



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CAMPUS UNIVERSITÁRIO REITOR JOÃO DAVID FERREIRA LIMA
PROGRAMA DE PÓS-GRADUAÇÃO EM MÉTODOS E GESTÃO EM AVALIAÇÃO

Cláudia Jung Prado

**Fatores associados ao desempenho em Matemática dos estudantes do 9º ano do Ensino
Fundamental, baseado no Saeb 2017**

Florianópolis/SC
2020

Cláudia Jung Prado

Fatores associados ao desempenho em Matemática dos estudantes do 9º ano do Ensino Fundamental, baseado no Saeb 2017

Dissertação submetida ao Programa de Pós-Graduação em Métodos e Gestão em Avaliação da Universidade Federal de Santa Catarina para a obtenção do título de Mestre em Métodos e Gestão em Avaliação.

Orientador: Prof^a. Silvana Ligia Vincenzi, Dr^a.

Coorientador: Prof. Pedro Alberto Barbetta, Dr.

Florianópolis/SC

2020

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Prado, Cláudia Jung

Fatores associados ao desempenho em Matemática dos
estudantes do 9º ano do Ensino Fundamental, baseado no
Saeb 2017 / Cláudia Jung Prado ; orientadora, Silvana
Ligia Vincenzi, coorientador, Pedro Alberto Barbeta, 2020.
115 p.

Dissertação (mestrado profissional) - Universidade
Federal de Santa Catarina, Centro Tecnológico, Programa de
Pós-Graduação em Métodos e Gestão em Avaliação, Florianópolis,
2020.

Inclui referências.

1. Métodos e Gestão em Avaliação. 2. Avaliação
educacional em larga escala. 3. Fatores associados. 4.
Proficiência em Matemática. 5. Regressão multinível. I.
Vincenzi, Silvana Ligia. II. Barbeta, Pedro Alberto. III.
Universidade Federal de Santa Catarina. Programa de Pós
Graduação em Métodos e Gestão em Avaliação. IV. Título.

Cláudia Jung Prado

Fatores associados ao desempenho em Matemática dos estudantes do 9º ano do Ensino Fundamental, baseado no Saeb 2017

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof^ª. Silvana Ligia Vincenzi, Dra.
Universidade Federal de Santa Catarina

Prof. Marcelo Menezes Reis, Dr.
Universidade Federal de Santa Catarina

Prof^ª. Raquel Regina Zmorzenski Valduga Schöninger, Dra.
Prefeitura Municipal de Florianópolis - PMF

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de mestre em Métodos e Gestão em Avaliação.

Coordenação do Programa de Pós-Graduação

Prof^ª. Silvana Ligia Vincenzi, Dra.
Orientadora

Prof. Pedro Alberto Barbetta, Dr.
Coorientador

Florianópolis, 2020.

Este trabalho é dedicado à minha família, meus pais, irmãs, meu marido e meus filhos.

AGRADECIMENTOS

Agradeço aos meus pais, Paulo e Edit, pelo exemplo que me deram, de pessoas íntegras e comprometidas, pelos ensinamentos ao longo da vida e pelo incentivo e apoio em todos os momentos.

Aos meus sogros, Antonio e Conceição, pelo apoio.

Ao meu esposo, André, pela parceria, paciência e bom humor nestes anos de companheirismo.

Aos meus filhos, Túlio e Diana, pela alegria de viver.

Aos colegas, amigos de curso, que me acolheram e me auxiliaram nesta árdua jornada de ampliação dos conhecimentos.

Aos professores do PPGMGA, Marcelo, Dilvo, Sílvia, Lizandra, Andreia, Silvana, Adriano, Dalton e Pedro, pelas aulas, nas quais nos apresentaram um “mundo novo”, com desafios e novas possibilidades.

À minha orientadora, professora Silvana, pela disponibilidade e parceria no desenvolvimento da dissertação.

Ao meu coorientador, professor Pedro, por ter aceitado o convite e por todas as contribuições ao longo dos nossos encontros, que foram determinantes para o desenvolvimento deste trabalho.

À professora Lizandra da Silva Menegon, pelo apoio, compreensão e palavras de incentivo no período do curso em que vivenciei o luto por uma perda pessoal.

Aos professores do Programa de Pós-Graduação em Educação, Jéferson Silveira Dantas e Ione Ribeiro Valle, pela acolhida nas disciplinas cursadas no Mestrado em Educação e por terem proporcionado a ampliação das lentes interpretativas a serem utilizadas tanto no campo acadêmico quanto no campo profissional.

Aos meus colegas da Escola Básica Municipal Professora Herondina Medeiros Zeferino, que me apoiaram e incentivaram ao longo do curso. Aos meus colegas do grupo de estudos da escola Herondina, que, nos anos de 2016 e 2017, através das discussões e aprofundamentos teóricos, contribuíram para a minha formação e no desenvolvimento do anseio por retomar os estudos acadêmicos e, à Prefeitura Municipal de Florianópolis, Secretaria Municipal de Educação, pelo incentivo ao aperfeiçoamento.

Enfim, a todos aqueles que, de alguma forma, contribuíram para o desenvolvimento desta dissertação, o meu muito obrigada!

Desta maneira, o educador já não é o que apenas educa, mas o que, enquanto educa, é educado, em diálogo com o educando que, ao ser educado, também educa. Ambos, assim, se tornam sujeitos do processo em que cresceram juntos. (FREIRE, 1987, p.44).

RESUMO

A dissertação tem como objetivo analisar os fatores associados à proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental, com base nos dados do Saeb 2017. Para a realização da análise, foram utilizadas variáveis extraídas do questionário contextual e também foram construídos, por meio da Teoria da Resposta ao Item, indicadores para mensurar o nível socioeconômico dos estudantes e as condições de infraestrutura das escolas. O modelo de regressão multinível foi a técnica estatística adotada para avaliação dos efeitos dos fatores associados à proficiência, porque considera a estrutura hierárquica dos dados, ou seja, possibilita levar em consideração o fato dos estudantes, que possuem características individuais e relativas ao seu contexto sociofamiliar, estarem agrupados em escolas, que também possuem características relativas a forma de organização didático-pedagógica e recursos disponíveis. Os resultados apontaram que medidas como a melhoria das condições de infraestrutura escolar, a diminuição das taxas de reprovação escolar e ações do professor como encaminhamento de tarefas para realização em casa e correção em sala destas tarefas, impactam positivamente no desempenho dos estudantes. Para a exemplificação de como a pesquisa desenvolvida pode contribuir para a melhoria das condições de ensino aprendizagem nas escolas, também foi realizada uma simulação, com base nos resultados de uma escola do município de Florianópolis/SC, de como mudanças em fatores associados às atividades de gestão podem impactar no desempenho dos estudantes.

Palavras-chave: Proficiência em Matemática. Fatores associados. Regressão multinível.

ABSTRACT

The dissertation had the objective to analyze the associated factors to Mathematics proficiency of the 9th year Elementary School students based on the data from Saeb 2017. For carrying out the analysis, were used variables extracted from the contextual questionnaire and also were built, through the Item Response Theory, indicators to measure the socioeconomics level of the students and the infrastructure conditions of the schools. The multilevel regression model was the statistics technique adopted for evaluation of the effects of associated factors to proficiency, because it considers the hierarchical structure of the data, that is, enables take in consideration the fact of the students, who has individual characteristics and relative to its sociofamiliar context, be grouped in schools, which also possesses the related characteristics to form of organization and available resources. The results pointed that measures such as improving the school infrastructure condition, the decrease of the scholar reproval rate and teacher's actions like forwarding tasks to be performed at home and correction of that tasks at class positively impacts on students performance. For the exemplification on how the developed research can contribute for the improval of the teaching-learning conditions at schools, a simulation was also performed, based in results from a school in the municipality of Florianópolis/SC, on how the changes in factors associated to management activities can impact on students performance.

Key-words: Mathematics Proficiency. Associated Factors. Multilevel Regression.

LISTA DE FIGURAS

Figura 1 – Descrição das etapas da pesquisa.....	44
Figura 2 - Variáveis de nível 1 (estudante) e nível 2 (escola) com impacto na proficiência dos estudantes nas avaliações educacionais em larga escala selecionados na revisão integrativa da literatura.....	45
Figura 3 - Alfa de Cronbach da escala de infraestrutura escolar via TCT.....	50
Figura 4 - Escore de infraestrutura das escolas.....	52
Figura 5 - Curva de informação do teste da escala de infraestrutura escolar.....	55
Figura 6 - Régua com os itens posicionados na escala de infraestrutura escolar.....	57
Figura 7 - Régua com os itens posicionados na escala de nível socioeconômico.....	63
Figura 8 - Histograma da variável dependente proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental no Saeb 2017.....	67
Figura 9 - Gráfico qqplot da variável dependente proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental no Saeb 2017.....	68
Figura 10 - Gráfico boxplot/diagrama em caixas da variável dependente proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental no Saeb 2017 agrupada pela média por escola da medida de nível socioeconômico dos estudantes.....	76
Figura 11 - Gráfico boxplot/diagrama em caixas da variável dependente proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental no Saeb 2017 agrupada pela medida de infraestrutura das escolas.....	76

LISTA DE QUADROS

Quadro 1 – Artigos identificados nas bases de dados Scielo e Scopus sobre fatores associados ao desempenho escolar/proficiência através de regressão multinível/regressão hierárquica ...	26
Quadro 2 - Trabalhos identificados na base de dados Portal de Periódicos da CAPES.....	32
Quadro 3 - Trabalhos identificados na base de dados BDTD.....	33
Quadro 4 - Variáveis utilizadas para montagem do banco de dados da escala de infraestrutura escolar.....	48
Quadro 5 - Variáveis utilizadas para montagem do banco de dados da escala de nível socioeconômico.....	59
Quadro 6 - Descrição das variáveis relacionadas aos estudantes utilizadas na regressão multinível.....	69
Quadro 7 - Descrição das variáveis relacionadas às escolas utilizadas na regressão multinível.	73
Quadro 8 – Artigos identificados bases de dados Scielo e Scopus.....	104
Quadro 9 – Trabalhos identificados na base de dados portal de periódicos da CAPES.....	105
Quadro 10 – Trabalhos identificados na base de dados BDTD.....	106

LISTA DE TABELAS

Tabela 1 – Frequência dos itens da escala de infraestrutura escolar	50
Tabela 2 - Descrição dos itens e valores estimados na escala de infraestrutura escolar.....	52
Tabela 3 - Descrição dos itens e valores estimados na escala de infraestrutura escolar, após exclusões.....	54
Tabela 4 - Probabilidade de ter o atributo em função do nível da escala de infraestrutura escolar.....	56
Tabela 5 - Descrição dos níveis da escala de infraestrutura escolar e a porcentagem de escolas por intervalo de proficiência.....	57
Tabela 6 - Itens posicionados com intervalo de 0,5 pontos na escala de nível socioeconômico.	62
Tabela 7 - Posicionamento e descrição dos itens na escala de nível socioeconômico.....	63
Tabela 8 - Descrição dos níveis socioeconômicos dos estudantes por intervalo de proficiência.....	64
Tabela 9 - Estatísticas descritivas das características/variáveis dos estudantes do 9º ano do Ensino Fundamental - Saeb 2017.....	70
Tabela 10 - Estatísticas descritivas das características/variáveis dos estudantes do 9º ano do Ensino Fundamental - Saeb 2017, após exclusões.....	71
Tabela 11 - Estatísticas descritivas das características/variáveis das escolas de Ensino Fundamental - Saeb 2017.....	74
Tabela 12 - Resumo dos efeitos aleatórios dos modelos de regressão multinível e redução percentual em relação ao modelo nulo.....	85
Tabela 13 - Valor predito/esperado da proficiência em Matemática do estudante x da escola X e do estudante y da escola Y.....	90
Tabela 14 - Efeito dos fatores associados na proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental, base no Saeb 2017.....	94

LISTA DE ABREVIATURAS E SIGLAS

ANA Avaliação Nacional da Alfabetização
ANEB Avaliação Nacional da Educação Básica
ANRESC Avaliação Nacional do Rendimento Escolar
BDE Bônus de Desempenho Educacional
BDTD Biblioteca Digital Brasileira de Teses e Dissertações
BU Biblioteca Universitária
CNTE Confederação Nacional dos Trabalhadores em Educação
CAPES Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
Game-UFMG Grupo de Avaliação e Medidas Educacionais da Universidade Federal de Minas Gerais
IDEB Índice de Desenvolvimento da Educação Básica
IDH Índice de Desenvolvimento Humano
IEE Instituto Estadual de Educação
INEP Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
Inse Indicador de Nível Socioeconômico das Escolas de Educação Básica
LDBEN Lei de Diretrizes e Bases da Educação Nacional
MEC Ministério da Educação
MLH Modelo Linear Hierárquico
ML2 Modelo Logístico Unidimensional de dois Parâmetros
MV Máxima Verossimilhança Marginal
NA Não Apresentada/Não Respondeu
PISA Programa Internacional de Avaliação de Estudantes
PNE Plano Nacional de Educação
PROEB Programa de Avaliação da Educação Básica
REML Método de Máxima Verossimilhança Restrita
Saeb Sistema de Avaliação da Educação Básica
SAEGO Sistema de Avaliação Educacional do Estado de Goiás
SAEPE Sistema de Avaliação Educacional do Pernambuco
SAERS Sistema de Avaliação do Rendimento Escolar do Rio Grande do Sul
Saresp Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo
SC Santa Catarina

SPAECE Sistema Permanente de Avaliação da Educação Básica do Ceará

SIM-Escolas Sistema de Indicadores para o Monitoramento das Escolas Estaduais Paulistas

SIMAVE Sistema Mineiro de Avaliação da Educação Pública

SisPAE Sistema Paraense de Avaliação Educacional

TCT Teoria Clássica dos Testes

TRI Teoria da Resposta ao Item

UFSC Universidade Federal de Santa Catarina

SUMÁRIO

1	INTRODUÇÃO	15
1.1	PROBLEMA DE PESQUISA	17
1.2	JUSTIFICATIVA	19
1.3	OBJETIVOS	22
1.3.1	Objetivo Geral.....	22
1.3.2	Objetivos Específicos	22
1.4	ESTRUTURA DO TRABALHO	23
2	REFERENCIAL TEÓRICO	24
2.1	AVALIAÇÃO EDUCACIONAL EM LARGA ESCALA NA EDUCAÇÃO BÁSICA.....	24
2.2	FATORES ASSOCIADOS AO DESEMPENHO ESCOLAR	25
2.3	A TEORIA DA RESPOSTA AO ITEM	28
2.4	REGRESSÃO MULTINÍVEL	29
2.5	APLICAÇÕES DO MODELO DE REGRESSÃO MULTINÍVEL PARA ANÁLISE DE FATORES ASSOCIADOS À PROFICIÊNCIA	31
3	METODOLOGIA.....	43
3.1	DESIGN DE PESQUISA	43
3.2	ANÁLISE DE DADOS	44
4	CONSTRUÇÃO DE INDICADORES.....	47
4.1	ESCALA DE INFRAESTRUTURA ESCOLAR.....	47
4.1.1	Análise via Teoria Clássica dos Testes.....	49
4.1.2	Análise via Teoria da Resposta ao Item.....	52
4.2	ESCALA DE NÍVEL SOCIOECONÔMICO	58
5	AVALIAÇÃO DO EFEITO DOS FATORES ASSOCIADOS POR REGRESSÃO MULTINÍVEL	66
5.1	DADOS	66

5.2	ESTIMAÇÃO DOS EFEITOS DOS FATORES ASSOCIADOS	76
5.3	MODELO	77
5.3.1	Modelo nulo.....	77
5.3.2	Modelo 1 – O efeito do nível socioeconômico familiar no desempenho do estudante.....	79
5.3.3	Modelo 2 – O efeito da localização da escola no desempenho do estudante... 80	
5.3.4	Modelo 3 – inclusão de todas as variáveis de nível 1 consideradas no estudo.....	81
5.3.5	Modelo 4 – inclusão das variáveis de nível 1 (estudante) e de nível 2 (escola) consideradas no estudo.....	83
5.4	DISCUSSÃO E CONTEXTUALIZAÇÃO DOS RESULTADOS.....	85
6	ANÁLISE COM OS DADOS DE UMA ESCOLA DO MUNICÍPIO DE FLORIANÓPOLIS/SC	89
7	CONSIDERAÇÕES	94
7.1	SUGESTÕES PARA TRABALHOS FUTUROS.....	98
	REFERÊNCIAS.....	99
	APÊNDICE A – Referencial Teórico.....	104
	APÊNDICE B – Trabalhos identificados no portal de periódicos CAPES..	105
	APÊNDICE C – Trabalhos identificados no portal da BDTD	106
	APÊNDICE D – Descrição da base de dados utilizada na aplicação dos modelos de regressão	107
	APÊNDICE E – Descrição dos passos para a simulação com os dados de uma escola do município de Florianópolis/SC	108
	ANEXO A – Protocolo de pesquisa BU/UFSC.....	112

1 INTRODUÇÃO

As avaliações educacionais em larga escala permitem a aferição do desempenho dos estudantes e subsidiam a elaboração e a reformulação de políticas públicas com vistas à melhoria da qualidade da educação. Em âmbito nacional estas avaliações da aprendizagem são coordenadas pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), cuja missão é promover estudos, pesquisas e avaliações sobre o sistema educacional brasileiro, com o objetivo de subsidiar a formulação e a implementação de políticas públicas para a área educacional.

A avaliação da educação brasileira está prevista em documentos legais como a Lei de Diretrizes e Bases da Educação Nacional (LDBEN), Lei nº 9.394/1996, que, em seu artigo 9º, inciso VI, estabelece que a União será responsável por “assegurar processo nacional de avaliação do rendimento escolar no ensino fundamental, médio e superior, em colaboração com os sistemas de ensino, objetivando a definição de prioridades e a melhoria da qualidade do ensino”. O Plano Nacional de Educação (PNE), também, desde a sua primeira edição, em 1996, prevê a criação de um Sistema Nacional de Avaliação. De acordo com a Constituição Federal, artigo 214, o PNE deve ser elaborado com periodicidade decenal e deve estabelecer metas e estratégias para o desenvolvimento do ensino. O PNE em vigência, Lei nº 13.005/2014, reafirma, em seu artigo 11, que:

O Sistema Nacional de Avaliação da Educação Básica, coordenado pela União, em colaboração com os Estados, o Distrito Federal e os Municípios, constituirá fonte de informação para a avaliação da qualidade da educação básica e para a orientação das políticas públicas desse nível de ensino (BRASIL, 2014, p.02).

Dentre as avaliações realizadas pelo INEP, o Sistema de Avaliação da Educação Básica (Saeb), que é composto por um conjunto de avaliações externas em larga escala, é utilizado para o cálculo do Índice de Desenvolvimento da Educação Básica (IDEB), que também considera os dados sobre a aprovação escolar, extraídos do Censo Escolar, para a obtenção de seus resultados. De acordo com informações do portal do Ministério da Educação (MEC), o IDEB mede a qualidade do aprendizado nacional e permite o estabelecimento de metas para a melhoria do ensino, ou seja, este índice “funciona como um indicador nacional que possibilita o monitoramento da qualidade da educação” (MEC, 2018).

Desde a sua criação em 1990 o Saeb passou por algumas reestruturações com relação as etapas avaliadas/público alvo (1^a, 3^a, 5^a e 7^a séries do Ensino Fundamental de 8 anos¹; 4^a e 8^a séries do Ensino Fundamental de 8 anos e 3^a série do Ensino Médio²; 5^o e 9^o anos do Ensino Fundamental de 9 anos e 3^a série do Ensino Médio³), abrangência (escolas públicas, privadas de forma censitária ou amostral) e áreas do conhecimento avaliadas (Língua Portuguesa, Matemática, Ciências Naturais, Física, Química, Biologia, História, Geografia⁴). Sendo que, a partir de 2019, as siglas ANA (Avaliação Nacional da Alfabetização), ANEB (Avaliação Nacional da Educação Básica) e ANRESC (Avaliação Nacional do Rendimento Escolar/Prova Brasil) passaram a ser identificadas apenas pela sigla Saeb, acompanhada da etapa, área de conhecimento e tipo de instrumento envolvido. É importante ressaltar que, mesmo com as alterações implementadas, até 2017, a sistemática avaliativa manteve as condições necessárias para a comparabilidade dos resultados ao longo do tempo.

Na edição de 2019 do Saeb outra novidade foi a participação amostral de turmas de escolas de Educação Infantil e turmas do 2^o ano do Ensino Fundamental, bem como a aplicação amostral de provas das áreas do conhecimento de Ciências Humanas e Ciências da Natureza com algumas turmas do 9^o ano do Ensino Fundamental⁵.

O levantamento de dados realizado pelo INEP com a aplicação dos testes cognitivos e com os questionários contextuais é disponibilizado no seu portal na forma de microdados e permite estudos e pesquisas sobre o rendimento educacional, as condições das escolas e das redes de ensino.

Além do IDEB, o INEP também disponibiliza outros indicadores educacionais, que são medidas construídas a partir de diferentes dados para atribuir um valor estatístico à qualidade da educação, considerando, além do desempenho dos estudantes, também o contexto econômico e social em que as escolas estão inseridas (INEP, 2018a). Desta forma, ao considerar o acesso, a permanência e a aprendizagem dos estudantes, os indicadores educacionais como a Complexidade de Gestão da Escola, Esforço Docente, Nível Socioeconômico, Média de Alunos por Turma, Média de Horas-Aula Diária, Taxas de Distorção Idade-Série, Taxas de Rendimento, Percentual de Docentes com Curso Superior, Adequação da Formação Docente e

¹ Edições de 1990, 1993, 1995.

² Edições de 1997, 1999, 2001, 2003, 2005, 2009, 2011.

³ Edições de 2013, 2015, 2017.

⁴ As áreas do conhecimento Língua Portuguesa e Matemáticas são as únicas que estiveram presentes em todas as edições da avaliação.

⁵ Fonte: Cartilha Saeb 2019, disponível no portal do INEP / Educação Básica / Saeb

(http://portal.inep.gov.br/informacao-da-publicacao/-/asset_publisher/6JYIsGMAMkWI/document/id/6734620)

a Regularidade do Corpo Docente, podem ser utilizados para o monitoramento dos sistemas educacionais e para subsidiar a elaboração de políticas públicas mais específicas, voltadas para a melhoria da educação e dos serviços oferecidos à sociedade pela escola.

1.1 PROBLEMA DE PESQUISA

A Prova Brasil caracteriza-se por ser um instrumento de avaliação composto por testes cognitivos nas áreas de Língua Portuguesa e Matemática e também por questionários contextuais⁶. É aplicada aos estudantes do 5º ano, ou seja, ao término do ciclo dos Anos Iniciais do Ensino Fundamental, e aos estudantes do 9º ano, ao término do ciclo dos Anos Finais do Ensino Fundamental, das escolas públicas com pelo menos 20 estudantes matriculados nas etapas avaliadas. Os questionários contextuais da Prova Brasil são dirigidos aos estudantes participantes da avaliação, aos professores das áreas avaliadas e aos diretores das escolas participantes, além disso, o aplicador externo designado pelo INEP também preenche um questionário sobre a infraestrutura, segurança e condições dos recursos pedagógicos disponíveis nas escolas avaliadas. Assim, os questionários contextuais servem como instrumentos de coleta de informações sobre aspectos que podem estar relacionados à proficiência, como informações sobre os fatores do contexto socioeconômico, cultural e escolar que podem influenciar o desempenho dos estudantes nos testes cognitivos.

Estudos sobre fatores associados ao desempenho dos estudantes em avaliações educacionais em larga escala vem sendo desenvolvidos buscando compreender quais características relacionadas aos estudantes e/ou às escolas ou aos sistemas de ensino influenciam a proficiência dos estudantes. Andrade e Laros (2007) construíram um modelo de desempenho escolar com análise multinível utilizando o banco de dados do Saeb 2001, referentes aos estudantes do 3º ano do Ensino Médio, nas disciplinas de Língua Portuguesa e Matemática, e constataram que a variável atraso escolar/reprovação é a que mais afeta negativamente o desempenho escolar dos estudantes, tanto no nível 1 (variáveis associadas aos estudantes) quanto no nível 2 (variáveis associadas às escolas). Soares *et al.* (2010), construíram modelos de regressão multinível utilizando informações dos questionários contextuais e dos testes cognitivos de uma avaliação educacional em larga escala de Minas

⁶ A Prova Brasil, também denominada Avaliação Nacional do Rendimento Escolar (ANRESC) entre os anos de 2005 e 2017, a partir de 2019, passou a ser denominada de Saeb/etapa/área de conhecimento/tipo de instrumento, por isso, ao longo da dissertação, os termos “Prova Brasil” e “Saeb” serão utilizados como sinônimos.

Gerais do ano de 2006, com estudantes da 4ª série do Ensino Fundamental, para estudar fatores associados ao desempenho dos estudantes nas disciplinas de Língua Portuguesa e Matemática. Palermo, Silva e Novellino (2014), utilizaram dados do Saeb 2007 para compreender fatores associados ao desempenho escolar em Matemática dos estudantes do 5º ano do Ensino Fundamental da rede municipal do Rio de Janeiro através de modelos hierárquicos multiníveis. Mas, não foram identificados estudos baseados nos resultados da edição de 2017 da Prova Brasil⁷, usando a proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental como variável dependente e utilizando, no mesmo modelo, as variáveis explicativas relacionadas ao estudante (nível 1): nível socioeconômico, sexo, cor, reprovação, trabalho fora de casa, incentivo dos pais aos estudos, o estudante declarar que gosta da disciplina e ter feito educação infantil e, como variáveis explicativas relacionadas à escola (nível 2): a localização, dependência administrativa, área, turno, proporção de estudantes por escola que declaram fazer as tarefas de casa da disciplina, a proporção de estudantes por escola que declaram que o professor corrige as tarefas de casa da disciplina, a proporção por escola de estudantes com defasagem escolar decorrente de reprovação, a média por escola do nível socioeconômico dos estudantes e a medida das condições de infraestrutura escolar.

Neste contexto, esta dissertação procura responder a seguinte questão: *quais são os fatores que estão mais consistentemente associados ao desempenho escolar em Matemática dos estudantes do 9º ano do Ensino Fundamental, de escolas brasileiras, baseado nos dados do Saeb 2017.*

Assim, espera-se que a análise realizada nesta pesquisa com os resultados do ano de 2017 do Saeb, por meio de modelos de regressão multinível, possa contribuir para a continuidade da reflexão acerca da influência de variáveis intra e extraescolares no desempenho cognitivo dos estudantes, na perspectiva de entender/compreender os fatores associados à proficiência que dependem de políticas públicas mais amplas e a longo prazo para que se chegue a um patamar de equidade ou justiça social na área da educação, mas, principalmente, no sentido de subsidiar o debate acerca do que é passível de intervenção/de mudança/de reestruturação por parte das escolas e/ou das redes de ensino para a melhoria da qualidade da educação.

⁷ Busca realizada no Portal de Periódicos da Capes e na Biblioteca Digital Brasileira de Teses e Dissertações com os descritores “Saeb 2017” ou “Prova Brasil 2017”.

1.2 JUSTIFICATIVA

Uma das formas de medição do sucesso escolar está dada nas avaliações educacionais em larga escala, que medem a proficiência, ou seja, aferem o desempenho dos estudantes por meio de testes cognitivos ao término de um ciclo de aprendizagem. Mas, para a compreensão dos fatores que estão associados a este desempenho é preciso que sejam consideradas as variáveis intra e extraescolares e, para analisar estatisticamente estas relações, neste trabalho, utilizou-se o modelo de regressão multinível, porque permite que seja levada em consideração a estrutura hierárquica de agrupamento dos dados: os estudantes (nível 1) agrupados em escolas (nível 2).

Ao buscar identificar os fatores associados ao desempenho em Matemática dos estudantes do 9º ano do Ensino Fundamental, baseado nos dados do Saeb 2017, utilizando o modelo de regressão multinível, o projeto de pesquisa previa que se trabalhasse apenas com os dados das escolas das redes públicas do município de Florianópolis/SC, mas, na organização dos dados para aplicação dos modelos de regressão, constatou-se uma significativa diminuição da base de dados, porque muitas escolas não tiveram os seus dados divulgados, por conta da Portaria nº 447/2017, que estabelece o critério de participação no dia da prova de, pelo menos, 80% dos estudantes matriculados. Desta forma, a base de dados, inicialmente composta por 4.016 estudantes, passou a ter apenas 886 estudantes com a informação do ID_ESCOLA válido, que é o campo de identificação do código de escola identificável que permitiria fazer a junção com informações de outras planilhas ou base de dados. As demais escolas tiveram a variável ID_ESCOLA substituída por uma máscara, isto é, por códigos fictícios iniciados pelo número 6. Desta forma, a pesquisa foi reestruturada e trabalhou-se com os dados de todas as escolas do Brasil e, no capítulo 6, deu-se ênfase aos resultados pertinentes à uma escola da rede pública municipal de Florianópolis/SC, como recorte para aprofundamento, por ser o contexto no qual a autora está inserida profissionalmente e por tratar-se de um estudo que poderá contribuir para reflexões acerca dessa realidade.

Por trabalhar como Administradora Escolar em uma escola básica da rede municipal de educação de Florianópolis/SC desde 2015, almeja-se que o aprofundamento teórico-metodológico proporcionado pelo mestrado profissional do Programa de Pós-Graduação em Métodos e Gestão em Avaliação contribua para a discussão acerca de possíveis investimentos e/ou ações para a melhoria da educação nessa rede de ensino.

A rede municipal de educação de Florianópolis/SC é composta por escolas básicas, que atendem estudantes dos Anos Iniciais (1º ao 5º ano) e Anos Finais (6º ao 9º ano) do Ensino Fundamental; por Núcleos de Educação Infantil, que atendem crianças de 0 a 5 anos de idade; e, por Núcleos de Educação de Jovens e Adultos, que atendem estudantes do 1º e 2º segmentos (Anos Iniciais e Anos Finais do Ensino Fundamental, respectivamente).

A atual *Proposta Curricular da Rede Ensino de Florianópolis* (2016) aponta que o currículo precisa contemplar o que está previsto nos documentos oficiais, mas, também, levar “em consideração os contextos histórico, político, econômico e de organização da sociedade em que o/a estudante está inserido/a [...] prevendo a possibilidade de ajustes às opções de organizações adotadas em cada unidade educativa e em cada sala de aula” (FLORIANÓPOLIS, 2016, p.22-23). Com relação a função social da escola, no contexto da rede municipal de educação de Florianópolis/SC,

A escola deve ser um local que oportuniza o desenvolvimento intelectual e social dos/das estudantes, oferecendo espaço para a dúvida e para a problematização do que é objeto de conhecimento. Dessa forma, garantir acesso ao conhecimento, no mencionado tensionamento dialético, não significa negar a dúvida como mola propulsora para se conhecer. Garantir o acesso ao conhecimento consiste em negar uma tradição conteudista excludente, daí a importante articulação entre o local e o global, entre o cotidiano e o grande tempo/grande espaço, sempre em uma perspectiva problematizadora, crítica, que foque na formação humana integral e que tenha a emancipação como meta, o que implica a inclusão social e a acolhida de vivências de grupos socialmente estigmatizados como constitutivas das escolhas curriculares na integração com o que já está historicizado como objeto de conhecimento escolar. (FLORIANÓPOLIS, 2016, p.23).

De acordo com o Plano Municipal de Educação de Florianópolis/SC, Lei Complementar nº 546, de janeiro de 2016, os dados censitários tem evidenciado um aumento populacional no município, com “um número significativo de imigrantes vindos de outras cidades catarinenses, de outros Estados brasileiros e também de outros países” e, conseqüentemente, um aumento da demanda por vagas nas escolas. Este aumento na demanda por matrículas na rede pública é uma realidade principalmente das escolas localizadas no norte da ilha de Florianópolis/SC. Como alternativa de curto prazo, o poder público municipal, em alguns momentos, tem adotado uma medida prevista nas Resoluções nº 01 e 02/2015 do Conselho Municipal de Educação de Florianópolis/SC, que é o acréscimo de estudantes por turma. Por exemplo, as turmas do 1º ano do Ensino Fundamental, que, a priori, deveriam ter 25 estudantes, passam a atender, até 30 estudantes, e assim sucessivamente, podendo chegar a 40 estudantes por turma nos Anos Finais do Ensino Fundamental. Empiricamente o que observa-

se no cotidiano escolar é que este aumento no número de estudantes por turma, principalmente nos Anos Iniciais, onde se dá o processo de alfabetização, repercute em prejuízo para o processo de ensino aprendizagem, uma vez que, com mais estudantes por turma, o tempo para uma atenção individualizada às necessidades específicas de cada estudante, fica menor.

Além disso, com a política de progressão continuada na rede municipal de Florianópolis/SC, regulamentada pela Resolução nº 02/2011 do Conselho Municipal de Educação, cada vez mais as turmas são compostas por estudantes em diferentes etapas/níveis de desenvolvimento cognitivo/aprendizagem em uma mesma sala de aula, uma vez que não há mais a reprovação por aproveitamento escolar e sim o acompanhamento pedagógico diferenciado no ano subsequente. Assim, há uma heterogeneidade no perfil dos estudantes e na composição das turmas, e, com o aumento do número de estudantes por sala, o desafio é ainda maior para o desenvolvimento do processo de ensino aprendizagem de acordo com as especificidades de cada estudante. Do ponto de vista das relações múltiplas e das trocas para o desenvolvimento do processo de ensino aprendizagem, há uma ampliação das possibilidades pedagógicas nesta perspectiva da heterogeneidade do perfil dos estudantes. E, com o acompanhamento pedagógico no contraturno, que é oferecido aos estudantes que não atingiram o aproveitamento escolar esperado para o ano/série, também se ampliam as possibilidades, os tempos e espaços para a consolidação da aprendizagem dos conteúdos/conceitos necessários para a continuidade da trajetória escolar do estudante.

Pesquisa realizada por Machado (2014), sobre análise de fatores associados ao desempenho dos estudantes no Sistema de Avaliação Educacional do Estado de Goiás (SAEGO), evidenciou carência de levantamento qualitativo no que diz respeito aos fatores associados à proficiência dos estudantes no Brasil e, para que se possa ser feita a análise qualitativa, inicialmente, é o levantamento quantitativo que pode fornecer subsídios/dados para essa análise. Andrade e Laros (2007), no estudo realizado sobre fatores associados ao desempenho escolar com dados do Saeb, utilizando o modelo de regressão multinível, concluíram que a construção de um modelo de desempenho escolar favorece a compreensão dos fatores associados ao desempenho dos estudantes e que as análises estatísticas podem contribuir para a melhoria da qualidade da educação. Laros, Marciano e Andrade (2012), em estudo sobre o desempenho em Língua Portuguesa de estudantes do Ensino Médio no Saeb 2001, concluíram que a ampliação do conhecimento sobre fatores associados ao desempenho escolar pode contribuir para a melhoria da qualidade do ensino.

Assim sendo, é de relevância social a realização e a continuidade de estudos para aprofundamento da discussão e análise dos fatores associados ao desempenho dos estudantes nas avaliações educacionais em larga escala, para subsidiar a proposição de investimentos e/ou ações que permitam a melhoria das condições físicas e estruturais nas escolas, das condições de trabalho dos professores e das condições de aprendizagem para os estudantes, o que repercutirá em melhoria na qualidade da educação. Sendo o foco desta pesquisa a realização da análise com os microdados das escolas brasileiras e, posteriormente, no capítulo 5, a análise de uma escola da rede pública municipal de Florianópolis/SC⁸.

