às elites cultural e econômica. Assim, as práticas culturais não são determinadas por fora (por algo que escape completamente a elas mesmas), mas participam de sua própria determinação. A estrutura das relações de classe está assentada em distinções de nobreza e julgamentos de gosto que precisam ser reafirmados justamente porque as relações de classe não consistem de uma estrutura rigidamente estabelecida, mas constantemente negociada e disputada (Bourdieu, 1984). As posições sociais a partir das quais as pessoas lutam são negociadas na própria luta. Assim, a análise de correspondência permite pensar a sociedade como um sistema autocorrelacionado e, portanto, aberto a transformações (Fowler, 2020). Disso resulta também que a noção de “estrutura” em Bourdieu foge à noção intuitiva de algo rígido e invariante, mas corresponde a um conjunto de relações determinantes que se transformam ao longo do tempo.

MÉTODOS

Para avaliar o efeito da origem social sobre o desempenho dos estudantes brasileiros em Ciências da Natureza (vide questões de pesquisa), nós importamos e empilhamos os dados de todas as edições do ENEM de 2012 até 2019. As variáveis de escolaridade e renda foram recodificadas para que houvesse coerência de interpretação (e.g., padronizando as faixas de renda em múltiplos do salário-mínimo em vez de valores absolutos em reais). Pela maneira como a escala do ENEM é construída, podemos assumir que pontuações iguais obtidas em anos diferentes designam o mesmo grau de proficiência. Após o empilhamento, a quantidade de participantes era muito grande para conduzir a análise multivariada em um computador pessoal (34,8 milhões de observações). Além disso, o ganho em confiabilidade de uma análise com essa quantidade de observações seria inexpressivo e não justificaria o esforço. Logo, optamos por selecionar uma amostra representativa desse universo (com $n = 489.167$). A amostra corresponde aos estudantes de 600 escolas tomadas aleatoriamente do universo das que têm participação consistente no certame (i.e., um mínimo de 30 estudantes concorrendo em cada edição do ENEM). Essa condição será importante para garantir a confiabilidade do que chamaremos “efeito social da escola”.

As informações de renda e escolaridade dos pais foram, enfim, submetidas a uma Análise de Correspondência Múltipla (ACM) Ajustada (Greenacre, 2007). Tomamos o cuidado de explicar algumas ideias-chave desse método de análise para que o leitor possa interpretar conosco seus resultados gráficos. Os mapas gerados por meio da ACM Ajustada permitiram determinar a “posição de classe individual” da família de cada estudante. Como toda nossa amostra é composta por estudantes localizados em escolas onde há vários outros estudantes, pudemos calcular, também, a “posição de classe da escola”. Essas duas informações foram inseridas em um modelo de regressão linear com o propósito de testar o poder preditivo da origem social dos estudantes sobre seu desempenho na prova. A variável “ano” foi inserida para avaliarmos como a associação entre origem social e desempenho evolui ao longo do tempo. Como a

regressão linear é um método mais familiar para pesquisadores com formação em Ciências da Natureza, não nos ocupamos tanto em introduzi-la ao longo da análise.

ANÁLISE

Tabelas de Contingência

A análise de correspondência tem como ponto de partida uma tabela de contingência, que apresenta o número de casos (ocorrências) nos cruzamentos dos níveis de variáveis categóricas. A saber, todas as variáveis codificadas em níveis (ordenados ou não) são consideradas categóricas (e.g., cor, sexo, faixas de renda, escolaridade e ocupação dos pais). A Tabela 1 apresenta o resultado da tabulação cruzada entre as variáveis de escolaridade do pai e da mãe para os participantes da nossa amostra. As células da tabela de contingência exibem, em primeiro lugar, a quantidade de casos que satisfazem simultaneamente às condições de linha e coluna (i.e., 7334 estudantes têm pai e mãe sem nenhuma escolaridade). Observe que as tabelas de contingência também informam ocorrências totais por linha, coluna e geral (todas em negrito).

Tabela 1 – Tabulação cruzada da escolaridade do pai (linhas) contra a escolaridade da mãe (colunas) em nossa amostra de participantes do ENEM (2012-2019)

	M.NEST	M.FUND1	M.FUND2	M.MÉDIO	M.GRAD	M.PGRAD	TOTAL
P.NEST	7334 (0,335)	10276 (0,469)	1729 (0,079)	2059 (0,094)	279 (0,013)	234 (0,011)	21911 [0,045]
P.FUND1	4375 (0,024)	97458 (0,537)	26549 (0,146)	40986 (0,226)	6325 (0,035)	5891 (0,032)	181584 [0,371]
P.FUND2	571 (0,008)	15169 (0,21)	23628 (0,327)	24327 (0,337)	5008 (0,069)	3529 (0,049)	72232 [0,148]
P.MÉDIO	467 (0,003)	17799 (0,123)	18520 (0,128)	79168 (0,547)	16231 (0,112)	12499 (0,086)	144684 [0,296]
P.GRAD	75 (0,002)	1794 (0,047)	2505 (0,066)	13218 (0,349)	13115 (0,346)	7169 (0,189)	37876 [0,077]
P.PGRAD	40 (0,001)	1331 (0,043)	1226 (0,04)	7161 (0,232)	6371 (0,206)	14751 (0,478)	30880 [0,063]
TOTAL	12862 (0,026)	143827 (0,294)	74157 (0,152)	166919 (0,341)	47329 (0,097)	44073 (0,09)	489167 [1,000]

Fonte: Autores.

Perfil é o nome dado a um conjunto de frequências relativas dos dados em uma tabela de contingência. Para calcular o perfil, devemos dividir as frequências observadas em cada linha (ou coluna) pelo total dessa linha (ou coluna). Os valores dos perfis-linha da Tabela 1 foram informados entre parênteses em cada célula. A saber, os valores entre parênteses podem ser interpretados como coordenadas em um espaço multidimensional (Greenacre, 2007). Em outras palavras, os perfis podem ser representados por uma posição no espaço. O perfil da última linha pode ser considerado a média ponderada dos demais perfis-linha. Por ser uma média ponderada de posições, o perfil médio pode ser interpretado como um centro de massa.

Além de uma posição, dizemos que cada perfil tem uma massa, que está relacionada à ocupação desse perfil. Perfis muito “massivos” carregam mais casos que os demais. Para obtermos a massa dos perfis-linha, devemos efetuar a divisão do total de cada linha pelo total geral da amostra. As massas dos perfis-linha foram publicadas entre colchetes na Tabela 1.

Desconsiderando os totais, podemos perceber que a Tabela 1 é uma matriz 6 x 6. A saber, uma representação espacial perfeita dos dados dessa matriz necessitaria de um espaço com 5 dimensões (Greenacre, 2007). Para dar conta de tantas dimensões em um mapa bidimensional, precisamos buscar uma saída. O processo usual é relativamente simples e análogo a uma fotografia, que permite visualizar, em um registro bidimensional, uma realidade com mais de duas dimensões. A chamada redução da dimensionalidade consiste em projetar as posições dos perfis multidimensionais em um plano. Essa redução

dimensional é o coração da análise de correspondência. Há, no entanto, um preço a pagar. Assim como nas fotografias, a redução de dimensionalidade acarreta perda de informação. Sendo assim, um dos principais desafios da análise de correspondência é encontrar a menor quantidade de dimensões capazes de representar os dados da maneira mais verossímil possível. A propósito, de maneira análoga ao que ocorre na análise de variância (Silveira, 1999), a análise de correspondência busca um modelo que explique, ao máximo, a dispersão dos dados. Aqui, a medida dessa dispersão é chamada inércia. Em geral, quando a análise de correspondência é bem-sucedida, o mapa produzido é capaz de representar grande parte da inércia total dos dados.

