

Explorando o *Gap* de Gênero em Matemática no ENEM em Municípios Brasileiros como Indicativo de Potenciais Candidatas para Carreiras em Computação

Tiago F. Tavares¹, Fábio R. Miranda¹

¹Inspere Instituto de Ensino e Pesquisa
Rua Quatá, 300, São Paulo – SP – Brazil

{tiagoft1, fabiomiranda}@insper.edu.br

Abstract. *In the fields of STEM (Science, Technology, Engineering, and Mathematics), there is typically a significant difference between the sizes of the male and female populations, known as the gender gap. This phenomenon is observed worldwide and is related to cultural aspects that can be measured from high school onwards. In this article, we estimate the gender gap by analyzing the math grades from the ENEM exam. Data analysis shows that there are cities where the gender gap is consistently small, indicating potential sources of further inspiration for actions to reduce this phenomenon.*

Resumo. *Nas ciências exatas e tecnológicas tipicamente há uma grande diferença entre o tamanho das populações masculina e feminina, o que é conhecido como gap de gênero. Esse fenômeno é observado mundialmente em disciplinas do campo STEM (Science, Technology, Engineering, and Math; ciência, tecnologia, engenharia e matemática), e é relacionado a aspectos culturais que são mensuráveis a partir do ensino médio. Neste artigo, estimamos o gap de gênero à partir da análise das notas de matemática de alunos no ENEM. A análise dos dados mostra que há cidades que consistentemente têm um gap de gênero pequeno, o que indica possíveis fontes de inspiração para ações futuras na redução deste fenômeno.*

1. Introdução

No campo STEM (*Science, Technology, Engineering, and Math*; ciência, tecnologia, engenharia e matemática), tipicamente observa-se que homens estão muito mais presentes que mulheres (Kanny et al. 2014; Alvarenga and Braga 2024). Essa disparidade, conhecida como *gap* de gênero, pode ser observada tanto na contagem das populações que optam por carreiras universitárias ligadas a STEM quanto na diferença média de desempenho entre homens e mulheres em testes padronizados anteriormente à escolha da universidade (Cimpian et al. 2020; Holanda et al. 2022). O *gap* de gênero geralmente está associado a questões culturais que impõem a adesão a certos papéis sociais (Makarova et al. 2019; Santos et al. 2019; Sakellariou and Fang 2021; Alvarenga and Braga 2024). Essas disparidades podem ser observadas após o ensino médio (Card and Payne 2020; Marcom and Aragão 2022), embora sejam menos evidentes ao final do ensino fundamental (Marcom and Aragão 2022).

O *gap* de gênero se relaciona a consequências negativas para a diversidade dos cursos nas áreas de exatas e pode levar a culturas organizacionais hostis para mu-

Iheres (Santos et al. 2019). Estes tipos de cultura organizacional podem evoluir para situações de diferenças salariais ligadas ao gênero (Alvarenga and Braga 2024).

Uma das maneiras de combater o *gap* de gênero é através de ações afirmativas na divulgação e no processo de recrutamento em cursos universitários. Para este fim, é relevante entender e medir as dinâmicas do *gap* de gênero entre as pessoas que são potenciais candidaturas para cursos universitários da área. Nesse contexto, as potenciais candidaturas podem ser encontradas dentre as notas do Exame Nacional do Ensino Médio (ENEM), uma vez que trata-se de um exame padronizado que mede o desempenho dos participantes em escala nacional.

O estudo do desempenho de matemática no ENEM é relevante para compreender a opção por carreiras de ciência e tecnologia, mais especificamente em computação. Num contexto canadense, Blotnicky (Blotnicky et al. 2018) mostrou que, no ensino médio, as variáveis que mais se correlacionam com propensão a seguir carreiras STEM são auto-eficácia (Bandura 1977) em matemática e conhecimento sobre o que fazem as carreiras de tecnologia. No Brasil, diversas iniciativas (Novaes et al. 2023; Lima et al. 2022) têm sido fundamentais para promover o entendimento entre mulheres do ensino médio sobre as carreiras em tecnologia, preditor importante de trajetórias nas áreas de STEM.

