

FUNDAÇÃO JOÃO PINHEIRO
Escola de Governo Professor Paulo Neves de Carvalho

Júlia Perdigão de Melo Couto

A ESCOLA PÚBLICA SEGUE IMPORTANTE PARA QUALIDADE E EQUIDADE: UM
ESTUDO SOBRE O 3º ANO DO ENSINO MÉDIO DE 2019

Belo Horizonte

2021

Júlia Perdigão de Melo Couto

A ESCOLA PÚBLICA SEGUE IMPORTANTE PARA QUALIDADE E EQUIDADE: UM
ESTUDO SOBRE O 3º ANO DO ENSINO MÉDIO DE 2019

Trabalho apresentado ao Curso de Graduação em Administração Pública da Escola de Governo Professor Paulo Neves de Carvalho da Fundação João Pinheiro, como requisito parcial para obtenção do título de Especialista em Políticas Públicas e Gestão Governamental.

Orientador: Marcos Arcanjo Assis

Belo Horizonte
2021

C871e Couto, Júlia Perdigão de Melo.
A escola pública segue importante para qualidade e equidade [manuscrito] : um estudo sobre o 3º ano do ensino médio de 2019 / Júlia Perdigão de Melo Couto. – 2021.
[7], 132 f. : il.

Trabalho de conclusão de Curso (Graduação em Administração Pública) – Fundação João Pinheiro, Escola de Governo Professor Paulo Neves de Carvalho, 2021.

Orientador: Marcos Arcanjo de Assis

Bibliografia: f. 90-94

1. Escola pública – Minas Gerais. 2. Avaliação escolar. 3. Ensino médio – Minas Gerais. 4. Estudante – Desempenho escolar. I. Assis, Marcos Arcanjo de. II. Título.

CDU 37(815.1)

Júlia Perdigão de Melo Couto

A escola pública segue importante para qualidade e equidade? Um estudo sobre O 3º ano do ensino médio de 2019

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Administração Pública da Escola de Governo Professor Paulo Neves de Carvalho, da Fundação João Pinheiro, como requisito parcial para a obtenção do título de bacharel em Administração Pública.

Aprovada na Banca Examinadora



Prof. Marcos Arcanjo de Assis (Orientador) – Fundação João Pinheiro

**Bruno Lazzarotti
Diniz Costa**

Assinado de forma digital por Bruno Lazzarotti Diniz Costa
DN: cn=Bruno Lazzarotti Diniz Costa, o, ou=Escola de Governo, email=bruno.diniz@fjp.mg.gov.br, c=BR
Dados: 2021.01.18 15:25:21 -03'00'

Prof. Bruno Lazzarotti Diniz Costa (Avaliador) – Fundação João Pinheiro



Assinado digitalmente por Victor Barcelos Ferreira
DN: C=BR, OU=DPP, O=FJP, CN=Victor Barcelos Ferreira, E=victor.ferreira@fjp.mg.gov.br
Data: 2021-01-18 16:34:43
Foxit Reader Versão: 10.0.0

Prof. Victor Barcelos Ferreira – Fundação João Pinheiro

Belo Horizonte, 18 de janeiro de 2021.

AGRADECIMENTOS

Este trabalho faz parte de uma longa jornada que sou muito grata pela oportunidade de ter percorrido. Agradeço imensamente a minha família, quem me criou, esteve ao meu lado e com quem compartilhei os momentos mais preciosos. Aos professores e servidores da Fundação João Pinheiro, que sempre lembrarei com muito carinho, por sua dedicação e por tudo que me ensinaram. Aos meus colegas pelo companheirismo que me foi fundamental, em especial àqueles que se tornaram meus amigos, por todas as experiências que vivemos e por todo apoio que sempre me deram e cito aqui a Mylena que foi uma grande parceira este ano.

Agradeço muito ao Marcos, meu orientador, pela confiança, por sempre ter me ajudado em tudo que eu precisei, sua orientação sem dúvida engrandeceu muito este trabalho. Aos servidores da Superintendência de Avaliação Educacional e da Diretoria de Avaliação dos Sistemas Educacionais, sobretudo ao Magno, meu tutor, por me receberem tão bem no estágio, por incitarem minha curiosidade sobre avaliações educacionais e por todo apoio que me deram. Ao Thiago e aos professores Juliana e Bruno, por toda a disponibilidade em tirar minhas dúvidas.

RESUMO

Esta monografia avalia a eficácia das escolas de ensino médio mineiras em 2019. A escola eficaz é aquela que consegue agregar ao aluno mais do que seria esperado, assim o objetivo do trabalho é identificar fatores intraescolares associados a um aumento do desempenho médio (dimensão da qualidade) das escolas de ensino médio mineiras. Além disso, é evidente o impacto das características do aluno em sua proficiência, dessa forma, objetiva-se analisar a distribuição dessas desigualdades de resultado e verificar, dos fatores promotores de qualidade, quais também são capazes de reduzir o efeito das desigualdades (dimensão da equidade), em específico as geradas pelas condições socioeconômicas. Para tanto, baseia-se na literatura de eficácia escolar a fim de definir os conceitos e contextualizar a pesquisa. São utilizados dados de proficiência em Língua Portuguesa e Matemática do 3º ano EM e o questionário contextual do PROEB 2019, bem como do Censo Escolar 2019. A análise é feita por meio de regressão hierárquica linear e considera tanto as características dos discentes quanto os processos escolares. Conclui-se que, as escolas de fato fazem diferença, mas a maior variabilidade é do nível dos alunos, inclusive foram identificadas desigualdades por condição social, gênero, raça e trajetória escolar. Para os dados analisados, os processos escolares, em especial os associados aos professores, são eficazes em aumentar o desempenho médio dos estudantes, mesmo quando controlado por suas características pessoais. No entanto, em geral, estes processos não são capazes de promover equidade, ou seja, reduzir o efeito de índice socioeconômico.

Palavras chave: eficácia escolar, qualidade educacional, equidade intraescolar, desigualdades educacionais, índice socioeconômico, regressão hierárquica linear.

ABSTRACT

This study assesses the effectiveness of high schools in Minas Gerais in 2019. The effective school is the one that can add to the student more than would be expected, so the objective of the work is to identify intraschool factors associated with an increase in the average educational attainment (quality dimension) of high schools in Minas Gerais. In addition, it is evident the impact of the students' characteristics on their attainment, thus, the objective is to analyze the distribution of these educational inequalities and verify, of the quality promoting factors, which are also able to reduce the effect of inequalities (equity dimension), especially those generated by socioeconomic conditions. To this end, it is based on the literature of school effectiveness in order to define the concepts and contextualize the research. It is used Portuguese and Mathematics performance in 3rd year of high school and the contextual questionnaire of PROEB 2019, as well as the Censo Escolar 2019. The analysis is performed by means of linear hierarchical regression and considers both the characteristics of the students and the school processes. It is concluded that schools actually make a difference, but the greatest variability is at the level of students, including inequalities by social condition, gender, race and school trajectory. For the data analyzed, school processes, especially those associated with teachers, are effective in increasing the average attainment of students, even when controlled by their personal characteristics. However, in general, these processes are not able to promote equity, that is, reduce the effect of socioeconomic index.

Keywords: school effectiveness, educational quality, intraschool equity, educational inequalities, socioeconomic index, linear hierarchical regression.