Correspondência múltipla

A análise de correspondência múltipla toma como ponto de partida uma tabela que representa, ao mesmo tempo, as relações entre todas as variáveis que desejamos analisar. Há duas tabelas que cumprem essa função (1) a matriz indicadora e (2) a matriz de Burt. Para alguns propósitos, a análise dessas matrizes produz resultados equivalentes. Neste artigo, escolhemos dar atenção à matriz de Burt, que pode ser definida como um conjunto de tabelas de contingência empilhadas. O Quadro 1 mostra como seria uma matriz de Burt composta pelo cruzamento de duas variáveis com quatro níveis cada. A partir do Quadro 1, o leitor poderá imaginar como seria uma matriz de Burt com uma quantidade maior de variáveis e níveis. A saber, a análise de correspondência múltipla pode ser definida como a análise de correspondência simples aplicada à matriz de Burt (Greenacre, 2007). Em outras palavras, calculamos os perfis-linha da matriz de Burt. O resultado da análise é o mapa bidimensional capaz de representar a maior parte da inércia total dos perfis.

Quadro 1 – Matriz de Burt hipotética envolvendo três variáveis de quatro níveis

	Variável A				Variável B			
	A:01	A:02	A:03	A:04	B:01	B:02	B:03	B:04
A:01	A11	0	0	0	A1B1	A1B2	A1B3	A1B4
A:02	0	A22	0	0	A2B1	A2B2	A2B3	A2B4
A:03	0	0	A33	0	A3B1	A3B2	A3B3	A3B4
A:04	0	0	0	A44	A4B1	A4B2	A4B3	A4B4
B:01	B1A1	B1A2	B1A3	B1A4	B11	0	0	0
B:02	B2A1	B2A2	B2A3	B2A4	0	B22	0	0
B:03	B3A1	B3A2	B3A3	B3A4	0	0	B33	0
B:04	B4A1	B4A2	B4A3	B4A4	0	0	0	B44

Fonte: Autores.

A correspondência múltipla é uma boa ferramenta para representarmos a associação entre variáveis categóricas de uma tabela de contingência. Porém, a inércia explicada costuma ser subestimada e a inércia total, superestimada (Greenacre, 2007). De fato, uma das razões para o baixo poder explicativo da análise de correspondência múltipla convencional está relacionada à maneira como a matriz de Burt foi definida. Ao aplicar a análise de correspondência simples à matriz de Burt como um todo, obrigamos o algoritmo a representar todas as associações presentes na matriz, enquanto as associações realmente importantes estão nas submatrizes que não tocam a diagonal principal (vide Quadro 1). As submatrizes diagonais da matriz de Burt são repletas de zeros e não trazem informação relevante para a análise. É preciso, portanto, encontrar uma maneira de neutralizá-las. Há dois métodos que contornam esse problema: as chamadas análises de correspondência conjunta e ajustada (Greenacre, 2007). A representação gráfica obtida nos dois métodos

não é exatamente igual, mas costuma ser semelhante. Optamos por mostrar aqui o mapa simétrico da ACM Ajustada (Gráfico 3), que responde por 85,4% da inércia total dos dados.

Há vários *outputs* gráficos da análise de correspondência. De todas as opções disponíveis, nós temos preferido trabalhar com os chamados “mapas simétricos”. Para interpretá-los, devemos projetar os pontos do gráfico (que representam níveis das variáveis categóricas) sobre os eixos. Essas projeções dão sentido conceitual aos eixos. Por exemplo, no Gráfico 3, percebemos que todas as categorias estão organizadas em torno de uma ferradura, sugerindo uma direção de leitura do mapa. No começo da nuvem de pontos (à esquerda e acima), encontram-se as famílias de pais sem escolaridade. Seguindo a ferradura a partir dessa posição, percorremos, ordenadamente, todos os estratos de escolaridade e renda. Na parte baixa da ferradura, encontram-se as famílias com poucos salários mínimos e escolaridade média. Ao final da nuvem de pontos, encontramos as famílias com renda superior a 9 salários mínimos e pós-graduação. Dito tudo isso, percebemos que a dimensão 1 representa mais diretamente a posição dos cidadãos na estrutura das relações de classe. Portanto, a coordenada dos indivíduos ao longo da Dimensão 1 foi salva e chamada “posição individual de classe” (IND).

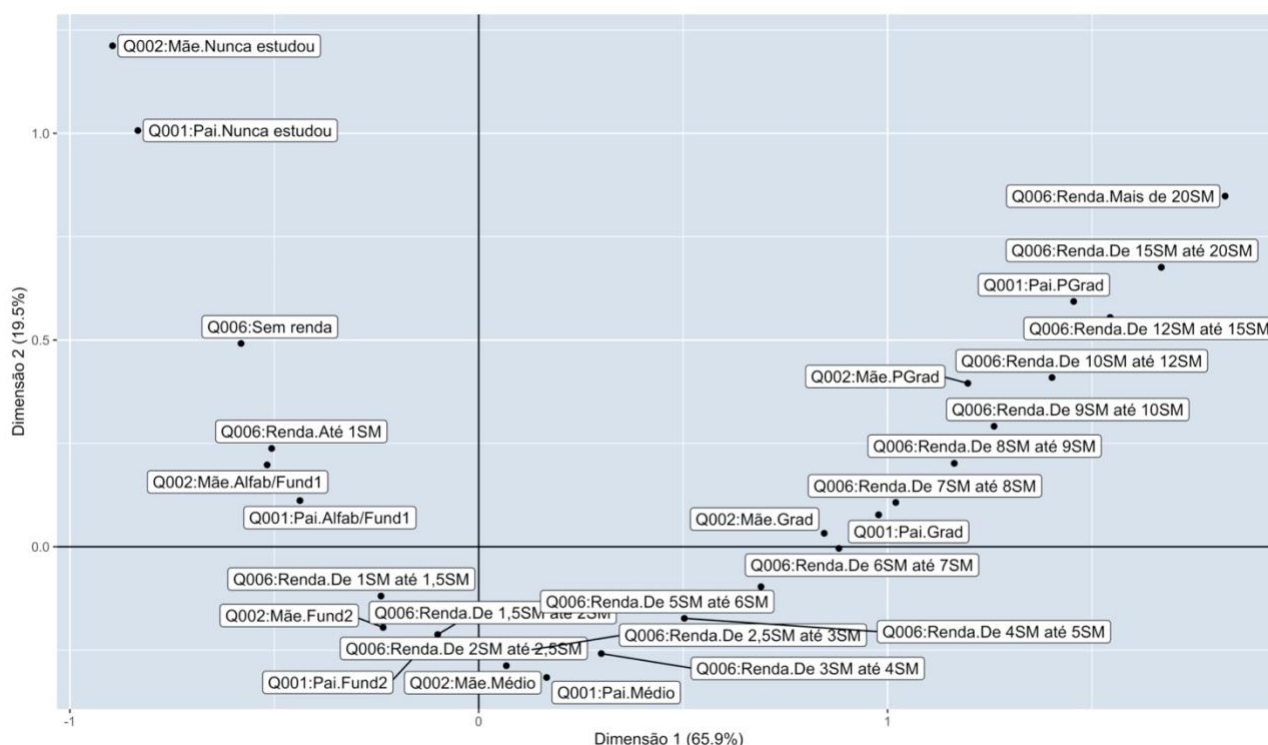


Gráfico 3 – Mapa simétrico da análise de correspondência ajustada dos dados socioeconômicos do ENEM 2012-2019 (renda e escolaridade dos pais). **Fonte:** Autores.