O conceito da autoeficácia (Bandura 1977) refere-se à crença de uma pessoa na sua habilidade em controlar suas atitudes e eventos que afetam sua própria vida. Resiliência face a adversidade e stress, desempenho e realizações no âmbito escolar estão associados à auto-eficácia. A autoeficácia em matemática está fortemente associada ao próprio desempenho na disciplina (Zakariya 2022). Essa variável prevê desempenho na área melhor que habilidade mental, conhecimento anterior de matemática, utilidade percebida de matemática, notas em outros testes de inteligência, traços de personalidade, extraversão e auto-estima. Portanto, a autoeficácia de matemática pode ser vista como preditor tanto de desempenho em matemática quanto de potencial opção por carreiras STEM. Neste estudo consideramos que, entre as mulheres, a opção por carreiras STEM e o desempenho em matemática no ENEM podem estar associados, e, portanto, o desempenho no ENEM pode ser um medidor do potencial para meninas optarem por cursos de ensino superior em áreas STEM.

Como se discutirá na Seção 2, já existem trabalhos anteriores sobre o *gap* de gênero no ENEM. Essa medição, realizada em nível nacional, aponta para um *gap* de gênero muito maior entre os matriculados em instituições de ensino superior que aquele potencialmente imposto pelos resultados do ENEM (da Justa Lemos 2019). Isso significa que há uma forte componente de opção própria, complementar ao desempenho, que determina as escolhas de carreira no fim do ensino médio.

Neste artigo, trazemos uma análise mais granular, mostrada na Seção 3, que identifica que o *gap* de gênero nas notas de matemática do ENEM é um problema maior em algumas cidades que em outras. Isso aponta para um fenômeno potencialmente causado por questões culturais e, por isso mesmo, afetando diferentes regiões do Brasil em medidas diferentes. A Seção 4 discutirá que estes conhecimentos são importantes para o planejamento de ações afirmativas na coordenação de cursos visando corrigir essas desigualdades.

2. Análises em nível nacional

O *gap* de gênero pode ser observado na diferença de desempenho entre homens e mulheres nas notas de matemática do Exame Nacional do Ensino Médio (ENEM) (Holanda et al. 2022). Essas notas estão disponíveis à partir de microdados disponibilizados pelo Instituto Nacional de Ensino e Pesquisa (INEP) e podem ser baixados livremente. Uma maneira de realizar essa medição é através do coeficiente d de Cohen (Cohen 1988). Essa medida estatística é usada em análise de dados para quantificar a magnitude de um efeito de interesse ou a diferença entre dois grupos em ciências sociais e psicologia. O d de Cohen é calculado pela seguinte expressão:

$$d_{\text{grupo}} = \frac{\mu_{\text{grupo}} - \mu_{\text{geral}}}{\sigma_{\text{grupo}}}. \quad (1)$$

Usando esses dados, é possível estimar a diferença entre as notas médias em matemática de homens e mulheres ($\Delta d = d_H - d_M$). Como mostra a Figura 1, essa diferença gira em torno do intervalo $0,4 \pm 0,04$, tendo aumentado levemente próximo a 2020 e depois caído no biênio 2021-2022.