SUMÁRIO

1	INTRODUÇÃO	8
2	EFICÁCIA ESCOLAR: tecendo conceitos, medidas e achados	11
2.1	Eficácia escolar: conceito e estudos iniciais.....	11
2.2	A diferença das escolas para a eficácia escolar.....	14
2.3	Estudos sobre eficácia escolar no Brasil.....	20
2.4	O panorama educacional brasileiro.....	23
3	O MODELO DE ANÁLISE DOS DETERMINANTES INTRAESCOLARES DA EFICÁCIA ESCOLAR EM MINAS GERAIS	34
3.1	Origem e técnicas de sistematização dos dados	34
3.2	Dados e modelo de análise.....	37
3.3	Abordagem analítica.....	42
4	AS ESCOLAS MINEIRAS FAZEM DIFERENÇA NO APRENDIZADO DE ESTUDANTES DO ENSINO MÉDIO?	46
4.1	Análise descritiva dos dados	46
4.2	A relação entre origem social e qualidade no ensino médio	56
4.3	A diferença da escola na proficiência do ensino médio e na promoção da equidade.....	66
4.4	A (in)eficácia das escolas mineiras de ensino médio	83
5	CONSIDERAÇÕES FINAIS	86
	REFERÊNCIAS.....	90
	APÊNDICE A – Resultados da análise fatorial.....	95
	APÊNDICE B – Análise dos dados ausentes	98
	APÊNDICE C – Sintaxes.....	101
	Anexo A – Questionário contextual PROEB	133
	Anexo B – Variáveis do Censo Escolar	138

1 INTRODUÇÃO

Os estudos na área da educação buscam evidências que contribuam para orientação de políticas e alocação de recursos públicos. O Relatório Coleman, publicado nos Estados Unidos em 1967, foi um desses estudos. Teve particular relevância, pois junto a pesquisas semelhantes da época evidenciou como as características pessoais dos alunos determinam em grande medida seus resultados, o que ratifica as persistentes desigualdades e aumenta os desafios das políticas educacionais na promoção da equidade.

Se, por um lado, estas evidências descortinavam que “as escolas não faziam diferença”, por outro, um conjunto de pesquisas se orientaram a entender como os processos escolares influenciam o alcance educacional. Inscritos nessa orientação, surgem os estudos sobre eficácia escolar, tema desta monografia. Eles evidenciam que que escolas tem sim um papel importante no desempenho dos alunos e que este não pode ser relativizado.

O desenvolvimento da pesquisa internacional permitiu consolidar os fatores correlatos da eficácia escolar, associando diversos processos escolares com a qualidade do aprendizado dos estudantes, o que foi de grande valia para os estudos nacionais a partir da década de 1990, concomitante às primeiras avaliações em larga escala no país. Hoje se sabe como as relações entre diretores, professores e estudantes ocorrem em nossas escolas, quais fatores escolares determinam o desempenho e quais podem diminuir as desigualdades educacionais ainda manifestas.

A partir dessa problemática, o presente trabalho busca responder, a partir de uma perspectiva regional, em que medida as escolas públicas mineiras foram eficazes no 3º ano do EM no ano de 2019. A análise da eficácia escolar se baseia na perspectiva da qualidade e da equidade educacionais, sendo a primeira entendida como desempenho cognitivo em Língua Portuguesa e Matemática, seguindo pesquisas anteriores, e a segunda entendida como “igualdade de capacidades”, à luz de Mokate (2001). Sobre estas perspectivas, reconhece-se, de saída, que a qualidade e os resultados educacionais não se restringem apenas ao aprendizado, mas dada as informações disponíveis para operacionalização da pesquisa, limitar-se-á a este construto. Além disso, o conceito de equidade não é consensual, mas está implicado à superação das desigualdades de oportunidades educacionais, ou seja, à diminuição

dos efeitos das características adscritas e condições sociais dos alunos em seu alcance educacional.

Assim, o objetivo geral desta pesquisa é analisar a eficácia das escolas de ensino médio da rede estadual de Minas Gerais, a partir da relação entre processos escolares, origem social e desempenho dos estudantes. Como objetivos específicos, o trabalho estabelece: i) construir um modelo de análise de eficácia escolar, a partir da revisão da literatura sobre o tema e dos dados disponíveis para a série e o ano selecionados; ii) analisar, num primeiro nível, as características de origem que determinam o desempenho médio dos estudantes; iii) analisar, em um segundo nível, os efeitos dos fatores intraescolares sobre a qualidade do aprendizado; iv) analisar quais desses fatores também afetam a equidade intraescolar, medida pelo impacto do índice socioeconômico no desempenho dos discentes.

É crucial para o sistema de ensino garantir o aprendizado de todos os estudantes independente de suas características pessoais. Portanto, este trabalho busca reforçar que a desigualdade das condições adscritas e socioeconômicas do alunado culmina em possibilidades díspares de usufruir das oportunidades educacionais e, portanto, deve ser foco de políticas públicas de redução das desigualdades. O trabalho também é relevante por utilizar dados do Programa de Avaliação da Rede Pública de Educação Básica (PROEB) de Minas Gerais ainda pouco explorada nas pesquisas do campo. Por fim, ele acrescenta ao discutir o impacto dos fatores intraescolares na equidade, relação ainda pouco tematizada na pesquisa de eficácia escolar brasileira. Assim, espera-se contribuir acadêmica e socialmente para o enfrentamento de problemas da educação pública mineira.

Serão utilizados os dados do PROEB de 2019 e do Censo Escolar 2019. As variáveis analisadas são dos alunos, seus resultados educacionais, proficiência em Língua Portuguesa e Matemática, e suas características pessoais e as das escolas que mensuram os seus processos e as relações entre os atores. A análise dos dados será feita por meio de regressão linear hierárquica a fim de identificar o impacto de cada variável na proficiência, verificar relações entre as variáveis e analisar como isso difere a cada escola.

Dessa maneira, este trabalho é composto por cinco capítulos, sendo o primeiro esta introdução. O segundo capítulo trata do campo de pesquisa em eficácia escolar, seus modelos conceituais, principais conclusões e o contexto do cenário educacional brasileiro. O terceiro capítulo apresenta a base de dados, as variáveis e

sua operacionalização, além da abordagem analítica utilizada. O capítulo quatro apresenta os resultados da pesquisa, evidenciando a eficácia dos processos escolares. Por fim, apresentam-se as considerações finais do estudo.

2 EFICÁCIA ESCOLAR: TECENDO CONCEITOS, MEDIDAS E ACHADOS

Esse capítulo tem como objetivo contextualizar o marco teórico do estudo e o cenário educacional brasileiro e definir os conceitos utilizados na pesquisa. Introduce-se um breve histórico da pesquisa em eficácia escolar, suas concepções metodológicas e questões de interesse. Em seguida, é introduzido ao modelo conceitual que fundamenta esta monografia, apresentam-se os principais achados da pesquisa nacional sobre a temática e, por fim, descreve-se a conjuntura nacional no âmbito da qualidade e equidade do ensino.

2.1 Eficácia escolar: conceito e estudos iniciais

Eficácia escolar refere-se à capacidade da escola em cumprir suas funções a partir da conclusão dos objetivos e metas fixados para ela. Inicialmente, essa medida de eficácia era a média de desempenho da escola em testes padronizados de habilidades básicas, mas reconheceu-se então que o conceito abrange, de forma mais ampla, medidas cognitivas e não-cognitivas múltiplas (BROOKE, 2010).

A definição de eficácia escolar nem sempre é tão convergente, na medida em que os estudos assumem objetivos e abordagens diversos como Economia, Administração, Sociologia da Educação e Pedagogia (SOARES, 2004b). Este estudo será baseado na concepção de Mortimore (1991 apud BROOKE, 2010), mais usual dentre os estudos com abordagem mais próximas à Sociologia da Educação. Define a escola eficaz como aquela que “faz seus alunos progredirem mais do que se esperaria, levando em consideração seu nível socioeconômico e ponto de partida em termos de desempenho”. Salienta-se ainda que a eficácia escolar será analisada pela perspectiva tanto da qualidade, quanto da equidade.

Na pauta da democratização da educação e da melhora dos seus índices de qualidade, foram desenvolvidos grandes *surveys* educacionais nos anos 1950 e 1960 (BARBOZA, 2006). O Relatório sobre Igualdade de Oportunidades Educacionais ou Relatório *Coleman*, como ficou conhecido, foi um grande marco da pesquisa educacional. Refere-se a um *survey*, encomendado pelo Congresso Americano e lançado em 1966, que observou os resultados de 645 mil alunos em mais de três mil estabelecimentos escolares norte-americanos para analisar a desigualdade de oportunidades educacionais entre os indivíduos por distinção do sexo, da raça, da religião ou da região de origem das instituições públicas de ensino (COLEMAN et al.,

1966). Os principais objetivos da pesquisa eram avaliar: a) o grau de segregação dos grupos raciais e culturais da escola pública; b) a infraestrutura e os recursos disponíveis; c) o quanto os alunos aprendiam na escola por meio de um teste padronizado e d) a relação entre os resultados e os insumos e recursos presentes na escola (MADAUS et al., 1980).