A posição individual de classe (IND) é, portanto, um escalar que distingue as famílias com respeito à renda familiar e à escolaridade dos pais. Por exemplo, posições individuais de classe inferiores a $-0,5$ indicam famílias sem renda e sem escolaridade. Valores entre $-0,5$ e 0 indicam famílias com ensino fundamental e renda flutuando um pouco acima de 1 salário mínimo. No segmento que vai de 0 a $0,5$ predominam famílias com formação média e renda entre 2 e 3 salários mínimos. Seguindo esse pensamento, o Quadro 2 foi elaborado para explicitar a correspondência entre os valores assumidos pela posição individual de classe (IND) e a interpretação que pode ser dada a essa posição.

Quadro 2 – Perfis das famílias correspondentes a cada posição individual de classe (cf. Dimensão 1 do Gráfico 3)

POSIÇÕES DE CLASSE	PERFIL DE FAMÍLIA MAIS FREQUENTE
$IND < -0,5$	Nenhuma escolaridade e sem nenhuma renda
$-0,5 < IND < 0$	Ensino Fundamental e renda inferior a 2 salários-mínimos
$0 < IND < 0,5$	Ensino Médio e renda de 2 a 5 salários-mínimos
$0,5 < IND < 1,0$	Graduação e renda de 5 a 7 salários-mínimos
$1,0 < IND < 1,5$	Pós-graduação e renda de 7 a 12 salários-mínimos
$IND > 1,5$	Renda superior a 12 salários-mínimos.

Essa posição individual de classe será o ponto de partida para nosso modelo preditivo do desempenho dos estudantes brasileiros em Ciências da Natureza. Todas as vezes que for necessário interpretar essa variável escalar, será preciso retornar ao Gráfico 3 e ao Quadro 2. Por exemplo, para cada valor assumido pela posição de classe ($IND = 0,1,2...$), o leitor deve ser capaz de identificar rapidamente o perfil da família designado por esse valor. Manter em mente essa informação tornará a análise mais significativa.

Especificando e avaliando o modelo

Tal como ocorre em qualquer processo de modelagem, a estimativa do efeito da origem social sobre o desempenho escolar é sensível às escolhas do analista (Heidemann, Araujo, & Veit, 2016). Segundo resultados do PISA (OECD, 2010), a origem social dos colegas da escola pode apresentar maior poder preditivo sobre o desempenho dos estudantes que suas próprias origens sociais. Esse fenômeno pode ser chamado “efeito social da escola” e há várias explicações para ele. Como nós sabemos, costuma haver uma distribuição desigual dos recursos (materiais, humanos) para as escolas das periferias (Garcia-Silva & Lima Junior, 2020) e essa desigualdade pode estar refletida na qualidade das experiências de aprendizagem proporcionada aos estudantes dessas escolas. No Brasil, a desigualdade entre escolas geralmente é abordada pela oposição entre rede pública e privada (Evolucional, 2020). Porém, as redes de ensino são internamente desiguais e as posições de classe do corpo discente podem ser um indicador mais sensível do privilégio de estudar nesta ou naquela instituição. Além disso, é possível que o convívio com colegas de outra origem social contribua para que os estudantes incorporem práticas típicas de uma origem diferente da sua. Com respeito a isso, também poderíamos argumentar que muitos jovens passam mais tempo com os colegas da escola que com seus familiares. Sendo assim, as amizades construídas na escola podem exercer sobre eles uma grande influência.

Para testar o efeito social da escola, nós estimamos a “posição de classe da escola” pela média das posições individuais dos seus estudantes. As posições de classe individuais serão designadas pela variável IND enquanto a posição de classe da escola será designada por ESC . A partir do Gráfico 3, percebemos que essas variáveis podem assumir qualquer valor na reta real, mas estão tipicamente entre $-1,0$ e $+2,0$. De início, essas duas variáveis foram inseridas em um modelo linear especificado da seguinte maneira:

$$NotaCN_i = \beta_0 + \beta_{ind}IND_i + \beta_{esc}ESC_i + \varepsilon_i$$

Na equação acima, as quantidades β representam os parâmetros constantes do modelo que serão ajustados aos dados. As quantidades ε_i designam os resíduos da regressão. Eles correspondem à diferença entre a nota obtida por cada estudante ($NotaCN_i$) e a nota esperada em vista da origem social individual e escolar ($\beta_0 + \beta_{ind}IND_i + \beta_{esc}ESC_i$). A saber, o modelo acima é uma generalização multivariada das regressões lineares típicas dos laboratórios de ensino de Física. Ele tenta ajustar retas aos dados, supondo que variações na origem social sejam capazes de produzir variações proporcionais no desempenho científico dos estudantes. Conforme veremos a seguir, o modelo permitirá quantificar a relação entre essas variáveis, determinando o desempenho esperado em Ciências da Natureza para cada perfil socioeconômico.

Ainda que a origem social seja sempre relevante, o tamanho do seu efeito sobre o desempenho escolar dos estudantes pode mudar entre diferentes países e épocas (OECD, 2010). Para avaliarmos como o efeito da origem social muda ao longo do tempo, incluímos termos de interação de IND e ESC com o ano de realização do ENEM.

Antes de observarmos o resultado do modelo acima, é importante examinar os chamados “gráficos diagnósticos” (Fahrmeir, Kneib, Lang, & Marx, 2013). Eles permitem avaliar se as premissas do modelo foram satisfeitas. Em caso negativo, as inferências feitas a partir do modelo poderão estar comprometidas. Sendo assim, essa é uma etapa muito importante para garantir a validade dos resultados que a análise está prestes a produzir. A Figura 1 apresenta os quatro gráficos diagnósticos simultaneamente. O analista familiarizado com modelos de regressão linear múltipla perceberá que as premissas do modelo foram todas satisfeitas¹. Portanto, a análise pode prosseguir.