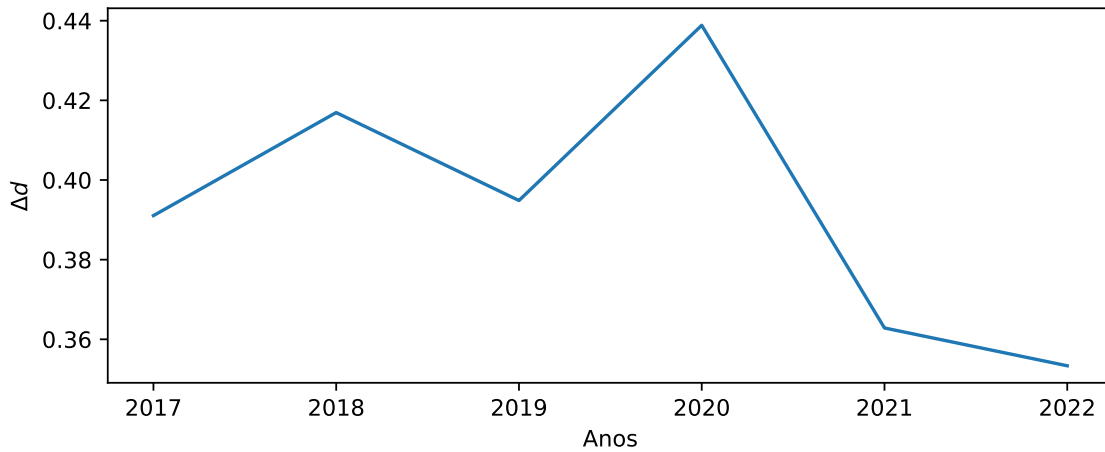


Figura 1. Diferença entre notas (d de Cohen) de matemática de homens e mulheres no ENEM entre 2017 e 2022.

O coeficiente Δd é uma ferramenta importante para a comparação de desempenho e evidencia uma diferença típica de quase meio desvio padrão ligada ao gênero. Porém, em sua relação com a universidade, o ENEM é usado como uma prova classificatória nacional através do sistema SISU, o que significa que apenas alunas e alunos cujas notas estão num restrito quantil superior serão aceitos para matrículas em universidades. Portanto, uma outra maneira de estimar o *gap* de gênero é através da comparação entre probabilidades de homens e mulheres atingirem notas de matemática superiores ao quantil superior, o que pode ser realizado usando o *risk ratio* r_o :

$$r_o = \frac{P(\text{grupo de interesse}|\text{homem})}{P(\text{grupo de interesse}|\text{mulher})} \quad (2)$$

Na análise das notas do ENEM, o grupo de interesse é aquele composto por participantes com nota entre as 10% melhores de todos os candidatos. A Figura 2 mostra

que a razão r_o , considerando uma nota de corte igual ao quantil 0,9 de cada ano, gira em torno de 2,1. Isso significa que, de acordo com os dados, os homens têm duas vezes mais chances que mulheres de conseguir uma nota entre as 10% melhores notas de matemática do ENEM.

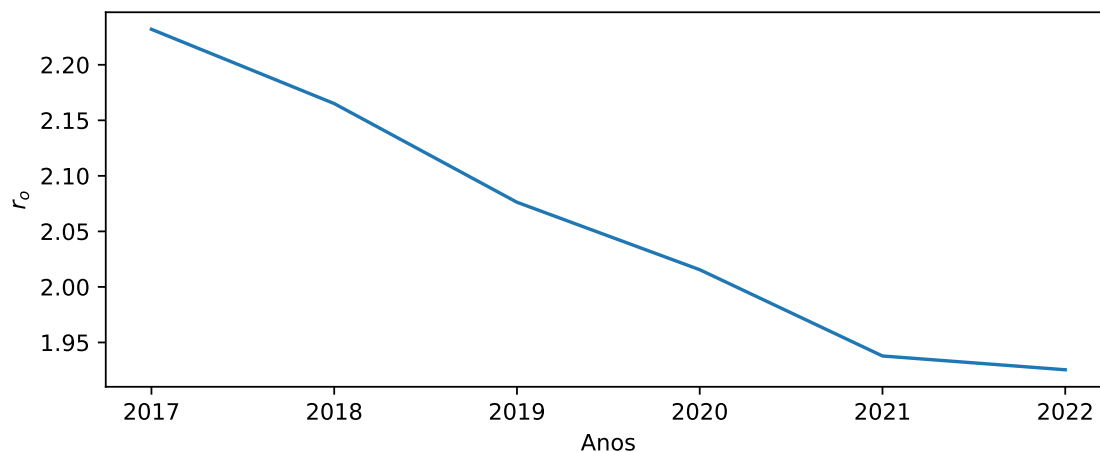


Figura 2. *risk ratio* entre homens e mulheres para notas de matemática acima do quantil 0,9 entre os anos de 2017 e 2022.