Os resultados do relatório apontaram que as escolas têm uma influência muito pequena no desempenho acadêmico dos alunos, sendo este, então explicado principalmente pelo contexto socioeconômico e antecedentes sociais dos alunos e suas famílias (COLEMAN et al., 1966). Os autores afirmam: "a escola não acrescenta muito ao desempenho de uma criança, independentemente do seu meio e contexto social geral" (COLEMAN et al., 1966 apud MADAUS et al., 1980, p.75). O relatório gerou muita polêmica e foi considerado um grande marco da pesquisa sociológica, não apenas pelos seus resultados, mas também pela extensão e tecnologia da pesquisa e pelo ineditismo das informações sobre as relações sociais dos Estados Unidos (BROOKE; SOARES, 2008).

Nesse contexto, vários outros estudos contemporâneos foram desenvolvidos e seus resultados corroboram a forte relação identificada por Coleman et al. (1966) entre desempenho acadêmico e origem social. Dentre eles, o Relatório Plowden, de 1967, solicitado pelo Ministério da Educação do Governo Britânico. Este estudo analisou o desempenho escolar de alunos do ensino primário com uso de uma metodologia semelhante ao do norte americano, além de incluir variáveis psicossociológicas, por exemplo, as atitudes dos pais em relação ao trabalho e ao futuro escolar dos filhos. Concluiu que a família, em especial as diferenças comportamentais dos pais, explicam mais as variações entre as crianças do que as diferenças entre as escolas (CACE, 1967; ALVES, 2006).

Ressalta-se também, um estudo demográfico longitudinal, conduzido pelo Instituto Nacional de Estudos Demográficos (INED) na França. Foram acompanhadas as trajetórias escolares de 17 mil alunos a partir do término do ensino primário durante dez anos (1962 - 1972) com o intuito de avaliar as desigualdades de acesso aos estudos. A pesquisa demonstra que havia maior probabilidade de exclusão dos filhos de categorias de menor prestígio social à medida que eles avançam nos níveis de ensino (BARBOZA, 2006; ALVES, 2006).

Tais pesquisas foram congruentes em demonstrar que fatores extraescolares explicavam mais as variações de desempenho dos alunos que fatores

intraescolares. Nos Estados Unidos, publicações posteriores ressaltaram os mesmos resultados (MOSTELER; MOYNIHAN, 1972; JENCKS, 1972). Essas conclusões geraram grande pessimismo àqueles que acreditavam que a escola era importante para o aprendizado e acabaram por ser sintetizadas na expressão “As escolas não fazem diferença” (SOARES, 2004b). Porém, é preciso destacar que a frase é uma conclusão simplista e mal interpretada. A escola faz muita diferença para o aluno, o que os relatórios demonstram, no entanto, é que, em relação às dimensões incluídas, as escolas não são muito diferentes entre si (BROOKE; SOARES, 2008), entendimento que também será questionado posteriormente.

Essas interpretações e críticas às pesquisas foram fundamentais para o desenvolvimento das pesquisas na área. A polêmica impulsionou vários estudos que tentavam identificar se as escolas realmente fazem diferença (EDMONDS, 1979; MORTIMORE et al. 1988; RUTTER et al. 1979). Dessa forma, constituiu-se a área de estudos de eficácia escolar e se desenvolveram importantes discussões sobre a equidade educacional e sobre como os fatores escolares e extraescolares interferem no desempenho acadêmico.

As principais críticas às pesquisas iniciais referem-se ao desenho teórico metodológico que as sustentou, o modelo de análise econômica *input-output* (BROOKE; SOARES, 2008). Este mensura a quantidade de recursos disponíveis na escola – insumos, como investimento econômico, contratação de professores, qualidade dos materiais, currículo, entre outros – e compara com os resultados obtidos nas avaliações padronizadas (MOSTELER; MOYNIHAN, 1972). Dessa forma, a metodologia baseia-se em uma conexão direta entre insumos e resultados, não levando em conta os processos escolares, ignorando assim, as especificidades de cada escola, suas características sociais e culturais. É dito então que os processos internos foram desconsiderados como se fossem uma “caixa-preta” a qual os pesquisadores não são capazes de analisar (BROOKE; SOARES, 2008).

Destaca-se ainda a limitação teórica dos estudos, a crítica é à ausência de uma teoria sólida que verificasse os diversos estudos empíricos realizados (VAN DEN EEDEN; HOX; HAUER, 1990). Carvallo-Pontón (2010) acrescenta que tais pesquisas utilizavam técnicas de regressão linear e, muitas vezes, havia alta colinearidade¹ entre

¹ A colinearidade existe quando há uma forte correlação entre dois ou mais precursores em um modelo de regressão e afeta o modelo por aumentar o erro padrão dos coeficientes e diminuem a significância das variáveis (FIELD, 2009).

as variáveis. Os resultados dessas pesquisas, então, não eram tão congruentes e identificavam efeitos fracos dos insumos devido tanto ao modelo utilizado – não considera os efeitos entre os níveis (escola e aluno, por exemplo) – quanto à escolha das variáveis, limitaram-se à fatores fáceis de serem medidos e não abrangeram muitos fatores contextuais (WILLMS, 1992).

Nesse sentido, no final da década de 1970 e na década de 1980 surgem novos estudos interessados em reavaliar os modelos teóricos e metodológicos e compreender os processos escolares, abrindo-se a “caixa-preta” das escolas. Destacam-se estudos como “Fifteen Thousand Hours” (RUTTER; MAUGHAN; MORTIMORE; OUSTON; SMITH, 1979), “Schools can make a difference” (BROOKOVER et al., 1979), “Effectiveness: a Reassessment of the Evidence” (MADAUS; AIRASIAN; KELLAGHAN, 1980) e “School Matters” (MORTIMORE; SAMMONS; STOLL; LEWIS; ECOB, 1988). Esses trabalhos evidenciaram que a variação entre as escolas era significativa e não poderia ser absolutizada na média (ALVES, 2006). Assim, identifica-se que a nova abordagem metodológica (entrada, processo, saída) revela as diferenças internas entre as escolas, portanto, as pesquisas passam a tentar compreender o que leva algumas escolas terem melhores resultados que outras para alunos de uma mesma origem social (LEE, BRYK; SMITH, 1993).

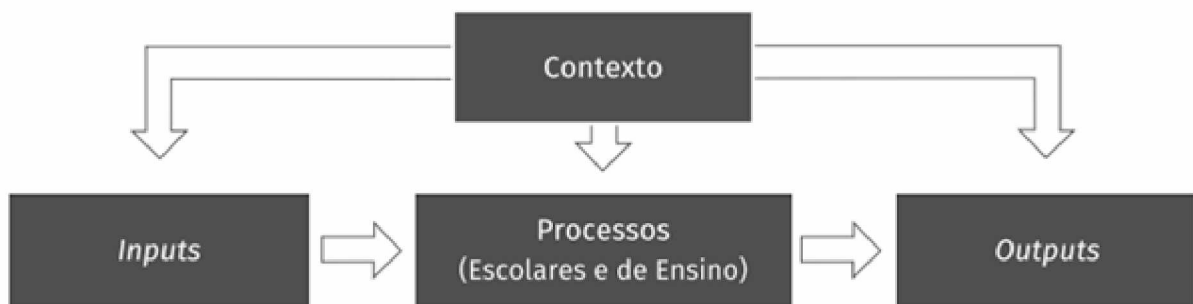
Dessa forma, compreende-se o Relatório Coleman como importante marco da pesquisa educacional, esta e outras pesquisas à época reforçaram a importância das condições socioeconômicas e culturais dos estudantes e suas famílias nos resultados educacionais. Tão importante também foram as pesquisas que surgiram em resposta ao entendimento de que “as escolas não fazem diferença”, dando origem ao campo de Eficácia Escolar. Escolas eficazes podem ser entendidas por aquelas que são capazes de cumprir seus objetivos educacionais, entendendo a influência dos processos escolares em tais objetivos. A subseção seguinte apresenta os modelos conceituais que passam a ser desenvolvidos no registro da eficácia escolar e que servem de referência para a presente pesquisa.