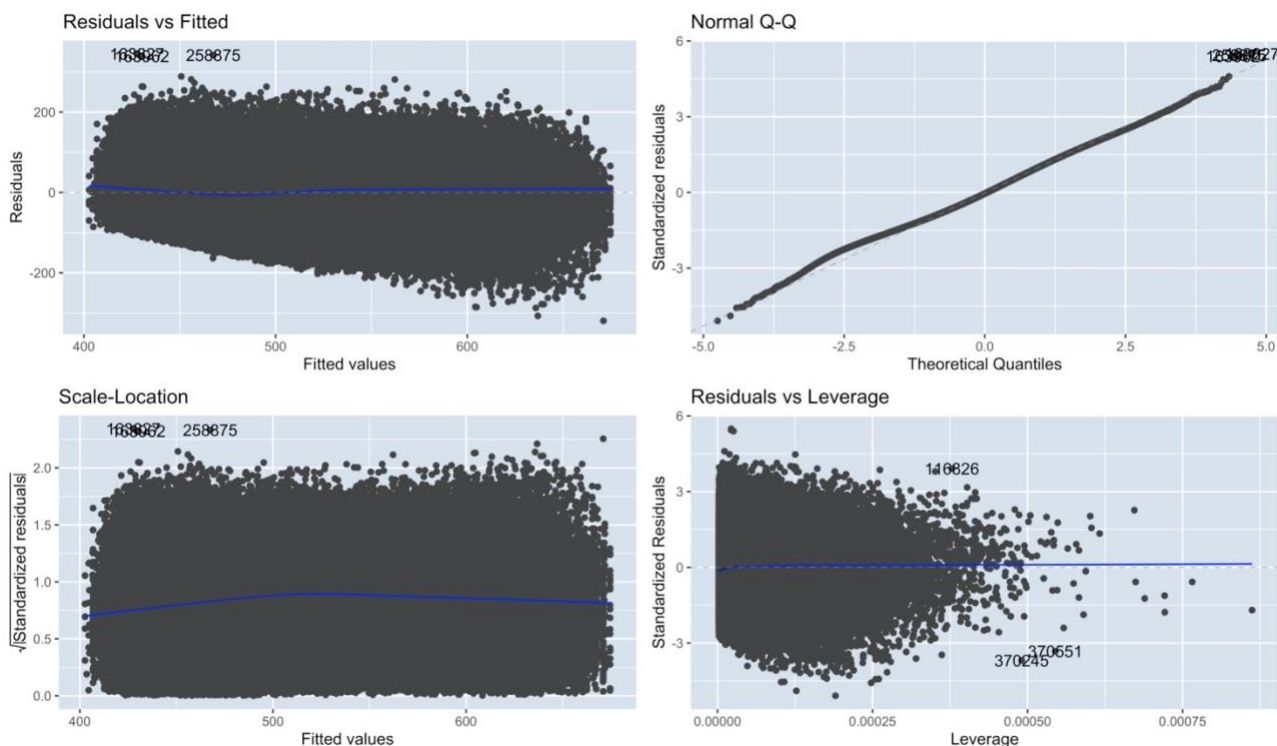


Figura 1 – Gráficos diagnósticos do modelo de regressão linear que prediz o desempenho em Ciências da Natureza a partir da origem social dos estudantes. **Fonte:** Autores.

O *output* do modelo final com efeitos principais e interações pode ser lido abaixo:

```
Call:
lm(formula = NU_NOTA_CN ~ (Dim1 + m.Dim1) * NU_ANO - NU_ANO,
    data = dados)

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  484.8068   0.0902 5374.619 < 2e-16 ***
Ind           9.5324   0.2540  37.533 < 2e-16 ***
Esc          36.5373   0.3278 111.453 < 2e-16 ***
Ind:NU_ANO2013 -0.8096   0.3485  -2.323 0.020174 *
Ind:NU_ANO2014 -3.0122   0.3465 -8.694 < 2e-16 ***
Ind:NU_ANO2015 -0.5182   0.3434 -1.509 0.131281
Ind:NU_ANO2016 -0.4023   0.3430 -1.173 0.240843
```

¹ A saber, as premissas do modelo linear clássico testadas nos gráficos diagnósticos são as seguintes (Fahrmeir et al., 2013): (1) os resíduos devem ser independentes e normalmente distribuídos; (2) a variância dos resíduos devem ser a mesma em todos os grupos e segmentos da amostra. Além disso, é importante a não ocorrência de *outliers* avançados. Todas essas condições podem ser avaliadas nos gráficos diagnósticos da seguinte maneira: (1) os pontos do gráfico “Normal Q-Q” devem estar alinhados, formando uma reta crescente; (2) as linhas de tendência (em cor azul) nos demais gráficos devem ser aproximadamente horizontais.

```

Ind:NU_ANO2017 -0.6102 0.3483 -1.752 0.079815 .
Ind:NU_ANO2018 -0.6262 0.3606 -1.736 0.082496 .
Ind:NU_ANO2019 -1.4007 0.3730 -3.755 0.000174 ***
Esc:NU_ANO2013 -1.3711 0.4836 -2.835 0.004580 **
Esc:NU_ANO2014 -3.2481 0.4834 -6.719 1.83e-11 ***
Esc:NU_ANO2015 -2.4048 0.4570 -5.262 1.42e-07 ***
Esc:NU_ANO2016 -2.9631 0.4528 -6.544 6.00e-11 ***
Esc:NU_ANO2017 -7.5721 0.4604 -16.448 < 2e-16 ***
Esc:NU_ANO2018 -4.7086 0.4757 -9.899 < 2e-16 ***
Esc:NU_ANO2019 -5.5465 0.4950 -11.204 < 2e-16 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 62.83 on 489150 degrees of freedom
Multiple R-squared: 0.3583, Adjusted R-squared: 0.3583
F-statistic: 1.707e+04 on 16 and 489150 DF, p-value: < 2.2e-16

```

Ao final do output, vemos que “*p-value*: < 2.2e-16”. Logo, o modelo é estatisticamente significativo e os resultados desta análise não podem ser atribuídos ao acaso. Porém, dada a quantidade de observações e a natureza do modelo, esse resultado era esperado: nós já sabíamos que a origem social é um preditor importante do desempenho dos estudantes em Ciências da Natureza (Kleinke, 2017; Nascimento *et al.*, 2018). A novidade aqui é determinar a magnitude do efeito da origem social sobre o desempenho dos estudantes brasileiros. No *output* acima, há diversas informações conclusivas com respeito às nossas questões de pesquisa. Nas seções seguintes, nós nos dedicamos a interpretar e discutir as informações disponíveis nesse *output*.

Analizando o tamanho do efeito

A primeira medida que pode ser interpretada como “tamanho do efeito” da origem social é o chamado *coeficiente de determinação* R^2 . Ele assume valores entre 0 e 1 e representa a fração de variância da variável dependente (NotaCN) explicada pelas demais variáveis do modelo (IND, ESC e suas interações). No *output*, vemos “*R-squared*: 0.3583”. Isso significa que o poder explicativo da origem social sobre o desempenho dos estudantes brasileiros em Ciências da Natureza pode ser estimado em 35,8%. Portanto, é correto afirmar que pouco mais de um terço do desempenho observado entre estudantes brasileiros pode ser explicado pela origem social desses estudantes. Uma questão importante agora é avaliar se 35,8% é muito ou pouco. De fato, se estivéssemos desenvolvendo uma vacina, uma eficiência da ordem de 30% a 40% representaria um grande fracasso. Por essa razão, alguns manuais de estatística aplicada informam que um efeito de 30% pode ser considerado pequeno. Porém, em se tratando de dados educacionais (que geralmente carregam muita variação individual) e de um efeito que não deveria acontecer de maneira alguma (i.e., a reprodução, na escola, de desigualdades sociais), esse percentual pode ser considerado elevadíssimo!