O *risk ratio* também pode ser usado para caracterizar a população de homens e mulheres entre os matriculados em cursos de computação. Também usando dados disponibilizados pelo INEP, é possível verificar que, entre os matriculados em cursos de computação presenciais em 2022, há 102086 homens e 15151 mulheres, o que significa um *risk ratio* de 6,74. Esse é um valor mais de 3 vezes superior a aquele calculado no ENEM. Para fins deste trabalho, consideramos como “de computação” todos os cursos que têm a palavra “computação” em seu nome.

É importante destacar que o *gap* de gênero varia significativamente entre diferentes cursos, conforme ilustrado na Figura 3. Enquanto alguns cursos apresentam um *risk ratio* extremamente alto, superando 100, outros exibem um *risk ratio* muito baixo, até mesmo inferior a 1.

Essa distribuição aponta para a ideia de que há diferenças no *gap* de gênero de acordo com aspectos regionais, e possivelmente socioeconômicos. Por isso, investigamos como se comporta o *gap* de gênero nas notas do ENEM quando realizamos análises de participantes agrupados pelo município onde frequentaram o ensino médio.

3. Análise por município

As análises a seguir calculam o *risk ratio*, conforme a Equação 2, considerando a nota de corte igual ao quantil 0,9 da nota de matemática nacional e o grupo de candidatos que frequentaram a escola em cada município. Essas análises partem do pressuposto de que cada município delimita um conjunto de valores culturais e condições socioeconômicas. É evidente que há desigualdades presentes dentro de cada município, porém elas foram deixadas de lado nesta análise.

Como há muitos municípios pequenos, foram excluídos da análise aqueles cujo total de alunos com nota superior ao quantil 0,9 daquele ano fosse menor que 100. Escolher

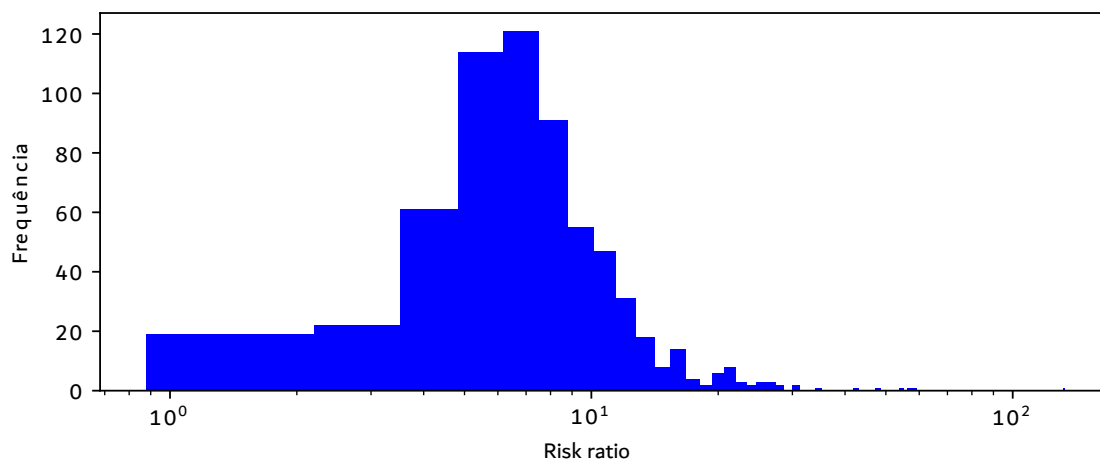


Figura 3. Histograma da razão entre as populações de homens e mulheres entre matriculadas e matriculados em cursos de computação no ano de 2022.

número menores pode levar à inclusão de grupos com maior variância, ao passo que escolher números maiores pode levar à exclusão excessiva de municípios, o que prejudicaria as análises seguintes.

O resultado desta análise é uma distribuição, ao longo dos anos, da *risk ratio* ligada ao *gap* de gênero de cada município. Como mostra a Figura 4, a mediana desse *risk ratio* se manteve praticamente estável em torno de 2,1. Porém, os intervalos interquartil revelam que os dados de municípios no geral podem tipicamente se distribuir entre pouco menos de 1,5 e pouco mais de 3,0. Essa diferença revela que, a cada ano, mesmo que haja municípios com o *gap* de gênero mais evidente e outros em que o *gap* é menor, trata-se de um fenômeno sempre presente em municípios brasileiros.