2.2 A diferença das escolas para a eficácia escolar

A partir da concepção de que as escolas poderiam sim fazer diferença, novos modelos teóricos foram desenvolvidos conforme será apresentado nessa subseção. Scheerens (1990) propõe o modelo de referência para os fatores que

exercem influência nos resultados educacionais e as categoriza por: contexto, insumos e processo, como demonstra a figura 1. Os fatores contextuais são aqueles que impactam o processo educativo, mas que a escola possui pouco ou nenhum gerenciamento sobre eles, como gestão administrativa, localização (rural ou urbana) e tamanho da escola. Os insumos são as entradas para o processo educativo, por exemplo a experiência e a formação do professor, a experiência do diretor e o suporte familiar. Os processos são os fatores intraescolares que permeiam a educação dos alunos e podem ser do nível da turma, como o estilo de ensino do professor ou a sua expectativa sobre os alunos, ou do nível da escola, tais quais a gestão escolar e o currículo. O resultado é o que se espera da escola, a formação dos alunos, em geral diz respeito ao desempenho escolar do aluno, controlado por fatores como o nível socioeconômico e a inteligência (KARINO; LAROS, 2017).

Figura 1 – Modelo integrado de avaliação escolar.



Fonte: Scheerens, 1990 apud Karino; Laros, 2017, p. 100.

As pesquisas mais usuais utilizam exclusivamente dos resultados dos testes padronizados de leitura e matemática (SAMMONS; HILLMAN; MORTIMORE, 1995). Não obstante, há autores que utilizaram outras medidas como RUTTER et al. (1979) que também avaliaram dimensões socio afetivas. A formação dos alunos de fato vai além dos aspectos cognitivos, no entanto, por conta da ausência de medidas sistemáticas de outros fatores que englobam a formação escolar, as pesquisas tendem a usar apenas o fator cognitivo de componentes curriculares específicos (KARINO; LAROS, 2017).

Um dos intuitos dos estudos em eficácia escolar é identificar fatores que explicam a variação de desempenho das escolas, controlando estatisticamente o contexto e condições socioeconômicas e culturais da escola e dos alunos. A partir disso, foi possível estabelecer um certo consenso de qual era o perfil das escolas

eficazes. Passou a ser comum descrever listas que enumeram essas características, estas foram se complexificando e tornando menos consensuais, na medida do desenvolvimento das pesquisas e do reconhecimento a importância do contexto para a formação dos fatores escolares (BROOKE, 2010).

No âmbito internacional, um importante estudo conduzido por Pam Sammons (1999), requisitado pelo serviço de inspeção escolar do governo britânico, fornece um resumo das evidências das pesquisas de eficácia escolar e a partir disso identifica 11 fatores-chave (ou correlatos) de eficácia². O estudo enfatiza que a lista não tem pretensão de ser exaustiva, apenas fornecer caminhos relevantes a partir das evidências anteriores, cada qual com contextos próprios e específicos. Além disso, destaca que existem várias associações entre os fatores que podem ajudar a compreensão de prováveis mecanismos de eficácia.

Os 11 fatores identificados nessa análise enfocam aspectos associados aos processos da escola como um todo (liderança, tomada de decisões, administração, metas expectativas e assim por diante) e aqueles associados ou diretamente relacionados, à organização da sala de aula e do ensino. (...) Dado isso, os processos escolares, incluindo a liderança profissional, permanecem altamente influentes, pois fornecem a estrutura global dentro da qual os professores e as salas de aula operam. (SAMMONS, 1999/2008; p. 378)

No Brasil, as pesquisas em eficácia escolar iniciam-se mais tarde, na década de 1990 com a internacionalização do campo e integração de resultados (REYNOLDS et. al, 2011, apud LAROS; KARINO, 2017). Este período é concomitante com o início das avaliações em larga escala que foram essenciais principalmente para os primeiros estudos publicados na época. É importante ter em conta que a literatura internacional é de grande relevância para o Brasil e contribui para compreensão das estruturas internas das escolas associadas ao aprendizado do aluno. No entanto, enfatiza-se a necessidade de pesquisas que possam compreender as peculiaridades locais levando em conta a realidade social do país (BROOKE; SOARES, 2008). Essa pesquisa, por focar no estado de Minas Gerais, usando os dados do PROEB, que

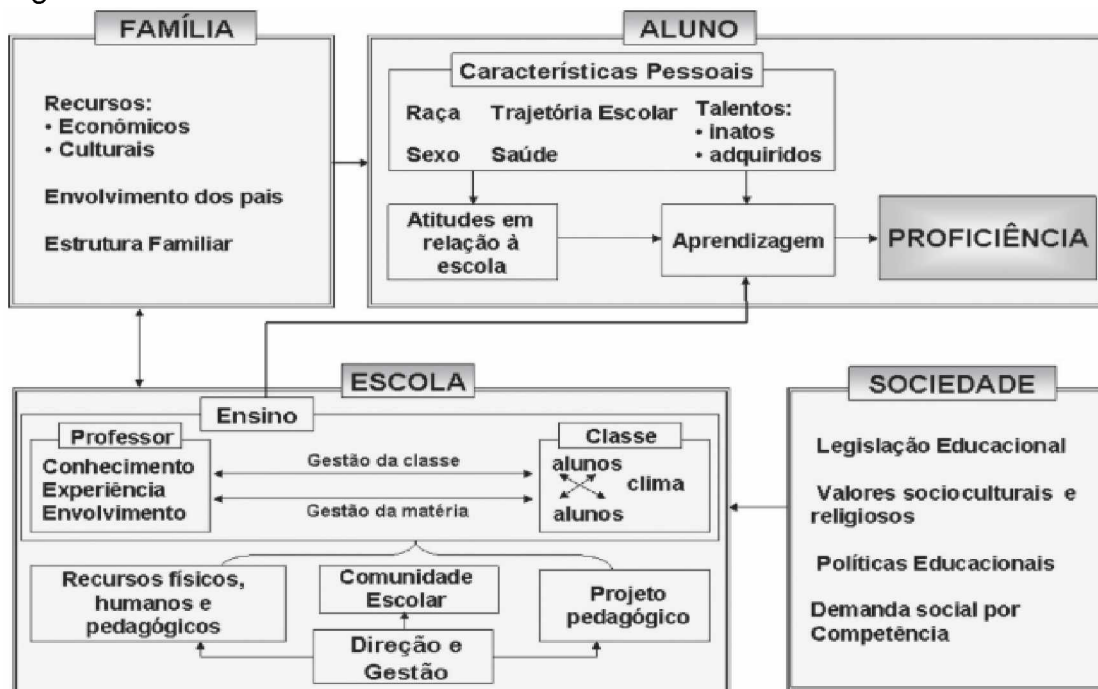
² 11 fatores das escolas-eficazes elencados por Sammons: 1- Liderança profissional; 2 – Objetivos e visões compartilhadas; 3 – Um ambiente de aprendizagem; 4 – Concentração no ensino e na aprendizagem; 5 – Ensino e objetivos claros; 6 – Altas expectativas; 7 – Incentivo positivo; 8 – Monitoramento do progresso; 9 – Direitos e responsabilidades dos alunos; 10 – Parceria casa-escola; e 11 – Uma organização orientada para a aprendizagem.

compõe seu sistema de avaliação, contribui para o aprofundamento das particularidades locais.

Ainda que a pesquisa brasileira seja mais recente, desde o início apropriou-se do rigor metodológico que a pesquisa internacional foi desenvolvendo ao longo dos anos. Ressalta-se aqui o reconhecimento da estrutura hierárquica do sistema educacional, ou seja, com os alunos agrupados em salas de aulas, que estão reunidas dentro de escolas e que, por sua vez, podem ser agrupadas em sistemas de ensino ou regiões geográficas. Estudos que tratam os indivíduos de modo independente e desconsiderando a interação desses grupos podem acabar por inflar os resultados (GOLDSTEIN, 2010; HOX, 2010). Além disso, é feito também o controle das análises pelas características pessoais dos alunos, como nível socioeconômico e cultural, raça, sexo e proficiência prévia, na medida em que, sem esse controle a análise tende a produzir resultados equivocados dado à alta correlação entre boas condições escolares e o alto nível socioeconômico médio das escolas. (FRANCO et al, 2007).

O modelo conceitual utilizado nesse estudo é de Soares (2004a), apresentado na figura 2, ele busca demonstrar como os fatores intra e extraescolares estão associados ao desempenho cognitivo dos alunos. Esses fatores são organizados em quatro grupos: características individuais do aluno; características familiares; características escolares; e características da sociedade. Sublinha-se que o modelo busca demonstrar a complexidade relacionada ao desempenho, para tanto, há de se levar em conta os efeitos tanto diretos quanto indiretos de cada fator, bem como a relação entre eles.