Optou-se, neste trabalho, pelo recorte analítico da variável dependente proficiência dos estudantes em Matemática, pela questão de tempo para a realização da pesquisa e aprofundamento da discussão acerca dos resultados obtidos e, também, fundamentando-se nos apontamentos realizados por Laros, Marciano e Andrade (2012) e por Palermo, Silva e Novellino (2014) de que esta área do conhecimento (Matemática) seria mais afetada por fatores pertinentes à qualidade das aulas e características das escolas, sendo a proficiência em Língua Portuguesa mais afetada por fatores do contexto sócio familiar do estudante.

No que se refere a opção pela análise dos dados dos estudantes do 9º ano do Ensino Fundamental, justifica-se a escolha por ser este o ano de encerramento do ciclo de aprendizagem que compreende os Anos Iniciais (1º ao 5º ano) e os Anos Finais (6º ao 9º ano).

1.3 OBJETIVOS

1.3.1 Objetivo Geral

Analisar os fatores associados ao desempenho em Matemática dos estudantes do 9º ano do Ensino Fundamental de escolas brasileiras no Saeb 2017.

1.3.2 Objetivos Específicos

- Fazer uma revisão sistemática da literatura sobre avaliação educacional e os fatores associados ao desempenho dos estudantes;

⁸ Em busca avançada nas bases de dados Scielo e Scopus, não foram localizados estudos de “fatores associados” ao “desempenho escolar” especificamente no município de Florianópolis/SC.

- Construir escalas/indicadores de Nível Socioeconômico e Condições de Infraestrutura Escolar, utilizando a Teoria da Resposta ao Item;
- Aplicar o modelo de regressão multinível para análise dos fatores associados ao desempenho dos estudantes do 9º ano do Ensino Fundamental, de escolas brasileiras, na disciplina de Matemática, no Saeb 2017;
- Identificar os possíveis fatores associados ao desempenho/proficiência dos estudantes do 9º ano do Ensino Fundamental, na disciplina de Matemática, no Saeb 2017, das escolas brasileiras;
- Realizar uma simulação de como mudanças em fatores associados às atividades de gestão podem incrementar o desempenho de estudantes em uma escola de Florianópolis/SC.

1.4 ESTRUTURA DO TRABALHO

Esta dissertação está organizada em uma primeira parte que consiste da introdução, onde são apresentados o problema de pesquisa, a justificativa para o desenvolvimento da pesquisa, os objetivos, geral e específicos, bem como a organização estrutural da pesquisa. O segundo capítulo consiste de uma revisão sistemática da literatura para composição do referencial teórico que contextualiza a pesquisa. No referencial teórico também são apresentados trabalhos nos quais o modelo de regressão foi utilizado para a análise de fatores associados à proficiência em avaliações educacionais em larga escala. Em seguida, no terceiro capítulo, sobre a metodologia, são apresentadas a metodologia da pesquisa e do modelo de análise. Depois, no capítulo 4, é descrito como se deu a organização dos dados da Prova Brasil 2017 para a aplicação da Teoria da Resposta o Item e construção das escalas/indicadores de Nível Socioeconômico e das Condições de Infraestrutura Escolar. Em seguida, no capítulo 5, são ajustados e apresentados os modelos de regressão multinível, utilizando as variáveis selecionadas e os indicadores construídos para posterior análise dos resultados e verificação do impacto dos fatores considerados no estudo na proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental das escolas brasileiras e, ainda, no capítulo 6, é feita uma análise do impacto que alterações relacionadas a algumas características/variáveis teriam no desempenho dos estudantes, utilizando, como exemplo, uma escola do município de Florianópolis/SC. Na conclusão, são apresentadas as considerações finais e sugestões para estudos futuros.

2 REFERENCIAL TEÓRICO

Este capítulo de referencial teórico apresenta os conceitos que subsidiaram o desenvolvimento da pesquisa e está organizado nas seguintes seções: a avaliação educacional em larga escala na Educação Básica; os fatores associados ao desempenho escolar; a teoria da resposta ao item e o modelo de regressão multinível.

2.1 A AVALIAÇÃO EDUCACIONAL EM LARGA ESCALA NA EDUCAÇÃO BÁSICA

A avaliação educacional em larga escala, avaliação padronizada ou avaliação externa pode ser definida como um instrumento de gestão para subsidiar a tomada de decisões e a definição de políticas públicas, a partir da aferição do desempenho dos estudantes, utilizado para o monitoramento do sistema educacional, por meio da coleta periódica de dados que permitem o acompanhamento e a comparação dos resultados.

Com relação ao processo avaliativo escolar e à avaliação educacional em larga escala, Andrade e Laros (2007), apontam que os resultados, ou seja, o desempenho estudantil, precisa ser contextualizado e que as avaliações educacionais em larga escala subsidiam adequações das políticas públicas e das diretrizes educacionais. Assim, com a identificação de fatores que impactam a proficiência, é possível o direcionamento de estratégias e medidas governamentais para a melhoria do processo de ensino aprendizagem.

No contexto internacional das avaliações educacionais em larga escala, Matos e Ferrão (2016), buscam analisar o fenômeno da repetência escolar por meio de dados do Programa Internacional de Avaliação de estudantes / *Programme for International Student Assessment* (PISA). No PISA, além dos testes cognitivos aplicados aos estudantes, também são coletadas informações através de questionários contextuais, direcionados aos estudantes e aos diretores das escolas que participam da avaliação.

Conforme Menezes *et al.* (2016), na América Latina, a maioria dos sistemas de avaliações educacionais em larga escala vêm sendo implementados a partir da década de 1990. No Brasil, o Sistema de Avaliação da Educação Básica (Saeb) é gerido pelo INEP, autarquia federal vinculada ao Ministério da Educação, que também é responsável pelas avaliações, exames e indicadores da educação superior.

Ainda, de acordo com Menezes *et al.* (2016), as avaliações educacionais em larga escala, além de aferir o desempenho dos estudantes por meio dos testes cognitivos, também

coletam informações dos contextos intraescolares e extraescolares, por meio de questionário, e, a partir da análise desses dados coletados e de suas interações e correlações com a proficiência, pode-se mensurar a eficácia escolar e subsidiar a formulação e reformulação de políticas.

Como mencionado na introdução, a avaliação da educação brasileira é amparada por um arcabouço normativo, ou seja, está prevista em diversos documentos legais, como, a Constituição Federal (BRASIL, 1988), que, no artigo 206, que trata dos princípios da educação, inciso VII, estabelece que o ensino será ministrado com base na garantia de padrão de qualidade, ou seja, demanda um processo contínuo de avaliação para a aferição e o monitoramento da qualidade da educação nas escolas brasileiras. A Lei de Diretrizes e Bases da Educação Nacional (BRASIL, 1996), que prevê o estabelecimento de um processo nacional de avaliação do rendimento escolar para a melhoria da qualidade do ensino. O Plano Nacional de Educação (BRASIL, 2014), que estabelece a implementação de um Sistema Nacional de Avaliação da Educação Básica, responsável por produzir indicadores de rendimento escolar dos estudantes e de avaliação institucional, com base em informações relativas às características de infraestrutura das escolas, dos recursos didático-pedagógicos, do perfil discente e docente, da gestão, entre outras informações. A Base Nacional Comum Curricular (BRASIL, 2016), que trata da padronização dos currículos escolares e, atualmente, subsidia a matriz de referência das avaliações do Sistema de Avaliação da Educação Básica - Saeb/INEP.

Neste sentido, para a consolidação da avaliação educacional em larga escala, faz-se necessário, além do levantamento de informações referentes à proficiência dos estudantes através dos testes cognitivos, também o levantamento de informações referentes às condições de ensino e de aprendizagem dos sistemas escolares, para o acompanhamento e monitoramento da qualidade da educação brasileira.

2.2 FATORES ASSOCIADOS AO DESEMPENHO ESCOLAR

De acordo com os autores Palermo, Silva e Novellino (2014), no cenário internacional, o marco inicial dos estudos sobre fatores que afetam o desempenho escolar, é o Relatório Coleman (*Equality of Educational Opportunity*), que foi publicado em 1966 apresentando estudos sobre os fatores associados às diferenças nos resultados entre escolas norte-americanas.

A partir de 1970, com os modelos hierárquicos, que consideram a estrutura dos dados escolares – variáveis ao nível da escola ou da turma e variáveis ao nível do estudante, de acordo

com Palermo, Silva e Novellino (2014), houve um avanço nos estudos sobre os fatores associados ao desempenho escolar.

A metodologia continuou a ser desenvolvida, principalmente na Inglaterra pelos pesquisadores Harvey Goldstein e Anthony Bryk, que reformularam os modelos de regressão para que fossem adequados aos dados que apresentassem essa estrutura de correlação. Na década de 1980, a modelagem hierárquica começou a ser utilizada na pesquisa educacional, permitindo quantificar os efeitos das escolas e avaliar a influência dos diversos fatores no desempenho dos alunos, por meio da utilização de variáveis relacionadas aos diferentes níveis de análise. (PALERMO; SILVA; NOVELLINO, 2014, p.369).

Tendo em vista situar algumas produções científicas já desenvolvidas sobre fatores associados à proficiência de estudantes em avaliações educacionais em larga escala e aplicações do modelo de regressão multinível na análise de resultados educacionais, realizou-se uma revisão sistemática da literatura nas bases de dados Scielo e Scopus. Esta revisão possibilitou um mapeamento da produção sobre o tema análise de fatores associados ao desempenho escolar através do modelo de regressão multinível.

Na busca foram utilizados os descritores “fatores associados” e (*and*) “desempenho escolar” ou (*or*) “proficiência” e (*and*) “regressão multinível” ou (*or*) “regressão hierárquica”. As bases de dados foram acessadas por autenticação (IP/IdUFSC/VPN/senha Pergamum), via Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) do Ministério da Educação (MEC), busca por base. Na base de dados Scopus, foram localizados 5 resultados e, na base de dados Scielo (periódicos em texto completo), foi encontrado apenas 1 resultado. A relação dos artigos identificados na busca nas bases de dados está apresentada abaixo, no Quadro 1.

Quadro 1 – Artigos identificados nas bases de dados Scielo e Scopus sobre fatores associados ao desempenho escolar/proficiência através de regressão multinível/regressão hierárquica

Título	Autores	Revista	Ano
Fatores associados ao desempenho escolar: estudo multinível com dados do Saeb/2001	Andrade, Josemberg M.; Laros, Jacob A.	Psicologia: teoria e pesquisa	2007
Fatores associados ao desempenho escolar em Português: um estudo multinível por regiões	Laros, Jacob Arie; Marciano João Luiz; Andrade, Josemberg Moura de	Revista Ensaio: avaliação e políticas públicas em Educação	2012
Fatores associados ao desempenho escolar: uma análise da	Palermo, Gabrielle A.; Silva, Denise Britz do	Revista Brasileira de Estudos de População	2014

Título	Autores	Revista	Ano
proficiência em matemática dos alunos do 5º ano do ensino fundamental da rede municipal do Rio de Janeiro	Nascimento; Novelino, Maria Salet Ferreira		
Repetência e indisciplina: evidências de Brasil e Portugal no Pisa 2012	Matos, Daniel Abud Seabra; Ferrao, Maria Eugenia	Cadernos de Pesquisa	2016
Policy implications of achievement testing using multilevel models: the case of Brazilian elementary schools	Menezes, Igor G.; Duran, Victor R.; Filho, Euclides J. Mendonça; Veloso, Taina J.; Sarmiento, Stella M. S.; Paget, Christine L.; Ruggeri, Kai	Frontiers in Psychology	2016
The relevance of the school socioeconomic composition and school proportion of repeaters on grade repetition in Brazil: a multilevel logistic modelo of PISA 2012	Ferrão, Maria Eugenia; Costa, Patricia Mota; Matos, Daniel Abud Seabra	Springer Open	2017

Fonte: Elaborado pela autora (2019).

Na aplicação de modelos estatísticos para identificar fatores associados ao desempenho escolar afim de subsidiar a definição de estratégias e políticas públicas para a área educacional, Andrade e Laros (2007), ressaltam a importância de identificar fatores passíveis de intervenção e que possam acarretar melhorias para o sistema educacional.

No estudo realizado por Ferrão; Costa; Matos (2017), com dados do PISA do ano de 2012, a respeito do fenômeno da repetência no Brasil, por meio do modelo de regressão multinível, os autores corroboram com a importância da investigação sobre os fatores associados ao desempenho escolar, uma vez que o Brasil tem conseguido, desde a década de 1990 avançar na questão do acesso ao Ensino Fundamental, mas ainda precisa avançar na questão da efetividade do aprendizado no sistema educacional brasileiro, ou seja, não basta garantir o acesso à escola, é preciso que sejam garantidas também as condições necessárias para a aprendizagem dos estudantes.

Com base na revisão sistemática da literatura, é possível inferir que as avaliações educacionais em larga escala, como o Saeb, geram extensas bases de dados com informações a

respeito da proficiência dos estudantes e também sobre os contextos intraescolares e extraescolares, que podem contribuir para a formulação e revisão de políticas públicas para melhoria da qualidade da educação. E, que o modelo de regressão multinível, ao considerar a estrutura hierárquica dos dados, é uma técnica estatística apropriada para a análise dos fatores associados à proficiência dos estudantes nas avaliações educacionais em larga escala.

2.3 A TEORIA DA RESPOSTA AO ITEM

A Teoria da Resposta ao Item (TRI) é um instrumento que permite medir traços latentes, ou seja, características que não podem ser medidas diretamente, mas podem ser inferidas a partir da observação de variáveis secundárias que estejam relacionadas a ela. De acordo com Andrade, Tavares e Valle (2000, p. 07),

a TRI é um conjunto de modelos matemáticos que procuram representar a probabilidade de um indivíduo dar uma certa resposta a um item como função dos parâmetros do item e da habilidade (ou habilidades) do respondente. Essa relação é expressa de tal forma que quanto maior a habilidade, maior a probabilidade de acerto no item.

A Teoria da Resposta ao Item possui vários modelos que permitem medir a característica que não pode ser observada diretamente e que precisa ser medida utilizando-se variáveis secundárias relacionadas a ela. Para a construção das escalas/indicadores das condições de infraestrutura escolar e do nível socioeconômico, nesta pesquisa, utilizou-se o modelo logístico unidimensional de dois parâmetros (ML2) e o modelo de resposta gradual, respectivamente.

O modelo logístico unidimensional de dois parâmetros define a probabilidade de um indivíduo j marcar a categoria 1 (presença do atributo) da questão i como:

$$P(U_{ij} = 1|\theta_j) = \frac{1}{1+e^{-a_i(\theta_j - b_i)}} \quad (1)$$

com questão $i = 1, 2, \dots, I$ e o indivíduo $j = 1, 2, \dots, n$ onde:

U_{ij} é uma variável dicotômica ou dicotomizada com valores 0 ou 1, sendo 1 quando o indivíduo j marcar para a questão i a alternativa que indica a presença do atributo e 0 a ausência.

θ_j é a medida de proficiência do traço latente do indivíduo j .

$P(U_{ij} = 1|\theta_j)$ é a probabilidade de um indivíduo j com uma proficiência θ marcar a categoria 1 da questão i .

a_i é o parâmetro de discriminação da questão i .

b_i é o parâmetro de dificuldade da questão i .

O chamado modelo de resposta gradual de Samejima, é apropriado quando as categorias de resposta de um item, conforme descrito acima, podem ser ordenadas entre si. A probabilidade de um indivíduo j escolher uma alternativa k ou superior do item i é representada pela equação:

$$P_{i,k}^+(\theta_j) = \frac{1}{1 + e^{-a_i(\theta_j - b_{i,k})}} \quad (2)$$

com questão $i = 1, 2, \dots, I$; o indivíduo $j = 1, 2, \dots, n$; e a alternativa $k = 1, 2, \dots, m_i$, onde:

$P_{i,k}^+(\theta_j)$ é a probabilidade de um indivíduo j com proficiência/habilidade θ marcar a categoria k ou superior da questão i .

θ_j é a medida da proficiência do indivíduo j .

a_i é o parâmetro de discriminação da questão i .

$b_{i,k}$ é o parâmetro de dificuldade da questão i em relacionado a alternativa k .

E, conforme Andrade, Tavares e Valle (2000, p. 19), deve-se ter necessariamente uma ordenação entre o nível de dificuldade das categorias de um dado, de acordo com a classificação de seus escores, ou seja:

$$b_{i,1} \leq b_{i,2} \leq \dots \leq b_{i,m_i} \quad (3)$$

2.4 REGRESSÃO MULTINÍVEL

Para avaliar os fatores associados à proficiência dos estudantes concluintes dos Anos Finais do Ensino Fundamental no Saeb 2017, utilizou-se o modelo de regressão multinível, que permite a investigação da relação entre variáveis explicativas e a variável de resposta. De acordo com Ferrão (2008, p. 29) em estudos sobre os sistemas de saúde, por exemplo, onde também são utilizados modelos de regressão multinível, podem ser identificados 4 níveis na estrutura de agrupamento da população: o indivíduo, a família, o conjunto de famílias que é

acompanhado por um médico da família e o grupo de médicos organizados em centros de saúde ou hospitais. Ou seja, o modelo de regressão multinível, pode ter tantos níveis quanto a estrutura hierárquica da população observada. Neste trabalho, com dados referentes aos sistema educacional, as variáveis explicativas ou variáveis independentes foram consideradas em dois níveis, sendo o nível 1 variáveis/características relacionadas aos estudantes e, o nível 2, variáveis/características relacionadas às escolas/redes de ensino, e a variável de resposta ou variável dependente, a proficiência em Matemática dos estudantes.

A seguir, é apresentado o modelo de regressão multinível de dois níveis, com a inclusão de variáveis explicativas de nível 1 e variáveis explicativas de nível 2:

Nível 1 (aluno)

$$y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \dots + \beta_{Kj}X_{Kij} + e_{ij}$$

(4)

Nível 2 (escola)

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_{1j} + \dots + \gamma_{0L}W_{Lj} + u_{0j}$$

$$\beta_{kj} = \gamma_{k0} + u_{kj} \quad k = 1, 2, \dots$$

Onde,

$i = 1, 2, \dots$, estudantes

$j = 1, 2, \dots$, escolas

X variáveis independentes de nível 1

W variáveis independentes de nível 2

y_{ij} o desempenho do estudante i da escola j

β_{0j} o desempenho esperado (desempenho médio) dos estudantes da escola j

γ_{00} o desempenho esperado (desempenho médio) das escolas

e_{ij} o termo aleatório associado ao estudante i da escola j

u_{0j} o termo aleatório associado à escola j

γ_{k0} efeito médio da variável independente X_k

u_{kj} efeito aleatório associado à variação do efeito da variável X_k

No modelo utilizado, com a componente aleatória associada ao intercepto, u_{0j} tem variância σ^2_{u0} , que representa a variabilidade entre escolas e a componente aleatória associada

ao coeficiente de inclinação tem variância σ_{ul}^2 , que representa a variabilidade dos coeficientes de inclinação entre escolas.

Muitas vezes, simplifica-se o modelo considerando somente o intercepto variando entre escolas, sendo os demais betas (coeficientes das variáveis de nível 1) considerados fixos. Esta simplificação será usada neste trabalho, porque assume-se a premissa de que apenas o intercepto varia entre as escolas. Outras vezes, são colocadas variáveis explicativas de nível 2 também nas equações dos β_{kj} , ou seja, procura-se explicar o porquê o efeito de alguma variável de nível 1 pode ser diferente dependendo da escola.

O erro de nível 1, e_{ij} , com variância σ_e^2 representa a variabilidade intra-escola.

Ainda sobre a opção pelo modelo de regressão multinível, “uma diferença marcante nesses modelos (*de regressão multinível*) em relação aos modelos de regressão linear múltipla é o fato de que leva em consideração a possível correlação existente entre os dados de um mesmo grupo, como nos diferentes níveis de hierarquia” (COELHO, 2017, p.23). Portanto, o modelo de regressão multinível é indicado para situações em que a população apresenta subgrupos e é suposto que as observações neles podem apresentar correlações.

2.5 APLICAÇÕES DO MODELO DE REGRESSÃO MULTINÍVEL PARA ANÁLISE DE FATORES ASSOCIADOS À PROFICIÊNCIA

Com o intuito de apresentar mais algumas aplicações do modelo estatístico de regressão para a análise de variáveis e/ou fatores associados à proficiência de estudantes em avaliações educacionais em larga escala, fez-se também uma revisão integrativa, que, de acordo com Botelho, Cunha e Macedo (2011) é uma revisão sistemática da literatura que possibilita a síntese e análise do conhecimento científico já produzido sobre o tema da pesquisa pretendida.

A revisão integrativa foi realizada utilizando-se o protocolo para busca sistemática da literatura disponível no portal da Biblioteca Universitária (BU) da Universidade Federal de Santa Catarina (UFSC), nas bases de dados portal de periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e no acervo da Biblioteca Digital Brasileira de Teses e Dissertações (BDTD). A seguir, nos quadros 2 e 3, são apresentados os trabalhos identificados na revisão integrativa, organizados por base de dados e, em seguida, são destacados os objetivos de cada estudo, as variáveis explicativas e a variável de resposta utilizadas nos respectivos modelos, organizados em sequência cronológica de publicação.

Quadro 2 – Trabalhos identificados na base de dados Portal de periódicos da CAPES ⁹

Nº	Título	Autor(es)	Ano
01	Construção de um método de regressão hierárquico para os dados do SIMAVE-2000	Tufi Machado Soares Márcia Cristina Meneghin Mendonça	2003
02	As pesquisas sobre o efeito das escolas: contribuições metodológicas para a sociologia da educação	Maria Tereza Gonzaga Alves José Francisco Soares	2007
03	Expectativas de futuro como efeito-escola: explorando possibilidades	Márcio da Costa Reginaldo Guedes	2009
04	Organização e contextualização de dados como subsídio para a compreensão dos resultados das avaliações educacionais	Margareth Watanabe Maria Candida Raizer Cardinalli Perez	2009
05	A expectativa do professor e o desempenho dos alunos	Tufi Machado Soares Neimar da Silva Fernandes Mariana Santos Botarro Ferraz Juliana de Lucena Ruas de Riani	2010
06	Contexto escolar e indicadores educacionais: condições desiguais para a efetivação de uma política de avaliação educacional	Maria Teresa Gonzaga Alves José Francisco Soares	2013
07	Indicadores educacionais e contexto escolar: uma análise das metas do Ideb	Daniel Abud Seabra Matos Erica Castilho Rodrigues	2016
08	Indisciplina no PISA: entre o intra e o extraescolar	Luciano Campos da Silva Daniel Abud Seabra Matos	2017
09	Estudo de fatores associados através de regressão quantílica hierárquica	Pedro Alberto Barbeta Dalton Francisco de Andrade Héilton Ribeiro Tavares	2018

Fonte: Elaborado pela autora (2019).

⁹ Trabalhos identificados no portal de periódicos da CAPES com os descritores “qualquer” / “contém” / “avaliação educacional” e (and) “qualquer” / “contém” / “regressão”. Como critério para seleção foi utilizado como período de tempo inicial o ano de 1996 (ano de promulgação da atual LDBEN) e como final o ano de 2018. Como tipo de material foi utilizada a opção todos os itens (livros, artigos, imagens e audiovisuais). Foram encontrados 22 resultados, dos quais foram considerados os 17 periódicos revisados por pares. Em seguida, outro critério de seleção foi a adequação do título e/ou resumo ao tema de interesse. Foram considerados para elaboração da matriz de análise publicações pertinentes ao Ensino Fundamental e que utilizaram modelos de regressão.

Quadro 3 – Trabalhos identificados na base de dados BDTD ¹⁰

Nº	Título	Autor(es)	Ano	Instituição
01	O negro e as políticas de ação afirmativa no contexto da desigualdade educacional	Julimar Constâncio	2009	Universidade Federal de Juiz de Fora (UFJF)
02	O aspecto socioeconômico e sua influência na qualidade do ensino fundamental público no Brasil	Raquel Dias Costa	2010	Fundação Getúlio Vargas (FGV)
03	Variáveis significativamente relacionadas com o desempenho dos alunos do 5º ano do Ensino Fundamental público no estado do Ceará no teste de Língua Portuguesa da Prova Brasil/2011	Antônia Bruna da Silva	2015	Universidade Federal do Ceará (UFC)
04	Responsabilização educacional em Pernambuco	Clayton Sirilo do Valle Furtado	2015	Universidade Federal de Juiz de Fora (UFJF)
05	Análise do Sistema Permanente de Avaliação da Educação Básica do Ceará (SPAECE): um estudo comparativo entre o efeito escola e os resultados de proficiência em escolas da rede estadual nos anos de 2012 a 2014	José Valmir Guimarães de Oliveira	2016	Universidade Federal do Ceará (UFC)
06	Seleção de modelos multiníveis para dados de avaliação educacional	Fabiano Rodrigues Coelho	2017	Universidade de São Paulo (USP)

Fonte: Elaborado pela autora (2019).

Soares e Mendonça (2003) utilizaram modelos de regressão hierárquicos para identificar a relação entre o desempenho escolar, as características técnico-pedagógicas das escolas e o perfil socioeconômico dos alunos das 4ª séries do Ensino Fundamental que participaram do Sistema Mineiro de Avaliação da Educação Pública (SIMAVE) no ano de 2000. Foram utilizados 29 fatores extraídos dos questionários dos professores que foram agregados para representarem características das escolas (**variáveis explicativas referentes às escolas – nível 2**) e 10 fatores dos questionários dos alunos, que foram agregadas para a turma (**variáveis explicativas referentes às turmas – nível 1**). A **variável de resposta** foi a proficiência média em Matemática e Português dos alunos das 4ª séries do Ensino Fundamental no SIMAVE 2000. Os resultados apontam que as variáveis defasagem, abandono escolar, a incidência de alunos negros na turma e o percentual de alunos do sexo masculino na turma tem maior impacto negativo sobre a proficiência dos alunos. O efeito da escolaridade da mãe, o

¹⁰ Trabalhos identificados no portal da Biblioteca Digital de Teses e Dissertações (BDTD) com a ferramenta de busca avançada, busca por “avaliação educacional” e “regressão”, no período de 1996 a 2018, em todos os idiomas, sem preferência com relação a ilustrações e teses e dissertações como tipo de documento. Foram localizados 16 resultados, dos quais seguem 07 com título e/ou resumo pertinentes ao tema de interesse e relacionados ao Ensino Fundamental. Seguem as publicações selecionadas que utilizaram o modelo de regressão hierárquica ou que utilizaram o modelo de regressão linear múltipla.

escore socioeconômico e a frequência a cultos religiosos tem maior impacto positivo na proficiência dos alunos. As variáveis do nível da escola com impacto positivo são a avaliação pelos professores da escola da participação escolar do aluno e a frequência com que os professores da escola passam lições de casa.

Alves e Soares (2007), no artigo “As pesquisas sobre o efeito das escolas: contribuições metodológicas para a Sociologia da Educação” apresentam os resultados da pesquisa longitudinal desenvolvida para medir o efeito das escolas no processo de aprendizagem dos alunos. A pesquisa foi realizada em sete escolas públicas localizadas na cidade de Belo Horizonte, no Estado de Minas Gerais, durante dois anos. Foram realizadas duas coletas de dados no primeiro ano, com os alunos na 5ª série e uma coleta de dados no final do segundo ano com os alunos na 6ª série. Também foram realizadas entrevistas com profissionais da escola e com os familiares. A medida da aprendizagem foi obtida através de testes respondidos pelos alunos, compostos por itens desenvolvidos pelos professores das escolas, itens públicos do banco de dados do Saeb e do projeto Nova Escola e itens pré-testados do Banco de Itens do Grupo de Avaliação e Medidas Educacionais da Universidade Federal de Minas Gerais (Game-UFGM). A partir do uso de itens públicos, foi possível a equalização da escala de desempenho dos alunos da pesquisa com a escala do Saeb. No modelo multinível de regressão múltipla a **variável de resposta** foi o desempenho dos alunos no início do estudo e o seu progresso ao final dos dois anos. As **variáveis explicativas referentes aos estudantes (nível 1)** foram: gênero, atraso escolar e nível socioeconômico. Pelo fato de serem apenas sete escolas consideradas na pesquisa, elas “não foram consideradas como um nível hierárquico, mas foram incluídas nos modelos como variáveis indicadoras” (p.455). Os resultados evidenciaram que as diferenças entre as escolas se mantiveram, mas as médias gerais se aproximaram ao final do período, além disso, escolas com resultados ruins no início conseguiram melhorar os resultados ao final da pesquisa e as que tiveram melhores resultados no início mantiveram os seus resultados.

Costa e Guedes (2009) analisaram as diferenças nas expectativas profissionais e de escolarização entre estudantes de escolas públicas do município do Rio de Janeiro, a partir de uma pesquisa realizada sobre diferenças de prestígio entre escolas de uma mesma rede de ensino. Foram utilizados dados de um *survey*¹¹ realizado com alunos de 5ª a 8ª séries e aproximadamente 100 docentes de seis escolas municipais do Rio de Janeiro. Após a análise

¹¹ *Survey*, método quantitativo para coleta de dados de determinado grupo de pessoas, geralmente através de questionário (FREITAS, OLIVEIRA, SACCOL, MOSCAROLA, 2000).

descritiva, foi utilizado o modelo de regressão logística, com **variáveis no nível do aluno** (escolaridade da mãe, nível socioeconômico, sexo, se considera aluno muito ruim/muito bom, esforço, hábitos de estudo, defasagem escolar) e a **variável no nível da escola** (prestígio da escola), sendo a **variável de resposta** a expectativa dos alunos de chegada ao ensino superior. Os resultados indicam que escolas consideradas de alto prestígio elevam a probabilidade de um aluno declarar sua expectativa de cursar o nível superior e que a defasagem idade/série reduz a cada ano de atraso esta expectativa.

Watanabe e Perez (2009), descrevem os aspectos metodológicos que orientam a construção do Sistema de Indicadores para Monitoramento das Escolas Estaduais Paulistas (SIM-Escolas). O SIM-Escolas é um sistema de informações com indicadores educacionais relevantes de cada escola, de acordo com o seu contexto socioeconômico. De acordo com os autores, o desempenho dos alunos é influenciado por fatores intraescolares (estrutura escolar, formação docente, equipamentos, número de alunos por sala, práticas pedagógicas, entre outros) e fatores extraescolares (condições socioeconômicas e educacionais da família, especificidades dos alunos). Para validação dos indicadores educacionais foram utilizados modelos de regressão multinível. A **variável de resposta** utilizada foi o desempenho dos alunos da 8ª série do Ensino Fundamental nas provas de Língua Portuguesa e Matemática no Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo (Saresp) do ano de 2007. As **variáveis explicativas referentes aos alunos (nível 1)** utilizadas foram: sexo do aluno; cor; defasagem idade-série; escolaridade da mãe; classe econômica; mais de 20 livros em casa; frequência a cursos extraescolares; condição de frequência às aulas; trajetória escolar; incentivo ao estudo; satisfação do aluno com os professores das diferentes matérias; satisfação do aluno com os professores de Matemática e de Português. As **variáveis explicativas referentes às escolas (nível 2)** foram: a proporção de docentes eventuais; proporção de professores que permaneceram na escola entre 2006 e 2007; idade do diretor; renda familiar do diretor; atitude do diretor diante das dificuldades. Das variáveis consideradas, a idade do diretor foi a única que apresentou valores não significativos tanto em Matemática quanto em Língua Portuguesa, as demais tiveram efeito positivo ou negativo sobre o desempenho dos alunos nas provas de Língua Portuguesa e Matemática.

Na dissertação “O negro e as políticas de ação afirmativa no contexto das desigualdades sociais” do Programa de Pós-graduação em Educação da UFJF, Constâncio (2009) utilizou modelos de regressão multinível para identificar como os fatores relacionados ao perfil do aluno e a composição da turma estão associados ao desempenho dos alunos. Foram

utilizados dados de avaliações censitárias estaduais, sendo do PROEB 2007 (Minas Gerais), do programa de avaliação Nova Escola 2006 (Rio de Janeiro), do Sistema de Avaliação do Rendimento Escolar do Rio Grande do Sul (SAERS) 2006, e do Saeb 2003 - que coleta dados nacionais de forma amostral. A **variável de resposta** foi a proficiência em Língua Portuguesa e em Matemática dos alunos. As **variáveis explicativas referentes aos alunos (1º nível)** consideradas no modelo foram gênero (masculino); indicador socioeconômico; cor (preta); cor (parda); defasagem. As **variáveis explicativas referentes à turma (2º nível)** foram o indicador socioeconômico médio da turma; percentual de alunos de cor preta na turma. Os resultados da pesquisa mostram que a diferença de desempenho dos alunos pretos e brancos é reduzida quando controladas as condições econômicas, mas o aluno preto ainda que controlada a variável econômica possui uma proficiência escolar inferior aos alunos brancos, e que este fato necessita de pesquisas complementares para maior entendimento das relações de ensino-aprendizagem destes e para com estes alunos. No aporte teórico o autor justifica a descaracterização da raça como uma hierarquização genética, mostrando-a como um “modelo de classificação social, em que os próprios indivíduos são dispostos por categoria de cor/raça, em razão de uma compleição física que estabeleça consonância com aspectos históricos e culturais” (p.94).

Costa (2010), em sua dissertação “O aspecto socioeconômico e sua influência na qualidade do ensino fundamental público no Brasil” utilizou o modelo de regressão múltipla para tentar demonstrar a relação do nível socioeconômico com o desempenho escolar de alunos da educação básica brasileira. Foram utilizados dados do ano de 2007 do IDEB, do Censo Escolar referentes a formação docente e da remuneração média dos professores disponibilizados pela Confederação Nacional dos Trabalhadores em Educação (CNTE) e dados do ano de 2005 referentes ao Índice de Desenvolvimento Humano (IDH). A técnica de análise de regressão múltipla “descreve o relacionamento entre variáveis [...] que permitem a avaliação da significância de determinadas variáveis na explicação do fenômeno estudado” (p.49). A **variável dependente ou de resposta** foi o IDEB 2007 de cada estado brasileiro, e, as **variáveis explicativas ou independentes** foram o IDH; a remuneração média dos professores; e, a proporção de docentes com formação superior. Das três variáveis selecionadas, a que apresentou relação mais significativa com o desempenho escolar foi o aspecto socioeconômico, demonstrando forte relação entre o IDH e o IDEB, ou seja, o nível social e econômico da população teve forte influência no desempenho escolar dos alunos.