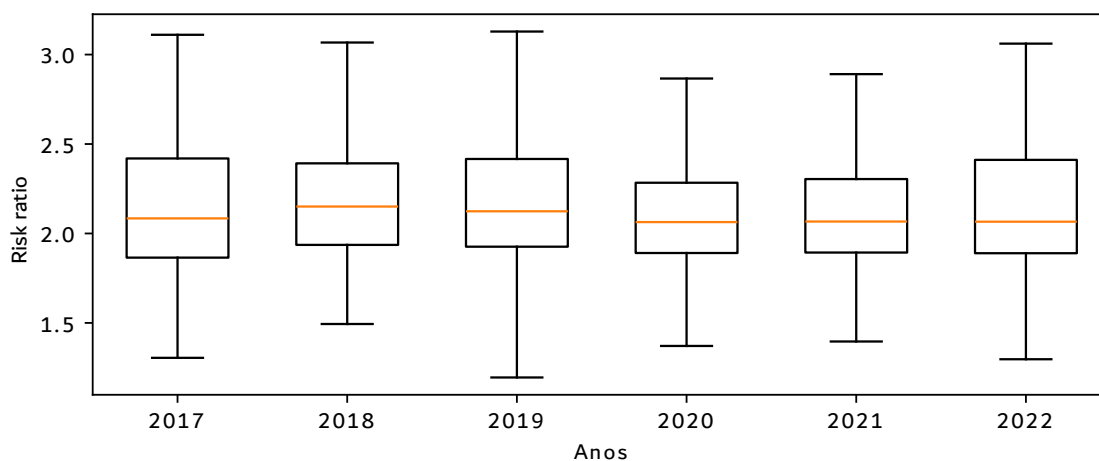


Figura 4. Distribuição de risk ratio (r_0) municipal ao longo dos anos.

Anteriormente, a Figura 2 havia mostrado um resultado calculado sobre todos os candidatos do ENEM de forma uniforme. Diferentemente, a Figura 4 mostra um resultado referente aos municípios brasileiros ao longo dos anos. É importante ressaltar que mesmo os maiores valores de r_0 não são capazes de explicar inteiramente a diferença de gênero

entre pessoas efetivamente matriculadas em cursos de computação.

Uma pergunta a ser levantada em relação à Figura 4 é se os municípios com r_o próximos a 3 ou próximos a 1,5 são sempre os mesmos. A Figura 5 mostra as trajetórias de r_o por município, sem agrupamentos. Embora seja impossível identificar municípios individuais (ao menos sem uma ferramenta de interatividade), é possível verificar que os municípios com r_o mais altos e mais baixos a cada ano não são os mesmos.

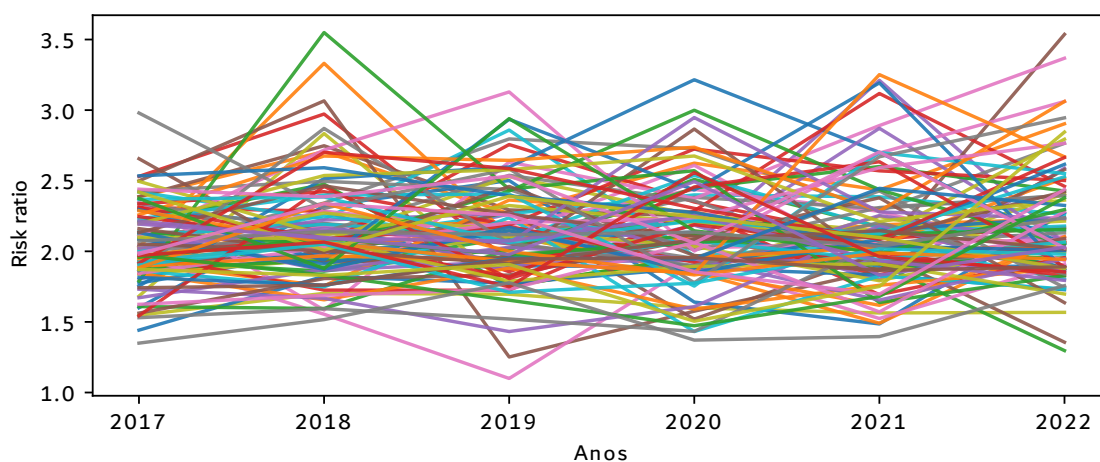


Figura 5. Progressão de *risk ratio* municipal ao longo dos anos. Cada linha corresponde a um município.