Figura 2 – Modelo conceitual dos fatores associados aos resultados educacionais



Fonte: Soares, 2004a, p. 85

Em primeiro lugar, o autor explicita as influências externas, fatores que limitam a atuação das escolas, sobre os quais ela não tem nenhuma ou pouca influência, como, por exemplo, as leis e regulamentos do sistema de ensino as quais a escola está sujeita ou a quantidade de alunos que deve atender. A localização geográfica da escola, seu contexto social, é um importante determinante externo, pois reflete a condição social dos alunos e indiretamente a disponibilidade de recursos da escola. Se as características pessoais dos alunos são tão cruciais, como demonstrou o Relatório Coleman (1966), quem estuda na escola, de certo, é um ponto que determina em grande medida o nível de desafio a ser enfrentado por ela. Outras questões, como o nível de conservação dos prédios, a qualificação dos professores e o índice de violência são importantes e estão desigualmente distribuídos pelo território. Estes são fatores que devem ser tratados fora dos muros das escolas, portanto, não serão objeto desse estudo, mas registra-se sua importância por interferirem e, por vezes, limitarem o aprendizado.

Ainda assim, existem fatores, mesmo os estruturais, que são tangíveis ou internos às escolas. Sobre os próprios recursos, mais importante que sua disponibilidade é sua utilização efetiva. Por exemplo, o autor aponta o livro didático, que além de disponível deve ter seu uso incorporado pelo professor na rotina da classe, servindo ao aluno como referencial. A relação com a comunidade é mais um

fator que não depende apenas da escola, mas influencia positivamente o desempenho, sendo relevante promoção dessa participação pela escola e a transparência de sua atuação. Ressalta-se ainda a gestão, na qual o diretor tem papel fundamental, legitimidade social e competência são destacadas. Ele deve ser capaz de induzir uma unidade de propósito na escola, com objetivos claros e bem definidos, desenvolvendo uma gestão colaborativa com o envolvimento dos professores. Outras características como, legitimidade social e competência do diretor também são relacionadas ao desempenho.

Associado a isso, Soares (2004a), citando outros autores, destaca também as relações sociais na escola, o que ele chama de efeitos contextuais, que podem ser positivos no desempenho dos alunos e principalmente para redução da diferença entre diferentes grupos de alunos (LEE; SMITH, 2001; BRYK, DRISCOLL, 1988 apud SOARES, 2004a). Trata-se da cultura da escola, do que se constrói sobre o papel da instituição, sobre os padrões de comportamento esperados e as expectativas quanto aos resultados dos alunos. Estes valores resultam em uma agenda que pode contribuir não só para o cumprimento desses objetivos esperados, como também para relação entre os atores e para a criação de estruturas comunitárias na escola. As interações sociais são muito importantes, em especial a relação entre os professores e deles com os alunos. Esta se dá principalmente na sala de aula, onde um ambiente adequado ao aprendizado e um clima de ordem estão na base de uma boa relação e são requisitos para o desenvolvimento adequado dos discentes.

Nesse modelo o professor tem um papel central, é preciso que ele seja qualificado, tenha experiência, comprometimento e um bom relacionamento com os alunos. Soares (2004a) citando Gauthier (1997) divide a atuação do professor em dois principais aspectos: a gestão da classe e a gestão da matéria. A primeira é menos técnica e mais complexa, refere-se a postura do professor, seu envolvimento, a forma que se porta aos alunos e sua expectativa sobre eles. Já a gestão da matéria está relacionada com as ferramentas didáticas que o professor utiliza para mediar os conhecimentos de forma mais efetiva, está muito relacionada a formação pedagógica do professor. Por fim, o autor explicita o projeto pedagógico, em especial as decisões acadêmicas, como o tamanho da classe, a forma como os alunos são alocados às turmas, a distribuição do uso do tempo escolar, a escolha e efetivação do currículo e a realização de avaliações.

O desenho desse modelo foi feito com base nas evidências internacionais e nacionais, inclusive na lista de Sammons (1999) apresentada anteriormente. Soares (2004a) consegue sistematizar os principais fatores que influenciam o desempenho considerando uma realidade nacional. Fica claro que se trata de uma problemática muito complexa e que o melhoramento educacional dispenderá de muito esforço. É importante que estas mudanças aconteçam nas escolas, pois é onde de fato ocorre o processo de ensino e aprendizagem e para que sejam efetivas precisam partir da escola. (BROKE; SOARES, 2008). A partir dessa construção teórica sobre quais fatores intra e extraescolares determinam o desempenho, passa-se à análise das evidências empíricas encontradas para as escolas do Brasil.

2.3 Estudos sobre eficácia escolar no Brasil

É consenso entre as pesquisas o quanto as características pessoais dos alunos influenciam seu desempenho cognitivo, o sistema educacional brasileiro sustenta desigualdades de desempenho e de oportunidades de acesso a recurso. É demonstrado, por exemplo, discrepâncias nas proficiências dos alunos por gênero, raça e nível socioeconômico. O grande desafio para a promoção de uma educação mais equitativa é a garantia que os alunos tenham um resultado educacional satisfatório independente de suas condições sociais. No entanto, o que tem sido evidenciado é que os alunos mais favorecidos tendem a usufruir mais das melhorias das condições escolares, ou seja, os fatores escolares que promovem uma melhora no desempenho acadêmico também agravam as desigualdades educacionais (LAROS; KARINO, 2017).

Há de se constatar que a desigualdade educacional envolve múltiplas dimensões. Dentre eles, aborda-se aqui o que Sampaio e Oliveira (2015) denominaram de desigualdade de conhecimento. A partir disso, os autores definem a equidade educacional como aquela que garante a mesma probabilidade de bom desempenho a todos os alunos. Em um mesmo sentido, Mokate (2001) a define como igualdade de capacidades, ou seja, seria "justo" ou equitativo que os alunos desfavorecidos por suas características pessoais tivessem compensações que os garantisse a possibilidade de alcançar os mesmos resultados que os demais. É por isso que também busca-se averiguar se e em que medida essa equidade é concretizada nas escolas de ensino médio de Minas Gerais.

Soares e Alves (2003) analisaram dados do SAEB 2001 e perceberam uma diferença de desempenho entre alunos brancos e pardos e maior ainda entre alunos brancos e negros. Além disso, constataram que os fatores produtores de eficácia escolar favorecem principalmente o desempenho escolar de alunos com maiores índices socioeconômico e não contribuem para reduzir as discrepâncias entre grupos raciais. Fatores como melhor qualificação docente, melhores salários e melhores equipamentos têm efeito significativo no sentido acirrar as desigualdades educacionais.

Franco et al (2007) buscaram identificar, a partir de dados do SAEB 2001, características escolares que eram ao mesmo tempo promotoras de equidade e qualidade escolar, ou seja, relacionam-se com o aumento médio de desempenho e a redução do efeito do nível socioeconômico. No entanto, nenhum fator, ainda que associado ao aumento de qualidade, reduzia o efeito do nível socioeconômico sobre o desempenho. Em relação ao gênero, Andrade, Franco e Carvalho (2003) constataram que os meninos tendem a ter proficiências maiores em matemática, enquanto meninas tendem a ter maior proficiência em leitura. Essas diferenças variam por escola e são maiores quanto menor o nível socioeconômico médio dos alunos da escola e, quando controlado pelo nível socioeconômico médio, nenhum fator escolar analisado reduziu as diferenças entre gêneros.

Soares (2004b) identifica diferenças de desempenho no SAEB 2001 para alunos de diferentes regiões do país, cores de pele, posições sociais, sexo, raça e se possuem ou não atraso escola. No entanto, das variáveis analisadas nenhuma diminuía essas diferenças. Soares interpreta, dessa forma, que a promoção da qualidade deve ser feita de forma concomitante à promoção da equidade explicitando esse esforço na ação da escola, das famílias, da comunidade e dos governos.