Soares, Fernandes, Ferraz e Riani (2010), no artigo “A expectativa do professor e o desempenho dos alunos” descrevem o estudo no qual cruzam informações dos questionários

contextuais do professor e do aluno com os resultados dos testes cognitivos do Programa de Avaliação da Educação Básica (PROEB) 2006 com alunos da 4ª série do Ensino Fundamental de escolas das redes Estaduais e Municipais de Minas Gerais. Foram construídos dois modelos de regressão hierárquico, um com a **variável de resposta/dependente** a proficiência em Língua Portuguesa e outro a proficiência em Matemática. No **primeiro nível/variáveis explicativas referentes aos alunos** foram consideradas as variáveis gênero (masculino); cor (preta); bolsa-família; defasagem do aluno. No **segundo nível/variáveis de resposta referentes às turmas** foram calculadas as médias e os percentuais das características individuais dos alunos – percentual de alunos que recebem bolsa-família; percentual de alunos pretos; percentual de alunos do gênero masculino; defasagem média dos alunos; além das variáveis interesse e dedicação do professor; barulho e bagunça em sala de aula; exigência do professor; falta de compromisso dos professores e alunos para com a aula; compromisso do professor com o dever; falta de colaboração dos pais; coesão e trabalho em equipe; indisciplina dos alunos; carência de materiais e instalações; expectativa; sintonia. Os resultados demonstram que as características do professor e da sala de aula são significativamente importantes no desempenho dos alunos. A sintonia e a expectativa foram as variáveis com efeito mais positivo sobre a proficiência, tanto em Matemática quanto em Língua Portuguesa.

Na pesquisa realizada por Alves e Soares (2013), foram utilizados dados referentes ao IDEB, Censo Escolar e Prova Brasil do ano de 2009, de escolas públicas. No modelo foram consideradas as dimensões escolares passíveis de serem aferidas empiricamente com os dados disponíveis. A **variável de resposta** foi o IDEB 2009, as **variáveis explicativas referentes aos alunos (nível 1)** foram a média do nível socioeconômico da escola (medida estimada via TRI a partir de informações dos questionários respondidos pelos alunos em avaliações educacionais do INEP); o percentual de alunos brancos (proporção de alunos que se declararam brancos, obtida no questionário da Prova Brasil); o percentual de alunos do sexo masculino (proporção de alunos do sexo masculino, obtida no questionário da Prova Brasil); já as **variáveis explicativas referentes às escolas (nível 2)** foram a infraestrutura da escola (medida obtida via TRI a partir de informações do Censo Escolar) e a complexidade da escola (medida obtida via TRI a partir de informações do Censo Escolar). Os resultados da pesquisa evidenciam que as escolas com a média do nível socioeconômico mais baixas são as que apresentam os menores resultados no IDEB, ou seja, o IDEB está fortemente associado ao nível socioeconômico médio de seus alunos. Entre os fatores das escolas, os resultados destacam o impacto positivo da infraestrutura, sendo maior nas séries iniciais do que nas séries finais. O artigo também discute

o conceito de equidade, *accountability* (políticas de responsabilização) e evidencia que se faz necessária uma análise mais ampla para a discussão de qualidade educacional do que apenas a sua síntese em um único indicador.

Silva (2015), também considerando que o desempenho escolar sofre influências de fatores intraescolares e extraescolares, em sua dissertação do Programa de Pós-Graduação em Educação Brasileira da Universidade Federal do Ceará, utilizou o **modelo de regressão linear múltipla com variáveis referentes ao aluno, ao professor, ao diretor e à escola**, para investigar a influência delas na proficiência em Língua Portuguesa dos alunos do 5º ano do Ensino Fundamental de escolas públicas municipais no estado do Ceará, na Prova Brasil do ano de 2011. No modelo explicativo, a variável relativa à reprovação e ao abandono escolar teve maior influência negativa sobre a proficiência e a participação dos pais nas reuniões escolares e o incentivo aos estudos teve influência positiva na proficiência desses alunos.

Furtado (2015), em sua tese de doutorado do Programa de Pós-Graduação em Educação da Universidade Federal de Juiz de Fora, intitulada “Responsabilização educacional em Pernambuco” buscou, por meio de regressões lineares verificar se o bônus de desempenho educacional (BDE¹²) produziu efeitos positivos no desempenho dos alunos nas disciplinas de Língua Portuguesa e Matemática entre os anos de 2008 e 2012, com dados do Saeb/Prova Brasil e do Sistema de Avaliação Educacional do Pernambuco (SAEPE). Para isso, a pesquisa trabalhou com dados referentes às escolas estaduais (que recebiam o bônus) e referentes às escolas municipais (que não recebiam o bônus). De acordo com o autor, há uma forte relação entre as políticas de responsabilização e as avaliações educacionais em larga escala, e uma forte relação entre as políticas de responsabilização e a eficácia escolar. O resultado por escola foi escolhido como **variável de resposta** pois a bonificação se faz pelo desempenho da escola, assim, as informações da proficiência dos alunos foram agregadas por escola. O autor pondera que “a verificação do desempenho do aluno é determinada por inúmeras variáveis, e, portanto, o desempenho da escola em si consiste em um emaranhado de variáveis” (p.139), por isso, foram comparadas escolas com características semelhantes. A metodologia utilizada foi o escore de propensão, o pareamento de escores entre escolas da rede estadual e da rede municipal e o método das diferenças em diferenças. Depois foram feitas análises descritivas da evolução do recebimento do bônus e, depois foram realizadas análises com **regressões lineares** considerando **variáveis explicativas referentes à escola** (depredação na escola; depredação

¹² Incentivo financeiro pago aos funcionários das escolas da rede estadual de Pernambuco mediante o cumprimento de metas educacionais pré-estabelecidas. (FURTADO, 2015).

nos banheiros), **referentes ao professor** (sensação de segurança do professor na escola; se apresentou a matriz do teste aos alunos; uso dos resultados da avaliação pela direção; expectativa pelo recebimento integral do bônus; conhecimento das regras de recebimento do bônus), **referentes ao aluno** (sensação de segurança na escola; percepção de bagunça na escola; se os alunos identificam falta de professor; se os alunos eram avisados com antecedência sobre o teste; se na véspera do teste eram passados mais exercícios de Matemática e Língua Portuguesa) e **referentes à direção** (sensação de segurança do diretor na escola; opinião sobre a importância do bônus aos professores; expectativa no recebimento integral do bônus pelo diretor). Além destas, em todas as regressões foram utilizadas as variáveis índice socioeconômico, tamanho da escola e tipo de ensino. Nas considerações finais o autor destaca que as variáveis: clima escolar e frequência de alunos e professores, escola limpa e segura, tiveram impacto positivo na proficiência dos alunos e, conseqüentemente, no recebimento da bonificação e que, no caso do estado do Pernambuco, a política de bônus por desempenho teve efeito positivo no desempenho dos alunos nas avaliações educacionais em larga escala.

Oliveira (2016), em sua dissertação, buscou demonstrar através da **regressão linear múltipla** a relação entre a **variável de resposta** proficiência dos alunos do Ensino Médio em Matemática e Português no Sistema Permanente de Avaliação da Educação Básica do Ceará (SPAECE) e as variáveis contextuais, nos anos de 2012, 2013 e 2014. As **variáveis contextuais ou variáveis explicativas** utilizadas foram: tipo de escola; desempenho da escola; situação socioeconômica do aluno; contexto familiar do aluno; práticas pedagógicas do professor; clima escolar; expectativa acadêmica/profissional do aluno; gestão democrática. A análise mostra que o efeito-escola da variável clima escolar tem mais significância que as demais variáveis consideradas, indo de encontro à revisão de literatura da dissertação que apontava o contexto familiar como a variável com maior influência no desempenho escolar do aluno em Português e Matemática.

A pesquisa realizada por Matos e Rodrigues (2016), teve como objetivo analisar a associação das metas do IDEB 2013 com outros indicadores educacionais e comparar os Anos Iniciais e os Anos Finais do Ensino Fundamental. Para isso, foram utilizadas a estatística descritiva, a TRI e a regressão logística, com dados do portal do INEP referentes às avaliações educacionais em larga escala e o Censo Escolar. A **variável de resposta** foi a meta do IDEB 2013 (assumindo valor zero se não atingiu e 1 se atingiu a meta), as **variáveis explicativas** foram nível socioeconômico (medida obtida via TRI com dados do Censo Escolar); índice de infraestrutura da escola (obtido via TRI com informações do Censo Escolar); índice de

complexidade da escola (obtido via TRI com dados do Censo Escolar, agrega informações sobre o número de alunos, funcionários e número de salas, os níveis de ensino ofertados e o número de alunos com necessidades especiais); média de alunos por turma da escola (variável disponibilizada pelo INEP); percentual de docentes da escola com curso superior. As variáveis foram padronizadas e foi ajustado o modelo de regressão logística utilizando o *software* R. Os resultados da pesquisa corroboram estudos anteriores de que existe associação entre a infraestrutura das escolas e o desempenho dos estudantes, e que este efeito é mais relevante nos Anos Iniciais do Ensino Fundamental. Com relação a complexidade da escola os resultados indicam que a complexidade escolar diminui a chance de a escola atingir a meta do IDEB, tendo impacto mais negativo nos Anos Finais do Ensino Fundamental. Na variável nível socioeconômico, para os Anos Iniciais, há um aumento de chance da escola atingir a meta do IDEB com o aumento do nível socioeconômico, mas, para os Anos Finais, o aumento do nível socioeconômico tem efeito negativo. A média de alunos por turma foi estatisticamente significativa apenas para os Anos Iniciais, sendo que o aumento do número de estudantes reduz as chances de a escola atingir a meta do IDEB. O impacto da variável docentes com curso superior foi maior nos Anos Iniciais, mas foi positivo nos dois níveis do Ensino Fundamental, como esperado. A pesquisa demonstra um funcionamento diferencial entre os Anos Iniciais e os Anos Finais do Ensino Fundamental nas variáveis nível socioeconômico, infraestrutura escolar, complexidade escolar e percentual de docentes com curso superior, mas os autores consideram que pesquisas complementares são necessárias assim como estudos com dados longitudinais, para o monitoramento das escolas ao longo dos anos.

Coelho (2017), na dissertação intitulada “Seleção de modelos multiníveis para dados de avaliação educacional”, divide a pesquisa em seleção de variáveis e seleção de modelos. A **variável de resposta** utilizada foi o desempenho dos alunos do 9º ano do Ensino Fundamental do estado de São Paulo em matemática, no Saeb 2011. As **variáveis explicativas** que influenciaram a proficiência dos alunos do 9º ano em matemática foram: grau de instrução da mãe; frequência de leitura de livros; tempo gasto com recreação em dia de aula; o fato de gostar de matemática; o desempenho em matemática global da escola; o desempenho em Língua Portuguesa do aluno; a dependência administrativa da escola; o sexo; o grau de instrução do pai; reprovações e distorção idade-série.

Silva e Matos (2017), analisaram o fenômeno da indisciplina escolar com dados do Programa Internacional de Avaliação de Estudantes (PISA) 2012, relativos aos estudantes brasileiros de 15 anos de idade, buscando identificar características dos estudantes e das escolas

(fatores intraescolares e extraescolares) associados à indisciplina. O PISA é constituído de testes cognitivos padronizados e questionários contextuais aplicados aos estudantes e diretores das escolas avaliadas. Na pesquisa desenvolvida por Silva e Matos, a **variável de resposta** foi o índice do clima escolar e as **variáveis explicativas referentes ao aluno** foram o gênero; a repetência; o índice do nível socioeconômico, social e cultural; o índice da relação professor-aluno; e, as **variáveis explicativas referentes à escola** foram o tipo de escola; o tamanho da escola; o índice de qualidade da infraestrutura; o nível socioeconômico médio da escola; e a proporção de repetentes da escola. O resultado da pesquisa evidencia que as variáveis nível socioeconômico dos alunos, gênero, nível socioeconômico da escola e tipo de escola não apresentaram resultados estatisticamente significativos, aponta ainda que, quanto maior a escola, maior a indisciplina, que a repetência implica um aumento na indisciplina e que quanto melhor a relação professor-aluno, menor a indisciplina. Neste estudo, a indisciplina estaria mais associada a fatores intraescolares do que extraescolares.

Barbetta, Andrade e Tavares (2018), utilizaram regressões quantílicas hierárquicas para o estudo de fatores associados ao desempenho dos estudantes nas provas de Língua Portuguesa e Matemática do 8º ano do Ensino Fundamental do Sistema Paraense de Avaliação Educacional (SisPAE) do ano de 2016. O estudo teve por objetivo avaliar o efeito dos diferentes fatores associados nos diversos quantis ou percentis da distribuição das proficiências, para identificar o efeito de cada fator em estudantes ou escolas de alto ou de baixo desempenho. A **variável de resposta** foi o desempenho dos estudantes do 8º ano do Ensino Fundamental no SisPAE 2016; as **variáveis explicativas referentes ao estudante (nível 1)** foram o nível socioeconômico da família do estudante; o nível de relacionamento escolar na percepção do estudante; o nível do clima escolar na percepção do estudante; o sexo (indicativo de sexo feminino); se o estudante teve reprovação; o tempo gasto até a escola maior que ½ hora; estudante aprende matéria com professor; o aprendizado depende mais do professor do que do estudante; o professor dá mais atenção aos estudantes bons; o professor se preocupa com o dever de casa; o professor costuma faltar; o estudante lê frequentemente; os pais incentivam a leitura; as **variáveis explicativas referentes à escola (nível 2)** foram o nível socioeconômico do ano/escola; o nível de relacionamento escolar da série/escola; o nível do clima escolar da série/escola. Os resultados do estudo indicam fatores que exercem mais influência em estudantes de alto e de baixo desempenho. Estudantes com alto desempenho são mais afetados/efeito positivo, por exemplo, pela preocupação do professor com os deveres de casa e as correções feitas pelo professor e com o nível socioeconômico. Nas considerações os autores

avaliam a regressão quantílica hierárquica como um instrumento que pode contribuir de forma objetiva para a análise dos efeitos das políticas públicas que visam atenuar as diferenças de desempenho entre os estudantes.

Com a revisão integrativa realizada, pode-se observar que vários trabalhos vem sendo desenvolvidos com o interesse em contribuir na análise dos fatores associados ao desempenho dos estudantes nas avaliações educacionais em larga escala e, com isso, buscar subsidiar a tomada de decisões por parte de gestores escolares, e também de agentes públicos na criação, reformulação e proposição de políticas públicas para a melhoria da qualidade da educação.

3 METODOLOGIA

Neste capítulo, são apresentadas as características metodológicas da pesquisa.

Do ponto de vista de sua natureza, caracteriza-se como uma pesquisa aplicada, ou seja, que “objetiva gerar conhecimentos para aplicação prática e dirigidos à solução de problemas específicos” (SILVA; MENEZES, 2005). E, com relação a forma de abordagem, é uma pesquisa quantitativa, utilizando dados secundários oriundos de avaliações educacionais em larga escala, especificamente, os microdados da Prova Brasil 2017, além de outros dados disponibilizadas pelo INEP, bem como técnicas estatísticas para análise destes dados. Como foco de pesquisa, os estudantes concluintes dos Anos Finais do Ensino Fundamental nas escolas brasileiras.

Para desenvolver uma análise dos fatores associados ao desempenho em Matemática dos estudantes do 9º ano do Ensino Fundamental das escolas brasileiras no Saeb 2017, foi necessária a construção de indicadores educacionais e, para isso, utilizou-se a Teoria da Resposta ao Item, por ser um instrumento que permite medir um traço latente.

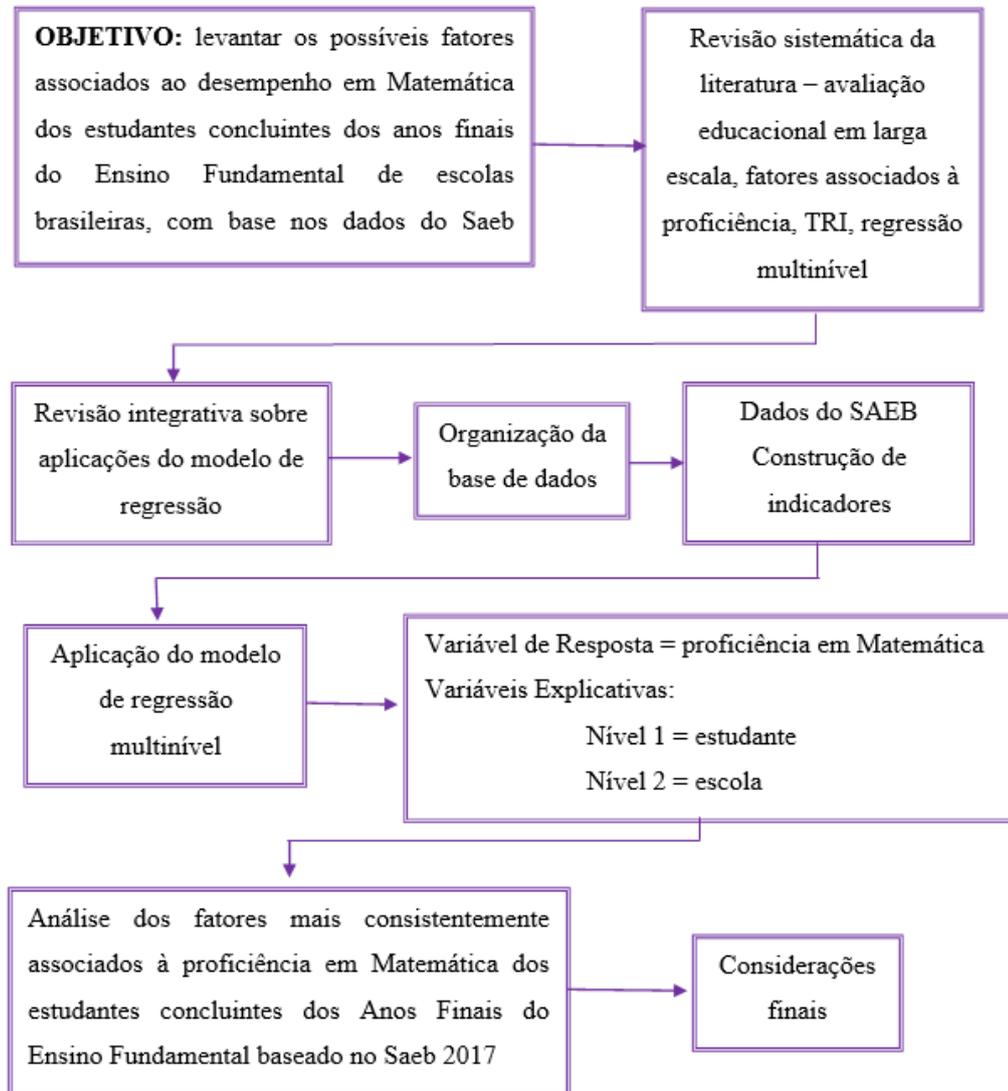
E, para a análise dos fatores associados à proficiência dos estudantes serão utilizados modelos de regressão multinível, pois este modelo,

possibilita investigar a relação de causa e efeito, levando em consideração o indivíduo dentro do grupo e dentro da sociedade, bem como a relação de influência do grupo no contexto e como esse grupo é influenciado pelo meio. [...] Assim, considera que o sistema educacional é organizado em níveis, ou seja, em hierarquia, onde os estudantes são agrupados em turmas e as turmas em escolas que, por sua vez, podem ser agrupadas em municípios e estes em unidades da federação. (PEREIRA, 2016, p. 51-52).

3.1 DESIGN DE PESQUISA

Com o intuito de descrever os passos percorridos para o desenvolvimento da pesquisa, foi elaborado o esquema a seguir, onde são apresentadas, de forma breve, as principais etapas do trabalho.

Figura 1 – Descrição das etapas da pesquisa



Fonte: Elaborado pela autora (2020).

3.2 ANÁLISE DE DADOS

A partir dos exemplos apresentados no capítulo Referencial Teórico, seção 2.5, de aplicações do modelo de regressão para análise de fatores associados ao desempenho de estudantes em avaliações educacionais em larga escala, pode-se inferir que há consenso entre os autores de que a aprendizagem é um fenômeno complexo que é afetado por fatores intraescolares e extraescolares. E também, considerando a revisão integrativa, foi possível a realização de um levantamento das variáveis que apresentaram impacto significativo na proficiência dos estudantes em pesquisas anteriores, subsidiando a definição de algumas

variáveis que serão consideradas na aplicação do modelo de regressão multinível para a análise dos fatores que estão mais consistentemente associados ao desempenho em Matemática dos estudantes concluintes dos anos finais do Ensino Fundamental de escolas brasileiras no Saeb 2017.

Na figura abaixo, estão representadas as variáveis relacionadas às escolas e as variáveis relacionadas aos estudantes, identificadas a partir da revisão integrativa, com impacto relacionado à proficiência de estudantes em avaliações educacionais em larga escala.

Figura 2: Variáveis de nível 1 (estudante) e nível 2 (escola) com impacto na proficiência dos estudantes nas avaliações educacionais em larga escala selecionadas na revisão integrativa da literatura



Fonte: Elaborado pela autora (2019).

A partir da Figura 2, elaborada com os achados da revisão integrativa da literatura, é importante destacar que, de acordo com a Diretoria de Avaliação da Educação Básica do INEP (2018b), a proficiência dos estudantes é uma das evidências a respeito da qualidade do ensino ministrado nas escolas, mas, nenhum fator determina isoladamente a qualidade do ensino. Mesmo a análise multivariada dos fatores associados ao desempenho dos estudantes, pretendida

nesta pesquisa, precisa ser reconhecida enquanto um recorte, uma perspectiva com base em uma ferramenta avaliativa (a avaliação educacional em larga escala/Saeb 2017) e por um modelo estatístico (modelo de regressão multinível).

A opção pelo modelo de regressão multinível¹³ para a análise dos fatores associados à proficiência dos estudantes na Prova Brasil 2017, deu-se pela estrutura hierárquica dos dados, ou seja, os estudantes estão inseridos em uma turma, que por sua vez está inserida em uma escola, que faz parte de uma rede de ensino e de determinado contexto socioeconômico e cultural. No modelo serão considerados dois níveis, sendo o nível 1 relacionado ao estudante e o nível 2 relacionado às escolas.

Pretendeu-se, com a aplicação do modelo de regressão multinível, identificar quais variáveis estão influenciando mais fortemente, positiva ou negativamente, a proficiência em Matemática dos estudantes concluintes dos anos finais do Ensino Fundamental das escolas brasileiras, considerando os dados do ano de 2017 da Prova Brasil. Ou seja, verificar a influência das variáveis relacionadas ao estudante (nível 1) e de variáveis relacionadas às escolas (nível 2) na proficiência dos estudantes (variável de resposta).

Inicialmente foi gerado o modelo nulo (sem variáveis explicativas). No modelo nulo, o preditor linear é composto apenas pelo intercepto. Conforme Ferrão (2003, p. 32), o “modelo nulo multinível é particularmente importante no estudo da distribuição da variância total da variável de resposta pelos níveis de agrupamento”. Com ele é possível o cálculo do coeficiente de correlação intraclasse, medida do grau de correlação entre estudantes de uma mesma escola e a proporção da variância das escolas em relação a variância total (FERRÃO, 2003, p. 34).

Depois foram introduzidas variáveis explicativas relacionadas ao estudante (nível 1). Com isso, de acordo com Hartmann (2017, p. 43) pode-se “observar a contribuição de cada variável explicativa relativa aos alunos, bem como estimar a contribuição de cada variável na redução da variância, deste nível, na proficiência ou variável de resposta”.

E, foram acrescentadas variáveis relacionadas às escolas (nível 2), para a verificação dos seus impactos no rendimento dos estudantes. Por fim, no modelo 4, foram incluídas todas as variáveis para construção/criação de um modelo mais explicativos dos fatores intra e extraescolares que impactam o desempenho em Matemática dos estudantes do 9º do Ensino Fundamental no Saeb 2017.

¹³ Hartmann (2017, p. 34), contextualiza que o modelo de regressão multinível também é denominado modelo linear de coeficientes randômicos, modelos lineares mistos, análise de regressão multinível, modelo hierárquico, modelo de coeficientes aleatórios. Nesta pesquisa adotou-se a nomenclatura modelo de regressão multinível para a técnica estatística de modelagem hierárquica.

4 CONSTRUÇÃO DE INDICADORES

Apesar do INEP disponibilizar alguns indicadores educacionais em seu portal, para a realização deste trabalho, optou-se pela construção de indicadores complementares para a investigação dos fatores associados à proficiência dos estudantes do 9º ano do Ensino Fundamental no Saeb 2017, pois, no processo de organização dos dados utilizando diferentes bases de dados para a aplicação da regressão multinível, observou-se uma significativa perda de dados pelo fato de muitas escolas que tiveram seus resultados divulgados com a utilização de máscaras, ou seja, códigos identificadores fictícios, porque não obtiveram o percentual mínimo de 80% de participação dos estudantes no dia da avaliação, e, por conta desses dados faltantes/*missings*, a análise estatística ficaria prejudicada. Desta forma, foram geradas as medidas da escala de infraestrutura escolar, tomando como referência a escala desenvolvida por Neto, Jesus, Karino e Andrade (2013), e, pelo fato da medida de nível socioeconômico disponibilizada pelo INEP em seu portal ser referente ao ano de 2015, optou-se por criar outra escala de nível socioeconômico com as informações do questionário contextual do Saeb 2017 dos estudantes e, posteriormente, agrupá-la por escola para obtenção da medida de nível 2, tomando por referência a escala desenvolvida pelo INEP, que classifica a medida de nível socioeconômico em 7 níveis.

Para a construção destes indicadores, que permitem mensurar características relativas a realidade educacional, foi utilizada a Teoria da Resposta ao Item (TRI), que é um instrumento que permite medir características que não podem ser medidas diretamente, mas podem ser inferidas a partir da observação de variáveis secundárias que estejam relacionadas a ela.

4.1 ESCALA DE INFRAESTRUTURA ESCOLAR

Com o objetivo de fazer uma análise dos dados sobre infraestrutura das escolas brasileiras, disponíveis no questionário contextual da escola (TS_ESCOLA) do Saeb do ano de 2017, foi construída uma escala por meio da Teoria da Resposta ao Item, tomando-se como referência a escala de infraestrutura escolar dos autores Neto, Jesus, Karino e Andrade (2013), que estabelece quatro categorias (elementar, básica, adequada e avançada) para as condições de infraestrutura das escolas a partir da utilização da TRI. Porém, nesta pesquisa, foram utilizados somente os dados do questionário contextual da escola do Saeb 2017 ao invés dos dados coletados no Censo Escolar 2017 como os autores Neto, Jesus, Karino e Andrade fizeram em

2013, também pelo fato de que em 2017, muitas escolas não obtiveram o percentual mínimo de 80% de participação dos estudantes no dia da avaliação, tendo estas escolas os seus resultados divulgados com o código identificador substituído por uma máscara, iniciada pelo dígito 6, o que, no momento de juntar as informações coletadas por diferentes instrumentos, acaba gerando dados faltantes/*missings*, que prejudicam a análise estatística. No caso da base de dados aqui utilizada, o questionário contextual da escola do Saeb 2017, foram aproximadamente 20% das escolas que tiveram os resultados divulgados com código identificador iniciado pelo dígito 6.

As respostas do questionário contextual da escola foram dicotomizadas para utilização do Modelo Logístico Unidimensional de Dois Parâmetros (ML2). Alternativas A (Bom) e B (Regular) foram recodificadas para 1 e C (Ruim) e D (Inexistente) foram recodificadas para 0, marcações em branco ou anuladas foram recodificadas para NA (não apresentadas)¹⁴.

No quadro a seguir, estão listados os itens referentes à infraestrutura escolar utilizados do questionário contextual da escola do Saeb 2017:

Quadro 4 – Variáveis utilizadas para montagem do banco de dados da escala de infraestrutura escolar

Questão	Descrição no Questionário Contextual da Escola - Saeb
	Questões referente a itens e equipamentos do prédio escolar
Q007	Telhado
Q008	Paredes
Q009	Piso
Q010	Entrada do Prédio
Q011	Pátio
Q012	Corredores
Q013	Salas de Aula
Q014	Portas
Q015	Janelas
Q016	Banheiros
Q017	Cozinha
Q018	Instalações Hidráulicas
Q019	Instalações Elétricas
	Questões referentes a recursos existentes nas escolas
Q037	Computadores para uso dos alunos
Q038	Acesso à internet para uso dos alunos
Q039	Computadores para uso dos professores
Q040	Acesso à internet para uso dos professores
Q041	Computadores exclusivamente para uso administrativo
Q042	Fitas de vídeo ou DVD (educativas)
Q043	Fitas de vídeo ou DVD (lazer)

¹⁴ As alternativas A (bom) e B (regular) foram agrupadas e recodificadas para 1 e as alternativas C (ruim) e D (inexistente) foram agrupadas e recodificadas para 0 porque alguns itens tratam de questões muito básicas, como telhado e parede, por exemplo, e, sendo o conceito de regular entendido, neste caso, como mais próximo da condição satisfatória.

Questão	Descrição no Questionário Contextual da Escola - Saeb
Q044	Máquina copiadora
Q045	Impressora
Q046	Retroprojektor
Q047	Projektor de slides/Datashow
Q048	Videocassete ou aparelho de DVD
Q049	Televisão
Q050	Mimeógrafo
Q051	Câmera fotográfica
Q052	Antena parabólica
Q053	Internet com conexão Banda Larga
Q054	Linha telefônica
Q055	Aparelho de fax
Q056	Aparelho de som
	Questões referentes a espaços da escola
Q057	Biblioteca
Q058	Sala de leitura
Q059	Quadra de esportes
Q060	Laboratório de informática
Q061	Laboratório de Ciências
Q062	Auditório
Q063	Sala para atividades de música
Q064	Sala para atividades de artes plásticas

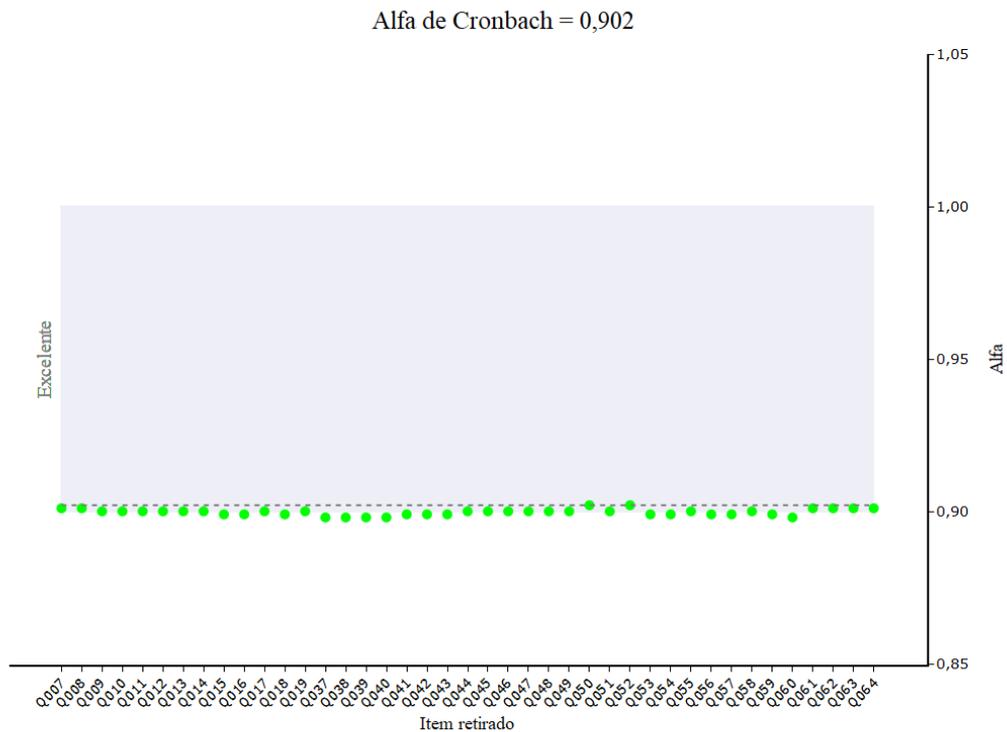
Fonte: Elaborado pela autora (2019).

Os dados foram analisados no *software* R Studio (R CORE TEAM, 2019), utilizando a sintaxe para a Teoria Clássica dos Testes (TCT) e a sintaxe da análise na Teoria da Resposta ao Item (TRI).

4.1.1 Análise via Teoria Clássica dos Testes

Na análise pela Teoria Clássica dos Testes (TCT), o Alfa de Cronbach é uma medida da consistência interna do conjunto dos itens que formam o teste. Seu valor varia entre 0 (zero) e 1 (um). Sendo que, quanto mais próximo de 1 for o seu valor, melhor. De acordo com Hair *et al.*, (2006, p. 100), valores de 0,60 a 0,70 são considerados o limite inferior de aceitabilidade.

Figura 3 – Alfa de Cronbach da escala de infraestrutura escolar via Teoria Clássica dos Testes



Fonte: Elaborado pela autora (2019).

O Alfa de Cronbach 0,902, apresentado na Figura 3, obtido na análise da TCT, indica uma boa medida de consistência interna do conjunto dos itens analisados.

A Tabela 1 e a Figura 4, apresentados na sequência, contém a frequência de escolas nas quais os itens de infraestrutura considerados neste estudo existem (Bom/Regular) ou não (Não Existem/Ruins), sendo 57.634 o total de escolas.