Outra maneira de mensurarmos o efeito das variáveis é examinar as estimativas dos parâmetros β representados na coluna “*Estimates*”. O primeiro parâmetro marca o valor de interceptação do modelo, que designamos por β_0 neste artigo. Ele corresponde ao valor esperado da variável independente (NotaCN) quando todas as demais são iguais a zero (i.e., quando $IND = 0$ e $ESC = 0$). Consultando o Gráfico 3 e o Quadro 2, podemos perceber que o valor de interceptação do modelo ($NotaCN = 484,8$) corresponde às famílias de escolaridade fundamental e média com renda em torno de 2 salários mínimos. Os dois parâmetros seguintes, designados por β_{ind} e β_{esc} , acompanham as variáveis IND e ESC. Da maneira como o modelo foi especificado (i.e., com interações), os valores de β_{ind} e β_{esc} devem ser atribuídos ao certame de 2012 (voltaremos a essa questão na próxima seção).

Visando contribuir para a interpretação dos parâmetros do modelo declarados no *output*, geramos um gráfico que representa a fração dos candidatos ao ENEM de 2012 com posição individual de classe $IND \cong 0$. Com respeito aos estudantes que satisfazem essas condições, geramos um diagrama de dispersão do desempenho na prova de Ciências da Natureza (NotaCN) contra a posição de classe da escola (ESC). Esse gráfico permite avaliar, portanto, o efeito social da escola quando o efeito individual de classe é mantido constante. O resultado foi publicado no Gráfico 4. Nele, a reta vermelha representa o modelo linear ajustado. Observe que essa reta acompanha a inclinação da nuvem de pontos: os estudantes de baixo desempenho científico ($NotaCN < 400$) estão sobrerrepresentados nas escolas de classe média-baixa ($ESC < 1$) enquanto estudantes posicionados em escolas relativamente privilegiadas ($ESC > 1$) tendem a estar mais bem posicionados na escala do exame.

Ao mesmo tempo, é flagrante uma grande dispersão dos indivíduos em torno da reta do modelo. O leitor não deve se surpreender com isso. De fato, essa dispersão individual é muito importante para entendermos adequadamente o estado usual das desigualdades sociais frente à escola: ainda que existam tendências estruturais que marcam privilégios de classe, não é razoável esperar que a estrutura social esteja identicamente incorporada a todos os indivíduos. O privilégio de classe frente à escola é uma tendência não desprezível em torno da qual se acumulam diversas exceções individuais. Por essa razão, é sempre possível encontrar casos de sucesso escolar em meios populares (Lahire, 1997). Essa possibilidade, por outro lado, em nada contradiz a tendência central explicitada no modelo linear que estamos analisando.

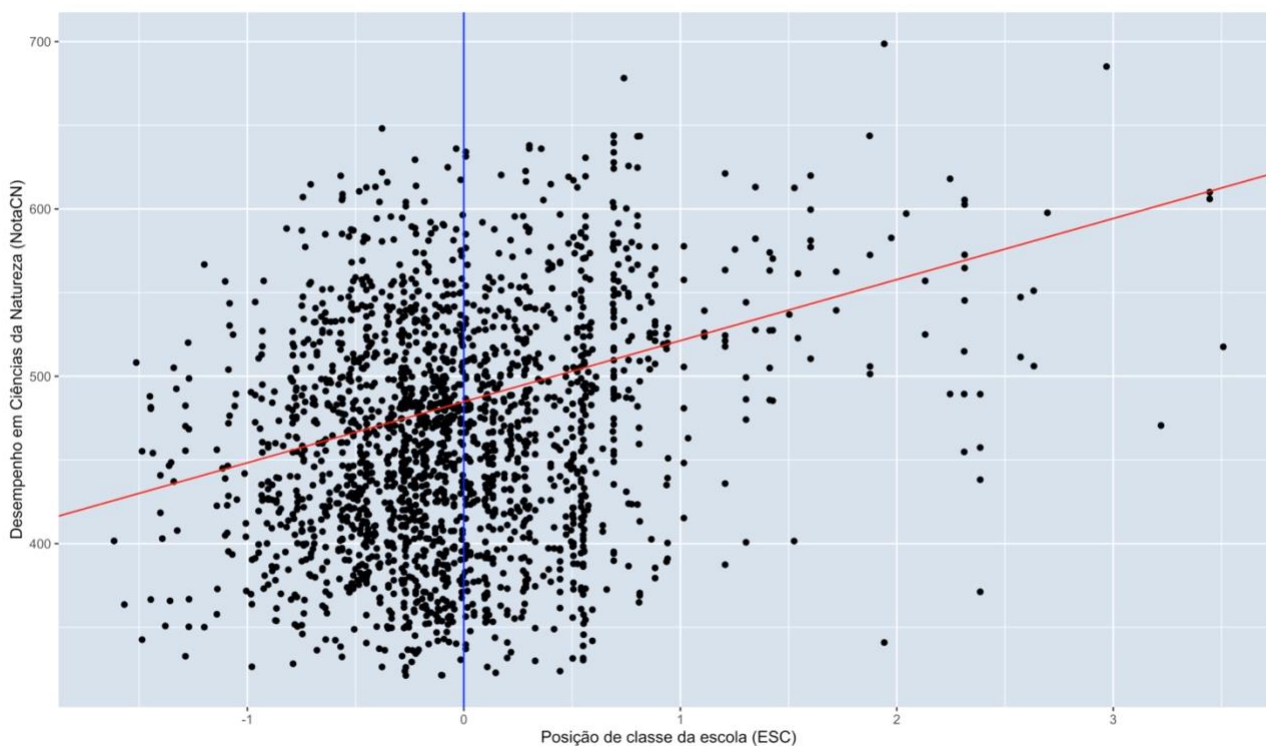


Gráfico 4 – Diagrama de dispersão do desempenho dos estudantes na prova de Ciências da Natureza (*NotaCN*) do ENEM de 2012 contra a posição social da escola (*ESC*) quando a posição individual de classe é aproximadamente nula ($IND \cong 0$). **Fonte:** Autores.

A saber, os parâmetros que determinam a reta representada no Gráfico 4 foram extraídos do *output*. O ponto de interceptação ($\beta_0 = 484,8$) corresponde ao ponto em que a reta do modelo (em vermelho) cruza o eixo vertical (em azul). O parâmetro do efeito social da escola ($\beta_{esc} = 36,5$) corresponde à declividade do gráfico. Portanto, quanto maior for esse parâmetro, mais inclinada será a reta do modelo e maior será o efeito social da escola sobre o desempenho dos estudantes em Ciências da Natureza. O leitor pode experimentar valores diferentes na equação do modelo para ter uma ideia prática do quanto a origem social muda o desempenho esperado dos estudantes na prova. Por exemplo, nos casos em que $IND \cong 2$ e $ESC \cong 2$ (ou seja, filhos de famílias com pós-graduação e renda superior a 20 salários mínimos que estudam em escolas da mesma posição de classe), o desempenho esperado \widehat{NotaCN} pode ser calculado fazendo:

$$\widehat{NotaCN} = 484,8 + 9,53 \cdot 2 + 36,53 \cdot 2 = 576,92$$

O leitor pode comparar esses resultados com outros valores de referência (e.g., a nota de corte para ingresso nos cursos de sua universidade) para perceber o quanto a diferença acima representa uma vantagem elevada. Finalmente, observe que, controlando o efeito das demais variáveis, o efeito social da escola ($\beta_{esc} = 36,5$) chega a ser quase quatro vezes superior ao efeito da origem social individual ($\beta_{ind} = 9,53$). Isso mostra que a posição média do corpo discente na estrutura das relações de classe pode ser considerada muito mais importante que a posição individual da família de cada estudante.