Grande parte dos estudos em eficácia escolar busca identificar os fatores escolares associados ao alto desempenho educacional dos estudantes. Diferente da literatura internacional, a evidência brasileira tem mostrado que os recursos escolares importam. Franco, Sztajn e Ortigão (2007) e Soares (2004a) identificaram a relevância da conservação dos equipamentos e do prédio escolar. Do mesmo modo, Albernaz, Ferreira e Franco (2002), a partir de dados do SAEB de 1999, identificaram o efeito negativo da falta de recursos financeiros e pedagógicos da escola sobre a qualidade escolar, mesmo resultado encontrado por Franco, Sztajn e Ortigão (2007) e Soares (2004a) com dados do SAEB 2001. Essa diferença em relação a literatura

internacional, provavelmente se dá por uma especificidade do sistema de ensino brasileiro que é a grande desigualdade na distribuição de recursos escolares (FRANCO et al, 2007).

A gestão da escola também tem se mostrado relevante. Franco, Sztajn e Ortigão (2007) e Soares e Alves (2003), a partir de dados do SAEB 2001, relataram que o reconhecimento por parte dos professores da liderança do diretor é característica associada ao desempenho cognitivo escolar. Franco et al (2007), baseados nos mesmos dados, identificaram a importância da liderança colaborativa para a proficiência. A liderança pedagógica do diretor, sobretudo em sua atuação nas atividades escolares, no planejamento pedagógico e na promoção da melhoria dos resultados foi identificado como importante fator nos estudos de caso de Schattner (2012), Viana (2013), Sales (2015), Rodrigues (2015) e Martins (2015) todos citados por Januário (2019). Outras pesquisas também citadas pelo autor demonstram que a gestão escolar está relacionada a outros componentes pedagógicos importantes, reportam a pertinência de um planejamento bem definido e a divulgação dos resultados à comunidade (RANGEL, 2013; RODRIGUES, 2015; SCHATTNER, 2012 apud JANUÁRIO, 2019).

O clima escolar, ou seja, um ambiente ordenado e focado na aprendizagem dos alunos, também é um fator relatado pelas pesquisas. Trata-se principalmente da relação interna entre os atores escolares. Sobre este fator, dentre as pesquisas citadas por Januário (2019) as seguintes se destacam: Abramovay (2003) ressalta a importância desse fator, Martins (2015) identificou que regras de convivência claramente estipuladas e a incisiva ação da gestão para controlar a disciplina colaboram para um bom clima escolar, com centralidade para o papel do professor. Rangel (2012) também acrescenta a contribuição da valorização dos resultados e méritos dos alunos pelos atores escolares. Martins (2015), Santos (2010) e Viana (2013) ressaltam a dimensão da disciplina, principalmente na relação professor aluno.

Inclusive, o papel do professor, sobretudo na dimensão acadêmica, para o desempenho dos alunos é reiterado na literatura. Machado Soares (2003, 2005), analisando dados do SIMAVE 2002, identificou que as características do professor e do ambiente em sala de aula afetam decisivamente a proficiência dos alunos e explicam grande parte da variabilidade das notas. Destaca fatores como o interesse e a dedicação do professor e seu nível de exigência sobre os alunos, bem como passar e corrigir dever de casa. Esse elemento também foi discutido por Franco, Sztajn e

Ortigão (2007), em estudo baseado no SAEB 2001. Lee, Franco e Albernaz (2007), a partir de dados do PISA 2000 também relataram a importância da exigência docente para o desempenho escolar, sendo que absenteísmo docente, por outro lado, tem efeito negativo, não só para o desempenho como também para equidade escolar. Machado Soares (2003, 2005) também reportou efeitos negativos do absenteísmo do professor (FRANCO et al, 2007).

Toda essa revisão dos estudos brasileiros sobre eficácia escolar mostra que a escola é muito importante para os resultados educacionais. Os recursos escolares, como recursos financeiros e pedagógicos podem ser considerados relevantes no Brasil, por conta da desigualdade de distribuição existente. A gestão da escola, também é um fator importante, em especial em relação a atuação do diretor e aspectos como gestão participativa e planejamento definido. Ressalta-se também o clima escolar que diz respeito a relação interna entre os fatores escolares e o papel do professor, afinal é dentro da sala de aula que ocorre de fato os processos de ensino e aprendizagem.

Este estudo busca verificar se os fatores relacionados a essas dimensões, em específico os processos escolares, também são importantes no contexto mineiro. Para tanto, faz sentido analisar o contexto educacional que está sendo tratado, levando em conta as desigualdades existentes. Dessa forma, a próxima seção traz uma breve contextualização da educação brasileira.

2.4 O panorama educacional brasileiro

A educação de qualidade é assegurada a todos pela Constituição Federal, definida como aquela que garante "o pleno desenvolvimento da pessoa, seu preparo para o exercício da cidadania e sua qualificação para o trabalho" (BRASIL, 1988). Evidentemente tem um papel importante que impacta os âmbitos econômico e sociais, no entanto, verifica-se um quadro de fortes desigualdades educacionais, seja pela falta de acesso ao sistema escolar, pela exclusão dentro do próprio sistema ou mesmo pelo acesso a diferentes níveis de qualidade (SAMPAIO; OLIVEIRA, 2015).

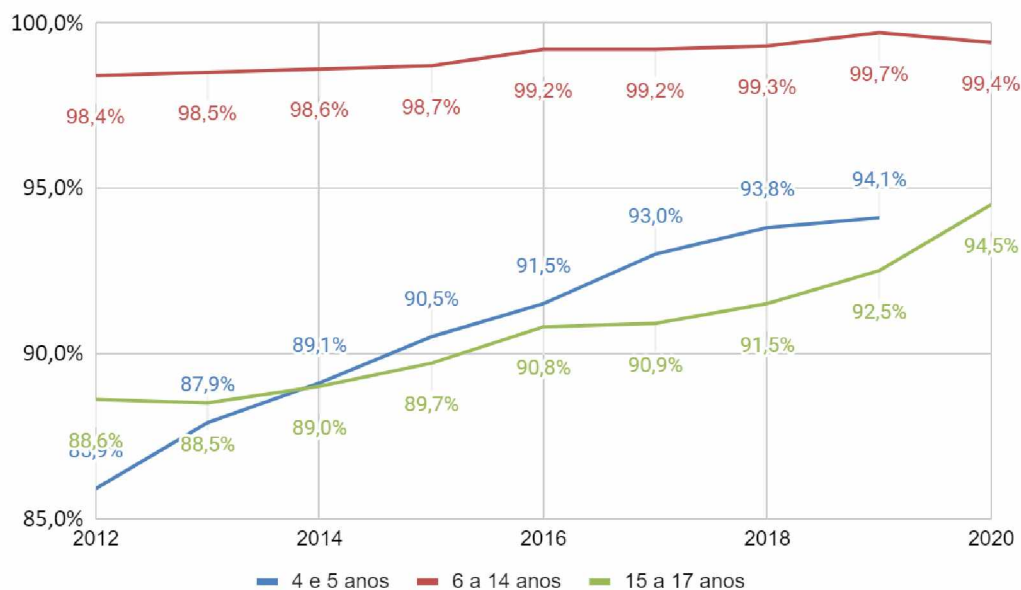
Destaca-se aqui que o conceito de qualidade educacional não é único ou consensual, para Oliveira e Araújo (2005) ele foi evoluindo na medida das condições de ensino de cada período, bem como os principais desafios em voga. A insuficiência da oferta foi o primeiro desafio, até a década de 1980, as políticas educacionais orientavam-se pela universalização das oportunidades educacionais e a qualidade era

entendida principalmente pelo acesso. Com sua expansão, a atenção foi dada as disfunções no fluxo, especialmente no ensino fundamental, então a qualidade é mensurada pela redução das taxas de repetência e evasão. A partir da expansão e consolidação dos testes padronizados em larga escala, nos anos 2000, a qualidade começa a ser entendida como os desempenhos cognitivos mensurados, perspectiva adotada por esse trabalho (OLIVEIRA; ARAÚJO, 2005; SAMPAIO; OLIVEIRA, 2015).