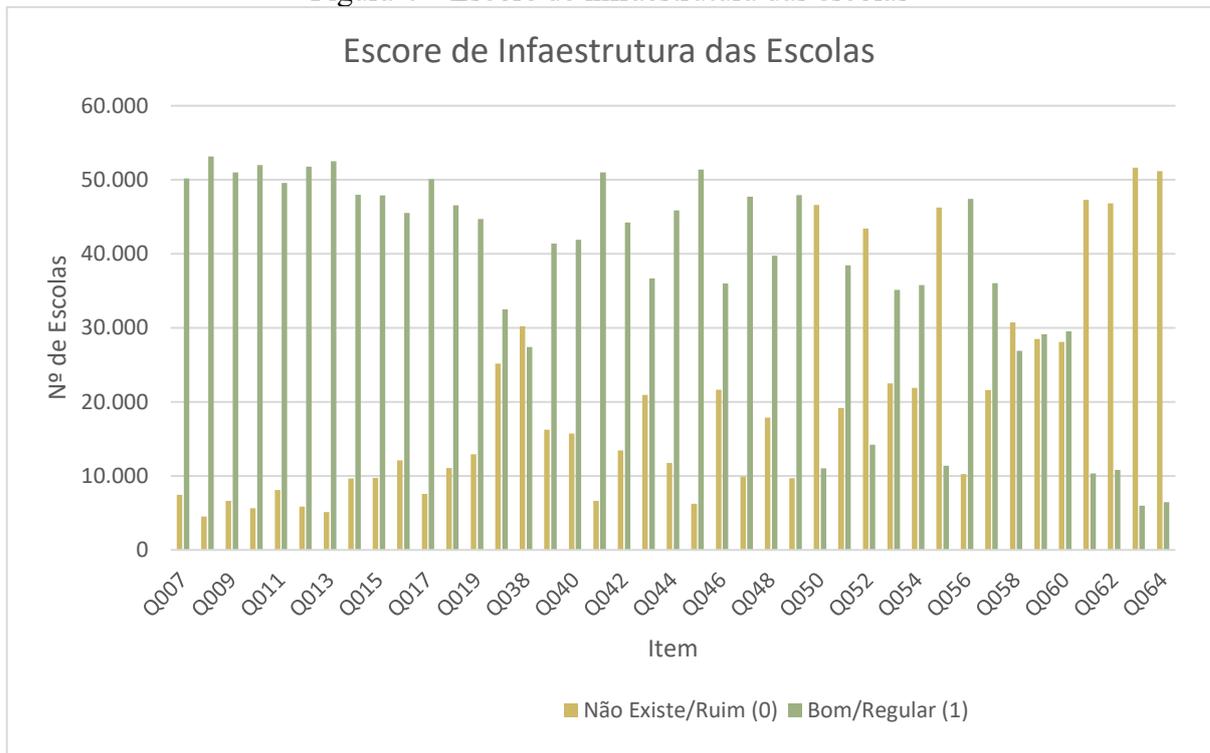
Tabela 1 - Frequência dos itens da escala de infraestrutura escolar

Item	Descrição	Não Existe/ Ruins (0)		Existente Bom/Regular (1)	
		Frequência	%	Frequência	%
Q007	Telhado	7.437	12,90%	50.197	87,10%
Q008	Paredes	4.502	7,81%	53.132	92,19%
Q009	Piso	6.650	11,54%	50.984	88,46%
Q010	Entrada do Prédio	5.643	9,79%	51.991	90,21%
Q011	Pátio	8.077	14,01%	49.557	85,99%
Q012	Corredores	5.843	10,14%	51.791	89,86%
Q013	Salas de Aula	5.112	8,87%	52.522	91,13%
Q014	Portas	9.662	16,76%	47.972	83,24%
Q015	Janelas	9.749	16,92%	47.885	83,08%
Q016	Banheiros	12.105	21,00%	45.529	79,00%
Q017	Cozinha	7.563	13,12%	50.071	86,88%

Item	Descrição	Não Existe/ Ruins (0)		Existente Bom/Regular (1)	
		Frequência	%	Frequência	%
Q018	Instalações Hidráulicas	11.063	19,20%	46.571	80,80%
Q019	Instalações Elétricas	12.918	22,41%	44.716	77,59%
Q037	Computadores para uso dos alunos	25.154	43,64%	32.480	56,36%
Q038	Acesso à internet para uso dos alunos	30.207	52,41%	27.427	47,59%
Q039	Computadores para uso dos professores	16.231	28,16%	41.403	71,84%
Q040	Acesso à internet para uso dos professores	15.740	27,31%	41.894	72,69%
Q041	Computadores exclusivamente para uso administrativo	6.621	11,49%	51.013	88,51%
Q042	Fitas de vídeo ou DVD (educativas)	13.422	23,29%	44.212	76,71%
Q043	Fitas de vídeo ou DVD (lazer)	20.949	36,35%	36.685	63,65%
Q044	Máquina copiadora	11.769	20,42%	45.865	79,58%
Q045	Impressora	6.242	10,83%	51.392	89,17%
Q046	Retroprojektor	21.645	37,56%	35.989	62,44%
Q047	Projektor de slides/Datashow	9.924	17,22%	47.710	82,78%
Q048	Videocassete ou aparelho de DVD	17.893	31,05%	39.741	68,95%
Q049	Televisão	9.701	16,83%	47.933	83,17%
Q050	Mimeógrafo	46.586	80,83%	11.048	19,17%
Q051	Câmera fotográfica	19.178	33,28%	38.456	66,72%
Q052	Antena parabólica	43.397	75,30%	14.237	24,70%
Q053	Internet com conexão Banda Larga	22.485	39,01%	35.149	60,99%
Q054	Linha telefônica	21.877	37,96%	35.757	62,04%
Q055	Aparelho de fax	46.274	80,29%	11.360	19,71%
Q056	Aparelho de som	10.234	17,76%	47.400	82,24%
Q057	Biblioteca	21.604	37,48%	36.030	62,52%
Q058	Sala de leitura	30.728	53,32%	26.906	46,68%
Q059	Quadra de esportes	28.492	49,44%	29.142	50,56%
Q060	Laboratório de informática	28.092	48,74%	29.542	51,26%
Q061	Laboratório de Ciências	47.290	82,05%	10.344	17,95%
Q062	Auditório	46.808	81,22%	10.826	18,78%
Q063	Sala para atividades de música	51.641	89,60%	5.993	10,40%
Q064	Sala para atividades de artes plásticas	51.161	88,77%	6.473	11,23%

Fonte: Elaborado pela autora (2019).

Figura 4 – Escore de infraestrutura das escolas



Fonte: Elaborado pela autora (2019).

4.1.2 Análise via Teoria da Resposta ao Item

De acordo com Junior e Szinvelski (2013, p. 01), “a TRI permite analisar os itens (conjuntamente e individualmente) por meio das Curvas Características e das Curvas de Informação de cada Item e do Instrumento, além de criar uma escala interpretável, onde os itens e os respondentes são posicionados”.

A Tabela 2 apresenta os parâmetros iniciais de discriminação (a) e de dificuldade (b) obtidos utilizando a Teoria da Resposta ao Item, na escala (0,1), ou seja, escala com média de valor zero e desvio padrão de valor um.

Tabela 2 - Descrição dos itens e valores estimados na escala de infraestrutura escolar

Item	Descrição	Parâmetro a discriminação	EP(a)	Parâmetro b dificuldade	EP(b)
Q007	Telhado	1,0138	0,0150	-2,2335	0,0274
Q008	Paredes	1,4800	0,0212	-2,1756	0,0219
Q009	Piso	1,4601	0,0187	-1,8320	0,0171
Q010	Entrada do Prédio	1,6362	0,0212	-1,8633	0,0165
Q011	Pátio	1,6188	0,0190	-1,5537	0,0133
Q012	Corredores	1,8384	0,0228	-1,7156	0,0139

Item	Descrição	Parâmetro <i>a</i> discriminação	EP(<i>a</i>)	Parâmetro <i>b</i> dificuldade	EP(<i>b</i>)
Q013	Salas de Aula	1,8995	0,0247	-1,8057	0,0147
Q014	Portas	1,4196	0,0166	-1,4834	0,0137
Q015	Janelas	1,5068	0,0172	-1,4262	0,0127
Q016	Banheiros	1,4784	0,0162	-1,2046	0,0111
Q017	Cozinha	1,3573	0,0172	-1,7890	0,0173
Q018	Instalações Hidráulicas	1,6009	0,0175	-1,2501	0,0109
Q019	Instalações Elétricas	1,2825	0,0145	-1,2433	0,0124
Q037	Computadores para uso dos alunos	1,8712	0,0175	-0,1923	0,0063
Q038	Acesso à internet para uso dos alunos	1,7875	0,0167	0,1046	0,0064
Q039	Computadores para uso dos professores	1,8693	0,0182	-0,7609	0,0076
Q040	Acesso à internet para uso dos professores	1,6821	0,0168	-0,8471	0,0083
Q041	Computadores exclusivamente para uso administrativo	1,9897	0,0234	-1,5541	0,0118
Q042	Fitas de vídeo ou DVD (educativas)	1,2474	0,0140	-1,1888	0,0122
Q043	Fitas de vídeo ou DVD (lazer)	1,2537	0,0129	-0,5632	0,0087
Q044	Máquina copiadora	1,1710	0,0139	-1,4459	0,0150
Q045	Impressora	1,7449	0,0214	-1,7203	0,0143
Q046	Retroprojektor	0,9096	0,0108	-0,6485	0,0114
Q047	Projektor de slides/Datashow	1,3111	0,0155	-1,5180	0,0146
Q048	Videocassete ou aparelho de DVD	1,0243	0,0118	-0,9102	0,0119
Q049	Televisão	1,2884	0,0154	-1,5658	0,0152
Q050	Mimeógrafo	0,5664	0,0108	2,7292	0,0494
Q051	Câmera fotográfica	1,0736	0,0119	-0,7593	0,0106
Q052	Antena parabólica	0,6445	0,0103	1,8816	0,0296
Q053	Internet com conexão Banda Larga	1,3455	0,0134	-0,4194	0,0079
Q054	Linha telefônica	1,4836	0,0143	-0,4155	0,0075
Q055	Aparelho de fax	1,2539	0,0145	1,4618	0,0144
Q056	Aparelho de som	1,5211	0,0169	-1,3556	0,0120
Q057	Biblioteca	1,2529	0,0128	-0,5093	0,0085
Q058	Sala de leitura	1,0471	0,0113	0,1646	0,0089
Q059	Quadra de esportes	1,3415	0,0131	0,0118	0,0075
Q060	Laboratório de informática	1,9338	0,0180	-0,0194	0,0062
Q061	Laboratório de Ciências	1,3446	0,0156	1,5079	0,0142
Q062	Auditório	1,1034	0,0135	1,6344	0,0173
Q063	Sala para atividades de música	1,4911	0,0192	1,9621	0,0179
Q064	Sala para atividades de artes plásticas	1,5894	0,0200	1,8391	0,0160

Fonte: Elaborado pela autora (2019).

Partiu-se do critério de exclusão de itens na análise da TRI parâmetro *a* menor que 0,55; parâmetro *b* fora do intervalo -4,5 e 4,5; itens com erro padrão alto. Sendo que o parâmetro *a*

indica a discriminação, o quanto o item separa bem unidades de ensino de baixa e de alta infraestrutura escolar. O parâmetro b indica a dificuldade, está relacionado ao percentual de respondentes que atendem ao padrão esperado para uma boa infraestrutura.

Após a primeira etapa, foi eliminado o item “Q050” mimeógrafo ($a = 0,5664$), apesar de apresentar o parâmetro a dentro do critério de exclusão, porque trata-se de um equipamento obsoleto e pouco frequente nas escolas.

A Tabela 3 apresenta a descrição dos itens e os valores estimados dos parâmetros de discriminação (a) e de dificuldade (b) dos itens após a exclusão do item Q050.

Tabela 3 – Descrição dos itens e valores estimados na escala de infraestrutura após exclusões

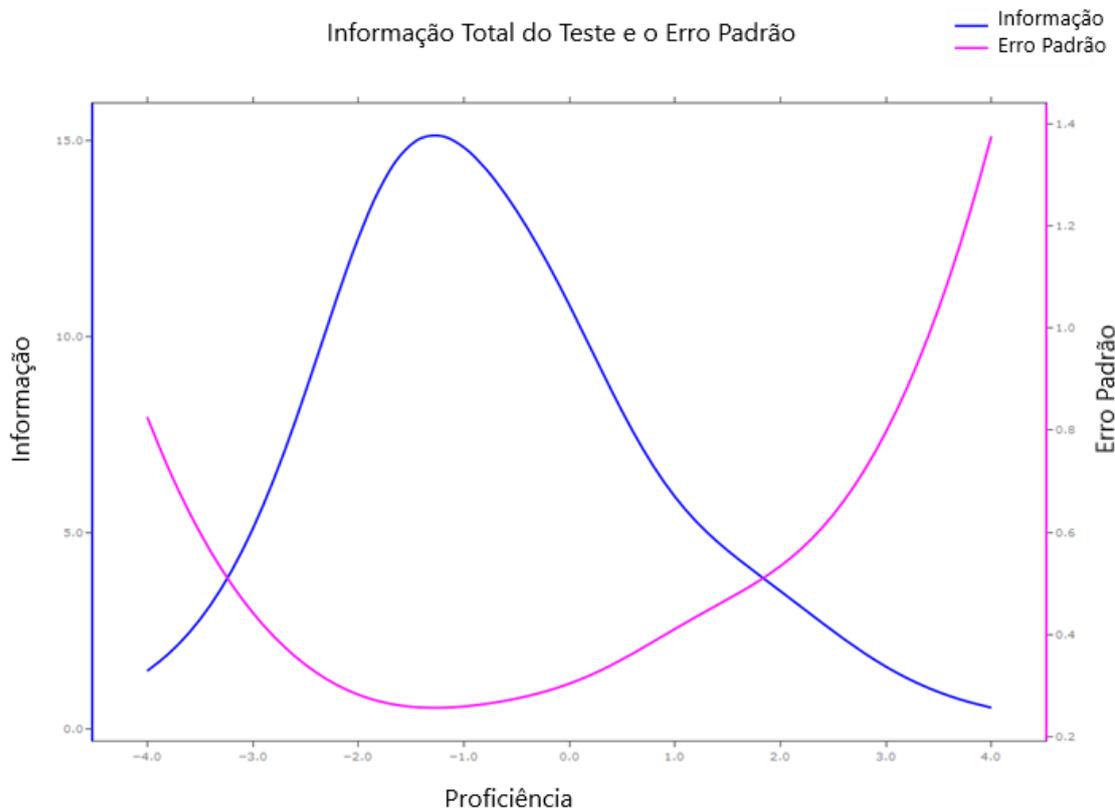
Item	Descrição	Parâmetro a discriminação	EP(a)	Parâmetro b dificuldade	EP(b)
Q007	Telhado	1,0131	0,0150	-2,2354	0,0274
Q008	Paredes	1,4797	0,0212	-2,1772	0,0219
Q009	Piso	1,4588	0,0187	-1,8340	0,0172
Q010	Entrada do Prédio	1,6351	0,0211	-1,8651	0,0165
Q011	Pátio	1,6191	0,0190	-1,5544	0,0133
Q012	Corredores	1,8404	0,0228	-1,7159	0,0139
Q013	Salas de Aula	1,8971	0,0247	-1,8080	0,0148
Q014	Portas	1,4165	0,0166	-1,4861	0,0137
Q015	Janelas	1,5034	0,0172	-1,4287	0,0127
Q016	Banheiros	1,4762	0,0161	-1,2061	0,0111
Q017	Cozinha	1,3535	0,0171	-1,7929	0,0173
Q018	Instalações Hidráulicas	1,5974	0,0175	-1,2521	0,0109
Q019	Instalações Elétricas	1,2795	0,0145	-1,2455	0,0125
Q037	Computadores para uso dos alunos	1,8748	0,0175	-0,1921	0,0063
Q038	Acesso à internet para uso dos alunos	1,7922	0,0168	0,1048	0,0064
Q039	Computadores para uso dos professores	1,8670	0,0182	-0,7616	0,0076
Q040	Acesso à internet para uso dos professores	1,6821	0,0168	-0,8474	0,0083
Q041	Computadores exclusivamente para uso administrativo	1,9897	0,0234	-1,5552	0,0118
Q042	Fitas de vídeo ou DVD (educativas)	1,2368	0,0139	-1,1956	0,0124
Q043	Fitas de vídeo ou DVD (lazer)	1,2421	0,0129	-0,5664	0,0088
Q044	Máquina copiadora	1,1669	0,0139	-1,4497	0,0151
Q045	Impressora	1,7388	0,0214	-1,7244	0,0143
Q046	Retroprojeter	0,8999	0,0108	-0,6539	0,0116
Q047	Projeter de slides/Datashow	1,3087	0,0154	-1,5203	0,0146
Q048	Videocassete ou aparelho de DVD	1,0103	0,0117	-0,9192	0,0121
Q049	Televisão	1,2777	0,0153	-1,5747	0,0154
Q051	Câmera fotográfica	1,0653	0,0119	-0,7634	0,0107
Q052	Antena parabólica	0,6351	0,0103	1,9052	0,0303
Q053	Internet com conexão Banda Larga	1,3448	0,0134	-0,4195	0,0080
Q054	Linha telefônica	1,4847	0,0143	-0,4154	0,0075
Q055	Aparelho de fax	1,2448	0,0144	1,4688	0,0145

Item	Descrição	Parâmetro <i>a</i> discriminação	EP(<i>a</i>)	Parâmetro <i>b</i> dificuldade	EP(<i>b</i>)
Q056	Aparelho de som	1,5131	0,0168	-1,3602	0,0120
Q057	Biblioteca	1,2557	0,0128	-0,5087	0,0085
Q058	Sala de leitura	1,0466	0,0113	0,1648	0,0089
Q059	Quadra de esportes	1,3455	0,0131	0,0119	0,0075
Q060	Laboratório de informática	1,9413	0,0181	-0,0192	0,0062
Q061	Laboratório de Ciências	1,3616	0,0157	1,4969	0,0140
Q062	Auditório	1,1128	0,0136	1,6251	0,0171
Q063	Sala para atividades de música	1,5006	0,0193	1,9551	0,0178
Q064	Sala para atividades de artes plásticas	1,6008	0,0201	1,8322	0,0159

Fonte: Elaborado pela autora (2019).

A Figura 5 mostra a informação do teste em função da “proficiência” da escola na escala de infraestrutura. Observa-se que a escala de infraestrutura escolar tem maior informação no intervalo de aproximadamente -3 a 1,2.

Figura 5 – Curva de informação do teste da escala de infraestrutura escolar



Fonte: Elaborado pela autora (2019).

Na Tabela 4 e na Figura 6, foram posicionados os itens, em uma escala que varia de 4 pontos negativos a 4 pontos positivos, com intervalo de 0,5 pontos, considerando que para o

modelo logístico de 2 parâmetros, o item será posicionado se a maior ou igual a 1,00 e $P(U=1|Z) \geq 0,50$.

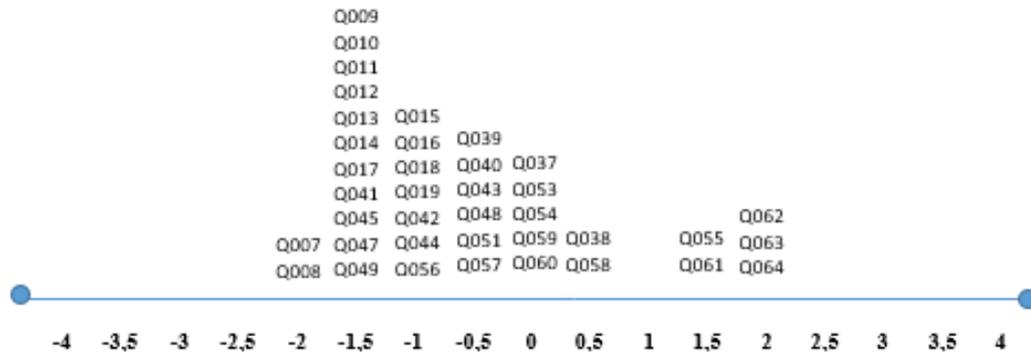
Obs.: os itens foram classificados pelo parâmetro b (quanto mais alto o valor de b , mais difícil é o item, ou seja, b alto significa que só escolas com alto grau de infraestrutura têm esse item).

Tabela 4 – Probabilidade de ter o atributo em função do nível da escala de infraestrutura escolar

Item	Parâmetros		Níveis da Escala (0,1)																		
	a	b	c	-4	-3,5	-3	-2,5	-2	-1,5	-1	-0,5	0	0,5	1	1,5	2	2,5	3	3,5	4	
Q007	1,0131	-2,2354		0,14	0,22	0,32	0,56	0,68	0,78	0,85	0,91	0,94	0,96	0,98	0,99	1,00	1,00	1,00	1,00	1,00	1,00
Q008	1,4797	-2,1772		0,06	0,12	0,23	0,38	0,57	0,73	0,85	0,92	0,96	0,98	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Q010	1,6351	-1,8651		0,03	0,06	0,14	0,26	0,45	0,64	0,80	0,90	0,95	0,98	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Q009	1,4588	-1,8340		0,04	0,08	0,15	0,27	0,44	0,62	0,77	0,88	0,94	0,97	0,98	0,99	1,00	1,00	1,00	1,00	1,00	1,00
Q013	1,8971	-1,8080		0,02	0,04	0,09	0,21	0,41	0,64	0,82	0,92	0,97	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Q017	1,3535	-1,7929		0,05	0,09	0,16	0,28	0,43	0,60	0,75	0,85	0,92	0,96	0,98	0,99	1,00	1,00	1,00	1,00	1,00	1,00
Q045	1,7388	-1,7244		0,02	0,04	0,10	0,21	0,38	0,60	0,78	0,89	0,95	0,98	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Q012	1,8404	-1,7159		0,01	0,04	0,09	0,19	0,37	0,60	0,79	0,90	0,96	0,98	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Q049	1,2777	-1,5747		0,04	0,08	0,14	0,23	0,37	0,52	0,68	0,80	0,88	0,93	0,96	0,98	0,99	1,00	1,00	1,00	1,00	1,00
Q041	1,9897	-1,5552		0,01	0,02	0,05	0,13	0,29	0,53	0,75	0,89	0,96	0,98	0,99	1,00	1,00	1,00	1,00	1,00	1,00	1,00
Q011	1,6191	-1,5544		0,02	0,04	0,09	0,18	0,33	0,52	0,71	0,85	0,93	0,97	0,98	0,99	1,00	1,00	1,00	1,00	1,00	1,00
Q047	1,3087	-1,5203		0,04	0,07	0,13	0,22	0,35	0,51	0,66	0,79	0,88	0,93	0,96	0,98	0,99	1,00	1,00	1,00	1,00	1,00
Q014	1,4165	-1,4861		0,03	0,05	0,10	0,19	0,33	0,50	0,67	0,80	0,89	0,94	0,97	0,99	1,00	1,00	1,00	1,00	1,00	1,00
Q044	1,1669	-1,4497		0,05	0,08	0,14	0,23	0,34	0,49	0,63	0,75	0,84	0,91	0,95	0,97	0,98	0,99	1,00	1,00	1,00	1,00
Q015	1,5034	-1,4287		0,02	0,04	0,09	0,17	0,30	0,47	0,66	0,80	0,90	0,95	0,97	0,99	1,00	1,00	1,00	1,00	1,00	1,00
Q056	1,5131	-1,3602		0,02	0,04	0,08	0,15	0,28	0,45	0,63	0,79	0,89	0,94	0,97	0,99	1,00	1,00	1,00	1,00	1,00	1,00
Q018	1,5974	-1,2521		0,01	0,03	0,06	0,12	0,23	0,40	0,60	0,77	0,88	0,94	0,97	0,99	1,00	1,00	1,00	1,00	1,00	1,00
Q019	1,2795	-1,2455		0,03	0,05	0,10	0,17	0,28	0,42	0,58	0,72	0,83	0,90	0,95	0,97	0,98	0,99	1,00	1,00	1,00	1,00
Q016	1,4762	-1,2051		0,02	0,03	0,07	0,13	0,24	0,39	0,58	0,74	0,86	0,93	0,96	0,98	0,99	1,00	1,00	1,00	1,00	1,00
Q042	1,2368	-1,1956		0,03	0,05	0,10	0,17	0,27	0,41	0,56	0,70	0,81	0,89	0,94	0,97	0,98	0,99	1,00	1,00	1,00	1,00
Q048	1,0103	-0,9192		0,04	0,07	0,11	0,17	0,25	0,36	0,48	0,60	0,72	0,81	0,87	0,92	0,95	0,97	0,98	0,99	0,99	0,99
Q040	1,6821	-0,8474		0,00	0,01	0,03	0,06	0,13	0,25	0,44	0,64	0,81	0,91	0,96	0,98	0,99	1,00	1,00	1,00	1,00	1,00
Q051	1,0653	-0,7634		0,03	0,05	0,08	0,14	0,21	0,31	0,44	0,57	0,69	0,79	0,87	0,92	0,95	0,97	0,98	0,99	0,99	0,99
Q039	1,8670	-0,7616		0,00	0,01	0,02	0,04	0,09	0,20	0,39	0,62	0,81	0,91	0,96	0,99	0,99	1,00	1,00	1,00	1,00	1,00
Q046	0,8999	-0,6539		0,05	0,07	0,11	0,16	0,23	0,32	0,42	0,53	0,64	0,74	0,82	0,87	0,92	0,94	0,96	0,98	0,98	0,99
Q043	1,2421	-0,5664		0,01	0,03	0,05	0,08	0,14	0,24	0,37	0,52	0,67	0,79	0,87	0,93	0,96	0,98	0,99	1,00	1,00	1,00
Q057	1,2557	-0,5087		0,01	0,02	0,04	0,08	0,13	0,22	0,35	0,50	0,65	0,78	0,87	0,93	0,96	0,98	0,99	0,99	1,00	1,00
Q053	1,3448	-0,4195		0,01	0,02	0,03	0,06	0,11	0,19	0,31	0,47	0,64	0,77	0,87	0,93	0,96	0,98	0,99	0,99	1,00	1,00
Q054	1,4847	-0,4154		0,00	0,01	0,02	0,04	0,09	0,17	0,30	0,47	0,65	0,80	0,89	0,94	0,97	0,99	0,99	1,00	1,00	1,00
Q037	1,8748	-0,1921		0,00	0,00	0,01	0,01	0,03	0,08	0,18	0,36	0,59	0,79	0,90	0,96	0,98	0,99	1,00	1,00	1,00	1,00
Q060	1,9413	-0,0192		0,00	0,00	0,00	0,01	0,02	0,05	0,13	0,28	0,51	0,73	0,88	0,95	0,98	0,99	1,00	1,00	1,00	1,00
Q059	1,3455	0,0119		0,00	0,01	0,02	0,03	0,06	0,12	0,20	0,33	0,50	0,66	0,79	0,88	0,94	0,97	0,98	0,99	1,00	1,00
Q038	1,7922	0,1048		0,00	0,00	0,00	0,01	0,02	0,05	0,12	0,25	0,45	0,67	0,83	0,92	0,97	0,99	0,99	1,00	1,00	1,00
Q058	1,0466	0,1648		0,01	0,02	0,04	0,06	0,09	0,15	0,23	0,33	0,46	0,59	0,71	0,80	0,87	0,92	0,95	0,97	0,98	0,98
Q055	1,2448	1,4688		0,00	0,00	0,00	0,01	0,01	0,02	0,04	0,08	0,14	0,23	0,36	0,51	0,66	0,78	0,87	0,93	0,96	0,96
Q061	1,3616	1,4969		0,00	0,00	0,00	0,00	0,01	0,02	0,03	0,06	0,12	0,20	0,34	0,50	0,66	0,80	0,89	0,94	0,97	0,97
Q062	1,1128	1,6251		0,00	0,00	0,01	0,01	0,02	0,03	0,05	0,09	0,14	0,22	0,33	0,47	0,60	0,73	0,82	0,89	0,93	0,93
Q064	1,6008	1,8322		0,00	0,00	0,00	0,00	0,00	0,00	0,01	0,02	0,05	0,11	0,21	0,37	0,57	0,74	0,87	0,94	0,97	0,97
Q052	0,6351	1,9052		0,02	0,03	0,04	0,06	0,08	0,10	0,14	0,18	0,23	0,29	0,36	0,44	0,52	0,59	0,67	0,73	0,79	0,79
Q063	1,5006	1,9551		0,00	0,00	0,00	0,00	0,00	0,01	0,01	0,02	0,05	0,10	0,19	0,34	0,52	0,69	0,83	0,91	0,96	0,96

Fonte: Elaborado pela autora (2019).

Figura 6 – Régua com os itens posicionados na escala de infraestrutura escolar



Fonte: Elaborado pela autora (2019).

É importante ressaltar que os níveis da escala de infraestrutura escolar tomaram como referência a escala elaborada pelos autores Neto, Jesus, Karino e Andrade (2013), porém, como o instrumento utilizado foi o questionário contextual da escola no Saeb e não as informações do Censo Escolar, os parâmetros obtidos e o posicionamento dos itens na escala acabou sendo diferente, mas, de certa forma, equivalente, em aspectos como a presença de laboratório de Ciências estar compreendido no nível avançado, o acesso à internet para os estudantes e laboratório de informática estarem no nível adequado, equipamentos de Videocassete ou DVD no nível básico e cozinha no nível elementar.

Em seguida, ainda pela Teoria da Resposta ao Item, foi gerado o relatório do escore das escolas na escala (0, 1) de infraestrutura e, depois, os escores foram transformados para a escala com média 50 e desvio padrão 10, para a observação da frequência por intervalo de proficiência, que segue descrito na Tabela 5:

Tabela 5 - Descrição dos níveis da escala de infraestrutura escolar e a porcentagem de escolas por intervalo de proficiência

Nível	Intervalo	Frequência %	Descrição dos níveis de infraestrutura escolar
Elementar	0 < 50	50,97%	Estão neste nível escolas que possuem somente aspectos de infraestrutura elementares para o funcionamento de uma escola em bom estado ou regular, tais como telhado, paredes, piso, entrada do prédio, pátio, corredores, salas de aula e portas. Além de terem cozinha, computadores exclusivamente para uso administrativo, impressora, projetor de slides e televisão.
Básica	50 < 60	33,13%	Além dos itens presentes no nível anterior, neste nível as escolas já possuem uma infraestrutura básica. Em geral, elas possuem em estado bom ou regular: janelas, banheiros, instalações hidráulicas e elétricas. Além de possuírem também em bom estado ou regular:

Nível	Intervalo	Frequência %	Descrição dos níveis de infraestrutura escolar
			computadores para uso dos professores, acesso à internet para uso dos professores, fitas de vídeo ou DVD (educativas e de lazer), máquina copiadora, aparelho de vídeo cassete ou DVD, câmera fotográfica, aparelho de som e biblioteca.
Adequada	60 < 70	13,33%	Além dos itens presentes nos níveis anteriores, as escolas deste nível, em geral, possuem uma infraestrutura mais adequada às condições de ensino aprendizagem. Essas escolas possuem em bom estado ou regular, por exemplo, computadores para uso dos alunos, acesso à internet para uso dos alunos, internet com conexão Banda Larga, linha telefônica, sala de leitura, quadra de esportes e laboratório de informática.
Avançada	> = 70	2,56%	As escolas neste nível, além dos itens presentes nos níveis anteriores, possuem uma infraestrutura escolar com a presença de laboratório de Ciências, auditório, sala para atividades de música e sala para atividades de artes plásticas.

Fonte: Elaborado pela autora (2019).

4.2 ESCALA DE NÍVEL SOCIOECONÔMICO

Para a construção do indicador de nível socioeconômico, nesta pesquisa, foram utilizados os dados dos questionários contextuais respondidos pelos estudantes do 9º ano do Ensino Fundamental no Saeb 2017, tomando como referência o Indicador de Nível Socioeconômico das Escolas de Educação Básica (Inse) do Ministério da Educação, Instituto de Estudos e Pesquisas Educacionais Anísio Teixeira e Diretoria de Avaliação da Educação Básica. Os microdados utilizados estão disponíveis para *download* no portal do INEP, em Dados Abertos, Microdados, Saeb (Aneb/Prova Brasil), 2017.

Inicialmente foi feita a organização da base de dados para a estimação dos parâmetros pelo método da Máxima Verossimilhança Marginal (MV), via Teoria da Resposta ao Item. Os dados foram rodados no *software* R Studio (R CORE TEAM, 2019; CHALMERS, 2012).

O score foi gerado na escala (0, 1), ou seja, com média 0 e desvio padrão 1 e, posteriormente deu-se a conversão para a escala (50, 10), com média 50 e desvio padrão 10.

Os critérios utilizados na estimação dos parâmetros foram:

- Parâmetro a (que mede a capacidade de discriminação do item) – eliminados itens com valor menor que 0,55;
- Parâmetro b (que mede o grau de dificuldade do item) - eliminados itens fora do intervalo -4,5 e 4,5;

Partindo-se destes critérios, não foram eliminados itens do banco de dados utilizado na construção da escala de nível socioeconômico.

O Indicador de Nível Socioeconômico das Escolas de Educação Básica (Inse) do Ministério da Educação, Instituto de Estudos e Pesquisas Educacionais Anísio Teixeira e Diretoria de Avaliação da Educação Básica, de acordo com a sua nota técnica, apresenta

informações sobre a escolaridade dos pais e a renda da família, porém, como as análises estatísticas realizadas com a matriz de correlações entre os itens apontaram a existência de um fator ou constructo latente predominante, tais informações foram agregadas em uma única medida de nível socioeconômico. [...] Assim, o nível socioeconômico é considerado um constructo latente, que sintetiza de maneira unidimensional informações sobre a escolaridade dos pais e sobre a renda familiar, e o Inse objetiva contextualizar o desempenho das escolas nas avaliações e exames realizados pelo Inep, bem como o seu esforço na realização do trabalho educativo, ao caracterizar, de modo geral, o padrão de vida de seu público, relacionados à respectiva posição na hierarquia social. (NOTA TÉCNICA, MEC/INEP/DAEB, 2015, p.05).

Para a construção deste indicador foi utilizada a Teoria da Resposta ao Item (TRI), técnica que possibilita a mensuração de traços latentes, ou seja, características que não podem ser medidas diretamente, mas inferidas a partir da observação de variáveis secundárias que estejam relacionadas a ela. As questões foram consideradas em categorias ordinais (Escala tipo Likert) para aplicação do modelo de resposta gradual (SAMEJIMA, 1969), onde, de acordo com Andrade, Tavares e Valle (2000, p. 19), “as categorias de resposta de um item podem ser ordenadas entre si”.

No quadro abaixo estão descritas as variáveis utilizadas inicialmente na composição do indicador, por possuírem informações relacionadas à renda e educação das famílias dos estudantes, e a recodificação feita para a construção da base de dados, necessária para ordenar as categorias de forma ordinal e crescente.

Quadro 5 – Variáveis utilizadas para a montagem do banco de dados da escala de nível socioeconômico

Questão	Descrição da Variável	Alternativas	Recodificação
Q005	Na sua casa tem televisão em cores?	A=não tem B=Sim, uma C=Sim, duas D=Sim, três E=Sim, quatro ou mais	A=1 B=2 C=3 D=4 E=4 Else=NA
Q006	Na sua casa tem aparelho de rádio?	A=não tem B=Sim, um C=Sim, dois D=Sim, três	A=1 B=2 C=3 D=4

Questão	Descrição da Variável	Alternativas	Recodificação
		E=Sim, quatro ou mais	E=4 Else=NA
Q007	Na sua casa tem videocassete e/ou DVD?	A=não tem B=Sim, um C=Sim, dois D=Sim, três E=Sim, quatro ou mais	A=1 B=2 C=3 D=4 E=4 Else=NA
Q008	Na sua casa tem geladeira?	A=não tem B=Sim, uma C=Sim, duas D=Sim, três E=Sim, quatro ou mais	A=1 B=2 C=3 D=4 E=4 Else=NA
Q009	Na sua casa tem freezer (parte da geladeira duplex)?	A=não tem B=Sim, um C=Sim, dois D=Sim, três E=Sim, quatro ou mais	A=1 B=2 C=2 D=2 E=2 Else=NA
Q010	Na sua casa tem freezer separado da geladeira?	A=não tem B=Sim, uma C=Sim, duas D=Sim, três E=Sim, quatro ou mais	A=1 B=2 C=2 D=2 E=2 Else=NA
Q011	Na sua casa tem máquina de lavar roupa (O tanquinho NÃO deve ser considerado)?	A=não tem B=Sim, uma C=Sim, duas D=Sim, três E=Sim, quatro ou mais	A=1 B=2 C=3 D=4 E=4 Else=NA
Q012	Na sua casa tem carro?	A=não tem B=Sim, um C=Sim, dois D=Sim, três E=Sim, quatro ou mais	A=1 B=2 C=3 D=4 E=4 Else=NA
Q013	Na sua casa tem computador?	A=não tem B=Sim, um C=Sim, dois D=Sim, três E=Sim, quatro ou mais	A=1 B=2 C=3 D=4 E=4 Else=NA
Q014	Na sua casa tem banheiro?	A=não tem B=Sim, um C=Sim, dois D=Sim, três E=Sim, quatro ou mais	A=1 B=2 C=3 D=4 E=5 Else=NA
Q015	Na sua casa tem quarto para dormir?	A=não tem B=Sim, um	A=1 B=2

Questão	Descrição da Variável	Alternativas	Recodificação
		C=Sim, dois D=Sim, três E=Sim, quatro ou mais	C=3 D=4 E=5 Else=NA
Q017	Em sua casa trabalha empregado(a) doméstico(a) pelo menos cinco dias por semana?	A=não B=Sim, um(a) empregado(a) C=Sim, dois(duas) empregados(as) D=Sim, três empregados(as) E=Sim, quatro ou mais empregados(as)	A=1 B=2 C=2 D=2 E=2 Else=NA
Q019	Até que série sua mãe, ou a mulher responsável por você, estudou?	A=nunca estudou B=não completou a 4ª série/5º ano do Ensino Fundamental C=completou a 4ª série/5º ano do Ensino Fundamental, mas não completou a 8ª série/9º ano do Ensino Fundamental D=completou a 8ª série/9º ano do Ensino fundamental, mas não completou o Ensino Médio E=completou o Ensino Médio, mas não completou a Faculdade F=completou a Faculdade G=não sei	A=1 B=2 C=2 D=3 E=4 F=5 G=NA Else=NA
Q023	Até que série seu pai, ou o homem responsável por você, estudou?	A=nunca estudou B=não completou a 4ª série/5º ano do Ensino Fundamental C=completou a 4ª série/5º ano do Ensino Fundamental, mas não completou a 8ª série/9º ano do Ensino Fundamental D=completou a 8ª série/9º ano do Ensino fundamental, mas não completou o Ensino Médio E=completou o Ensino Médio, mas não completou a Faculdade F=completou a Faculdade G=não sei	A=1 B=2 C=2 D=3 E=4 F=5 G=NA Else=NA

Fonte: Elaborado pela autora (2019).