Transformações ao longo do tempo

Outra questão que nós nos fizemos no início desse artigo diz respeito às transformações do efeito da origem social. Já sabemos que as posições dos indivíduos e da escola na estrutura das relações de classe explicam, juntas, 35,8% do desempenho dos estudantes brasileiros em Ciências Naturais. Mas esse poder explicativo é mantido constante ao longo dos anos? Para avaliar essa questão, consideramos os termos de interação do modelo linear que ajustamos aos dados. Os termos de interação informam como os coeficientes lineares β_{ind} e β_{esc} devem ser ajustados anualmente, tomando 2012 como ano de referência. Por exemplo, ao final do *output*, lemos o seguinte:

```
Esc:NU_ANO2018 -4.7086 0.4757 -9.899 < 2e-16 ***
Esc:NU_ANO2019 -5.5465 0.4950 -11.204 < 2e-16 ***
```

Isso significa que o coeficiente linear associado à variável ESC (β_{esc}) deve sofrer uma redução de 4,7 pontos no ano de 2018 e de 5,5 pontos no ano de 2019 com relação ao valor estimado no ano de 2012. Feitas essas considerações, nós somamos os efeitos de interação e geramos um diagrama de dispersão que nos permite avaliar a série temporal dos efeitos combinados da origem social (individual e escolar). O resultado está disponível no Gráfico 5.

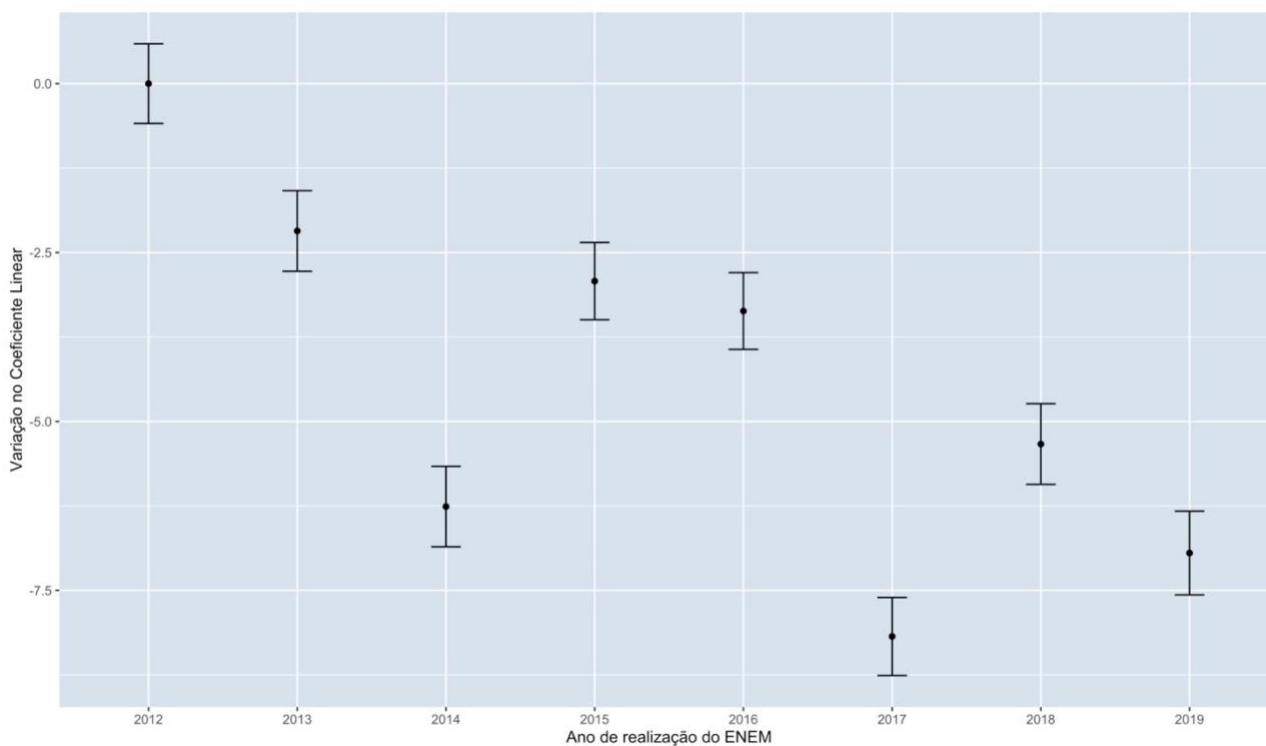


Gráfico 5 – Variações combinadas nos coeficientes lineares do modelo devidas aos termos de interação.

Fonte: Autores.

Como é possível perceber, o Gráfico 5 indica uma tendência de queda. Ou seja, o poder preditivo da origem social sobre o desempenho dos brasileiros em Ciências da Natureza esteve diminuindo no período que vai de 2012 a 2019. Os anos de 2014 e 2018 apresentam resultados discrepantes que não prejudicam a tendência geral de queda. As razões dessa queda, no entanto, ainda precisam ser examinadas. Talvez sejam devidas a uma recomposição da prova que tenha incluído progressivamente uma quantidade maior de questões com baixa discriminação social (Nascimento *et al.*, 2018). Isso pode não ser positivo, visto que questões de baixa discriminação são, também, questões com menor taxa de acerto global (Kleinke, 2017). Talvez a redução do poder discriminatório da prova reflita o efeito tardio de políticas de bem-estar social implementadas nos anos anteriores. De qualquer maneira, a queda observada aqui está longe de neutralizar os efeitos de classe sobre o resultado escolar, correspondendo a um recuo máximo de 7 pontos em um coeficiente conjunto ($\beta_{ind} + \beta_{esc}$) de aproximadamente 46 pontos.

CONSIDERAÇÕES FINAIS

No Brasil, o uso de métodos estatísticos ainda é um tema polêmico para a pesquisa educacional. Análises quantitativas são geralmente consideradas positivistas e, a partir daí, descartadas. Enquanto isso, o sentido real do positivismo em ciências sociais quase nunca é aprofundado. Acredita-se que os métodos quantitativos foram amplamente empregados no passado e superados pela pesquisa qualitativa contemporânea. No entanto, as evidências indicam que nunca houve, no Brasil, uma tradição quantitativa consistente na pesquisa educacional (Gatti, 2004). De fato, a resistência em quantificar fenômenos educacionais está amplamente baseada em uma relação ingênua e não-crítica com essas quantificações, que sempre estão inscritas em um processo de modelagem. Portanto, as conclusões de análises quantitativas não são mais definitivas ou confiáveis que aquelas que poderíamos obter com outras metodologias. Enquanto ferramentas de pesquisa, métodos qualitativos e quantitativos abrem caminhos distintos, geralmente irreduzíveis uns aos outros.