Para confirmar essa observação, calculamos a intersecção dos conjuntos dos 30 municípios com maior r_o de cada ano. Essa intersecção é vazia, mostrando que, de fato, há um “revezamento” desses municípios. Porém, ao fazermos o mesmo procedimento com os 30 municípios de menor r_o a cada ano, encontramos três municípios do estado de Minas Gerais: *Viçosa*, *Belo Horizonte* e *Juiz de Fora*. Estes municípios estão consistentemente entre os com menor diferença entre as populações masculina e feminina levando em conta notas de matemática entre as 10% melhores do ENEM.

Ainda, a Figura 5 sugere que, apesar das oscilações anuais, não há uma tendência geral de subida ou descida do *gap* de gênero tal qual medido pelo *risk ratio*. Essa observação fica mais clara ao observarmos os componentes da projeção PCA dessas trajetórias, mostradas na Figura 6. A PCA (D’Agostino Sr. 2005), ou análise de componentes principais, é um método para transformar um conjunto de variáveis correlacionadas em um conjunto menor de variáveis não correlacionadas. Na análise PCA, como é possível observar na figura, não há nenhuma componente principal que claramente mostre uma tendência de subida ou descida; pelo contrário, todas as componentes mudam de direção, sugerindo comportamentos mais ligados a variações anuais que a tendências de longo prazo.

A seção a seguir traz discussões sobre os dados apresentados.

4. Discussões

Os dados mostrados nas seções 2 e 3 mostram que o *gap* de gênero é um fenômeno presente, em menor ou maior grau, em todos os municípios brasileiros. Os dados mos-

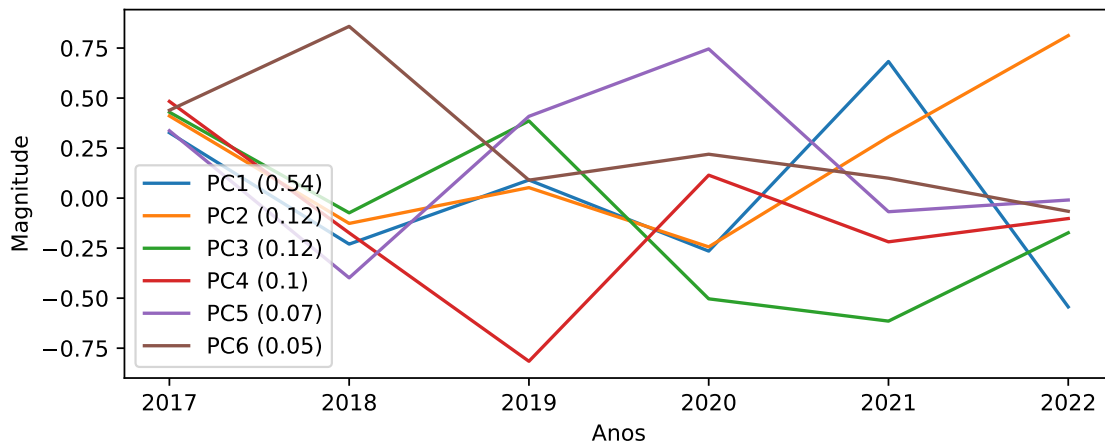


Figura 6. Componentes principais da progressão da *risk ratio* municipal ao longo dos anos. O valor mostrado entre parênteses na legenda é a fração da variância explicada por cada componente.

tram também que a diferença de desempenho no ENEM, embora importante, não é capaz de explicar toda a disparidade populacional entre pessoas matriculadas em cursos de computação. Ainda, os dados apontam para a inexistência de uma tendência global de aumento ou diminuição do *gap* de gênero.

Apesar disso, esta investigação evidenciou que há locais específicos – alguns cursos de computação e alguns municípios – em que o *gap* de gênero é diminuído. É relevante investigar o processo de formação das populações de interesse desses locais específicos para buscar planos de ação efetivos para a redução do *gap* em outros lugares.