De fato, as concepções da qualidade educacional, são refletidas pelo contexto das condições de ensino. O Brasil inicia-se os anos 1980 após um rápido crescimento econômico que sustentou a expansão da oferta de ensino, o que não foi suficiente, no entanto, para que o país se destacasse em relação a América Latina. O que se observou nas duas décadas seguintes, devido aos investimentos públicos em políticas educacionais, ao aumento da renda familiar e à transição demográfica, foi um aumento médio dos anos de estudo da população e uma diminuição das diferenças por região, raça, gênero e condição econômica (SILVA; HASENBALG, 2000). A ampliação do acesso continuou nos anos seguintes, como é demonstrado pelo gráfico 1 que apresenta a evolução da taxa de atendimento por grupos de idade de 2012 a 2020.

Considerando a idade adequada, crianças de 4 e 5 anos devem estar na educação infantil, 6 a 14 no ensino fundamental e 15 a 17 no ensino médio. O gráfico evidencia que o acesso à escola das crianças de 6 a 14 anos está praticamente universalizado, mesmo em 2012, a pior taxa do período, 98,4% estavam na escola. As outras faixas de idade apresentam um desafio maior, passaram a ter um acesso obrigatório um pouco mais tarde, em 2016, ainda assim, apresentam um aumento progressivo das taxas de atendimento, principalmente as crianças da educação na educação infantil. Entre elas a taxa de atendimento era de 85,9% em 2012 e aumentou 8,2 pontos percentuais até 2019 (94,1%), para a faixa dos 15 a 17 anos o aumento foi de 5,9 pontos até 2020, tendo uma taxa de 94,5% em 2020.

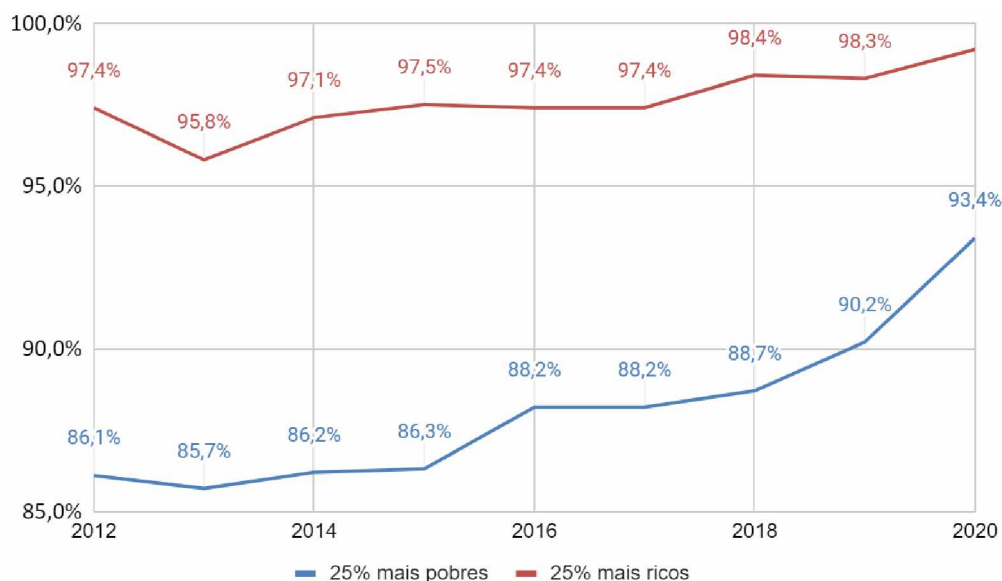
Gráfico 1 – Porcentagem de crianças entre 4 e 17 anos na escola entre 2012 e 2020 no Brasil – Taxa de Atendimento.



Fonte: Dados: Observatório do Plano Nacional de Educação, 2020.
Elaboração própria, 2020.

O gráfico acima mostra um cenário promissor, ainda assim, é importante fazer algumas ressalvas. Em primeiro lugar, consta dizer que as condições de oferta não são equânimes; o gráfico 2 mostra a taxa de atendimento entre jovens de 15 a 17 anos com uma distinção entre os 25% mais ricos e os 25% mais pobres, se por um lado fica claro que a política educacional tem um atendimento desigual, do mesmo modo, nota-se uma redução dessas discrepância a partir de 2016. Enquanto em 2012 os jovens do grupo 25% mais ricos tinham uma taxa 11,3% maior, já em 2020 a diferença diminuiu para 5,8% graças ao aumento do acesso dos 25% mais pobres.

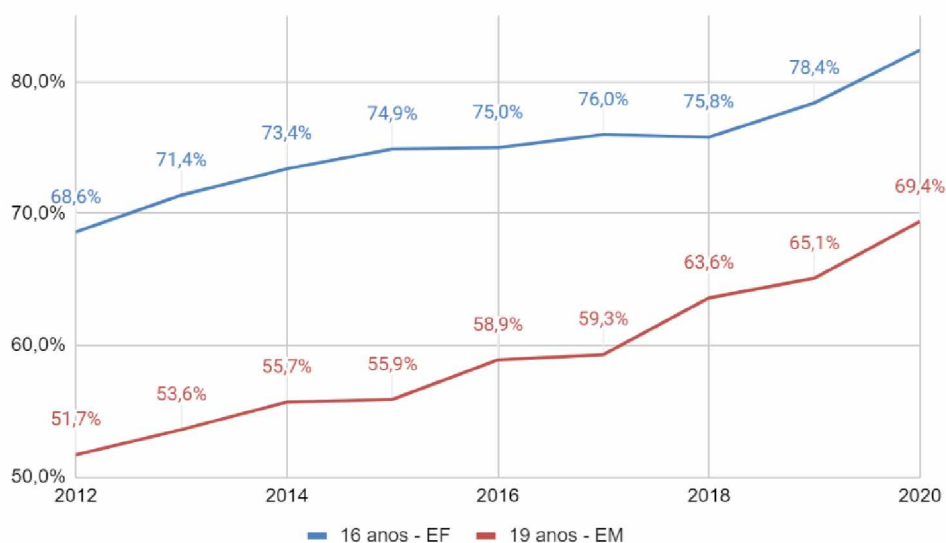
Gráfico 2 – Porcentagem de jovens entre 15 e 17 anos na escola entre 2012 e 2020 no Brasil por quartil de renda familiar per capita– Taxa de Atendimento



Fonte: Dados: Observatório do Plano Nacional de Educação, 2020.
Elaboração própria, 2020.

Em segundo lugar, é preciso reforçar a evolução do conceito de qualidade; na medida da evolução do acesso ao ensino, novas questões passam a ser consideradas com mais atenção. O fato de a criança estar na escola não é o bastante, é preciso que ela esteja na etapa adequada à sua idade e continue na escola até a conclusão da formação básica, ou seja, o fluxo adequado e a permanência também precisam ser considerados. O gráfico 3 apresenta a porcentagem de jovens de 16 anos que concluíram o Ensino Fundamental e de jovens de 19 anos que concluíram o Ensino Médio. Percebe-se que ainda que quase todos os alunos de 6 a 14 anos estejam na escola, em 2020, apenas 82,4% deles havia concluído o nível de ensino e quanto ao Ensino Médio, só havia sido concluído por 69,4% dos jovens de 19 anos. Ainda assim, a perspectiva é positiva dada a melhora progressiva das taxas desde 2012.

Gráfico 3 – Porcentagem de jovens de que concluíram o nível de ensino na idade correta de 2012 a 2020.

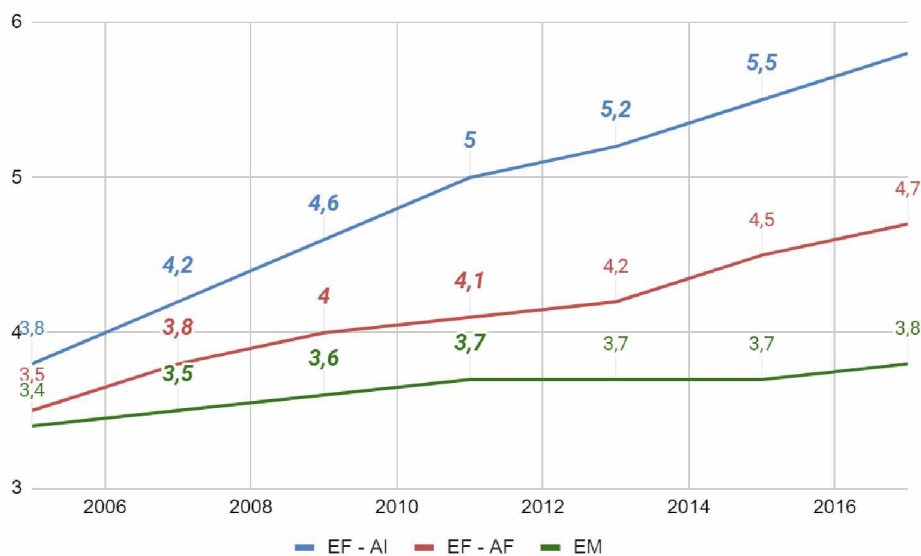


Fonte: Dados: Observatório do Plano Nacional de Educação, 2020.
Elaboração própria, 2020.