Outra ingenuidade consiste em acreditar que é possível trabalhar com desigualdades sociais e, ao mesmo tempo, descartar o legado histórico das metodologias quantitativas. Aparentemente, alguns pesquisadores em Educação (em Ciências) esquecem o quanto as perspectivas críticas que temos hoje com respeito às contribuições da escola para a manutenção/superação das desigualdades sociais são devidas a estudos quantitativos do passado e do presente. Ao longo das últimas décadas, a pesquisa em Educação em Ciências tem se tornado mais sensível às desigualdades sociais (Lemke, 2001). Essa sensibilidade tem contribuído para que alguns pesquisadores se voltem aos métodos tradicionalmente empregados pelas áreas de conhecimento que, há muito tempo, investigam as relações entre escolaridade e desigualdade social (Bourdieu, 1984). Esse retorno crítico aos métodos quantitativos (retorno ao que nunca esteve efetivamente estabelecido na pesquisa educacional) expande as oportunidades de transferência de conhecimentos das Ciências da Natureza para o campo da Educação. Neste artigo, empregamos uma ferramenta estatística que permite a análise gráfica de dados em uma tabela de contingência (Greenacre, 2007). Em seguida, construímos um modelo linear capaz de prever o desempenho científico de estudantes brasileiros em função da sua origem social. Esse modelo linear nos permitiu responder às duas questões de pesquisa enunciadas na abertura do artigo.

Em primeiro lugar, observamos que a origem social (determinada somente pelas variáveis renda e escolaridade dos pais) é capaz de explicar mais de 35% do desempenho científico dos estudantes brasileiros. Considerando a noção contemporânea de classe social (Bourdieu, 1984), essas variáveis representam informação muito limitada. Justamente por isso, é impressionante que um modelo baseado nelas atinja um poder explicativo tão elevado. Por exemplo, ao analisar dados do vestibular da UFRGS, Silveira (1999) elaborou um modelo com poder explicativo semelhante ($R^2 = 34.2\%$) mas contendo uma quantidade muito superior de variáveis ($k = 15$). Uma vantagem de trabalhar com dados nacionais é sua variabilidade. Exames vestibulares realizados por universidades distintas tendem a sofrer de um processo que Bourdieu e Passeron (2009) designaram por “superseleção”. A superseleção resulta de uma seleção anterior (no caso, uma autoeliminação) que contribui para que a amostra não represente o comportamento médio da população sobre a qual nós gostaríamos de fazer nossas inferências. Enquanto muitos cidadãos optam por não participar do vestibular, o Exame Nacional do Ensino Médio conta com a participação de uma porção muito mais numerosa e diversa da população brasileira. Isso favorece que desigualdades sociais não fiquem tão mascaradas quanto nos estudos restritos a uma única instituição.

Em segundo lugar, o chamado “efeito social da escola” (i.e., o efeito da origem social média dos estudantes de cada escola) resultou quase quatro vezes superior ao efeito da origem social individual em nosso modelo. Esse resultado está alinhado às análises do PISA (OECD, 2010) e algumas explicações foram antecipadas. É possível, por exemplo, que a predominância do efeito social da escola reflita a distribuição desigual de recursos humanos e materiais entre instituições de ensino que atendem a públicos diferentes. De fato, a desigualdade na distribuição de recursos é uma preocupação típica e importante na educação científica das periferias urbanas (Garcia-Silva & Lima Junior, 2020). Outra explicação possível para a prevalência do efeito social da escola atribui ao convívio com os colegas a capacidade de modificar as estratégias de investimento escolar dos estudantes. Com respeito a esse ponto de vista, alguma evidência pode ser encontrada na análise de história de vida de estudantes de classe popular que, convivendo com colegas de origem social privilegiada, passam a adotar ambições escolares semelhantes (Lima Junior, 2013). No entanto, mais pesquisas precisam ser feitas a respeito.

Enfim, ao contrário do que foi observado no período de 2006 a 2012 (Lima Junior, 2015), a associação entre origem social e desempenho científico dos estudantes brasileiros apresentou tendência de queda de 2012 a 2019. Ela pode ser um efeito tardio das políticas de bem-estar social e de valorização do corpo docente implementadas em anos anteriores. Essa redução também poderia ser explicada por uma recomposição do

próprio ENEM, incluindo menos itens de alta discriminação social (Kleinke, 2017; Nascimento *et al.*, 2018). De qualquer maneira, as razões dessa queda ainda estão bastante nebulosas e demandam pesquisas posteriores para serem compreendidas.

Agradecimentos

Os autores agradecem ao apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq (processo 436910/2018-7).

REFERÊNCIAS

- Almeida, M. P. de, & Watanabe, G. (2019). As concepções de graduandos sobre o seu capital cultural: perspectivas e persistências históricas nos contextos educacionais. *Revista Brasileira de Ensino de Ciência e Tecnologia*, 12(3). <https://doi.org/10.3895/rbect.v12n3.8538>
- Anderhag, P., Hamza, K. M., & Wickman, P.-O. (2015). What Can a Teacher Do to Support Students' Interest in Science? A Study of the Constitution of Taste in a Science Classroom. *Research in Science Education*, 45(5), 749–784. <https://doi.org/10.1007/s11165-014-9448-4>
- Archer, L., Dawson, E., DeWitt, J., Seakins, A., & Wong, B. (2015). "Science capital": A conceptual, methodological, and empirical argument for extending bourdieusian notions of capital beyond the arts. *Journal of Research in Science Teaching*, 52(7), 922–948. <https://doi.org/10.1002/tea.21227>
- Bourdieu, P. (1984). *Distinction: a social critique of the judgement of taste*. London, England: Routledge.
- Bourdieu, P. (1990). *The logic of practice*. Stanford, United States of America: Stanford University Press.
- Bourdieu, P. (2004). *A produção da crença: contribuição para uma economia dos bens simbólicos*. Porto Alegre, RS: Zouk.
- Bourdieu, P. (2007). *A Economia das Trocas Simbólicas* (6a. ed.). São Paulo, SP: Perspectiva.
- Bourdieu, P. (2013). *Outline of a theory of practice*. Cambridge, England: Cambridge University Press.
- Bourdieu, P., & Passeron, J.-C. (2009). *A reprodução: elementos para uma teoria do sistema de ensino* (2a. ed.). Petrópolis, RJ: Vozes.
- Claussen, S., & Osborne, J. (2013). Bourdieu's notion of cultural capital and its implications for the science curriculum. *Science Education*, 97(1), 58–79. <https://doi.org/10.1002/sce.21040>
- Evolucional. (2020). *Consolidação das notas por escola do Enem 2019*. Campinas, SP: Evolucional. <https://enem2019.evolucional.com.br>
- Fahrmeir, L., Kneib, T., Lang, S., & Marx, B. (2013). *Regression: models, methods and applications*. Berlin, Germany: Springer.
- Fernandes, J., Ueno Guimarães, M. H., Robert, A., & Passos, M. M. (2020). Estudo da evasão dos estudantes de Licenciatura e Bacharelado em Física: uma análise à luz da Teoria do Sistema de Ensino de Bourdieu. *Caderno Brasileiro de Ensino de Física*, 37(1), 105–126. <https://doi.org/10.5007/2175-7941.2020v37n1p105>
- Fowler, B. (2020). Pierre Bourdieu on social transformation, with particular reference to political and symbolic revolutions. *Theory and Society*, 49(3), 439–463. <https://doi.org/10.1007/s11186-019-09375-z>
- Garcia-Silva, S., & Lima Junior, P. (2020). A Educação Científica das Periferias Urbanas: Uma Revisão sobre o Ensino de Ciências em Contextos de Vulnerabilidade Social (1985–2018). *Revista Brasileira de Pesquisa Em Educação Em Ciências*, 221–243. <https://doi.org/10.28976/1984-2686rbpec2020u221243>