As questões Q009 e Q010 foram agrupadas considerando se a família do estudante possui ou não freezer em casa.

Após a calibração dos itens via Teoria da Resposta ao Item, a medida individual de cada estudantes neste indicador foi convertida para uma escala contínua, com média igual a 50 e desvio padrão igual a 10 e, os resultados foram classificados em sete níveis, seguindo padronização similar a utilizada pelo INEP na construção do seu Inse. A classificação nos níveis foi feita a partir do posicionamento dos itens pelos parâmetros de discriminação, sendo “a \geq

1,00” e “ $P(U=1|\theta=Z) \geq 0,50$ ” (modelo logístico de 2 parâmetros), (ANDRADE, TAVARES, VALLE, 2000).

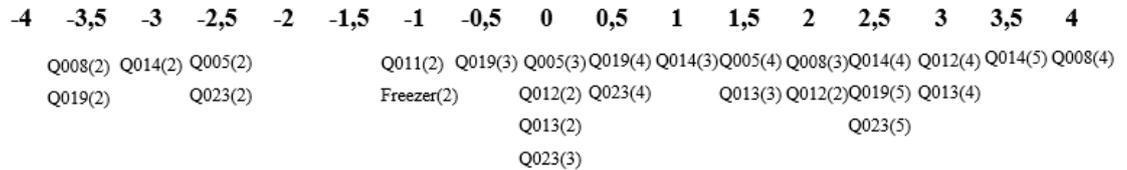
Os itens foram posicionados em uma escala que varia de 4 pontos negativos a 4 pontos positivos, com intervalo de 0,5 pontos, conforme tabela abaixo:

Tabela 6: Itens posicionados com intervalo de 0,5 pontos na escala de nível socioeconômico

Item	Categoria de resposta	Parâmetros			Níveis da Escala (0.1)																			
		a	b	c	-4	-3.5	-3	-2.5	-2	-1.5	-1	-0.5	0	0.5	1	1.5	2	2.5	3	3.5	4			
ITX RESP_Q015	2	0.8033	-6.52		-0.88	0.92	0.94	0.96	0.97	0.98	0.99	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ITX RESP_Q008	2	1.1431	-4.148		0.54	0.68	0.79	0.87	0.92	0.95	0.97	0.98	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ITX RESP_Q019	2	1.0585	-3.745		0.43	0.56	0.69	0.79	0.86	0.91	0.95	0.97	0.98	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ITX RESP_Q014	2	1.6619	-3.477		0.30	0.49	0.69	0.84	0.92	0.96	0.98	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ITX RESP_Q015	3	0.8033	-3.33		0.37	0.47	0.57	0.66	0.74	0.81	0.87	0.91	0.94	0.96	0.97	0.98	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00
ITX RESP_Q005	2	1.3296	-2.979		0.20	0.33	0.49	0.65	0.79	0.88	0.93	0.96	0.98	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ITX RESP_Q023	2	1.0847	-2.939		0.24	0.35	0.48	0.62	0.73	0.83	0.89	0.93	0.96	0.98	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ITX RESP_Q011	2	1.5155	-1.192		0.01	0.03	0.06	0.10	0.12	0.23	0.39	0.57	0.74	0.86	0.93	0.97	0.98	0.99	1.00	1.00	1.00	1.00	1.00	1.00
freezer	2	1.1434	-1.035		0.03	0.06	0.10	0.16	0.25	0.37	0.51	0.65	0.77	0.85	0.91	0.95	0.97	0.98	0.99	0.99	1.00	1.00	1.00	1.00
ITX RESP_Q019	3	1.0585	-0.722		0.03	0.05	0.08	0.13	0.21	0.30	0.43	0.56	0.68	0.78	0.86	0.91	0.95	0.97	0.98	0.99	0.99	1.00	1.00	1.00
ITX RESP_Q023	3	1.0847	-0.318		0.02	0.03	0.05	0.09	0.14	0.22	0.32	0.45	0.59	0.71	0.81	0.88	0.93	0.96	0.97	0.98	0.99	0.99	1.00	1.00
ITX RESP_Q013	2	1.7728	-0.185		0.00	0.00	0.01	0.02	0.04	0.09	0.19	0.36	0.58	0.77	0.89	0.95	0.98	0.99	1.00	1.00	1.00	1.00	1.00	1.00
ITX RESP_Q012	2	1.5734	-0.094		0.00	0.00	0.01	0.02	0.05	0.10	0.19	0.35	0.54	0.72	0.85	0.92	0.96	0.98	0.99	1.00	1.00	1.00	1.00	1.00
ITX RESP_Q005	3	1.3296	-0.087		0.01	0.01	0.02	0.04	0.07	0.13	0.23	0.37	0.53	0.69	0.81	0.89	0.94	0.97	0.98	0.99	1.00	1.00	1.00	1.00
ITX RESP_Q015	4	0.8033	-0.063		0.04	0.06	0.09	0.12	0.17	0.24	0.32	0.41	0.51	0.61	0.70	0.78	0.84	0.89	0.92	0.95	0.95	0.96	0.96	0.96
ITX RESP_Q019	4	1.0585	0.1454		0.01	0.02	0.03	0.06	0.09	0.15	0.23	0.34	0.46	0.59	0.71	0.81	0.88	0.92	0.95	0.97	0.98	0.98	0.98	0.98
ITX RESP_Q023	4	1.0847	0.4942		0.01	0.01	0.02	0.04	0.06	0.10	0.17	0.25	0.37	0.50	0.63	0.75	0.84	0.90	0.94	0.96	0.96	0.96	0.96	0.96
ITX RESP_Q014	3	1.6619	0.6		0.00	0.00	0.00	0.01	0.01	0.03	0.07	0.14	0.27	0.46	0.66	0.82	0.91	0.96	0.98	0.99	1.00	1.00	1.00	1.00
ITX RESP_Q013	3	1.7728	1.4615		0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.03	0.07	0.15	0.31	0.52	0.72	0.86	0.94	0.97	0.99	0.99	0.99	0.99
ITX RESP_Q005	4	1.3296	1.468		0.00	0.00	0.00	0.01	0.01	0.02	0.04	0.07	0.12	0.22	0.35	0.51	0.67	0.80	0.88	0.94	0.97	0.97	0.97	0.97
ITX RESP_Q012	3	1.5734	1.6313		0.00	0.00	0.00	0.00	0.00	0.01	0.02	0.03	0.07	0.14	0.27	0.45	0.64	0.80	0.90	0.95	0.95	0.95	0.95	0.95
ITX RESP_Q008	3	1.1431	1.8255		0.00	0.00	0.00	0.01	0.01	0.02	0.04	0.07	0.11	0.18	0.28	0.41	0.55	0.68	0.79	0.87	0.92	0.92	0.92	0.92
ITX RESP_Q014	4	1.6619	2.0731		0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.03	0.07	0.14	0.28	0.47	0.67	0.82	0.91	0.96	0.96	0.96	0.96
ITX RESP_Q019	5	1.0585	2.0816		0.00	0.00	0.00	0.01	0.01	0.02	0.04	0.06	0.10	0.16	0.24	0.35	0.48	0.61	0.73	0.82	0.88	0.88	0.88	0.88
ITX RESP_Q023	5	1.0847	2.3334		0.00	0.00	0.00	0.01	0.01	0.02	0.03	0.04	0.07	0.12	0.19	0.29	0.41	0.55	0.67	0.78	0.86	0.86	0.86	0.86
ITX RESP_Q013	4	1.7728	2.5162		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.06	0.14	0.29	0.49	0.70	0.85	0.93	0.93	0.93	0.93	0.93
ITX RESP_Q015	5	0.8033	2.6457		0.00	0.01	0.01	0.02	0.02	0.03	0.05	0.07	0.11	0.15	0.21	0.28	0.37	0.47	0.57	0.67	0.75	0.75	0.75	0.75
ITX RESP_Q012	4	1.5734	2.8085		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.03	0.05	0.11	0.22	0.38	0.57	0.75	0.87	0.87	0.87	
ITX RESP_Q014	5	1.6619	3.1352		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.03	0.06	0.13	0.26	0.44	0.65	0.81	0.81	0.81	
ITX RESP_Q008	4	1.1431	3.9699		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.03	0.06	0.10	0.16	0.25	0.37	0.51	0.51	0.51	

Fonte: Elaborada pela autora (2019).

Figura 7 – Régua com os itens posicionados na escala de nível socioeconômico



Fonte: Elaborado pela autora (2019).

Tabela 7: Posicionamento e descrição dos itens na escala de nível socioeconômico

Nível I	Nível II	Nível III	Nível IV	Nível V	Nível VI	Nível VII
Item	Item	Item	Item	Item	Item	Item
Q008 (2) Casa tem 1 geladeira	Q005 (2) Casa tem 1 TV	Q005 (3) Casa tem 2 TVs	Q014 (3) Casa tem 2 banheiros	Q005 (4) Casa tem 3 TVs	Q012 (4) Família tem 3 ou mais carros	Q008 (4) Casa tem 3 ou mais geladeiras
Q014 (2) Casa tem 1 banheiro	Q011 (2) Casa tem 1 máquina de lavar	Q012 (2) Família tem 1 carro	Q019 (4) Mãe, ou mulher responsável, completou o EM	Q008 (3) Casa tem 2 geladeiras	Q013 (4) Casa tem 3 computadores	Q014 (5) Casa tem 4 ou mais banheiros
Q019 (2) Mãe, ou mulher responsável, completou ou não a 4ª série/5º ano EF	Q023 (2) Pai, ou homem responsável, completou ou não a 4ª série/5º ano EF	Q013 (2) Casa tem 1 computador	Q023 (4) Pai, ou homem responsável, completou o EM	Q012 (3) Família tem 2 carros	Q014 (4) Casa tem 3 banheiros	
	Freezer (2) Casa tem 1 freezer	Q019 (3) Mãe, ou mulher responsável, completou a 8ª série/9º ano EF		Q013 (3) Casa tem 2 computadores	Q019 (5) Mãe, ou mulher responsável completou a Faculdade	
		Q023 (3) Pai, ou homem responsável, completou a 8ª série/9º ano EF			Q023 (5) Pai, ou homem responsável completou a faculdade	

Legenda: TV/TVs = Aparelhos de televisão; EF = Ensino Fundamental; EM = Ensino Médio

Fonte: Elaborada pela autora (2019).

Tabela 8 - Descrição dos níveis socioeconômicos dos estudantes por intervalo de proficiência

Nível	Intervalo	Frequência %	Descrição dos níveis socioeconômicos
Nível I “Mais baixo”	0 < 30	29.321 1,65%	Este é o menor nível da escala e os estudantes, de modo geral, indicaram que há em sua casa bens elementares, como uma geladeira, um banheiro e que a mãe ou mulher responsável completou ou não a 4ª série/5º ano do Ensino Fundamental, mas não completou a 8ª série/9º ano do Ensino Fundamental.
Nível II “Baixo”	30 < 40	256.929 14,43%	Neste, os estudantes, de modo geral, indicaram que, além dos itens descritos no nível anterior, há em sua casa bens elementares, como uma televisão em cores, máquina de lavar roupas, um freezer, e, que o pai ou homem responsável completou ou não a 4ª série/5º ano do Ensino Fundamental, mas não completou a 8ª série/9º ano do Ensino Fundamental.
Nível III “Médio baixo”	40 < 50	603.747 33,91%	Neste, os estudantes, de modo geral, indicaram que, além dos itens descritos nos níveis anteriores, há em sua casa bens elementares, como duas televisões em cores, um computador, a família possui um carro e, que a mãe ou mulher responsável e o pai ou homem responsável completaram a 8ª série/9º ano do Ensino fundamental, mas não completaram o Ensino Médio.
Nível IV “Médio”	50 < 60	613.116 34,44%	Já neste nível, os estudantes, de modo geral, indicaram que, além dos itens descritos nos níveis anteriores, há em sua casa bens elementares, como dois banheiros, e, que a mãe ou mulher responsável e o pai ou homem responsável completaram o Ensino Médio, mas não completaram a Faculdade.
Nível V “Médio alto”	60 < 70	231.819 13,02%	Neste, os estudantes, de modo geral, indicaram que, além dos itens descritos nos níveis anteriores, há em suas casas três televisões em cores, duas geladeiras, dois computadores e que a família possui dois carros.
Nível VI “Alto”	70 < 80	41.006 2,30%	Neste nível, os estudantes, de modo geral, indicaram que, além dos itens descritos nos níveis anteriores, há em sua casa um quantitativo alto de bens elementares, como três ou mais carros, três ou mais computadores, que a casa tem três banheiros e que a mãe ou mulher responsável e o pai ou homem responsável completaram a Faculdade.
Nível VII “Mais alto”	> = 80	4.259 0,24%	Este é o maior nível da escala e os estudantes, de modo geral, indicaram que há em sua casa, além dos itens descritos nos níveis anteriores, um quantitativo alto de bens elementares, como três ou mais geladeiras e que a casa tem quatro ou mais banheiros.

Fonte: Elaborado pela autora (2019).

A utilização da nomenclatura dos níveis socioeconômicos (mais baixo, baixo, médio baixo, médio, médio alto, alto e mais alto) baseou-se no portal QEdU Academia.

Para criação da medida de nível socioeconômico das escolas, foi feito o agrupamento por escola da média do nível socioeconômico dos estudantes do 9º ano do Ensino Fundamental construído a partir do questionário contextual do Saeb 2017.

5 AVALIAÇÃO DO EFEITO DOS FATORES ASSOCIADOS POR REGRESSÃO MULTINÍVEL

Neste capítulo são apresentados os modelos de regressão multinível obtidos por meio do *software* R (R CORE TEAM, 2019), pacotes Rcmdr (FOX; BOUCHET-VALAT, 2019) e lme4 (BATES; MAECHLER; BOLKER; WALKER, 2015), utilizando-se a base de dados com informações do Saeb 2017 e também dos indicadores educacionais descritos no capítulo 4.

5.1 DADOS

Os dados utilizados na pesquisa estão disponíveis no portal do INEP, em microdados 2017. De acordo com informações do portal do INEP (2018b),

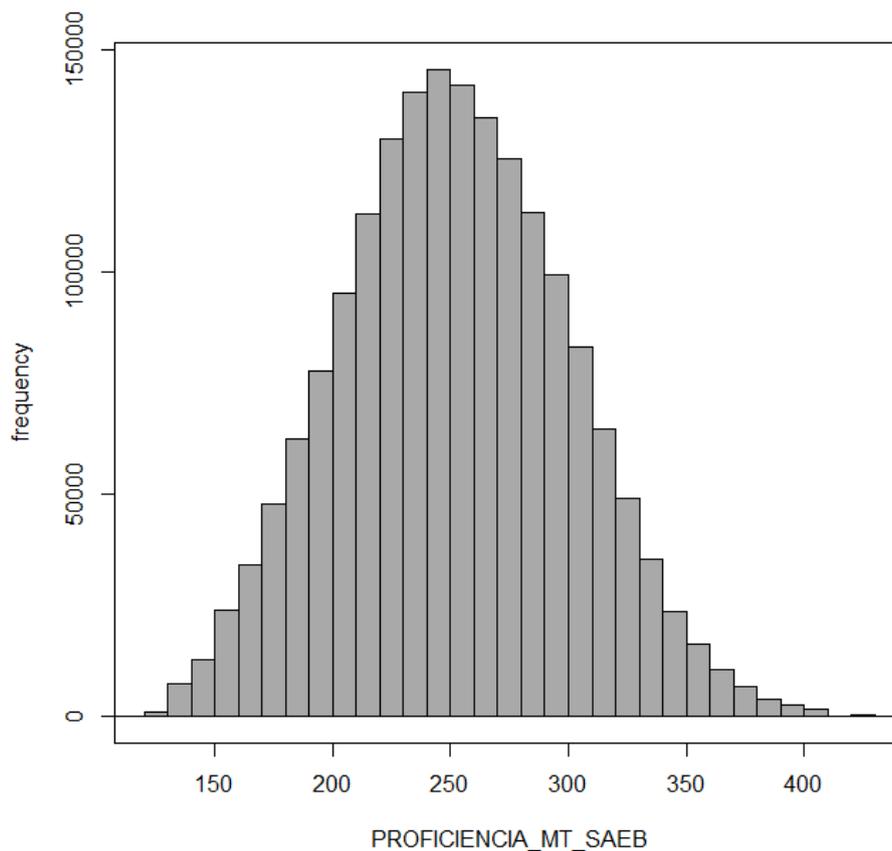
O Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) disponibilizou o Boletim das Escolas de todas as instituições de ensino que atenderam aos critérios para divulgação dos resultados, como previsto na Portaria nº 447/2017: ter tido pelo menos dez alunos presentes na aplicação dos instrumentos e ter alcançado a taxa de participação de, pelo menos, 80% dos alunos matriculados, conforme dados declarados pela escola ao Censo Escolar 2017. [...] As escolas que não cumpriram os requisitos da Portaria nº 447/2017 não têm os resultados divulgados.

À medida que foram sendo reunidas as informações dos questionários contextuais com os indicadores educacionais disponibilizados pelo INEP, observou-se uma grande ocorrência de dados perdidos/*missings*, mesmo utilizando-se como critério para junção das planilhas o código identificador da escola (ID_ESCOLA). Este fato ocorreu devido ao critério de participação na avaliação do Saeb de, no mínimo, 80% dos estudantes matriculados, muitas escolas não tiveram os seus dados divulgados, bem como, não tiveram calculados e divulgados o seu IDEB, nestas situações. O INEP divulgou as informações referentes às escolas com a utilização de máscaras, ou seja, o código identificador da escola verdadeiro foi substituído por um código fictício, iniciado pelo número 6. Foi o caso, no município de Florianópolis/SC, por exemplo, das escolas como o Instituto Estadual de Educação (IEE) - maior escola da rede pública estadual, e da Escola Básica Municipal Professora Herondina Medeiros Zeferino – maior escola de ensino fundamental da rede pública municipal. A partir desta constatação, optou-se por fazer análises apenas com informações obtidas diretamente dos instrumentos de coleta de dados do Saeb 2017.

Para a realização das regressões, foi utilizada a base de dados (microdados), inicialmente com 2.341.459 estudantes do 9º ano do Ensino Fundamental distribuídos em 37.837 escolas brasileiras no Saeb 2017. A proficiência em Matemática dos estudantes será utilizada como a **variável dependente**.

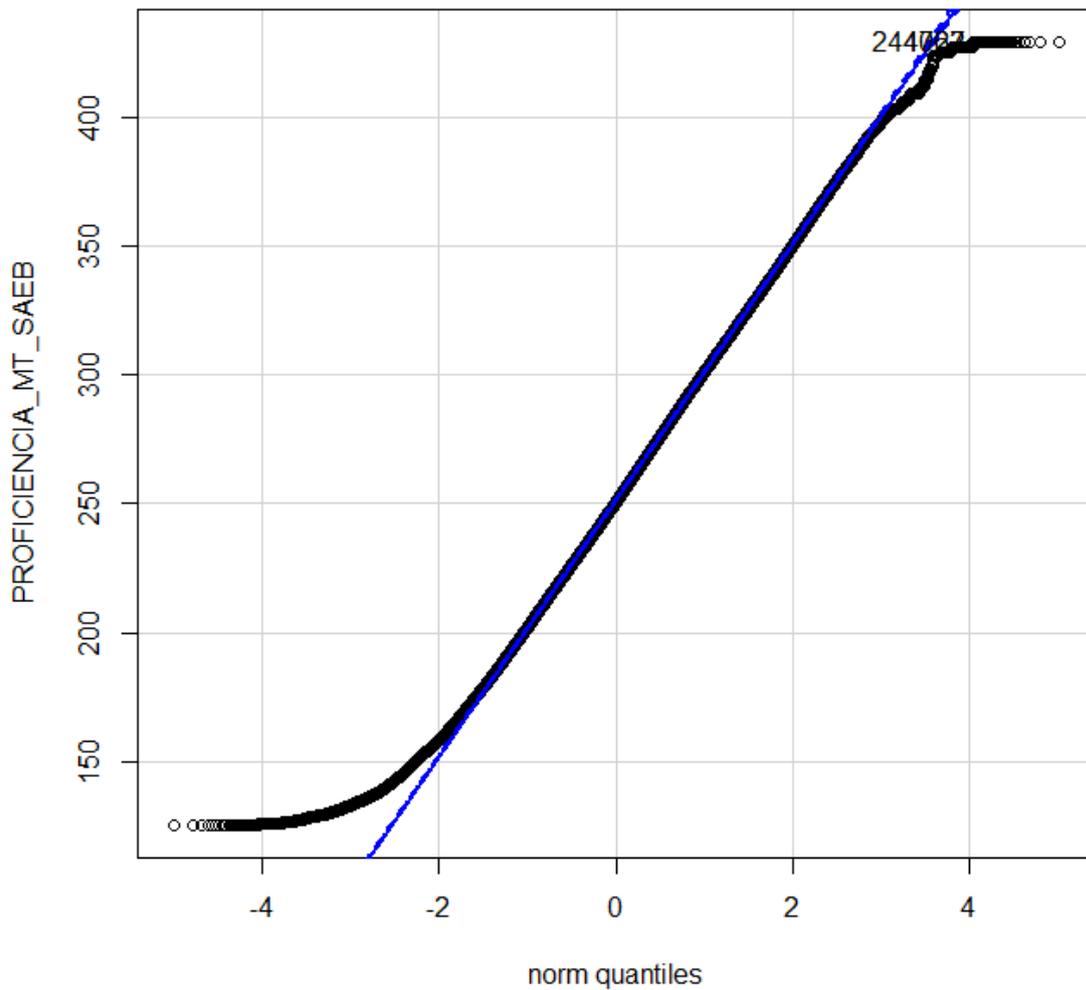
A Figura 8 apresenta o histograma da proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental no Saeb 2017 e sinaliza que a distribuição dos dados é aproximadamente simétrica. O gráfico qqplot da variável dependente (Figura 9) sugere uma distribuição aproximadamente normal dos dados, embora o gráfico mostre desvio da normalidade nas caudas da distribuição.

Figura 8 - Histograma da variável dependente proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental no Saeb 2017



Fonte: Elaborada pela autora (2020).

Figura 9 - Gráfico qqplot da variável dependente proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental no Saeb 2017



Fonte: Elaborada pela autora (2020).

Como variáveis independentes no nível 1 (aluno), foram utilizadas as variáveis que seguem descritas no Quadro 6, tendo sido estas variáveis escolhidas a partir da revisão de literatura, conforme descrito na seção 3.2, e passíveis de serem obtidas/extraídas nos questionários contextuais do Saeb 2017:

Quadro 6 - Descrição das variáveis relacionadas aos estudantes utilizadas na regressão multinível

Variáveis Nível 1	Descrição e categorias	Fonte da variável
nse	nível socioeconômico do estudante, escala transformada com média 50 e desvio padrão 10	Variável construída a partir do questionário contextual via TRI
sexo	gênero do estudante A=Masculino; B=Feminino, recodificado para 1, 0 (1 = masculino; 0 = feminino)	Questionário do Aluno – 9º Ano / questão 01
cor	raça/cor autodeclarada pelo estudante A=Branca; B=Preta; C=Parda; D=Amarelo; E=Indígena; F=Não quero declarar, recodificado para 1, 0 (1 = branca; 2 = Preta; 2 = Parda; 4 = Amarelo; 5 = Indígena; 0 = Não quero declarar)	Questionário do Aluno – 9º Ano / questão 02
reprov	Se o estudante já teve alguma reprovação A=Não; B=Sim, uma vez; C=Sim, duas vezes ou mais, recodificado para 1, 2, 0 (1 = Sim, uma vez; 2 = Sim, duas vezes ou mais; 0 = Não)	Questionário do Aluno – 9º Ano / questão 48
trab_fora	Se o estudante trabalha fora de casa A=Sim; B=Não, recodificado para 1, 0 (1 = Sim; 0 = Não)	Questionário do Aluno – 9º Ano / questão 45
part_pais_incentivo_estudos	Se os pais ou responsáveis incentivam a estudar A=Sim; B=Não, recodificado para 1, 0 (1 = Sim; 0 = Não)	Questionário do Aluno – 9º Ano / questão 27
gosta_mat	Se o estudante gosta de estudar Matemática A=Sim; B=Não, recodificado para 1, 0 (1 = Sim; 0 = Não)	Questionário do Aluno – 9º Ano / questão 53
fez_ed_infantil	Quando o estudante entrou na escola A=Na creche (0 a 3 anos); B=Na pré-escola (4 a 5 anos); C=Na 1ª série ou 1º ano (6 a 7 anos); D=Depois da 1ª série, recodificado para 0, 1, 2 (1 = Sim, desde 0 a 3 anos/creche; 2 = Sim, a partir de 4 a 5 anos/pré-escola; 0 = Não)	Questionário do Aluno – 9º Ano / questão 46

Fonte: Elaborada pela autora (2020).

A Tabela 9 contém as estatísticas descritivas das variáveis independentes relacionadas aos estudantes (nível 1).

Tabela 9 - Estatísticas descritivas das características/variáveis dos estudantes do 9º ano do Ensino Fundamental – Saeb 2017

Nível socioeconômico dos estudantes			Frequência	
Nível 1		Mais baixo ¹⁵	29.321	1,65%
Nível 2		Baixo	256.929	14,43%
Nível 3		Médio baixo	603.747	33,91%
Nível 4		Médio	613.116	34,44%
Nível 5		Médio alto	231.819	13,02%
Nível 6		Alto	41.006	2,30%
Nível 7		Mais alto ¹⁶	4.259	0,24%
TOTAL			1.780.197	
Sexo			Frequência	
A	1	Masculino	847.220	36,18%
B	0	Feminino	886.930	37,88%
NA	NA	Não respondeu	607.309	25,94%
TOTAL			2.341.459	
Cor/Raça			Frequência	
A	1	Branca	506.309	21,62%
B	2	Preta	206.216	8,81%
C	3	Parda	807.175	34,47%
D	4	Amarelo ¹⁷	62.894	2,69%
E	5	Indígena	42.875	1,83%
F	0	Não quero declarar	122.403	5,23%
NA	NA	Não respondeu	593.587	25,35%
TOTAL			2.341.459	
Reprovação			Frequência	
A	0	Não	1.223.078	52,24%
B	1	Sim, uma vez	363.806	15,54%
C	2	Sim, duas vezes ou mais	151.872	6,49%
NA	NA	Não respondeu	602.703	25,74%
TOTAL			2.341.459	
Trabalha fora			Frequência	
A	1	Sim	235.991	10,08%
B	0	Não	1.508.619	64,43%
NA	NA	Não respondeu	596.849	25,49%
TOTAL			2.341.459	
Participação dos pais/ incentivo nos estudos			Frequência	
A	1	Sim	1.731.562	73,95%
B	0	Não	29.485	1,26%
NA	NA	Não respondeu	580.412	24,79%
TOTAL			2.341.459	
Gosta de estudar Matemática			Frequência	
A	1	Sim	1.113.636	47,56%
B	0	Não	596.021	25,46%
NA	NA	Não respondeu	631.802	26,98%
TOTAL			2.341.459	
Fez Educação Infantil			Frequência	
A	1	Sim, creche (0 a 3 anos)	726.765	31,04%

¹⁵ O nível 1 (Mais baixo) será agrupado com o nível 2 (Baixo) devido ao reduzido número de ocorrências.

¹⁶ O nível 7 (Mais alto) será agrupado com o nível 6 (Alto), devido ao reduzido número de ocorrências.

¹⁷ A categoria cor/raça Amarelo será agrupada com a categoria cor/raça Indígena.

B	2	Sim, pré-escola (4 a 5 anos)	714.350	30,51%
C	0	Não (entrou na escola no 1º ano)	244.944	10,46%
D	0	Não (entrou na escola depois do 1º ano)	49.392	2,11%
NA	NA	Não respondeu	606.008	25,88%
TOTAL			2.341.459	

Fonte: Elaborada pela autora (2020).

Nesta primeira análise com estatísticas descritivas, observou-se que em torno de 25% dos estudantes não responderam o questionário contextual (NA). Por isto, foi feita a exclusão destes casos para a aplicação dos modelos de regressão multinível, a fim de não prejudicar a análise, e a base de dados passou a ser composta por 1.796.628 estudantes do 9º ano do Ensino Fundamental participantes do Saeb 2017.

Além disto, também foram agrupados os níveis 1 e 2 e os níveis 6 e 7 da medida de nível socioeconômico dos estudantes, por apresentarem um percentual reduzido de indivíduos nos níveis mais baixos (1,65%) e mais alto (0,24%). E, foram agrupadas as categorias de cor/raça Amarelo e Indígena.

A Tabela 10 representa as estatísticas descritivas das variáveis independentes relacionadas aos estudantes (nível 1), após a exclusão dos casos onde o estudante não respondeu ao questionário contextual e o agrupamento da medida de nível socioeconômico dos estudantes e das categorias de cor/raça amarela e indígena.

Tabela 10 - Estatísticas descritivas das características/variáveis dos estudantes do 9º ano do Ensino Fundamental – Saeb 2017, após exclusões

Nível socioeconômico dos estudantes			Frequência	
Nível 2	Baixo		286.250	16,08%
Nível 3	Médio baixo		603.747	33,91%
Nível 4	Médio		613.116	34,44%
Nível 5	Médio alto		231.819	13,02%
Nível 6	Alto		45.265	2,54%
TOTAL			1.780.197	
Sexo			Frequência	
A	1	Masculino	847.220	47,16%
B	0	Feminino	886.930	49,37%
NA	2	Não declarado	62.478	3,48%
TOTAL			1.796.628	
Cor/Raça			Frequência	
A	1	Branca	506.309	28,18%
B	2	Preta	206.216	11,48%
C	3	Parda	807.175	44,93%
D	4	Indígena/Amarelo	105.769	5,89%
F	0	Não quero declarar	122.403	6,81%
NA	NA	Não respondeu	48.756	2,71%
TOTAL			1.796.628	

Reprovação			Frequência	
A	0	Não	1.223.078	68,08%
B	1	Sim, uma vez	363.806	20,25%
C	2	Sim, duas vezes ou mais	151.872	8,45%
NA	NA	Não respondeu	57.872	3,22%
TOTAL			1.796.628	
Trabalha fora			Frequência	
A	1	Sim	235.991	13,14%
B	0	Não	1.508.619	83,97%
NA	NA	Não respondeu	52.018	2,90%
TOTAL			1.796.628	
Participação dos pais/ incentivo nos estudos			Frequência	
A	1	Sim	1.731.562	96,38%
B	0	Não	29.485	1,64%
NA	NA	Não respondeu	35.581	1,98%
TOTAL			1.796.628	
Gosta de estudar Matemática			Frequência	
A	1	Sim	1.113.636	61,98%
B	0	Não	596.021	33,17%
NA	NA	Não respondeu	86.971	4,84%
TOTAL			1.796.628	
Fez Educação Infantil			Frequência	
A	1	Sim, creche (0 a 3 anos)	726.765	40,45%
B	2	Sim, pré-escola (4 a 5 anos)	714.350	39,76%
C	0	Não (entrou na escola no 1º ano)	244.944	13,63%
D	0	Não (entrou na escola depois do 1º ano)	49.392	2,75%
NA	NA	Não respondeu	61.177	3,41%
TOTAL			1.796.628	

Fonte: Elaborada pela autora (2020).

Na tabela acima, é possível observar que os níveis socioeconômicos com maior concentração de estudantes do 9º ano do Ensino Fundamental que participaram do Saeb 2017 são o nível 3 (médio baixo) e o nível 4 (médio), que há uma distribuição próxima entre o número de estudantes do sexo masculino (47,16%) e estudantes do sexo feminino (49,37%), que os estudantes que se autodeclararam pardos são a maioria nesta população considerada, representando quase 50% do total, que a maioria dos estudantes não tem histórico de reprovação escolar (68,08%) e também não trabalha fora de casa (83,97%). No questionário contextual, quando perguntado se os pais ou responsáveis o incentivam a estudar, 96,38% dos estudantes responderam que sim. Quando perguntado se gosta de estudar Matemática, 61,98% dos estudantes responderam que sim e, na questão sobre quando entrou na escola, aproximadamente 80% dos estudantes respondeu ter feito creche ou pré-escola, ou seja, fez Educação Infantil.

No Quadro 7 são apresentadas as **variáveis independentes no nível 2 (escola)**. Neste trabalho foram utilizadas as variáveis: a localização da escola, se é urbana ou rural, a dependência administrativa da escola, se é federal, estadual, municipal ou privada, se a escola

está localizada na capital do estado ou no interior, o turno em que o estudante frequenta a escola, se é matutino, vespertino ou noturno, a medida por escola de estudantes que declaram fazer as atividades/tarefas de casa de matemática, a medida agrupada por escola de professor que corrige as atividades/tarefas de casa de matemática, a medida agrupada por escola de estudantes com reprovação, a média do nível socioeconômico dos estudantes da escola e a medida/escala das condições de infraestrutura das escolas, sendo as duas últimas construídas a partir da utilização da Teoria da Resposta ao Item.