5. Limitações do estudo

A população do estudo escolhida como quem está no corte acima do quantil 0.9 na nota de matemática porque corresponde aproximadamente aos alunos com notas do ENEM suficiente para optarem por cursos de Computação em universidades públicas. Poderia-se estudar melhor a variação deste parâmetro para entender o que acontece com o *risk ratio*.

Foram escolhidos apenas cursos superiores que têm a palavra “computação” no nome. Pode-se ter um trabalho mais detalhado e uma caracterização melhor do fenômeno adotando o catálogo completo do MEC para incluir todos os cursos superiores na área de tecnologia de informação.

A análise foi realizada calculando o *gap* de gênero por município. Essa divisão assume que dentro de um município há uma uniformidade do *gap* de gênero, o que pode não ser real, especialmente em cidades maiores. Uma análise mais detalhada poderia ser feita por bairros ou mesmo por escolas.

6. Conclusão

A escolha de carreira está ligada a diversas questões, como o mercado de trabalho, a disponibilidade de vagas e o desejo pessoal. O conjunto dessas motivações se manifesta em escolhas que passam pelo desempenho no Ensino Médio, mas que culminam na escolha de cursos específicos. Neste trabalho, observamos que o desempenho no Ensino Médio

tem apenas um pequeno papel neste processo, e as opções e percepções pessoais são mais relevantes para tal.

Também, verificamos que o *gap* de gênero entre estudantes com maior desempenho de matemática no ENEM é um fenômeno que varia muito em cada cidade ao longo dos anos. Apesar disso, também verificamos que há locais em que o *gap* de gênero é consistentemente menor. Isso aponta para potenciais fontes de inspiração para planos de ação visando a redução do *gap* de gênero em outros lugares.

Os dados expostos neste trabalho acrescentam um recorte regional onde possivelmente o *gap* de gênero de matemática no ENEM é menor, detalhe que pode apontar novos rumos de investigação para entender quais ações bem-sucedidas são realizadas nas cidades de Belo Horizonte, Juiz de Fora e Viçosa para replicar no ensino médio em outras regiões. Um outro desdobramento é entender se nestas regiões o potencial representado pela menor diferença de desempenho em matemática está se convertendo em matrículas no ensino superior de carreiras STEM. A sugestão deste trabalho é gerar derivados alinhados às abordagens de Medeiros (Medeiros et al. 2022), Ramos (Ramos and Araújo 2022) e Aires (Aires et al. 2018) com foco nas localidades de menor *gap* de gênero.

Este artigo tratou do *gap* de gênero, porém uma metodologia semelhante pode ser aplicada para o estudo de diferenças de raça ou de renda. Pode ainda ser relevante estudar como características socioeconômicas de cada cidade influenciam as desigualdades evidenciadas pelas notas do ENEM. Esperamos que as contribuições deste artigo possam inspirar pesquisas no sentido de entender e propor soluções futuras para atenuar as desigualdades no acesso ao ensino e ao trabalho.

Referências

- [Aires et al. 2018] Aires, J., Mattos, G., Oliveira, C., Brito, A., Aragão, A. F., Alves, S., Coelho, T., and Moreira, G. (2018). Barreiras que Impedem a Opção das Meninas pelas Ciências Exatas e Computação: Percepção de Alunas do Ensino Médio. In *Anais do Women in Information Technology (WIT)*. SBC. ISSN: 2763-8626.
- [Alvarenga and Braga 2024] Alvarenga, C. and Braga, C. (2024). Assessing the effects of gender stereotype in stem in a brazilian university. *Economia*.
- [Bandura 1977] Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review.*, 84(2):191–215.
- [Blotnicky et al. 2018] Blotnicky, K. A., Franz-Odenaal, T., French, F., and Joy, P. (2018). A study of the correlation between STEM career knowledge, mathematics self-efficacy, career interests, and career activities on the likelihood of pursuing a STEM career among middle school students. *International Journal of STEM Education*, 5(1):22.
- [Card and Payne 2020] Card, D. and Payne, A. A. (2020). High school choices and the gender gap in STEM . *Economic Inquiry*, 59(1):9–28.
- [Cimpian et al. 2020] Cimpian, J. R., Kim, T. H., and McDermott, Z. T. (2020). Understanding persistent gender gaps in stem. *Science*, 368(6497):1317–1319.
- [Cohen 1988] Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Routledge, New York, 2 edition.
- [da Justa Lemos 2019] da Justa Lemos, M. (2019). The effect of gender on college major choice: evidence from brazil. Master's thesis, Fundação Getulio Vargas, Rio de Janeiro, Brazil.