- Gatti, B. A. (2004). Estudos quantitativos em educação. *Educação e Pesquisa*, 30(1), 11–30. <https://doi.org/10.1590/S1517-97022004000100002>
- Genovese, L. G. R., Queiroz, J. R. de O., & Castilho, D. D. (2015). Incorporação do habitus de homo magister no interior do estágio colaborativo em Física: um olhar sobre a relação entre professor supervisor e estagiário. *Ensino Em Re-Vista*, 22(2), 311–332. <https://doi.org/10.14393/ER-v22n2a2015-5>
- Greenacre, M. (2007). *Correspondence analysis in practice*. New York, United States of America: Chapman & Hall.
- Heidemann, L. A., Araujo, I. S., & Veit, E. A. (2016). Modelagem Didático-científica: integrando atividades experimentais e o pro-cesso de modelagem científica no ensino de Física. *Caderno Brasileiro de Ensino de Física*, 33(1), 3. <https://doi.org/10.5007/2175-7941.2016v33n1p3>
- Kleinke, M. U. (2017). Influência do status socioeconômico no desempenho dos estudantes nos itens de física do Enem 2012. *Revista Brasileira de Ensino de Física*, 39(2), e2402. <https://doi.org/10.1590/1806-9126-RBEF-2016-0081>
- Klüger, E. (2018). Análise de correspondências múltiplas: fundamentos, elaboração e interpretação. *Revista Brasileira de Informação Bibliográfica Em Ciências Sociais*, 86(2), 68–97. <https://doi.org/10.17666/bib8604/2018>
- Lahire, B. (1997). *Sucesso escolar em meios populares: as razões do improvável*. São Paulo, SP: Ática.
- Lemke, J. L. L. (2001). Articulating communities: Sociocultural perspectives on science education. *Journal of Research in Science Teaching*, 38(3), 296–316. [https://doi.org/10.1002/1098-2736\(200103\)38:3<296::AID-TEA1007>3.0.CO;2-R](https://doi.org/10.1002/1098-2736(200103)38:3<296::AID-TEA1007>3.0.CO;2-R)
- Lima Junior, P., Andrade, V. C. de, Fraga Junior, J. C., Silva, J. A. e, Goulart, F. M., & Araújo, I. M. (2020). Excelência, evasão e experiências de integração dos estudantes de graduação em Física. *Ensaio Pesquisa em Educação em Ciências (Belo Horizonte)*, 22, e12165. <https://dx.doi.org/10.1590/21172020210140>
- Lima Junior, P., Fraga Junior, J. C., Andrade, V. C. de, & Bernardino, P. R. P. (2020). A Integração dos Estudantes de Periferia no Curso de Física: razões institucionais da evasão segundo a origem social. *Ciência & Educação (Bauru)*, 26, e20030. <https://dx.doi.org/10.1590/1516-731320200030>
- Lima Junior, P., Ostermann, F., & Rezende, F. (2018). *Razões para desistir: análise sociológica da evasão no curso de Física*. Curitiba, PR: Appris Editora.
- Lima Junior, P. (2015). Crítica sociológica do Exame Nacional do Ensino Médio: uma análise bourdiana. *Atas Do X Encontro Nacional de Pesquisa Em Educação Em Ciências*, 1–8. Águas de Lindóia: ABRAPEC. Recuperado de <http://www.abrapecnet.org.br/enpec/x-enpec/anais2015/resumos/R1971-1.PDF>
- Lima Junior, P., Ostermann, F., & Rezende, F. (2013). Análise dos condicionantes sociais do sucesso acadêmico em cursos de graduação em física à luz da sociologia de Bourdieu. *Ensaio Pesquisa Em Educação Em Ciências*, 15(1), 113–129. <https://dx.doi.org/10.1590/1983-21172013150108>
- Lima Junior, P., Pinheiro, N., & Ostermann, F. (2012). Bourdieu en la educación científica: consecuencias para la enseñanza y la investigación. *Didáctica de Las Ciencias Experimentales y Sociales*, 160(26), 145–160. <https://doi.org/10.7203/dces.26.1925>
- Massi, L., Muzzeti, L. R., & Suficier, D. M. (2017). A pesquisa sobre trajetórias escolares no Brasil. *Revista Ibero-Americana de Estudos Em Educação*, 12(3), 1854–1873. <https://doi.org/10.21723/riaee.v12.n.3.2017.10364>
- Massi, L., & Villani, A. (2015). Um caso de contratendência: baixa evasão na licenciatura em química explicada pelas disposições e integrações. *Educação e Pesquisa*, 41(4), 975–992. <https://doi.org/10.1590/s1517-9702201512135667>

- Mello, A. C. R. de, & Higa, I. (2018). Busca por capitais no campo da escola e sua relação com o desenvolvimento profissional docente de professores supervisores de estágio de Ciências e Biologia. *Ciência & Educação (Bauru)*, 24(2), 301–317. <https://doi.org/10.1590/1516-731320180020004>
- Moreno, A. C., & Oliveira, E. (2019). Brasil cai em ranking mundial de educação em matemática e Ciências; e fica estagnado em leitura. G1. Recuperado de <https://g1.globo.com/educacao/noticia/2019/12/03/brasil-cai-em-ranking-mundial-de-educacao-em-matematica-e-ciencias-e-fica-estagnado-em-leitura.ghtml>
- Nascimento, M. M., Cavalcanti, C., & Ostermann, F. (2018). Uma busca por questões de Física do ENEM potencialmente não reprodutoras das desigualdades socioeconômicas. *Revista Brasileira de Ensino de Física*, 40(3). <https://doi.org/10.1590/1806-9126-rbef-2017-0237>
- Nogueira, C. M. M., & Nogueira, M. A. (2009). *Bourdieu e a Educação*. Belo Horizonte, MG: Autêntica.
- OECD. (2010). *PISA 2009 Results: Overcoming Social Background*. Paris, France: OECD. <https://doi.org/10.1787/9789264091504-en>
- Salej, S. (2005). Quarenta anos do Relatório Coleman : capital social e educação. *Educação Unisinos*, 9(2), 116–129. Recuperado de <http://revistas.unisinos.br/index.php/educacao/article/view/6310>
- Silva, F. P., & Pereira, A. P. de. (2019). Uma revisão sobre fracasso escolar no ensino de Ciências A review on school failure in science education. *Atas Do XII Encontro Nacional de Pesquisa Em Educação Em Ciências*, 1–7. Natal: ABRAPEC. Recuperado de <http://abrapecnet.org.br/enpec/xii-enpec/anais/resumos/1/R1393-1.pdf>
- Silva, T. T. da. (2014). *Documentos de Identidade*. Belo Horizonte, MG: Autêntica.
- Silveira, F. L. da. (1999). Um exemplo de análise multivariada aplicada à pesquisa quantitativa em ensino de Ciências: explicando o desempenho dos candidatos ao concurso vestibular de 1999 da Universidade Federal do Rio Grande do Sul. *Investigações Em Ensino de Ciências*, 4(2), 161–180. Recuperado de <https://www.if.ufrgs.br/cref/ojs/index.php/ienci/article/view/607>
- Watanabe, G., & Gurgel, I. (2017). As Marcas Sociais Deixadas Pelas Escolas Em Nossos Professores De Ciências: a Questão Da Violência Simbólica. *Revista Contexto & Educação*, 31(99), 116. <https://doi.org/10.21527/2179-1309.2016.99.116-148>

Recebido em: 26.05.2020

Aceito em: 21.02.2021