Quadro 7 - Descrição das variáveis relacionadas às escolas utilizadas na regressão multinível

Variáveis Nível 2	Descrição e categorias	Fonte da variável
ID_LOCALIZACAO	1 = urbana; 2 = rural	Questionário do Aluno
ID_DEPEND_ADM	1 = Federal; 2 = Estadual; 3 = Municipal; 4 = Privada	Questionário do Aluno
ID_AREA	1 = Capital; 2 = Interior ou não capital	Questionário do Aluno
ID_TURNO	1 = Matutino; 2 = Vespertino; 3 = Noturno	Questionário do Aluno
faz_tarefa_mtm_prop_escola	<p>A proporção de estudantes por escola que declaram fazer as tarefas de casa de Matemática.</p> <p>A=Sempre ou quase sempre; B=De vez em quando; C=Nunca ou quase nunca; D=O(A) professor(a) não passa dever de casa, recodificado para 0, 1</p> <p>(1 = Sempre ou quase sempre; 1 = De vez em quando; 0 = Nunca ou quase nunca; 0 = Professor não passa tarefa)</p> <p>variável “faz_tarefa_mtm” foi agregada por escola com a função aggregate:</p> <pre>faz_tarefa_mtm_agregada<- aggregate(faz_tarefa_mtm ~ID_ESCOLA,data=base7_saeb_2017,FUN= mean)</pre>	<p>Questionário do Aluno – 9º Ano / questão 54</p> <p>Função aggregate da questão 54</p>
prof_corrige_tarefa_mtm_prop_escola	<p>A proporção de estudantes por escola que declaram que o professor corrige a tarefa de casa de Matemática.</p> <p>A=Sempre ou quase sempre; B=De vez em quando; C=Nunca ou quase nunca; D=O(A) professor(a) não passa dever de casa, recodificado para 0, 1</p> <p>(1 = Sempre ou quase sempre; 1 = De vez em quando; 0 = Nunca ou quase nunca; 0 = Professor não passa tarefa)</p> <p>variável “prof_corrige_tarefa_mtm” foi agregada por escola com a função aggregate:</p> <pre>prof_corrige_tarefa_mtm_agregada<- aggregate(prof_corrige_tarefa_mtm</pre>	<p>Questionário do Aluno – 9º Ano / questão 55</p> <p>Função aggregate da questão 55</p>

Variáveis Nível 2	Descrição e categorias	Fonte da variável
	~ID_ESCOLA,data=base8_saeб_2017,FUN=mean)	
reprov.prop_escola	A proporção da escola de estudantes com reprovação Obs.: TX_RESP_Q048 foi recodificada para 0, 1 ("A"="0"; "B"="1"; "C"="1"; else="NA") variável “reprov.x” foi agregada por escola com a função aggregate: reprov_agregada<- aggregate(reprov.x~ID_ESCOLA,data=base5_saeб_2017,FUN=mean)	Questionário do Aluno – 9º Ano / questão 48 Função aggregate da variável reprov
NSE_ESCOLA	A média da medida de nível socioeconômico dos estudantes da escola	Função aggregate da variável nse_50_10, recodificação para os níveis da escala e conversão para fator dos níveis agrupados por escola
infra_escolar	Medida da infraestrutura escolar, escala transformada com média 50 e desvio padrão 10	Variável construída a partir do questionário contextual via TRI

Fonte: Elaborada pela autora (2020).

A Tabela 11 apresenta as estatísticas descritivas das variáveis de nível 2 (escola), exceto das variáveis que foram agregadas por escola, por ser a média por escola considerando a medida/escala 0, 1 que são: a proporção de estudantes por escola que declaram fazer as tarefas de casa de Matemática (faz_tarefa_mtm_prop_escola), a proporção de estudantes por escola que declaram que o professor corrige a tarefa de casa de Matemática (prof_corrige_tarefa_mtm_prop_escola) e a proporção da escola de estudantes com reprovação (reprov.prop_escola).

Tabela 11 - Estatísticas descritivas das características/variáveis das escolas de Ensino Fundamental – Saeb 2017

Localização		Frequência	
1	Urbana	1.593.361	88,69%
2	Rural	203.267	11,31%
	TOTAL	1.796.628	
Dependência Administrativa		Frequência	
1	Federal	3.418	0,19%
2	Estadual	979.684	54,53%
3	Municipal	792.700	44,12%
4	Privada	20.826	1,16%
	TOTAL	1.796.628	

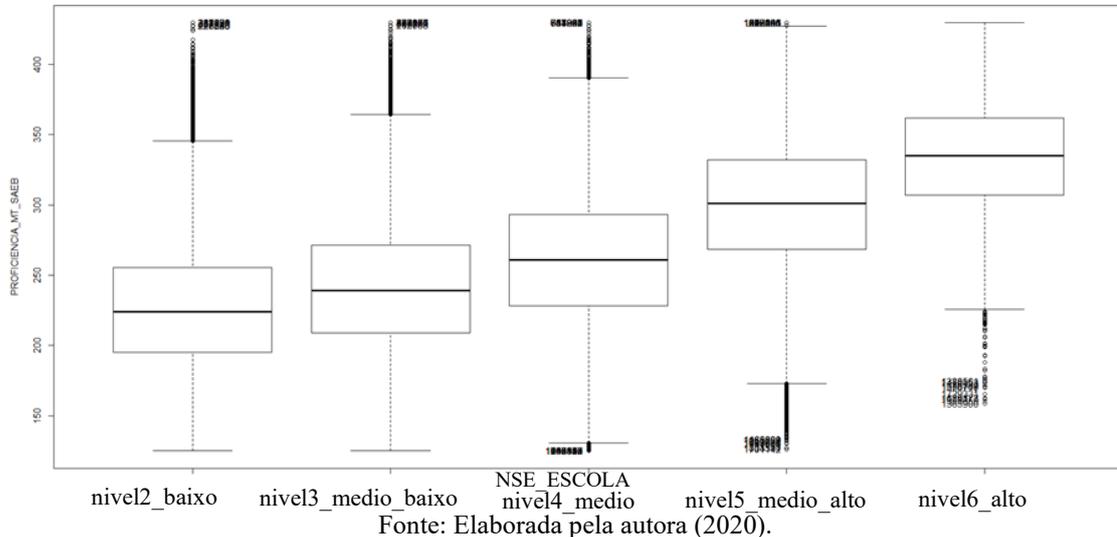
Área		Frequência	
1	Capital	317.959	17,70%
2	Interior	1.478.669	82,30%
	TOTAL	1.796.628	
Turno		Frequência	
1	Matutino	1.079.315	60,07%
2	Vespertino	703.246	39,14%
3	Noturno	14.067	0,78%
	TOTAL	1.796.628	
Nível socioeconômico dos estudantes agrupados por escola		Frequência	
Nível 2	Baixo	3.734	9,87%
Nível 3	Médio baixo	16.875	44,60%
Nível 4	Médio	16.333	43,17%
Nível 5	Médio alto	839	2,22%
Nível 6	Alto ¹⁸	57	0,15%
	TOTAL	37.838	
Nível infraestrutura escolar		Frequência	
Nível 1	Elementar	37.069	50,97%
Nível 2	Básica	24.096	33,13%
Nível 3	Adequada	9.696	13,33%
Nível 4	Avançada	1.861	2,56%
	TOTAL	72.722	

Fonte: Elaborada pela autora (2020).

Ao fazer a representação da distribuição da variável proficiência em Matemática dos estudantes concluintes do Ensino Fundamental no Saeb 2017 pela média por escola da medida de nível socioeconômico dos estudantes, observa-se, graficamente, a tendência da relação das mesmas, onde os estudantes que atingiram desempenhos mais elevados estão concentrados nas escolas com média de nível socioeconômico mais alto.

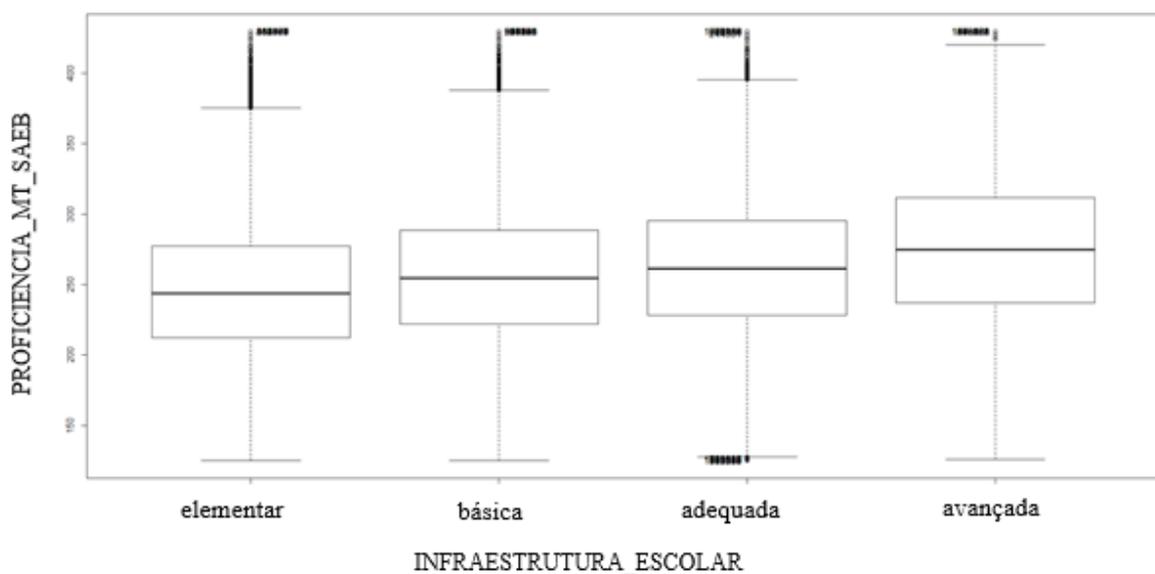
¹⁸ O nível 6 (Alto) será agrupado com o nível 5 (Médio_alto), devido ao reduzido número de ocorrências.

Figura 10 - Gráfico boxplot/diagrama em caixas da variável dependente proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental no Saeb 2017 agrupada pela média por escola da medida de nível socioeconômico dos estudantes



Também, com a medida de infraestrutura das escolas, observa-se a relação com o desempenho dos estudantes, sendo a proficiência dos estudantes concluintes dos Anos Finais na avaliação de Matemática do Saeb 2017 mais alta nas escolas com melhores condições de infraestrutura.

Figura 11 - Gráfico boxplot/diagrama em caixas da variável dependente proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental no Saeb 2017 agrupada pela medida de infraestrutura das escolas



No Apêndice D consta a relação de todas as variáveis que compõem a base de dados utilizada, denominada “dados_Claudia”, que contém 1.796.628 linhas e 22 colunas.

5.2 ESTIMAÇÃO DOS EFEITOS DOS FATORES ASSOCIADOS

Para a estimação dos efeitos dos fatores associados à proficiência em Matemática dos estudantes concluintes dos Anos Finais do Ensino Fundamental, com base nos dados do Saeb 2017, foi realizada a análise de regressão/modelos hierárquicos, considerando a estrutura hierárquica de agrupamento da população em estudo.

As estimativas dos parâmetros são feitas, usualmente, pelo método da máxima verossimilhança, pois “tem a vantagem de produzir estimativas que são eficientes, assintoticamente eficientes e consistentes, quando as condições de regularidade estão satisfeitas” (COELHO, 2017, p. 38). Neste estudo utilizou-se o método da verossimilhança restrita (REML), que leva em conta os graus de liberdade associados ao nível 1 (estudantes) e ao nível 2 (escola) e, onde os componentes da variância serão incluídos na função de verossimilhança e os coeficientes da regressão estimados numa segunda etapa, ou seja, uma referente aos efeitos fixos e outra aos efeitos aleatórios.

5.3 MODELO

A análise foi feita com o objetivo de verificar se as variáveis selecionadas afetam o desempenho dos estudantes concluintes dos Anos Finais do Ensino Fundamental no Saeb 2017.

No ajuste dos modelos, utilizou-se o *software* R (R CORE TEAM, 2019), com os pacotes *lme4* (BATES; MAECHLER; BOLKER; WALKER, 2015), *lmerTest* (KUZNETSOVA; BROCKHOFF; CHRISTENSEN, 2017), sendo os ajustes obtidos via método da verossimilhança restrita.

5.3.1 Modelo nulo

Como estratégia de análise, primeiramente foi ajustado o modelo nulo, sem variáveis independentes, conforme descrito abaixo pela saída computacional do R:

```
library(lme4)
```

```
library(lmerTest)
m0<-lmer(PROFICIENCIA_MT_SAEB~1+(1|ID_ESCOLA), data = dados_Claudia)
summary(m0)
```

REML criterion at convergence: 18711048

Scaled residuals:

Min	1Q	Median	3Q	Max
-4.8448	-0.6726	-0.0044	0.6732	4.3656

Random effects:

Groups	Name	Variance	Std.Dev.
ID_ESCOLA	(Intercept)	536	23.15
	Residual	1859	43.12

Number of obs: 1796213, groups: ID_ESCOLA, 37838

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	249.0679	0.1262	36617.4031	1973	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

A partir do modelo nulo, pode-se calcular o coeficiente de correlação intraclasse:

$$\text{Var}(u_{0j}) = \sigma_{u_0}^2 \text{ variância entre as escolas} = 536$$

$$\text{Var}(e_{ij}) = \sigma_e^2 \text{ variância dos alunos em suas escolas} = 1.859$$

$$(\gamma_{00}) \text{média das escolas} = 249,0679$$

Coefficiente de correlação intraclasse:

$$\rho = \frac{\sigma_{u_0}^2}{\sigma_e^2 + \sigma_{u_0}^2} \quad (5)$$

$$\rho = \frac{536}{1.859 + 536}$$

$$\rho = \frac{536}{2.395}$$

$$\rho = 0,22$$

Pelo modelo nulo, a correlação entre o desempenho dos estudantes de uma mesma escola é de 0,22. Ou seja, 22% da variância da proficiência em Matemática dos estudantes do

9º ano do Ensino Fundamental no Saeb 2017, é entre escolas, e 78% da variância no desempenho dos estudantes concluintes dos Anos Finais do Ensino Fundamental no Saeb 2017, é entre os estudantes em suas escolas.

5.3.2 Modelo 1 – O efeito do nível socioeconômico familiar no desempenho do estudante

Neste modelo, foi incluída a variável independente nível socioeconômico, nível 1 (estudante).

```
library(lme4)
library(lmerTest)
m1<-lmer(PROFICIENCIA_MT_SAEB~1+nse+(1|ID_ESCOLA), data = dados_Claudia )
summary(m1)
```

REML criterion at convergence: 18509688

Scaled residuals:

Min	1Q	Median	3Q	Max
-4.9019	-0.6708	-0.0021	0.6727	4.4210

Random effects:

Groups	Name	Variance	Std.Dev.
ID_ESCOLA	(Intercept)	430.7	20.75
	Residual	1833.3	42.82

Number of obs: 1779862, groups: ID_ESCOLA, 37837

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	239.5613	0.1389	75711.8250	1725.23	<2e-16 ***
nse[T.nivel3_medio_baixo]	6.5838	0.1034	1779726.5835	63.67	<2e-16 ***
nse[T.nivel4_medio]	14.9102	0.1098	1763246.9349	135.84	<2e-16 ***
nse[T.nivel5_medio_alto]	20.1648	0.1350	1770027.3476	149.36	<2e-16 ***
nse[T.nivel6_alto]	19.3394	0.2320	1778798.7144	83.34	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:

	(Intr)	n[T.3_	n[T.4_	n[T.5_
ns[T.nv3__]	-0.514			
ns[T.nv4__]	-0.517	0.711		
ns[T.nv5__]	-0.433	0.587	0.635	
ns[T.nv6__]	-0.257	0.343	0.375	0.343

Efeitos aleatórios:

Estimativa da variância do intercepto (entre escolas) (σ_{u0}^2): 430,7

Estimativa da variância residual (dos estudantes nas escolas) (σ_e^2): 1.833,3

Estimativa da média das escolas: 239,5613

No modelo 1, com a inclusão da variável independente nível socioeconômico (nível 1), a redução da variância entre escolas, em relação ao modelo nulo, é de 19,65% e, a redução da variância dos estudantes em suas escolas é de 1,38%, em relação ao nível socioeconômico baixo (nível 2). Então, no Brasil, o nível socioeconômico explica melhor a diferença de desempenho escolar entre escolas do que entre estudantes de uma mesma escola.

Conforme as estimativas dos coeficientes dos diferentes níveis socioeconômicos, pode-se dizer que, em média, estudantes com nível socioeconômico familiar alto (nível 6) e médio-alto (nível 5) têm desempenho cerca de 20 pontos maior que estudantes com nível socioeconômico familiar baixo, estudantes com nível socioeconômico familiar médio (nível 4) têm desempenho cerca de 15 pontos maior que estudantes com nível socioeconômico familiar baixo e, estudantes com nível socioeconômico familiar médio-baixo (nível 3) têm o desempenho em Matemática no Saeb 2017 cerca de 6 pontos acima que estudantes com nível socioeconômico familiar baixo.

5.3.3 Modelo 2 – O efeito da localização da escola no desempenho do estudante

No modelo que segue, tem-se a inclusão de uma variável de nível 1, medida do nível socioeconômico familiar, e de uma variável do nível 2, localização da escola (rural ou urbana), com o objetivo de explicar a variação entre escolas a partir da localização (rural ou urbana) da escola, considerando estudantes de mesmo nível socioeconômico familiar. Modelo com efeito aleatório só no intercepto (podendo variar entre escolas).

```
library(lme4)
library(lmerTest)
m2<-lmer(PROFICIENCIA_MT_SAEB~1+nse+ID_LOCALIZACAO+(1|ID_ESCOLA),
data = dados_Claudia)
summary(m2)
```

REML criterion at convergence: 18507662

Scaled residuals:

Min	1Q	Median	3Q	Max
-4.9008	-0.6707	-0.0019	0.6727	4.4110

```

Random effects:
  Groups      Name      Variance Std.Dev.
ID_ESCOLA (Intercept)  408.1   20.20
Residual                1833.0   42.81
Number of obs: 1779862, groups: ID_ESCOLA, 37837

Fixed effects:
              Estimate   Std. Error      df t value Pr(>|t|)
(Intercept)    242.6996     0.1535  69692.9300 1581.33 <2e-16 ***
nse[T.nivel3_medio_baixo]    6.2872     0.1036 1777625.9280   60.68 <2e-16 ***
nse[T.nivel4_medio]         14.4781     0.1102 1776917.2047  131.34 <2e-16 ***
nse[T.nivel5_medio_alto]    19.6900     0.1355 1778511.5147  145.31 <2e-16 ***
nse[T.nivel6_alto]          18.8553     0.2323 1776985.9633   81.16 <2e-16 ***
ID_LOCALIZACAO[T.rural]    -12.1122     0.2664  42229.3821  -45.47 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Efeitos aleatórios:

Estimativa da variância do intercepto (entre escolas) (σ_{u0}^2): 408,1

Estimativa da variância residual (dos estudantes nas escolas) (σ_e^2): 1.833,0

Estimativa da média das escolas: 242,6996

Neste modelo é possível observar que, considerando a mesma localização da escola, a elevação no nível socioeconômico do nível 2 (baixo) para o nível 3 (médio baixo) agrega de 6,2872 pontos e, para o nível 5 (médio_alto) agrega 19,69 pontos a mais na proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental, considerando os dados do Saeb 2017. E, a variável localização, reduz, aproximadamente 12 pontos na proficiência em Matemática, quando escola localizada em área rural, mesmo considerando escolas com o mesmo nível socioeconômico.

5.3.4 Modelo 3 – inclusão de todas as variáveis de nível 1 consideradas no estudo

No modelo 3 são incluídas todas as variáveis no nível 1 (estudantes) em estudo: a medida de nível socioeconômico do estudante, o sexo/gênero do estudante, a cor/raça autodeclarada pelo estudantes, se o estudante já teve alguma reprovação durante a trajetória escolar, se o estudante trabalha fora de casa, se os pais ou responsáveis incentivam o estudante nos estudos, se o estudante declara gostar da disciplina de Matemática, se o estudante fez a etapa da educação infantil (creche e/ou pré-escola).

```
library(lme4)
```

```
library(lmerTest)
m3<-lmer(PROFICIENCIA_MT_SAEB~1+ nse + sexo + cor_reord + reprov + trab_fora +
part_pais_incentivo_estudos + gosta_mat + fez_ed_infantil +(1 |ID_ESCOLA), data =
dados_Claudia)
summary(m3)
```

REML criterion at convergence: 14874707

Scaled residuals:

Min	1Q	Median	3Q	Max
-5.7107	-0.6556	0.0137	0.6721	4.5361

Random effects:

Groups	Name	Variance	Std.Dev.
ID_ESCOLA	(Intercept)	352.5	18.77
	Residual	1588.4	39.85

Number of obs: 1449662, groups: ID_ESCOLA, 37828

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	227.37770	0.32744	1111277.25557	694.41	<2e-16 ***
nse[T.nivel3_medio_baixo]	4.42396	0.10919	1449612.95659	40.52	<2e-16 ***
nse[T.nivel4_medio]	10.39696	0.11604	1434717.62420	89.60	<2e-16 ***
nse[T.nivel5_medio_alto]	14.59546	0.14181	1441514.13784	102.92	<2e-16 ***
nse[T.nivel6_alto]	14.70032	0.24110	1448647.49295	60.97	<2e-16 ***
sexo[T.masculino]	11.94748	0.06918	1424004.21050	172.69	<2e-16 ***
cor_reord[T.indigena]	-4.23742	0.14811	1435996.31852	-28.61	<2e-16 ***
cor_reord[T.parda]	-3.70333	0.08141	1446062.16664	-45.49	<2e-16 ***
cor_reord[T.preta]	-9.26351	0.11829	1440405.62340	-78.31	<2e-16 ***
reprov[T.sim_uma]	-21.08804	0.08830	1434993.81363	-238.82	<2e-16 ***
reprov[T.sim_duas_ou_mais]	-22.18473	0.13002	1436805.37435	-170.63	<2e-16 ***
trab_fora[T.sim]	-4.63226	0.10340	1430623.37122	-44.80	<2e-16 ***
part_pais_incentivo_estudos[T.sim]	5.79833	0.28470	1424363.41323	20.37	<2e-16 ***
gosta_mat[T.sim]	18.38672	0.07248	1433297.14610	253.66	<2e-16 ***
fez_ed_infantil[T.creche]	3.17115	0.10254	1445400.56620	30.93	<2e-16 ***
fez_ed_infantil[T.pre_escola]	7.41893	0.10118	1440723.05777	73.33	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Efeitos aleatórios:

Estimativa da variância do intercepto (entre escolas) (σ_{u0}^2): 352,5

Estimativa da variância residual (dos estudantes nas escolas) (σ_e^2): 1.588,4

Estimativa da média das escolas: 227,37770

No modelo 3, com a inclusão de todas as variáveis de nível 1 consideradas no estudo, a redução da variância em relação ao modelo nulo foi de 34,24% entre escolas e de 14,56% a redução da variância entre estudantes nas suas escolas. Dentre estas variáveis, pode-se destacar o impacto positivo gerado pelo indicador de gostar da disciplina sobre a proficiência do estudante em Matemática, 18,39 pontos em relação aos estudantes que declaram não gostar da disciplina, bem como o impacto positivo associado ao fato do estudante ter tido a oportunidade

de fazer a Educação Infantil (etapa da creche 3,17 pontos e etapa da pré-escola 7,42 pontos) em relação aos que não tiveram esta oportunidade.

5.3.5 Modelo 4 – inclusão das variáveis de nível 1 (estudante) e de nível 2 (escola) consideradas no estudo

Modelo com a todas as variáveis de nível 1 (estudante) e inclusão de todas as variáveis de nível 2 (escola) consideradas neste estudo: localização, dependência administrativa, área, turno, a proporção de estudantes por escola que declaram fazer as tarefas de casa da disciplina de Matemática, a proporção de estudantes por escola que declaram que o professor corrige as tarefas de casa da disciplina de Matemática, a proporção de estudantes da escola com alguma reprovação na trajetória escolar, a média por escola da medida de nível socioeconômico dos estudantes, a medida de infraestrutura escolar de cada escola.