- [D'Agostino Sr. 2005] D'Agostino Sr., R. B. (2005). Principal Components Analysis. In *Encyclopedia of Biostatistics*. John Wiley & Sons, Ltd. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/0470011815.b2a13069>.
- [Holanda et al. 2022] Holanda, M., M. Júnior, A. L., Silva, E. H. M. d., Laterza, J., Araujo, A., Castanho, C., Koike, C., and Oliveira, R. B. (2022). Uma análise comparativa do desempenho em matemática entre gêneros nas provas do enem. In *Anais do XVI Women in Information Technology (WIT 2022)*, WIT 2022. Sociedade Brasileira de Computação - SBC.
- [Kanny et al. 2014] Kanny, M. A., Sax, L. J., and Riggers-Piehl, T. A. (2014). Investigating forty years of stem research: How explanations for the gender gap have evolved over time. *Journal of Women and Minorities in Science and Engineering*, 20(2):127–148.
- [Lima et al. 2022] Lima, W. G. d., Maciel, C., Casagrande, A. L., Sassi, S. B., and Nunes, M. F. A. F. (2022). STEAM, Gênero e Ensino Médio: ações da extensão em parceria com o Meninas Digitais Mato Grosso. In *Anais do Women in Information Technology (WIT)*, pages 251–256. SBC. ISSN: 2763-8626.
- [Makarova et al. 2019] Makarova, E., Aeschlimann, B., and Herzog, W. (2019). The gender gap in stem fields: The impact of the gender stereotype of math and science on secondary students' career aspirations. *Frontiers in Education*, 4.
- [Marcom and Aragão 2022] Marcom, G. S. and Aragão, T. Z. B. (2022). Gender inequality and science education: Comparison with brazilian students in pisa and enem. *Acta Scientiae*, 24(4):57–80.
- [Medeiros et al. 2022] Medeiros, A., Ferreira, I. B. M. C., Fonseca, L., and Rolim, C. (2022). Percepções sobre a tecnologia da informação por alunas de ensino médio: um estudo sobre gênero e escolhas profissionais. In *Anais do Women in Information Technology (WIT)*, pages 122–132. SBC. ISSN: 2763-8626.
- [Novaes et al. 2023] Novaes, T. S., Lins, K. D. S., Neto, A. G. S. S., Setti, M. d. O. G., and Emer, M. C. F. P. (2023). Despertando o Interesse de Mulheres para os Cursos em STEM. In *Anais do Women in Information Technology (WIT)*, pages 103–112. SBC. ISSN: 2763-8626.
- [Ramos and Araújo 2022] Ramos, A. I. M. and Araújo, F. O. (2022). Questões de Gênero e a Evasão de Mulheres nos Cursos de Computação: Um Estudo de Caso na Região Metropolitana de Belém. In *Anais do Women in Information Technology (WIT)*, pages 239–244. SBC. ISSN: 2763-8626.
- [Sakellariou and Fang 2021] Sakellariou, C. and Fang, Z. (2021). Self-efficacy and interest in stem subjects as predictors of the stem gender gap in the us: The role of unobserved heterogeneity. *International Journal of Educational Research*, 109:101821.
- [Santos et al. 2019] Santos, M. E. S. d., Rocha, T. S., Brasileiro, V. L. J., and Souza, C. C. d. (2019). What computing brazilian community is saying about gender gap. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, 14(4):162–167.
- [Zakariya 2022] Zakariya, Y. F. (2022). Improving students' mathematics self-efficacy: A systematic review of intervention studies. *Frontiers in Psychology*, 13. Publisher: Frontiers.