```
library(lme4)
library(lmerTest)
m4<-lmer(PROFICIENCIA_MT_SAEB~1+nse+
sexo+cor_reord+reprov+trab_fora+part_pais_incentivo_estudos+gosta_mat+fez_ed_infantil+
ID_AREA+ID_LOCALIZACAO+ID_TURNO+ID_DEPEND_ADM+infra_escolar+reprov_
prop_escola+faz_tarefa_mtm_prop_escola+prof_corrige_tarefa_mtm_prop_escola+NSE_ES
COLA_AGRUP+(1 |ID_ESCOLA), data = dados_Claudia)
summary(m4)
```

REML criterion at convergence: 13333659

Scaled residuals:

Min	1Q	Median	3Q	Max
-5.7430	-0.6555	0.0141	0.6719	4.6014

Random effects:

Groups	Name	Variance	Std.Dev.
ID_ESCOLA	(Intercept)	233	15.26
	Residual	1587	39.84

Number of obs: 1300770, groups: ID_ESCOLA, 32479

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	181.65742	1.43564	35292.47972	126.534	< 2e-16 ***
nse[T.nivel3_medio_baixo]	3.38297	0.11577	1291078.81352	29.220	< 2e-16 ***
nse[T.nivel4_medio]	8.63393	0.12425	1293020.74399	69.489	< 2e-16 ***
nse[T.nivel5_medio_alto]	12.34752	0.15167	1290556.23540	81.412	< 2e-16 ***
nse[T.nivel6_alto]	11.83997	0.25470	1286157.55440	46.485	< 2e-16 ***
sexo[T.masculino]	12.05174	0.07292	1281891.13545	165.269	< 2e-16 ***
cor_reord[T.indigena]	-3.82691	0.15594	1291607.35498	-24.541	< 2e-16 ***

cor_reord[T.parda]	-3.33566	0.08590	1297972.67753	-38.830	< 2e-16	***
cor_reord[T.preta]	-8.91614	0.12516	1295555.29303	-71.240	< 2e-16	***
reprov[T.sim_uma]	-20.60874	0.09387	1274768.39362	-219.536	< 2e-16	***
reprov[T.sim_duas_ou_mais]	-21.37921	0.13888	1282237.14176	-153.935	< 2e-16	***
trab_fora[T.sim]	-4.59259	0.10947	1288154.60990	-41.952	< 2e-16	***
part_pais_incentivo_estudos[T.sim]	5.26120	0.30121	1282277.79833	17.467	< 2e-16	***
gosta_mat[T.sim]	18.57433	0.07651	1288357.64479	242.783	< 2e-16	***
fez_ed_infantil[T.creche]	3.11031	0.10894	1299970.51458	28.550	< 2e-16	***
fez_ed_infantil[T.pre_escola]	7.38004	0.10757	1297382.84660	68.608	< 2e-16	***
ID_AREA[T.interior]	1.10588	0.29549	26702.72557	3.743	0.000183	***
ID_LOCALIZACAO[T.rural]	-3.76335	0.27718	36416.14774	-13.577	< 2e-16	***
ID_TURNO[T.vespertino]	-3.39353	0.10552	615545.63707	-32.159	< 2e-16	***
ID_TURNO[T.noturno]	-9.40636	0.60285	506125.91842	-15.603	< 2e-16	***
ID_DEPEND_ADM[T.federal]	34.26248	2.88536	23633.16641	11.875	< 2e-16	***
ID_DEPEND_ADM[T.municipal]	2.05054	0.21562	29163.74974	9.510	< 2e-16	***
ID_DEPEND_ADM[T.privada]	7.37683	0.84312	32354.11261	8.749	< 2e-16	***
infra_escolar[T.basica]	3.88984	0.22536	29580.22241	17.260	< 2e-16	***
infra_escolar[T.adequada]	5.21906	0.30675	27968.69579	17.014	< 2e-16	***
infra_escolar[T.avancada]	7.70459	0.63459	27704.09800	12.141	< 2e-16	***
reprov_prop_escola	-20.57463	0.65657	35311.21312	-31.336	< 2e-16	***
faz_tarefa_mtm_prop_escola	5.70180	1.75172	34138.09606	3.255	0.001135	**
prof_corrige_tarefa_mtm_prop_escola	37.57380	2.00587	33181.07017	18.732	< 2e-16	***
NSE_ESCOLA_AGRUP[T.nivel3_medio_baixo]	6.05313	0.39871	41434.79378	15.182	< 2e-16	***
NSE_ESCOLA_AGRUP[T.nivel4_medio]	16.55818	0.45186	39449.38751	36.644	< 2e-16	***
NSE_ESCOLA_AGRUP[T.nivel5_medio_alto]	38.64242	0.88311	31167.99223	43.757	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Obs.: o t-value menor do que 2 indica que as variáveis não são significantes e devem ser excluídas.

Vale ressaltar que em todos os modelos testados, as características/fatores mostraram-se estatisticamente significantes, ou seja, $\Pr(> |t|) < 0,05$.

Efeitos aleatórios:

Estimativa da variância do intercepto (entre escolas) (σ_{u0}^2): 233

Estimativa da variância residual (dos estudantes nas escolas) (σ_e^2): 1.587

Estimativa da média das escolas: 181,65742

No modelo final, com a inclusão de todas as variáveis consideradas no estudo, tanto de nível 1 (estudante) quanto de nível 2 (escola), pode-se observar que houve uma diminuição da variância do desempenho dos estudantes em suas escolas de 14,63% em relação ao modelo nulo e, uma redução de 56,53% da variância entre escolas.

A Tabela 12 apresenta o resumo dos efeitos aleatórios entre todos os modelos, nela, observa-se que a medida em que são incluídas as variáveis explicativas nos níveis 1 e 2, reduz-se a variância entre as escolas, e, os coeficientes do modelo indicam que os indicadores considerados foram significantes para explicar a variação de desempenho em Matemática dos estudantes do 9º ano do Ensino Fundamental no Saeb 2017.

Tabela 12 - Resumo dos efeitos aleatórios dos modelos de regressão multinível e redução percentual em relação ao modelo nulo

	M0 Modelo nulo	M1 Modelo com 01 variável no nível 1 (nse)		M2 Modelo com variáveis nível 1 (nse) e nível 2 (localização escola – rural ou urbana)		M3 Modelo com todas as variáveis no nível 1		M4 Modelo com todas as variáveis no nível 1 e todas as variáveis no nível 2	
Variância entre escolas (intercepto)	536	430,7	Redução: 19,65%	408,1	Redução: 23,86%	352,5	Redução: 34,24%	233	Redução: 56,53%
Variância dos estudantes em suas escolas (residual)	1.859	1.833,3	Redução: 1,38%	1.833,0	Redução: 1,40%	1.588,4	Redução: 14,56%	1.587	Redução: 14,63%
Variância total	2.395	2.264,0	Redução: 5,47%	2.241,1	Redução: 6,43%	1.940,9	Redução: 18,96%	1.820	Redução: 24,01%

Fonte: Elaborada pela autora (2020).

Conforme demonstrado na tabela 12, os fatores associados usados neste estudo explicam melhor as diferenças entre escolas do que as diferenças entre estudantes de uma mesma escola.

5.4 DISCUSSÃO E CONTEXTUALIZAÇÃO DOS RESULTADOS

Com a aplicação dos modelos de regressão multinível, sendo o estudante a unidade de nível 1, a escola a unidade de nível 2 e a proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental no Saeb 2017 a variável dependente, observou-se que todas as variáveis testadas foram significantes e ajudam a explicar a variação no desempenho dos estudantes.

Diante dos valores calculados (coeficiente de correlação intraclasse), pode-se afirmar que, para os estudantes do 9º ano do Ensino Fundamental no Saeb 2017, 22% da variância das proficiências em Matemática tiveram como motivo as diferenças entre as escolas, ou redes de ensino, e podem ser explicadas por fatores como a infraestrutura escolar – questões relacionadas a equipamentos, recursos e espaços disponíveis nas escolas, como computadores, acesso à internet e laboratórios (laboratórios de Ciências e Informática, por exemplo), o fato dos estudantes fazerem as tarefas de casa passadas pelos professores e o fato do professor ter o

hábito de corrigir as tarefas que são encaminhadas para os estudantes realizarem em casa, entre outros.

Assim, de acordo com o modelo 4, os fatores que afetam a proficiência dos estudantes concluintes do Ensino Fundamental conforme o Saeb 2017, mantendo-se as demais condições inalteradas, são:

- **variáveis independentes no nível 1 (aluno):**

- ✓ nível socioeconômico dos estudantes (nse): estudantes no nível 3 (médio_baixo) têm, em média, 3,38 pontos a mais na proficiência em Matemática do que estudantes no nível 2 (baixo); estudantes no nível 4 (médio) têm, em média, 8,63 pontos a mais na proficiência em Matemática; estudantes no nível 5 (médio alto) têm, em média, 12,35 pontos a mais na proficiência em Matemática e, estudantes no nível 6 (alto) têm, em média, 11,84 pontos a mais na proficiência em Matemática, tomando como referência os estudantes no nível 2 (baixo).
- ✓ sexo/gênero do estudante (sexo): os estudantes do sexo masculino têm, em média, proficiência de 12,05 pontos a mais que estudantes do sexo feminino, na disciplina de Matemática, no Saeb 2017.
- ✓ cor/raça do estudante (cor_reord): em relação aos estudantes que se autodeclararam brancos, os estudantes indígenas/amarelos têm, em média, 3,83 pontos a menos na proficiência em Matemática; estudantes que se autodeclararam pardos uma redução média de 3,34 pontos; e, estudantes que se autodeclararam pretos uma redução média de 8,92 pontos na proficiência em Matemática no Saeb 2017.
- ✓ estudantes que já tiveram uma reprovação têm, em média, 20,61 pontos a menos na proficiência em Matemática que os estudantes que nunca reprovaram durante a trajetória escolar e, estudantes que já tiveram duas ou mais reprovações têm, em média, 21,38 pontos a menos na proficiência em Matemática comparado aos estudantes sem nenhuma reprovação.
- ✓ estudantes que trabalham fora de casa (trab_fora), têm, em média 4,59 pontos a menos na proficiência em Matemática, em relação aos que não trabalham fora de casa.

- ✓ a participação e/ou o incentivo dos pais ou responsáveis com relação aos estudos do(a) filho(a) (`part_pais_incentivo_estudos`), declarada pelo estudante no questionário contextual, é uma variável que agrega 5,26 pontos na proficiência em Matemática dos estudantes concluintes dos Anos Finais do Saeb 2017.
 - ✓ se o estudante declara gostar da disciplina de Matemática (`gosta_mtm`), em média, ele tem 18,57 pontos a mais na proficiência do que os estudantes que declaram não gostar da disciplina.
 - ✓ se frequentou a Educação Infantil (`fez_ed_infantil`), na etapa creche (0 a 3 anos de idade), o estudante tem, em média, 3,11 pontos a mais do que os estudantes que não fizeram a etapa da Educação Infantil e, se o estudante frequentou a pré-escola (4 a 5 anos de idade), o estudante tem, em média, 7,38 pontos a mais na proficiência em Matemática.
- **variáveis independentes no nível 2 (escola):**
 - ✓ localização da escola, se é urbana ou rural (`ID_LOCALIZACAO`): o fato de a escola estar localizada em área rural, diminui, em média, 3,76 pontos no desempenho em Matemática dos estudantes do 9º ano do Ensino Fundamental no Saeb 2017.
 - ✓ dependência administrativa da escola, se é federal, estadual, municipal ou privada (`ID_DEPEND_ADM`): com relação à dependência administrativa, comparando com escolas estaduais, estudantes de escolas federais obtiveram, em média, 34,26 pontos a mais na proficiência em Matemática, estudantes de escolas municipais, em média, 2,05 pontos a mais e, estudantes de escolas privadas, 7,38 pontos a mais na proficiência em Matemática no Saeb 2017.
 - ✓ se a escola está localizada na capital do estado ou no interior (`ID_AREA`): o fato de a escola não estar localizada na capital, agrega, em média, 1,11 pontos na proficiência em Matemática dos estudantes concluintes do Ensino Fundamental.
 - ✓ o turno em que o estudante frequenta a escola, se é matutino, vespertino ou noturno (`ID_TURNO`): com relação a esta variável, o modelo evidencia que o turno noturno representa uma redução, em média, de 9,41 pontos na proficiência dos estudantes em relação aos estudantes do turno matutino e, o turno vespertino,

uma redução média de 3,39 pontos na proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental no Saeb 2017, também em relação ao turno matutino.

- ✓ a proporção por escola de estudantes que declaram fazer as atividades/tarefas de casa de Matemática (*faz_tarefa_mtm_prop_escola*), agrega, em média 5,7 pontos, na proficiência em Matemática.
- ✓ a proporção por escola de professores que corrigem as atividades/tarefas de casa de Matemática (*prof_corrige_tarefa_mtm_prop_escola*), agrega, em média, 37,57 pontos na proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental no Saeb 2017.
- ✓ a proporção por escola de estudantes com reprovação (*reprov_prop_escola*), indica que, quando há histórico de reprovações entre os estudantes da escola, a média da proficiência em Matemática dos estudantes, tem um impacto negativo de 20,57 pontos.
- ✓ a média da medida de nível socioeconômico dos estudantes da escola (*NSE_ESCOLA_AGRUP*) demonstra que, em média, escolas com nível socioeconômico médio baixo, têm, impacto positivo de 6,05 pontos na proficiência dos estudantes, quando comparado a escolas do média de nível socioeconômico baixo; escolas com medida de nível socioeconômico médio, agregam 16,56 pontos na proficiência dos estudantes e, escolas com média de nível socioeconômico médio alto, têm impacto de 38,64 pontos na proficiência de Matemática dos estudantes concluintes do Ensino Fundamental, no Saeb 2017, quando comparado a escolas com nível socioeconômico baixo (nível 2).
- ✓ a medida/escala das condições de infraestrutura das escolas (*infra_escolar*), demonstra que, quanto melhores as condições de infraestrutura das escolas, melhor o desempenho dos estudantes. Em relação ao nível elementar da medida de infraestrutura escolar, escolas com nível básico de infraestrutura, agregam, em média, 3,89 pontos na proficiência em Matemática dos estudantes; escolas com nível adequado de infraestrutura, agregam, em média, 5,22 pontos a mais no desempenho dos estudantes e, escolas com nível de infraestrutura avançado, são responsáveis por um acréscimo de 7,7 pontos na proficiência em Matemática do estudantes do 9º ano do Ensino Fundamental, no Saeb 2017, quando comparadas às escolas com nível elementar de infraestrutura.

6 ANÁLISE COM OS DADOS DE UMA ESCOLA DO MUNICÍPIO DE FLORIANÓPOLIS/SC

Como mencionado na justificativa e, conforme o último objetivo específico da pesquisa, também foi feita uma análise com ênfase nos resultados pertinentes a uma escola do município de Florianópolis/SC, como recorte para aprofundamento, por ser o contexto no qual estou inserida profissionalmente e na tentativa de contribuir para reflexões acerca dessa realidade.

Para a realização desta análise, a predição dos efeitos aleatórios foi baseada no modelo final (modelo 4) e foi feita a predição do intercepto de cada escola, para simulação dos efeitos que a melhoria dos indicadores que são passíveis de intervenção pela escola/gestão teriam na proficiência de um estudante fictício (estudante y /escola Y). Para cálculo destes efeitos, foi utilizada a equação de regressão, que é a representação algébrica da linha de regressão, e segue descrita na equação 6:

Nível 1 (estudante)

$$y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \dots + \beta_{Kj}X_{Kij} + e_{ij} \quad (6)$$

Nível 2 (escola)

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01}W_{1j} + \dots + \gamma_{0L}W_{Lj} + u_{0j} \\ \beta_{kj} &= \gamma_{k0} \quad k = 1, 2, \dots \end{aligned}$$

Onde,

$i = 1, 2, \dots$, estudantes

$j = 1, 2, \dots$, escolas

X variáveis independentes de nível 1

W variáveis independentes de nível 2

y_{ij} o desempenho do estudante i da escola j

β_{0j} o desempenho esperado (desempenho médio) dos estudantes da escola j

γ_{00} o desempenho esperado (desempenho médio) das escolas

e_{ij} o termo aleatório associado ao estudante i da escola j

u_{0j} o termo aleatório associado à escola j

Também foi simulada uma situação na qual houve a piora de alguns indicadores, para verificação do impacto destas alterações na proficiência de um estudante fictício (estudante x/escola X).

O valor predito pelo modelo 4, da proficiência em Matemática no Saeb 2017, do estudante real da escola é 251,22, sendo que o intercepto da escola em estudo está 0,2 pontos acima do intercepto médio das escolas brasileiras (181,66), ou seja, é de 181,86. Desta forma, para a simulação dos efeitos gerados pela piora das condições (estudante x/escola X) e pela melhoria das condições (estudante y/escola Y), partiu-se deste valor de intercepto: 181,86.

Na tabela 13, seguem descritas as variáveis explicativas de nível 1 (do estudante) e as variáveis explicativas de nível 2 (da escola), e as estimativas/coeficientes gerados pelo modelo 4. Foram simuladas alterações em variáveis passíveis de intervenção por parte da gestão escolar e/ou da rede de ensino e, também, aquelas que, a partir de um trabalho em parceria com as famílias, podem influenciar no desempenho dos estudantes. As variáveis que foram modificadas estão destacadas em cinza.

Tabela 13 – Valor predito/esperado da proficiência em Matemática do estudante x da escola X e do estudante y da escola Y

Nível	Variável	Coefficientes	Estudante real		Estudante x		Estudante y	
	Intercepto			181,66		181,66		181,66
	Efeito aleatório escola real			0,2		0,2		0,2
Estudante	nse[T.nivel2 baixo]	0	0	0	0	0	0	0
Estudante	nse[T.nivel3 medio baixo]	3,38	1	3,38	1	3,38	1	3,38
Estudante	nse[T.nivel4 medio]	8,63	0	0	0	0	0	0
Estudante	nse[T.nivel5 medio alto]	12,35	0	0	0	0	0	0
Estudante	nse[T.nivel6 alto]	11,84	0	0	0	0	0	0
Estudante	sexo[T.feminino]	0	1	0	1	0	1	0
Estudante	sexo[T.masculino]	12,05	0	0	0	0	0	0
Estudante	cor reord[T.branca]	0	0	0	0	0	0	0
Estudante	cor reord[T.indigena/amarela]	-3,83	0	0	0	0	0	0
Estudante	cor reord[T.parda]	-3,34	1	-3,34	1	-3,34	1	-3,34
Estudante	cor reord[T.preta]	-8,92	0	0	0	0	0	0
Estudante	reprov[T.nao]	0	1	0	0	0	1	0
Estudante	reprov[T.sim uma]	-20,61	0	0	0	0	0	0
Estudante	reprov[T.sim duas ou mais]	-21,38	0	0	1	-21,38	0	0
Estudante	trab fora[T.nao]	0	1	0	1	0	1	0
Estudante	trab fora[T.sim]	-4,59	0	0	0	0	0	0
Estudante	part pais incentivo estudos[T.nao]	0	0	0	1	0	0	0
Estudante	part pais incentivo estudos[T.sim]	5,26	1	5,26	0	0	1	5,26
Estudante	gosta mat[T.nao]	0	1	0	1	0	0	0
Estudante	gosta mat[T.sim]	18,57	0	0	0	0	1	18,57
Estudante	fez ed infantil[T.nao]	0	0	0	1	0	0	0
Estudante	fez ed infantil[T.creche]	3,11	0	0	0	0	0	0
Estudante	fez ed infantil[T.pre escola]	7,38	1	7,38	0	0	1	7,38
Escola	ID AREA[T.capital]	0	1	0	1	0	1	0
Escola	ID AREA[T.interior]	1,11	0	0	0	0	0	0

Nível	Variável	Coeficientes	Estudante real		Estudante x		Estudante y	
Escola	ID LOCALIZACAO[T.urbana]	0	1	0	1	0	1	0
Escola	ID LOCALIZACAO[T.rural]	-3,76	0	0	0	0	0	0
Escola	ID TURNO[T.matutino]	0	0	0	0	0	1	0
Escola	ID TURNO[T.vespertino]	-3,39	1	-3,39	0	0	0	0
Escola	ID TURNO[T.noturno]	-9,41	0	0	1	-9,41	0	0
Escola	ID DEPEND ADM[T.estadual]	0	0	0	0	0	0	0
Escola	ID DEPEND ADM[T.federal]	34,26	0	0	0	0	0	0
Escola	ID DEPEND ADM[T.municipal]	2,05	1	2,05	1	2,05	1	2,05
Escola	ID DEPEND ADM[T.privada]	7,38	0	0	0	0	0	0
Escola	infra escolar[T.elementar]	0	0	0	1	0	0	0
Escola	infra escolar[T.basica]	3,89	1	3,89	0	0	0	0
Escola	infra escolar[T.adequada]	5,22	0	0	0	0	0	0
Escola	infra escolar[T.avancada]	7,7	0	0	0	0	1	7,7
Escola	reprov prop escola	-20,57	0,20	-4,11	1	-20,57	0	0
Escola	faz tarefa mtm prop escola	5,7	0,92	5,24	0		1	5,7
Escola	prof corrige tarefa mtm prop escola	37,57	0,97	36,44	0	0	1	37,57
Escola	NSE ESCOLA AGRUP[T.nivel2 baixo]	0	0	0	0	0	0	0
Escola	NSE ESCOLA AGRUP[T.nivel3 medio baixo]	6,05	0	0	0	0	0	0
Escola	NSE ESCOLA AGRUP[T.nivel4 medio]	16,56	1	16,56	1	16,56	1	16,56
Escola	NSE ESCOLA AGRUP[T.nivel5 medio alto]	38,64	0	0	0	0	0	0
	Valor predito da proficiência em Matemática			251,22		149,15		282,68

Fonte: Elaborado pela autora (2020).

Nas simulações realizadas, não foram alteradas as características iniciais de nível socioeconômico familiar (nível 3), sexo (feminino), cor (parda), trabalhar fora (não), área (capital) e localização da escola (urbana), dependência administrativa da escola (municipal) e média da escola do nível socioeconômico dos estudantes (nível 4), partindo-se do entendimento que estas são características que não são passíveis de intervenção direta das ações da escola e/ou rede de ensino.

E, nas simulações realizadas, foram alteradas as características de reprovação escolar na trajetória individual do estudante, a participação dos pais e incentivo aos estudos, a indicação do estudante de gostar ou não da disciplina de Matemática, o estudante ter feito ou não a Educação Infantil, o turno frequentado pelo estudante do 9º ano do Ensino Fundamental, o nível de infraestrutura escolar, a proporção de reprovação dos estudantes da escola, a proporção da escola de estudantes que costumam fazer as tarefas de casa da disciplina de Matemática e a proporção da escola de professores que costumam corrigir em sala de aula as tarefas que são enviadas para casa, partindo-se do entendimento de que estas características, direta ou indiretamente, podem ser impactadas por ações desenvolvidas no âmbito escolar e com o envolvimento das famílias nas atividades escolares.

No caso do estudante x da escola X, a intencionalidade das alterações/simulações foi no sentido de agravamento/piora dos indicadores. E, neste caso, a proficiência esperada em Matemática no Saeb 2017, do estudante x, foi de 149,15 pontos.

No caso do estudante y da escola Y, a intencionalidade das alterações/simulações foi no sentido de incremento/melhora dos indicadores. E, neste caso, a proficiência esperada em Matemática no Saeb 2017, do estudante y, foi de 282,68 pontos.

Considerando a hipótese de que o estudante não teve reprovações na sua trajetória escolar e que a proporção de reprovação da escola é igual a 0, ou seja, a escola, em parceria com as famílias, consegue desenvolver estratégias que garantem a aprendizagem dos estudantes no ano letivo e não há incidência de reprovações escolares; a escola, em parceria com as famílias, consegue mobilizar os pais a incentivarem os estudos dos(as) filhos(as); a escola, em parceria com as famílias, consegue desenvolver estratégias para que o estudante passe a gostar da disciplina de Matemática; ao estudante foi oportunizada a condição de frequentar a Educação Infantil, tendo sido ofertada a vaga pela rede de ensino e a matrícula ter sido realizada pela família, principalmente na etapa da pré-escola; a escola oferece a opção do período matutino para o 9º ano do Ensino Fundamental e o estudante está matriculado neste turno; considerando também que houveram melhorias das condições de infraestrutura escolar, passando para o nível de infraestrutura avançada (com Laboratório de Ciências, auditório, sala para atividades de Música e Artes Plásticas na escola, por exemplo); considerando, ainda, que a proporção de estudantes que fazem as tarefas de casa da disciplina de Matemática é igual a 1, ou seja, a escola, em parceria com as famílias, consegue a mobilização para que todos os estudantes façam as tarefas de casa; e que a proporção por escola de professores de Matemática que corrigem as tarefas de casa é igual a 1, ou seja, todos os professores corrigem as tarefas encaminhadas para casa, teve-se um acréscimo de 133,53 pontos do valor predito/esperado da proficiência em Matemática no Saeb 2017.

Pondera-se que as simulações foram realizadas pressupondo extremos nos indicadores modificados, a título de ênfase do impacto que as alterações podem ter no desempenho dos estudantes e, que alguns dos indicadores que tiveram seus valores alterados, não dependem exclusivamente de ações diretas da escola e/ou da rede de ensino. Mas, também, é importante destacar a importância da intencionalidade pedagógica das ações dos professores e dos gestores escolares, no sentido de conhecer a realidade do estudante, seu contexto sócio/histórico/familiar e buscar estratégias para concretização da função social da escola de disseminação do

conhecimento historicamente desenvolvido, primando pelo êxito no processo de ensino aprendizagem.

Esta simulação, que aqui foi realizada, pode ser replicada, estimando os efeitos que as modificações de cenários, com investimentos específicos em algum/alguns dos indicadores considerados no modelo podem gerar na aprendizagem dos estudantes e mensurar o impacto que estratégias / investimentos / políticas públicas na melhoria da qualidade da educação pode gerar na proficiência dos estudantes.

7 CONSIDERAÇÕES

A proposta da pesquisa foi a análise de fatores associados ao desempenho em Matemática dos estudantes concluintes dos Anos Finais do Ensino Fundamental no Saeb 2017, para isso, foi utilizado o modelo de regressão multinível e foram consideradas variáveis/características relacionadas aos estudantes (nível 1) e às escolas (nível 2), por entender a estrutura hierárquica de agrupamento da população observada – estudantes agrupados em turmas, que estão agrupadas em escolas, pertencentes a uma rede de ensino, que estão inseridas em comunidades com características específicas, e assim por diante. Como sinalizado na revisão de literatura, na análise, cabe destaque às variáveis passíveis de intervenção, sem deixar de considerar que as demais características, como o nível socioeconômico, são variáveis de controle importantes para a explicação da variância do desempenho dos estudantes. Assim, a ênfase aqui é dada aos indicadores que podem ser impactados por ações da gestão escolar e/ou das redes de ensino.

Dentre as variáveis consideradas no modelo final (modelo 4), as que apresentaram impacto significativo na proficiência em Matemática dos estudantes concluintes dos Anos Finais do Ensino Fundamental, baseado nos resultados do Saeb 2017, seguem descritas na tabela 14:

Tabela 14 – Efeito dos fatores associados na proficiência em Matemática dos estudantes do 9º ano do Ensino Fundamental, com base no Saeb 2017

Nível	Variável	Categoria	Estimativa	Erro Padrão	Pr(> t)
	Intercepto		181,66	1,44	<2e-16***
Estudante (1)	Nível socioeconômico	Mais baixo/baixo	0		
Estudante (1)	Nível socioeconômico	Médio baixo	3,38	0,12	<2e-16***
Estudante (1)	Nível socioeconômico	Médio	8,62	0,12	<2e-16***
Estudante (1)	Nível socioeconômico	Médio alto	12,33	0,15	<2e-16***
Estudante (1)	Nível socioeconômico	Alto/mais alto	11,82	0,25	<2e-16***
Estudante (1)	Sexo/gênero	Feminino	0		
Estudante (1)	Sexo/gênero	Masculino	12,05	0,07	<2e-16***
Estudante (1)	Cor/raça	Branca	0		
Estudante (1)	Cor/raça	Indígena/amarela	-3,83	0,16	<2e-16***
Estudante (1)	Cor/raça	Parda	-3,34	0,09	<2e-16***
Estudante (1)	Cor/raça	Preta	-8,92	0,13	<2e-16***
Estudante (1)	Reprovação	Não	0		
Estudante (1)	Reprovação	Sim, uma	-20,61	0,09	<2e-16***
Estudante (1)	Reprovação	Sim, duas ou mais	-21,38	0,14	<2e-16***
Estudante (1)	Trabalha fora de casa	Não	0		

Nível	Variável	Categoria	Estimativa	Erro Padrão	Pr(> t)
Estudante (1)	Trabalha fora de casa	Sim	-4,59	0,11	<2e-16***
Estudante (1)	Participação dos pais, incentivo aos estudos	Não	0		
Estudante (1)	Participação dos pais, incentivo aos estudos	Sim	5,26	0,3	<2e-16***
Estudante (1)	Gosta de estudar Matemática	Não	0		
Estudante (1)	Gosta de estudar Matemática	Sim	18,57	0,08	<2e-16***
Estudante (1)	Fez Educação Infantil	Não	0		
Estudante (1)	Fez Educação Infantil	Sim, creche	3,11	0,11	<2e-16***
Estudante (1)	Fez Educação Infantil	Sim, pré-escola	7,38	0,11	<2e-16***
Escola (2)	Área	Capital	0		
Escola (2)	Área	Interior/Não capital	1,11	0,3	0,000183***
Escola (2)	Localização	Urbana	0		
Escola (2)	Localização	Rural	-3,76	0,28	<2e-16***
Escola (2)	Turno	Matutino	0	0	0
Escola (2)	Turno	Vespertino	-3,4	0,11	<2e-16***
Escola (2)	Turno	Noturno	-9,41	0,6	<2e-16***
Escola (2)	Dependência Administrativa	Estadual	0		
Escola (2)	Dependência Administrativa	Federal	34,26	2,89	<2e-16***
Escola (2)	Dependência Administrativa	Municipal	2,05	0,22	<2e-16***
Escola (2)	Dependência Administrativa	Privada	7,38	0,84	<2e-16***
Escola (2)	Infraestrutura escolar	Elementar	0		
Escola (2)	Infraestrutura escolar	Básica	3,89	0,23	<2e-16***
Escola (2)	Infraestrutura escolar	Adequada	5,22	0,31	<2e-16***
Escola (2)	Infraestrutura escolar	Avançada	7,7	0,63	<2e-16***
Escola (2)	Proporção de reprovação na escola		-20,57	0,66	<2e-16***
Escola (2)	Proporção de estudantes por escola que fazem as tarefas de Matemática		5,70	1,75	0,001135**
Escola (2)	Proporção de professores que corrigem as tarefas de Matemática		37,57	2	<2e-16***
Escola (2)	Nível socioeconômico agrupado por escola	Baixo	0		
Escola (2)	Nível socioeconômico agrupado por escola	Médio baixo	6,05	0,4	<2e-16***
Escola (2)	Nível socioeconômico agrupado por escola	Médio	16,56	0,45	<2e-16***
Escola (2)	Nível socioeconômico agrupado por escola	Médio alto	38,64	0,88	<2e-16***

Fonte: Elaborado pela autora (2020).

A pesquisa evidenciou que o fato de o professor(a) passar tarefas/deveres/atividades para os estudantes realizarem em casa, a título de complementação das atividades desenvolvidas na sala de aula, e, principalmente o fato de o professor(a) corrigir em sala de aula as tarefas que são passadas para os estudantes realizarem em casa, têm um impacto positivo significativo no desempenho dos estudantes. Sendo essa uma informação importante para qualificar o planejamento docente no sentido de ampliação das estratégias para o desenvolvimento de seus planos de ensino.

Ao ressaltar a importância da correção em sala das atividades/tarefas/deveres que são encaminhados para casa, também pode-se destacar este momento de troca como mais uma oportunidade para o esclarecimento de eventuais dúvidas dos estudantes e para a consolidação de aprendizagens. Além deste ser um momento para o protagonismo do estudante, onde ele pode expor/compartilhar as suas hipóteses a respeito da solução dos problemas e, por vezes, a partir deste relato do estudante acerca do seu percurso para chegar à determinada solução/resultado, podem surgir evidências de aspectos que precisam ser retomados e/ou esclarecidos individualmente ou coletivamente para auxiliar o aprendizado dos estudantes. Portanto, um momento rico de troca de experiências e desenvolvimento cognitivo.

Além disso, políticas/estratégias para redução das taxas de reprovação nas escolas, voltadas para a garantia do direito de aprendizagem dos estudantes, podem ser destacadas como importantes contribuições no sentido de melhoria das condições de ensino aprendizagem.

Também, a obrigatoriedade da etapa da Educação Infantil para estudantes de 4 e 5 anos, implementada pela lei nº 12.796, de 04/04/2013, é uma medida que, conforme demonstrado no modelo explicativo, tende a representar um impacto positivo no desempenho dos estudantes.

O investimento, por parte das redes de ensino e também das próprias escolas, em melhorias das condições de infraestrutura escolar, também podem contribuir para o aumento da proficiência dos estudantes. A melhoria das condições de infraestrutura escolar, incluindo características de estrutura e conservação do prédio e mobiliário escolar, da qualidade e quantidade dos equipamentos disponíveis, bem como dos espaços e dos recursos de aprendizagem existentes nas escolas e acessíveis aos estudantes, além de repercutir em aumento no desempenho dos estudantes, como demonstrado no modelo estatístico, também proporcionam aos estudantes com condições socioeconômicas mais baixas acesso a condições que no convívio familiar e/ou comunitário, talvez, não estejam disponíveis, o que pode

contribuir para a perspectiva de equidade das condições de ensino aprendizagem nos sistemas de ensino.

A avaliação educacional em larga escala, como a desenvolvida pelo INEP através do Saeb, fornece uma extensa base de dados com informações referentes à proficiência dos estudantes nas áreas e etapas avaliadas, e também, informações sobre o contexto no qual as escolas e os estudantes estão inseridos, possibilitando o desenvolvimento de pesquisas como esta que foi realizada aqui, sobre fatores associados ao desempenho e, entre outras possibilidades, o monitoramento dos sistemas de ensino e de políticas públicas voltadas para a melhoria da qualidade da educação nacional. Neste sentido, é importante que os questionários contextuais, aplicados aos estudantes, professores, diretores/gestores, e aplicadores externos que são designados para acompanhamento da aplicação dos testes cognitivos nas escolas, sejam respondidos com seriedade, para retratarem a realidade e, que os instrumentos de coleta destes dados, sejam constantemente reavaliados, para que tenham a qualidade psicométrica necessária para mensuração de variáveis/características relevantes e também de traços latentes que permitam a mensuração de característica que não possa ser observadas diretamente, mas sim inferida a partir da observação de características secundárias que estejam relacionadas a ela, como, por exemplo, o nível socioeconômico familiar e a infraestrutura escolar.

Os resultados da pesquisa corroboram com os achados na revisão de literatura e ampliam a discussão para compreensão de fatores associados ao desempenho dos estudantes nas avaliações educacionais em larga escala, tanto aqueles que dependem de políticas públicas mais amplas e a longo prazo, como a obrigatoriedade e a ampliação da obrigatoriedade da Educação Infantil, como aqueles dos quais as próprias escolas e/ou redes de ensino podem intervir, como a melhoria das condições de infraestrutura escolar, o planejamento de atividades que aproximem as famílias do cotidiano escolar, a ampliação da discussão a respeito do desenvolvimento da alfabetização matemática pautada em jogos e materiais de apoio que ampliem as possibilidades de entendimento dos estudantes dos conceitos matemáticos numa perspectiva de aproximação dos estudantes com esta área do conhecimento e, talvez, com isto, o aumento da incidência de respostas positivas no questionário contextual sobre o estudantes gostar de Matemática, diminuindo o impacto negativo que o “não gostar” acarreta no processo de aprendizagem.

7.1 SUGESTÕES PARA TRABALHOS FUTUROS

É importante destacar que, como levantado no referencial teórico, o processo de ensino aprendizagem é um fenômeno complexo que sofre influências de fatores intra e extraescolares, e que a pesquisa aqui desenvolvida é um recorte analítico baseado em uma ferramenta avaliativa e numa perspectiva de modelo estatístico, que apresenta contribuições, mas que também possui limitações. Assim, sugere-se para o desenvolvimento de trabalhos futuros:

- a inclusão de outros conjuntos de variáveis no modelo, que contemplem mais aspectos que podem estar relacionados ao desempenho dos estudantes;
- a análise da variável dependente proficiência em Língua Portuguesa;
- a realização da análise com os resultados dos estudantes concluintes dos Anos Iniciais do Ensino Fundamental (5º ano), para contribuir nas discussões sobre os fatores associados à proficiência dos estudantes nas avaliações educacionais em larga escala;
- a realização de outros estudos acerca do impacto que a melhoria das condições de ensino aprendizagem nas escolas gera na proficiência dos estudantes, a partir de demandas das próprias escolas e/ou redes de ensino.

REFERÊNCIAS

ALVES, Maria Teresa Gonzaga; SOARES, José Francisco. As pesquisas sobre o efeito das escolas: contribuições metodológicas para a Sociologia da Educação. **Sociedade e Estado**, Brasília, v. 22, n. 2, p. 435-473, mai.-ago. 2007.

ALVES, Maria Tereza Gonzaga; SOARES, José Francisco. Contexto escolar e indicadores educacionais: condições desiguais para a efetivação de uma política de avaliação educacional. **Educação e Pesquisa**, São Paulo, v. 39, n. 1, p. 177-194, jan.-mar. 2013.

ANDRADE, Dalton Francisco de; TAVARES, Heliton Ribeiro; VALLE, Raquel da Cunha. **Teoria da resposta ao item: conceitos e aplicações**. SINAPE, 2000.

ANDRADE, Josemberg M.; LAROS, Jacob A.. Fatores associados ao desempenho escolar: um estudo multinível com dados do SAEB/2001. **Psicologia: Teoria e Pesquisa**. Brasília, vol. 23, n. 1, p. 033-042, jan.-mar. 2007.

BARBETTA, Pedro Alberto; ANDRADE, Dalton Francisco de; TAVARES, Héilton Ribeiro. Estudos de fatores associados através de regressão quantílica hierárquica. **Estudos em Avaliação Educacional**. vol. 29, n. 71, p. 320-349, mai.-ago. 2018. Disponível em: <http://publicacoes.fcc.org.br/ojs/index.php/eae/article/view/4973>. Acesso em: 26 fev. 2019.

BATES, Douglas; MAECHLER, Martin; BOLKER, Bem; WALKER, Steve (2015). Fitting Linear Mixed-Effects Models Using **lme4**. *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01

BOTELHO, Louise Lira Roedel; CUNHA, Cristiano Castro de Almeida; MACEDO, Marcelo. O método da revisão integrativa nos estudos organizacionais. **Gestão e Sociedade**. Belo Horizonte. vol. 5, n. 11, p. 121-136. mai./ago. 2011. Disponível em: <https://www.gestoesociedade.org/gestoesociedade/article/view/1220>. Acesso em: 27 jan. 2019.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**: promulgada em 5 de outubro de 1988. Brasília. Disponível em: <http://www.planalto.gov.br>. Acesso em: 22 jan. 2018.

BRASIL. Lei n. 9.394, de 20 de dezembro de 1996. **Lei de Diretrizes e Bases da Educação Nacional**. Estabelece as diretrizes e bases da educação nacional. Disponível em: <http://www.planalto.gov.br>. Acesso em: 22 jan. 2018.

BRASIL. **Lei n. 12.796, de 04 de abril de 2013**. Altera a Lei nº 9.394, de 20 de dezembro de 1996, que estabelece as diretrizes e bases da educação nacional, para dispor sobre a formação dos profissionais da educação e dar outras providências. Disponível em: <https://www2.camara.leg.br/legin/fed/lei/2013/lei-12796-4-abril-2013-775628-publicacaooriginal-139375-pl.html>. Acesso em: 31 jan. 2019.

BRASIL. Lei n. 13.005, de 25 de junho de 2014. **Plano Nacional de Educação**. Aprova o Plano Nacional de Educação e dá outras providências. Disponível em: <http://www.planalto.gov.br>. Acesso em: 22 jan. 2018.

BRASIL, Ministério da Educação, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, Diretoria de Avaliação da Educação Básica (2015) Nota Técnica: **Indicador de Nível Socioeconômico das Escolas Básicas (Inse) participantes da Avaliação Nacional da Alfabetização (ANA)**. Disponível em: http://download.inep.gov.br/educacao_basica/saeb/ana/resultados/2014/nota_tecnica_inse.pdf. Acesso em: 21 nov. 2019.

BRASIL. Ministério da Educação. Secretaria da Educação Básica. **Base Nacional Comum Curricular**. Brasília, DF, 2016. Disponível em: http://basenacionalcomum.mec.gov.br/images/BNCC_EI_EF_110518-versaofinal_site.pdf. Acesso em: 31 jan. 2019.

BRASIL. Ministério da Educação. Instituto de Estudos e Pesquisas Educacionais Anísio Teixeira. **Portaria nº 447**, de 24 de maio de 2017. Estabelece diretrizes para o planejamento e a operacionalização do SAEB no ano de 2017. Disponível em: http://www.lex.com.br/legis_27428172_PORTARIA_N_447_DE_24_DE_MAIO_DE_2017.aspx. Acesso em: 21 nov. 2019.

BRASIL. Ministério da Educação. **Ideb – Apresentação**. Disponível em: <http://portal.mec.gov.br/conheca-o-Ideb>. Acesso em: 26 jan. 2018.

CHALMERS, R. Philip (2012). Mirt: A multidimensional item response theory package for the R Environment. **Journal of statistical software**, 48 (6), 1-29. doi: 10.18637/jss.v048.i06.

COELHO, Fabiano Rodrigues. **Seleção de modelos multiníveis para dados de avaliação educacional**. Dissertação (Mestrado em Estatística) - Programa Interinstitucional de Pós-graduação em Estatística. USP, 2017, 171 p. Disponível em: <http://www.teses.usp.br/teses/disponiveis/104/104131/tde-05122017-145244/pt-br.php>. Acesso em: 19 fev. 2019.

CONSTANCIO, Julimar. **O negro e as políticas de ação afirmativa no contexto da desigualdade educacional**. Dissertação (Mestrado em Educação) - Programa de Pós-Graduação em Educação, UFJF, 2009, 105 p.

COSTA, Marcio da, GUEDES, Reginaldo. Expectativas de futuro como efeito-escola: explorando possibilidades. **São Paulo em Perspectiva**, v. 23, n. 1, p. 101-114, mai.-jun. 2009.

COSTA, Raquel Dias. **O aspecto socioeconômico e a sua influência na qualidade do ensino fundamental público no Brasil**. Dissertação (Mestrado Profissional Executivo em Gestão Empresarial) - Escola Brasileira de Administração Pública e de Empresas, FGV, 2010, 65 p.

FERRÃO, Maria Eugénia. **Introdução aos Modelos de Regressão Multinível em Educação**. Campinas, SP: Editora Komedi, 2003.

FERRÃO, Maria Eugénia; COSTA, Patrícia Mota; MATOS, Daniel Abud Seabra. **The relevance of the school proportion of repeaters on grade repetition in Brazil: a multilevel logistic modelo of PISA 2012**. Springer Open. Large-scale Assess Educ, 2017.

FLORIANÓPOLIS. **Resolução do Conselho Municipal de Educação nº 02/2011**. Dispõe sobre o processo de avaliação, recuperação, promoção, colegiado de classe e recursos do ato avaliativo para o Ensino Fundamental da Rede Municipal de Ensino de Florianópolis. Florianópolis, 2011.

FLORIANÓPOLIS. Secretaria Municipal de Educação. **Proposta Curricular da Rede Municipal de Ensino de Florianópolis – 2016**. Florianópolis: PMF/SME, 2016.

FLORIANÓPOLIS. Lei Complementar nº 546, de 12/01/2016. **Plano Municipal de Educação de Florianópolis 2015 – 2025**. Florianópolis: PMF/SME, 2016.

FOX, J., and BOUCHET-VALAT, M. (2019). **Rcmdr: R Commander**. R package version 2.6-1. Fox, J. (2017). *Using the R Commander: A Point-and-Click Interface for R*. Boca Raton FL: Chapman and Hall/CRC Press. Fox, J. (2005). *The R Commander: A Basic Statistics Graphical User Interface to R*. *Journal of Statistical Software*, 14(9): 1--42.

FREIRE, Paulo. **Pedagogia do oprimido**. 17ª ed. Rio de Janeiro: Paz e Terra, 1987. Disponível em: <http://lelivros.love/book/download-pedagogia-do-oprimido-paulo-freire-em-epub-mobi-e-pdf/>. Acesso em: 19 mai. 2019.

FREITAS, Henrique; OLIVEIRA, Mírian; SACCOL, Amarolinda Zanela; MOSCAROLA, Jean. O método de pesquisa survey. **Revista de Administração**. São Paulo, v.35, n.3, p.105-112, jul/set 2000. Disponível em: http://www.ufrgs.br/gianti/files/artigos/2000/2000__RAUSP.PDF. Acesso em: 05 mar 2019.

FURTADO. Clayton Sirilo do Valle. **Responsabilização educacional em Pernambuco**. Tese (Doutorado em Educação) - Programa de Pós-Graduação em Educação, UFJF, 2015, 249 p.

HAIR, Joseph F. Jr.; BLACK, Willian C.; BABIN, Barry J.; ANDERSON, Rolph E.; TATHAM, Ronald L.. **Análise multivariada de dados**. 6ª ed. Porto Alegre: Artmed, 2006.

HARTMANN, Fábio Cadore. **Fatores associados ao desempenho escolar em Santa Catarina (Ensino Fundamental – 5º e 9º ano) – 2013**. Tese (Doutorado em Sociologia Política) – Programa de Pós-graduação em Sociologia Política da UFSC, Florianópolis, 2017, 395 p.

INEP, Portal: **Indicadores Educacionais**. 23 out. 2018a. Disponível em: <http://portal.inep.gov.br/indicadores-educacionais>. Acesso em: 17 mai. 2019.

INEP, Portal: Resultados finais das escolas no Saeb 2017 já estão disponíveis no Portal do Inep. 01 out. 2018b. Disponível em: http://portal.inep.gov.br/artigo/-/asset_publisher/B4AQV9zFY7Bv/content/resultados-finais-das-escolas-no-saeb-2017-ja-estao-disponiveis-no-portal-do-inep/21206. Acesso em: 28 jul. 2019.

JUNIOR, Fernando de Jesus Moreira; SZINVELSKI, Charles Rogerio Paveglio. **Utilização da teoria da resposta ao item na análise dos dados da autoavaliação institucional da Universidade Federal de Santa Maria: vantagens e oportunidades.** Disponível em: http://download.inep.gov.br/educacao_superior/avaliacao_institucional/seminarios_regionais/trabalhos_regiao/2013/sul/eixo_2/utilizacao_teorias_resposta_item.pdf. Acesso em: 15 ago. 2018.

KUZNETSOVA, A; BROCKHOFF, PB; CHRISTENSES, RHB (2017). “**lmerTest** Package: Tests in Linear Mixed Effects Models.” *Journal of Statistical Software*, *82*(13), 1-26. doi: 10.18637/jss.v082.i13 (URL: <https://doi.org/10.18637/jss.v082.i13>).

LAROS, Jacob Arie; MARCIANO, João Luiz; ANDRADE, Josemberg Moura de. Fatores associados ao desempenho escolar em Português: um estudo multinível por regiões. **Ensaio: aval. pol. públ. Educ.**. Rio de Janeiro, v. 20, n. 77, p. 623-646, out./dez. 2012.

MACHADO, Denys Cristiano de Oliveira. **Análise de fatores associados ao desempenho escolar de alunos do quinto ano do ensino fundamental com base na construção de indicadores.** Dissertação (Mestrado Profissional em Métodos e Gestão em Avaliação) – Programa de Pós-graduação em Métodos e Gestão em Avaliação, Universidade Federal de Santa Catarina, Florianópolis, 2014, 227 p.

MATOS, Daniel Abud Seabra; FERRÃO, Maria Eugénia. Repetência e indisciplina: evidências de Brasil e Portugal no PISA 2012. **Cadernos de Pesquisa**, v. 46, n. 161, p. 614-636, jul.-set. 2016.

MATOS, Daniel Abud Seabra; RODRIGUES, Erica Castilho. Indicadores educacionais e contexto escolar: uma análise das metas do IDEB. **Estudos em Avaliação Educacional**. São Paulo, v. 27, n. 66, p. 662-688, set.-dez. 2016.

MENEZES, Igor G.; DURAN, Victor R.; FILHO, Euclides J. Mendonça; VELOSO, Tainã J.; SARMENTO, Stella M. S.; PAGET, Christine L.; RUGGERI, Kai. Policy implications of achievement testing using multilevel models: the case of Brazilian elementary schools. **Frontiers in Psychology**. v. 7, nov. 2016.

MICRODADOS. 2017. Disponível em: http://portal.inep.gov.br/micro_dados. Acesso em: 24 jan. 2019.

NETO, Joaquim José Soares; JESUS, Girlene Ribeiro de; KARINO, Camila Akemi; ANDRADE, Dalton Francisco de. Uma escala para medir a infraestrutura escolar. **Estudos em Avaliação Educacional**. São Paulo. v.24, n.54, p. 78-99, jan.-abr. 2013.

OLIVEIRA, José Valmir Guimarães de. **Análise do sistema permanente de avaliação da educação básica do Ceará (SPAECE): um estudo comparativo entre o efeito escola e os resultados de proficiência em escolas da rede estadual nos anos de 2012 a 2014.** Dissertação (Mestrado em Educação) - Programa de Pós-Graduação em Educação, UFC, 2016. 131 p.

PALERMO, Gabrielle A.; SILVA, Denise Britz do Nascimento; NOVELLINO, Maria Salet Ferreira. Fatores Associados ao desempenho escolar: uma análise da proficiência em

matemática dos alunos do 5º ano do ensino fundamental da rede municipal do Rio de Janeiro. **Revista Brasileira Est. Pop.**, Rio de Janeiro, v. 31, n. 2, p. 376-394, jul.-dez. 2014.

PEREIRA, Márcia Maria de Carvalho. **Análise dos fatores associados ao desempenho em matemática dos estudantes do 5º ano da rede estadual de educação do estado de Goiás no Saego de 2013**. Dissertação (Mestrado Profissional em Métodos e Gestão em Avaliação) – Programa de Pós-graduação em Métodos e Gestão em Avaliação, Universidade Federal de Santa Catarina, Florianópolis, 2016, 120 p.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

SAMEJIMA, F. (1969). Estimation of latent ability using a response pattern of graded scores (Psychometric Monograph nº 17). Richmond, VA: Psychometric Society. Disponível em: <https://www.psychometricsociety.org/sites/default/files/pdf/MN17.pdf>. Acesso em: 26 fev. 2019.

SILVA, Antônia Bruna da. **Variáveis significativamente relacionadas com o desempenho dos alunos do 5º ano do ensino fundamental público do estado do Ceará no teste de Língua Portuguesa da Prova Brasil/2011**. Dissertação (Mestrado em Educação) - Programa de Pós-Graduação em Educação Brasileira, UFC, 2015, 89 p.

SILVA, Edna Lúcia da; MENEZES, Eстера Muszkat. **Metodologia da pesquisa e elaboração de dissertação**. 4ª ed. rev. atual. Florianópolis: UFSC, 2005.

SILVA, Luciano Campos da; MATOS, Daniel Abud Seabra. Indisciplina no PISA: entre o intra e o extraescolar. **Estudos em Avaliação Educacional**. São Paulo, v.28, n.68, p.382-416, mai.-ago. 2017. Disponível em: <http://publicacoes.fcc.org.br/ojs/index.php/eae/article/view/4590>. Acesso em: 26 fev. 2019.

SOARES, Tufi Machado; MENDONÇA, Márcia Cristina Meneghin. Construção de um modelo de regressão hierárquico para os dados do SIMAVE-2000. **Pesquisa Operacional**. v. 23, n. 3, p. 421-441, set.-dez. 2003.

SOARES, Tufi Machado; FERNANDES, Neimar da Silva; FERRAZ, Mariana Santos Botarro; RAIANI, Juliana de Lucena Ruas de. A expectativa do professor e o desempenho dos alunos. **Psicologia: Teoria e Pesquisa**. v. 26, n. 1, p. 157-170. jan.-mar. 2010.

WATANABE, Margareth; PEREZ, Maria Cândida Raizer Cardinalli. Organização e contextualização de dados como subsídio para a compreensão dos resultados das avaliações educacionais. **São Paulo em Perspectiva**. v. 23, n. 1, p. 149-164, jan.-jun. 2009.

APÊNDICE A – Referencial Teórico

Artigos científicos identificados na busca nas bases de dados Scielo e Scopus com os termos “fatores associados” e (*and*) “desempenho escolar” ou (*or*) “proficiência” e (*and*) “regressão multinível” ou (*or*) “regressão hierárquica”.

Quadro 8 – Artigos identificados bases de dados Scielo e Scopus

Título	Autores	Revista	Ano
Fatores associados ao desempenho escolar em Português: um estudo multinível por regiões	Laros, Jacob Arie; Marciano João Luiz; Andrade, Josemberg Moura de	Revista Ensaio: avaliação e políticas públicas em Educação	2012
Fatores associados ao desempenho escolar: estudo multinível com dados do Saeb/2001	(Andrade, Josemberg M.; Laros, Jacob A.	Psicologia: teoria e pesquisa	2007
Fatores associados ao desempenho escolar: uma análise da proficiência em matemática dos alunos do 5º ano do ensino fundamental da rede municipal do Rio de Janeiro	Palermo, Gabrielle A.; Silva, Denise Britz do Nascimento; Novelino, Maria Salet Ferreira	Revista Brasileira de Estudos de População	2014
Repetência e indisciplina: evidências de Brasil e Portugal no Pisa 2012	Matos, Daniel Abud Seabra; Ferrao, Maria Eugenia	Cadernos de Pesquisa	2016
Policy implications of achievement testing using multilevel models: the case of Brazilian elementary schools	Menezes, Igor G.; Duran, Victor R.; Filho, Euclides J. Mendonça; Veloso, Taina J.; Sarmiento, Stella M. S.; Paget, Christine L.; Ruggeri, Kai	Frontiers in Psychology	2016
The relevance of the school socioeconomic composition and school proportion of repeaters on grade repetition in Brazil: a multilevel logistic modelo of PISA 2012	Ferrão, Maria Eugenia; Costa, Patricia Mota; Matos, Daniel Abud Seabra	Springer Open	2017

Fonte: Elaborada pela autora (2019).

APÊNDICE B – Trabalhos identificados no portal de periódicos da CAPES

Trabalhos identificados no portal de periódicos da CAPES com os descritores “qualquer” / “contém” / “avaliação educacional” e (*and*) “qualquer” / “contém” / “regressão”. Como critério para seleção foi utilizado como período de tempo inicial o ano de 1996 (ano de promulgação da atual LDBEN) e como final o ano de 2018. Como tipo de material foi utilizada a opção todos os itens (livros, artigos, imagens e audiovisuais). Foram encontrados 22 resultados, dos quais foram considerados os 17 periódicos revisados por pares. Em seguida, outro critério de seleção foi a adequação do título e/ou resumo ao tema de interesse. Foram considerados para elaboração da matriz de análise publicações pertinentes ao Ensino Fundamental e que utilizaram modelos de regressão

Quadro 9 – Trabalhos identificados na base de dados portal de periódicos da CAPES

Nº	Título	Autor(es)	Ano
01	Construção de um método de regressão hierárquico para os dados do SIMAVE-2000	Tufi Machado Soares Márcia Cristina Meneghin Mendonça	2003
02	As pesquisas sobre o efeito das escolas: contribuições metodológicas para a sociologia da educação	Maria Tereza Gonzaga Alves José Francisco Soares	2007
03	Expectativas de futuro como efeito-escola: explorando possibilidades	Márcio da Costa Reginaldo Guedes	2009
04	Organização e contextualização de dados como subsídio para a compreensão dos resultados das avaliações educacionais	Margareth Watanabe Maria Candida Raizer Cardinalli Perez	2009
05	A expectativa do professor e o desempenho dos alunos	Tufi Machado Soares Neimar da Silva Fernandes Mariana Santos Botarro Ferraz Juliana de Lucena Ruas de Riani	2010
06	Contexto escolar e indicadores educacionais: condições desiguais para a efetivação de uma política de avaliação educacional	Maria Teresa Gonzaga Alves José Francisco Soares	2013
07	Indicadores educacionais e contexto escolar: uma análise das metas do Ideb	Daniel Abud Seabra Matos Erica Castilho Rodrigues	2016
08	Indisciplina no PISA: entre o intra e o extraescolar	Luciano Campos da Silva Daniel Abud Seabra Matos	2017
09	Estudo de fatores associados através de regressão quantílica hierárquica	Pedro Alberto Barbeta Dalton Francisco de Andrade Héliton Ribeiro Tavares	2018

Fonte: Elaborada pela autora (2019).

APÊNDICE C – Trabalhos identificados no portal da BDTD

Trabalhos identificados no portal da BDTD com a ferramenta de busca avançada, busca por “avaliação educacional” e “regressão”, no período de 1996 a 2018, em todos os idiomas, sem preferência com relação a ilustrações e teses e dissertações como tipo de documento. Foram localizados 16 resultados, dos quais seguem 07 com título e/ou resumo pertinentes ao tema de interesse e relacionados ao Ensino Fundamental. Seguem no quadro 3 as publicações selecionadas que utilizaram o modelo de regressão hierárquica ou que utilizaram o modelo de regressão linear múltipla:

Quadro 10 – Trabalhos identificados na base de dados BDTD

Nº	Título	Autor(es)	Ano	Instituição
01	O negro e as políticas de ação afirmativa no contexto da desigualdade educacional	Julimar Constâncio	2009	Universidade Federal de Juiz de Fora (UFJF)
02	O aspecto socioeconômico e sua influência na qualidade do ensino fundamental público no Brasil	Raquel Dias Costa	2010	Fundação Getúlio Vargas (FGV)
03	Variáveis significativamente relacionadas com o desempenho dos alunos do 5º ano do Ensino Fundamental público no estado do Ceará no teste de Língua Portuguesa da Prova Brasil/2011	Antônia Bruna da Silva	2015	Universidade Federal do Ceará (UFC)
04	Responsabilização educacional em Pernambuco	Clayton Sirilo do Valle Furtado	2015	Universidade Federal de Juiz de Fora (UFJF)
05	Análise do Sistema Permanente de Avaliação da Educação Básica do Ceará (SPAECE): um estudo comparativo entre o efeito escola e os resultados de proficiência em escolas da rede estadual nos anos de 2012 a 2014	José Valmir Guimarães de Oliveira	2016	Universidade Federal do Ceará (UFC)
06	Seleção de modelos multiníveis para dados de avaliação educacional	Fabiano Rodrigues Coelho	2017	Universidade de São Paulo (USP)

Fonte: Elaborada pela autora (2019).

APÊNDICE D – Descrição da base de dados utilizada na aplicação dos modelos de regressão

Bases de dados:

```
> names(dados_Claudia)
[1] "cor_reord"                "infra_escolar"
[3] "faz_tarefa_mtm_prop_escola" "fez_ed_infantil"
[5] "gosta_mat"               "ID_AREA"
[7] "ID_DEPEND_ADM"          "ID_ALUNO"
[9] "ID_ESCOLA"              "ID_LOCALIZACAO"
[11] "ID_MUNICIPIO"           "ID_TURNO"
[13] "nse"                    "NSE_ESCOLA"
[15] "part_pais_incentivo_estudos" "prof_corrige_tarefa_mtm_prop_escola"
[17] "PROFICIENCIA_MT_SAEB"   "reprov"
[19] "reprov_prop_escola"     "sexo"
[21] "trab_fora"              "NSE_ESCOLA_AGRUP"
```

A base de dados “dados_Claudia” tem 1.796.628 linhas e 22 colunas.

```
> summary(dados_Claudia)
  cor_reord      infra_escolar  faz_tarefa_mtm_prop_escola  fez_ed_infantil
Length:1796628  elementar:686705    Min. :0.0000                nao :294336
Class :character  basica :591108      1st Qu.:0.8649              creche :726765
Mode :character  adequada :278487    Median :0.9189              pre_escola:714350
                    avancada : 51417      Mean :0.9014                NA's : 61177
                    NA's :188911      3rd Qu.:0.9583
                    Max. :1.0000
                    NA's :8

gosta_mat      ID_AREA      ID_DEPEND_ADM      ID_ALUNO      ID_ESCOLA
nao : 596021    capital : 317959    Length:1796628      Min. :21362287    Min. :11000260
sim :1113636    interior:1478669    Class :character     1st Qu.:23102418    1st Qu.:27221733
NA's: 86971      Mode :character     Median :24767928      Median :35036006
                    Mean :24817001      Mean :39516209
                    3rd Qu.:26532545    3rd Qu.:61182921
                    Max. :28424144      Max. :61250070

ID_LOCALIZACAO  ID_MUNICIPIO      ID_TURNO      nse      NSE_ESCOLA
urbana:1593361  Min. :1100015      matutino :1079315  Length:1796628  Length:1796628
rural : 203267  1st Qu.:2513703    vespertino: 703246  Class :character  Class :character
                    Median :3169356      noturno : 14067    Mode :character  Mode :character
                    Mean :3227011
                    3rd Qu.:3552205
                    Max. :6311029

part_pais_incentivo_estudos  prof_corrige_tarefa_mtm_prop_escola  PROFICIENCIA_MT_SAEB
nao : 29485                    Min. :0.0000                Min. :125.0
sim :1731562                    1st Qu.:0.9254              1st Qu.:218.1
NA's: 35581                    Median :0.9685              Median :250.8
                    Mean :0.9453                Mean :251.9
                    3rd Qu.:1.0000            3rd Qu.:285.2
                    Max. :1.0000                Max. :429.6
                    NA's :9                    NA's :415

reprov      reprov_prop_escola      sexo      trab_fora
nao :1223078  Min. :0.0000      feminino :886930    nao :1508619
sim_uma : 363806  1st Qu.:0.1875    masculino:847220    sim : 235991
sim_duas_ou_mais: 151872  Median :0.2821    NA's : 62478      NA's: 52018
NA's : 57872      Mean :0.2972
                    3rd Qu.:0.3898
                    Max. :1.0000
                    NA's :8

NSE_ESCOLA_AGRUP
Length:1796628
Class :character
Mode :character
```

APÊNDICE E – Descrição dos passos para simulação com os dados de uma escola do município de Florianópolis/SC

Passo 1:

```
dados_Claudia<-
subset(base9_saeb_2017,select=c(cor_recod,escala_infra_niveis_fator,faz_tarefa_mtm_prop_
escola,fez_educ_infantil_fator,gosta_mat,ID_AREA,ID_DEPENDENCIA_ADM_recod,ID_
ALUNO,ID_ESCOLA,ID_LOCALIZACAO,ID_MUNICIPIO,ID_TURNO_fator,nse_50_10
_niveis_fator_agrup,NSE_ESCOLA_NIVEIS_FATOR_AGRUP,part_pais_incentivo_estudos
,prof_corrige_tarefa_mtm_prop_escola,PROFICIENCIA_MT_SAEB,reprov.x,reprov_prop_e
scola,sexo_fator,trab_fora))
sapply(dados_Claudia, function(x)(sum(is.na(x)))) # NA counts
summary(dados_Claudia)
```

Passo 2:

```
library(lme4)
library(lmerTest)
mod<-lmer(PROFICIENCIA_MT_SAEB~1+nse+sexo+cor_reord+reprov+trab_fora+
part_pais_incentivo_estudos+gosta_mat+fez_ed_infantil+
ID_AREA+ID_LOCALIZACAO+ID_TURNO+ID_DEPEND_ADM+infra_escolar+reprov_
prop_escola+faz_tarefa_mtm_prop_escola+prof_corrige_tarefa_mtm_prop_escola+NSE_ES
COLA_AGRUP+(1|ID_ESCOLA), data = dados_Claudia)
summary(mod)
```

REML criterion at convergence: 13333659

Scaled residuals:

Min	1Q	Median	3Q	Max
-5.7430	-0.6555	0.0141	0.6719	4.6014

Random effects:

Groups	Name	Variance	Std.Dev.
ID_ESCOLA	(Intercept)	233	15.26
	Residual	1587	39.84

Number of obs: 1300770, groups: ID_ESCOLA, 32479

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	181.65742	1.43564	35292.47972	126.534	< 2e-16 ***
nse[T.nivel3_medio_baixo]	3.38297	0.11577	1291078.81352	29.220	< 2e-16 ***
nse[T.nivel4_medio]	8.63393	0.12425	1293020.74399	69.489	< 2e-16 ***
nse[T.nivel5_medio_alto]	12.34752	0.15167	1290556.23540	81.412	< 2e-16 ***
nse[T.nivel6_alto]	11.83997	0.25470	1286157.55440	46.485	< 2e-16 ***
sexo[T.masculino]	12.05174	0.07292	1281891.13545	165.269	< 2e-16 ***
cor_reord[T.indigena]	-3.82691	0.15594	1291607.35498	-24.541	< 2e-16 ***
cor_reord[T.parda]	-3.33566	0.08590	1297972.67753	-38.830	< 2e-16 ***
cor_reord[T.preta]	-8.91614	0.12516	1295555.29303	-71.240	< 2e-16 ***
reprov[T.sim_uma]	-20.60874	0.09387	1274768.39362	-219.536	< 2e-16 ***
reprov[T.sim_duas_ou_mais]	-21.37921	0.13888	1282237.14176	-153.935	< 2e-16 ***
trab_fora[T.sim]	-4.59259	0.10947	1288154.60990	-41.952	< 2e-16 ***
part_pais_incentivo_estudos[T.sim]	5.26120	0.30121	1282277.79833	17.467	< 2e-16 ***
gosta_mat[T.sim]	18.57433	0.07651	1288357.64479	242.783	< 2e-16 ***
fez_ed_infantil[T.creche]	3.11031	0.10894	1299970.51458	28.550	< 2e-16 ***
fez_ed_infantil[T.pre_escola]	7.38004	0.10757	1297382.84660	68.608	< 2e-16 ***

```

ID_AREA[T.interior] 1.10588 0.29549 26702.72557 3.743 0.000183 ***
ID_LOCALIZACAO[T.rural] -3.76335 0.27718 36416.14774 -13.577 < 2e-16 ***
ID_TURNO[T.vespertino] -3.39353 0.10552 615545.63707 -32.159 < 2e-16 ***
ID_TURNO[T.noturno] -9.40636 0.60285 506125.91842 -15.603 < 2e-16 ***
ID_DEPEND_ADM[T.federal] 34.26248 2.88536 23633.16641 11.875 < 2e-16 ***
ID_DEPEND_ADM[T.municipal] 2.05054 0.21562 29163.74974 9.510 < 2e-16 ***
ID_DEPEND_ADM[T.privada] 7.37683 0.84312 32354.11261 8.749 < 2e-16 ***
infra_escolar[T.basica] 3.88984 0.22536 29580.22241 17.260 < 2e-16 ***
infra_escolar[T.adequada] 5.21906 0.30675 27968.69579 17.014 < 2e-16 ***
infra_escolar[T.avancada] 7.70459 0.63459 27704.09800 12.141 < 2e-16 ***
reprov_prop_escola -20.57463 0.65657 35311.21312 -31.336 < 2e-16 ***
faz_tarefa_mtm_prop_escola 5.70180 1.75172 34138.09606 3.255 0.001135 **
prof_corrige_tarefa_mtm_prop_escola 37.57380 2.00587 33181.07017 18.732 < 2e-16 ***
NSE_ESCOLA_AGRUP[T.nivel3_medio_baixo] 6.05313 0.39871 41434.79378 15.182 < 2e-16 ***
NSE_ESCOLA_AGRUP[T.nivel4_medio] 16.55818 0.45186 39449.38751 36.644 < 2e-16 ***
NSE_ESCOLA_AGRUP[T.nivel5_medio_alto] 38.64242 0.88311 31167.99223 43.757 < 2e-16 ***
---
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Passo 3:

Para cálculo do efeito aleatório utilizou-se o seguinte script:

```

ef_aleatorio <- ranef(mod)
ef_aleatorio <- as.data.frame(ef_aleatorio)
head(ef_aleatorio)
tail(ef_aleatorio)
```

Que apresentou o seguinte resultado:

```

> ef_aleatorio <- ranef(mod)
> ef_aleatorio <- as.data.frame(ef_aleatorio)
> head(ef_aleatorio)

grpvar      term      grp  condval  condsd
1 ID_ESCOLA (Intercept) 11000260 -1.192546 3.784900
2 ID_ESCOLA (Intercept) 11000317 22.019440 3.524214
3 ID_ESCOLA (Intercept) 11000376  7.312080 4.838306
4 ID_ESCOLA (Intercept) 11000457 13.713773 4.987642
5 ID_ESCOLA (Intercept) 11000589 12.501710 9.440217
6 ID_ESCOLA (Intercept) 11000597 -1.776062 3.968181

> tail(ef_aleatorio)

      grpvar      term      grp  condval  condsd
32474 ID_ESCOLA (Intercept) 61249901 -5.799852 6.089227
32475 ID_ESCOLA (Intercept) 61249916 -30.760460 7.842075
32476 ID_ESCOLA (Intercept) 61250022 -5.230965 9.716906
32477 ID_ESCOLA (Intercept) 61250030 -7.188358 8.164757
32478 ID_ESCOLA (Intercept) 61250038 -3.840345 5.535165
32479 ID_ESCOLA (Intercept) 61250070 -11.685305 5.482511
```

Passo 4:

Para a seleção dos dados referentes a uma escola da rede pública municipal de Florianópolis/SC, utilizou-se o seguinte script:

```

ef_aleatorio_escola_x <- subset(ef_aleatorio, subset=grp == 42001005)
```

```
ef_aleatorio_escola_x
escola_x <- subset(dados_Claudia, subset=ID_ESCOLA == 42001005)
summary(escola_x)
escola_x$preditos <- with(escola_x, predict(mod, escola_x))
summary(escola_x)
```

Que apresentou como resultado:

```
> ef_aleatorio_escola_x
      grpvar      term      grp  condval  condsd
20554 ID_ESCOLA (Intercept) 42001005 0.2040744 4.73471

> summary(escola_x)

cor_reord      infra_escolar faz_tarefa_mtm_prop_escola  fez_ed_infantil gosta_mat
Length:77      elementar: 0      Min.      :0.9178      nao      :22      nao :17
Class :character  basica      :77      1st Qu.:0.9178      creche   :37      sim  :53
Mode  :character  adequada : 0      Median :0.9178      pre_escola:11      NA's : 7
                          avancada : 0      Mean   :0.9178      NA's     : 7
                          3rd Qu.:0.9178
                          Max.   :0.9178

      ID_AREA  ID_DEPEND_ADM      ID_ALUNO      ID_ESCOLA      ID_LOCALIZACAO
capital :77   Length:77      Min.      :25883544   Min.      :42001005   urbana:77
interior: 0   Class :character  1st Qu.:25883570   1st Qu.:42001005   rural : 0
                          Mode  :character  Median :25883603   Median :42001005
                          3rd Qu.:25883625   3rd Qu.:42001005
                          Max.   :25883646   Max.   :42001005

      ID_MUNICIPIO      ID_TURNO      nse      NSE_ESCOLA
Min.      :4205407   matutino :29   Length:77   Length:77
1st Qu.:4205407   vespertino:48   Class :character   Class :character
Median :4205407   noturno   : 0   Mode  :character   Mode  :character
Mean   :4205407
3rd Qu.:4205407
Max.   :4205407

part_pais_incentivo_estudos prof_corrige_tarefa_mtm_prop_escola PROFICIENCIA_MT_SAEB
nao : 1      Min.      :0.9726      Min.      :154.5
sim :73      1st Qu.:0.9726      1st Qu.:231.1
NA's : 3      Median :0.9726      Median :271.1
                          Mean   :0.9726      Mean   :266.9
                          3rd Qu.:0.9726      3rd Qu.:299.0
                          Max.   :0.9726      Max.   :370.8

      reprov      reprov_prop_escola      sexo      trab_fora NSE_ESCOLA_AGRUP
nao      :59      Min.      :0.1918      feminino :39      nao :64      Length:77
sim_uma  :13      1st Qu.:0.1918      masculino:36      sim : 9      Class :character
sim_duas_ou_mais: 1      Median :0.1918      NA's      : 2      NA's : 4      Mode  :character
NA's     : 4      Mean   :0.1918
                          3rd Qu.:0.1918
                          Max.   :0.1918

      preditos
Min.      :235.2
1st Qu.:262.8
Median :274.4
Mean   :272.4
3rd Qu.:282.2
Max.   :297.8
NA's    :13
```

Passo 5:

Tendo as informações referentes a escola X, foi utilizado o seguinte script para a obtenção das informações acerca de um estudante real:

```
escola_x <- with(escola_x, escola_x[order(preditos, decreasing=FALSE), ])
head(escola_x)
um_aluno_escola_x <- subset(escola_x, subset=ID_ALUNO == 25883646)
write.table(um_aluno_escola_x, "G:/Dinha/Dissertação/2020/Dados para aplicação
modelo/um_aluno_escola_x.txt", sep="\t", col.names=TRUE, row.names=FALSE,
quote=FALSE, na="NA")
```

Com as informações de um estudante real da escola X, foram simuladas duas condições referentes a estudantes fictícios, uma, estudante fictício x, considerando condições com maior impacto negativo e outra, estudante fictício y, melhorando os indicadores da escola e do estudante (níveis 1 e 2).

ID_ALUNO	ID_ESCOLA	ID_LOCALIZAÇÃO	ID_MUNICIPIO	ID_TURNO	ID_AREA
Estudante real	Escola real	urbana	4205407 (Florianópolis/SC)	vespertino	capital
Estudante fictício x	Escola X	urbana	4205407 (Florianópolis/SC)	noturno	capital
Estudante fictício y	Escola Y	urbana	4205407 (Florianópolis/SC)	matutino	capital

ID_ALUNO	ID_DEPEND_ADM	cor_reord	infra_escolar	faz_tarefa_mtm_prop_escola
Estudante real	Municipal	Parda	basica	0,9178082
Estudante fictício x	Municipal	Parda	elementar	0
Estudante fictício y	Municipal	Parda	avancada	1

ID_ALUNO	fez_ed_infantil	gosta_mat	nse	NSE_ESCOLA_AGRUP	part_pais_incentivo_estudos
Estudante real	pre_escola	nao	nivel3 medio baixo	nivel4 medio	sim
Estudante fictício x	nao	nao	nivel3 medio baixo	nivel4 medio	nao
Estudante fictício y	pre_escola	sim	nivel3 medio baixo	nivel4 medio	sim

ID_ALUNO	prof_corrige_tarefa_mtm_prop_escola	PROFICIENCIA_MT_SAEB	reprov	reprov_prop_escola	sexo
Estudante real	0,9726027	256,4334	nao	0,1917808	feminino
Estudante fictício x	0	NA	Sim_duas_ou_mais	1	feminino
Estudante fictício y	1	NA	nao	0	feminino

ID_ALUNO	trab_fora	predito
Estudante real	nao	251,487
Estudante fictício X	nao	149,157
Estudante fictício Y	nao	282,63

Fonte: Elaborada pela autora (2020).

```
predict(mod, aluno_ficticio)
```

```
1          2          3
```

```
251.4868 149.1569 282.6347
```

ANEXO A – Protocolo de pesquisa BU/UFSC

Protocolo para busca sistemática da literatura utilizado para levantamento de aplicações do modelo estatístico de regressão para identificação de fatores associados à proficiência em avaliações educacionais em larga escala.

	UNIVERSIDADE FEDERAL DE SANTA CATARINA BIBLIOTECA UNIVERSITÁRIA SERVIÇO DE REFERÊNCIA PROGRAMA DE CAPACITAÇÃO																														
PROTOCOLO PARA BUSCA SISTEMÁTICA DA LITERATURA																															
1. QUESTÃO OU PROBLEMA DE PESQUISA																															
Quais as variáveis utilizadas para análise dos fatores associados ao desempenho/proficiência dos estudantes nas avaliações educacionais em larga escala, com impacto positivo ou negativo, em estudos/pesquisas anteriores?																															
1.1 OBJETIVOS DA PESQUISA																															
Identificar trabalhos que utilizaram modelos de regressão para análise de fatores associados ao desempenho/proficiência dos estudantes nas avaliações educacionais em larga escala e quais as variáveis utilizadas nos modelos que apresentaram impactos significativos na proficiência dos estudantes.																															
2. BUSCA NA LITERATURA																															
2.1 SELEÇÃO DOS TÓPICOS																															
<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: center;">Tópico 1: Avaliação educacional</td> <td style="text-align: center;">Tópico 2: Fatores associados</td> <td style="text-align: center;">Tópico 3: Regressão</td> <td style="text-align: center;">Tópico 4: Questionários contextuais</td> </tr> <tr> <td style="text-align: center;">termos alternativos/ sinônimos</td> <td style="text-align: center;">termos alternativos/ sinônimos</td> <td style="text-align: center;">termos alternativos/ sinônimos</td> <td style="text-align: center;">termos alternativos/ sinônimos</td> </tr> <tr> <td style="text-align: center;">_____</td> <td style="text-align: center;">_____</td> <td style="text-align: center;">_____</td> <td style="text-align: center;">_____</td> </tr> <tr> <td style="text-align: center;">avaliação externa</td> <td style="text-align: center;">_____</td> <td style="text-align: center;">_____</td> <td style="text-align: center;">_____</td> </tr> <tr> <td style="text-align: center;">avaliação educacional</td> <td style="text-align: center;">_____</td> <td style="text-align: center;">_____</td> <td style="text-align: center;">_____</td> </tr> <tr> <td style="text-align: center;">testes padronizados</td> <td style="text-align: center;">_____</td> <td style="text-align: center;">_____</td> <td style="text-align: center;">_____</td> </tr> <tr> <td style="text-align: center;">_____</td> <td style="text-align: center;">_____</td> <td style="text-align: center;">_____</td> <td style="text-align: center;">_____</td> </tr> </table>	Tópico 1: Avaliação educacional	Tópico 2: Fatores associados	Tópico 3: Regressão	Tópico 4: Questionários contextuais	termos alternativos/ sinônimos	termos alternativos/ sinônimos	termos alternativos/ sinônimos	termos alternativos/ sinônimos	_____	_____	_____	_____	avaliação externa	_____	_____	_____	avaliação educacional	_____	_____	_____	testes padronizados	_____	_____	_____	_____	_____	_____	_____			
Tópico 1: Avaliação educacional	Tópico 2: Fatores associados	Tópico 3: Regressão	Tópico 4: Questionários contextuais																												
termos alternativos/ sinônimos	termos alternativos/ sinônimos	termos alternativos/ sinônimos	termos alternativos/ sinônimos																												
_____	_____	_____	_____																												
avaliação externa	_____	_____	_____																												
avaliação educacional	_____	_____	_____																												
testes padronizados	_____	_____	_____																												
_____	_____	_____	_____																												
2.2 CRITÉRIOS DE INCLUSÃO/EXCLUSÃO																															
<table border="1" style="width: 100%; border-collapse: collapse;"><tr><td style="width: 25%;">tipo de estudo</td><td style="width: 45%;">Sem restrição</td><td style="width: 15%;">incluir</td><td style="width: 15%;">excluir</td></tr><tr><td>área geográfica</td><td>Sem restrição</td><td>incluir</td><td>excluir</td></tr><tr><td>período de tempo</td><td>1996 a 2018</td><td>Incluir X</td><td>excluir</td></tr><tr><td>idioma</td><td>Português, Inglês e Espanhol</td><td>incluir X</td><td>excluir</td></tr><tr><td>outros</td><td></td><td>incluir</td><td>excluir</td></tr></table>	tipo de estudo	Sem restrição	incluir	excluir	área geográfica	Sem restrição	incluir	excluir	período de tempo	1996 a 2018	Incluir X	excluir	idioma	Português, Inglês e Espanhol	incluir X	excluir	outros		incluir	excluir											
tipo de estudo	Sem restrição	incluir	excluir																												
área geográfica	Sem restrição	incluir	excluir																												
período de tempo	1996 a 2018	Incluir X	excluir																												
idioma	Português, Inglês e Espanhol	incluir X	excluir																												
outros		incluir	excluir																												
2.3 FONTES DE INFORMAÇÃO																															
2.3.1 Fontes de informação eletrônica (base de dados, bibliotecas digitais, mecanismos de busca, repositórios, etc.)																															
<table border="1" style="width: 100%; border-collapse: collapse;"><tr><td style="width: 50%;">Tipo de fonte</td><td style="width: 50%;">Nome</td></tr><tr><td>Base de Dados</td><td>Periódicos CAPES</td></tr><tr><td>Base de Dados</td><td>Biblioteca Digital Brasileira de Teses e Dissertações (BDTD)</td></tr></table>	Tipo de fonte	Nome	Base de Dados	Periódicos CAPES	Base de Dados	Biblioteca Digital Brasileira de Teses e Dissertações (BDTD)																									
Tipo de fonte	Nome																														
Base de Dados	Periódicos CAPES																														
Base de Dados	Biblioteca Digital Brasileira de Teses e Dissertações (BDTD)																														
2.3.2 Outras fontes de informação																															

2.4 Estratégia de busca de acordo com o recurso utilizado

Estratégia da Fonte A (Periódicos CAPES)	Busca avançada: qualquer/contém: avaliação educacional AND qualquer/contém: regressão Data de publicação: 1996 – 2018 Idioma: Português, Inglês, Espanhol Resultados: 22
Estratégia da Fonte B (BDTD)	Busca avançada: qualquer/contém: avaliação educacional AND qualquer/contém: regressão Data de publicação: 1996 – 2018 Idioma: Português, Inglês, Espanhol Resultados: 16
Estratégia da Fonte C	

Observação: foram considerados na busca, estudos que utilizaram o modelo de regressão multinível/hierárquica e o modelo de regressão linear múltipla.

3. AVALIAÇÃO DOS RESULTADOS

Vide capítulo 2.

4. APRESENTAÇÃO DOS RESULTADOS (RELATÓRIO)

Vide capítulo 2.