Dessa forma, para além desses resultados passou-se a verificar também o desempenho dos alunos. Um importante indicador é o índice de desenvolvimento da educação (IDEB), mensurado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) a partir das dimensões do fluxo e do desempenho médio. Desde 2007, esse indicador é utilizado como medida de qualidade pelo poder público, foram estabelecidas metas bianuais para cada escola para que em 2022 a média nacional seja equiparada a dos países desenvolvidos.

O gráfico 4 apresenta a evolução do IDEB desde 2005, esclarece-se que ele é calculado em uma escala de 0 a 10 para cada nível de ensino e nas edições em que a média nacional atingiu a meta os valores no gráfico foram colocados de negrito. Verifica-se que desde o início conseguiu-se manter as metas estabelecidas para o Ensino Fundamental Anos iniciais, no entanto, para os Anos Finais e para o Ensino Médio as metas não são cumpridas desde 2011. No caso do Fundamental, para os dois níveis observa-se uma melhora do indicador, quanto ao Ensino Médio a média em 2005 foi 3,4 e em 2017 apenas quatro décimos maior, 3,8, sendo que a meta desse ano é 4,7, sinalizando níveis de qualidade abaixo do esperado.

Gráfico 4 – Índice de Desenvolvimento da Educação Básica (IDEB) do Brasil entre os anos de 2005 e 2017.



Fonte: Dados: Observatório do Plano Nacional de Educação, 2020.
Elaboração própria, 2020.

A melhora do sistema de ensino é progressiva, na medida em que as questões são superadas novos problemas entram na pauta para serem solucionados. Um dos grandes desafios atuais é garantir não só que as crianças tenham acesso à escola e nela permaneçam até a conclusão do Ensino Médio, mas também que esse processo signifique de fato uma aprendizagem para o aluno. A preocupação com o desempenho dos alunos é concomitante a preocupação de avaliar esse desempenho. A instituição do Sistema Nacional de Avaliação de Educação Básica (SAEB) em 1990 foi muito importante para consolidação das avaliações educacionais em larga escala no Brasil. O SAEB é conduzido pelo INEP e é composto por um conjunto de avaliações educacionais. O Sistema já passou por diversas estruturações, mas atualmente é constituído por testes e questionários que são aplicados na rede pública e de forma amostral na rede privada a cada biênio (INEP, 2020).

Desde 1995 é adotada a Teoria de Resposta ao Item (TRI), metodologia de construção do teste e análise dos resultados que permite a comparabilidade entre os resultados das avaliações ao longo do tempo. Neste ano também passou a aplicar-se também os questionários que trazem um levantamento de dados contextuais. As provas contemplam as últimas etapas de cada nível de ensino e a partir de 2019, também o 2º ano do Ensino Fundamental e a Educação Infantil. Seu objetivo é

fornecer elementos para a formulação e o monitoramento de políticas direcionadas à melhoria da qualidade da educação (INEP, 2020).

Os resultados de desempenho dos testes de proficiência são expressos em uma escala de 0 a 500 pontos, com média de 250 e desvio padrão de 50 (INEP, 2020). É preciso destacar que todas as etapas utilizam a mesma escala, de forma que a expectativa é que haja uma progressão na proficiência em cada nível de ensino. Nessa escala são definidos níveis de proficiência, o intervalo entre cada nível é equivalente a meio desvio padrão, 25 pontos e para cada nível são descritas as habilidades que se espera terem sido desenvolvidas. A partir dessa escala, a interpretação do Ministério da Educação (MEC) é que os níveis 0 a 3 representam uma proficiência insuficiente, 4 a 6 básica e 7 em diante adequada. Cada etapa tem um valor específico de proficiência para cada nível como apresentado no quadro 2 a seguir (INEP, 2018).

Quadro 2 - Escala de proficiência por etapa de ensino

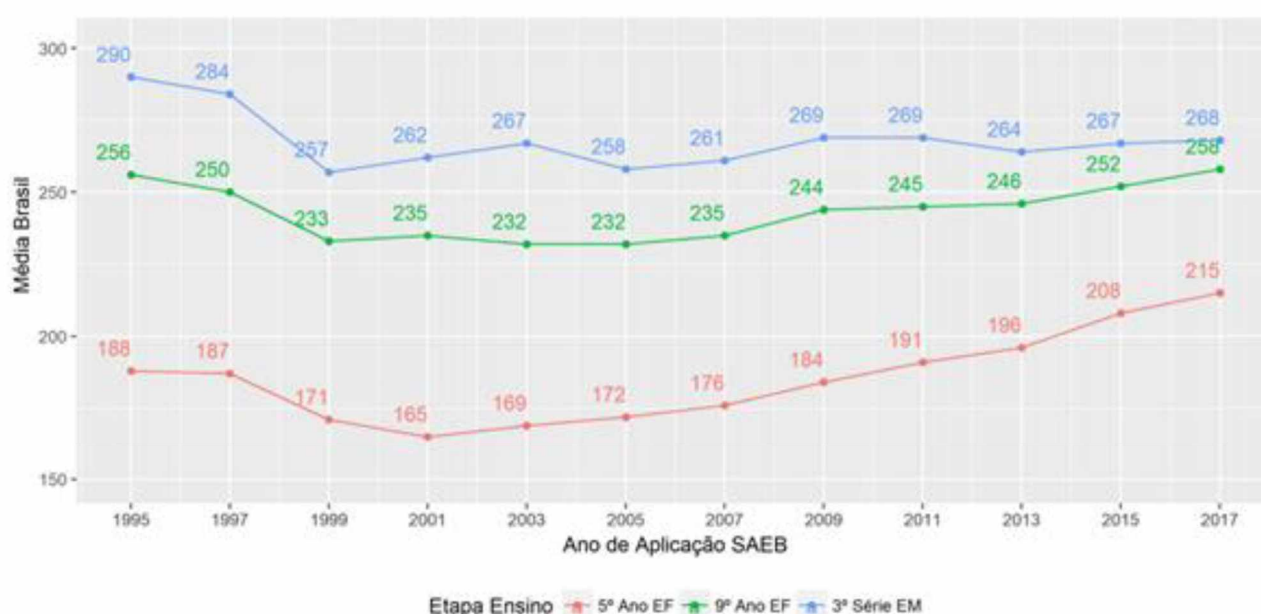
Faixas de desempenho	5° ano EF		9° ano EF		3° ano EM	
0 a 125	nível 0	Insuficiente	nível 0	Insuficiente	nível 0	Insuficiente
125 a 150	nível 1					
150 a 175	nível 2					
175 a 200	nível 3					
200 a 225	nível 4	Básico	nível 1	Insuficiente	nível 1	Insuficiente
225 a 250	nível 5		nível 2		nível 2	
250 a 275	nível 6		nível 3		nível 3	
275 a 300	nível 7	Adequado	nível 4	Básico	nível 4	Básico
300 a 325	nível 8		nível 5		nível 5	
325 a 350	nível 9		nível 6		nível 6	
350 a 375	nível 10		nível 7	Adequado	nível 7	Adequado
375 a 400	nível 11		nível 8		nível 8	
400 a 425	nível 12		nível 9		nível 9	
425 a 450	nível 13		nível 10		nível 10	

475 a 500	nível 14	nível 11	nível 10
-----------	----------	----------	----------

Fonte: INEP, 2020.

Os gráficos 5 e 6 apresentam a evolução das médias de proficiências em Língua Portuguesa e Matemática nas provas do SAEB de 1995 até 2017 para cada etapa de ensino avaliada: 5º e 9º ano do Ensino Fundamental e 3º ano do Ensino Médio.

Gráfico 5 – Evolução das proficiências médias demonstradas pelos estudantes brasileiros em Língua Portuguesa 1995-2017.

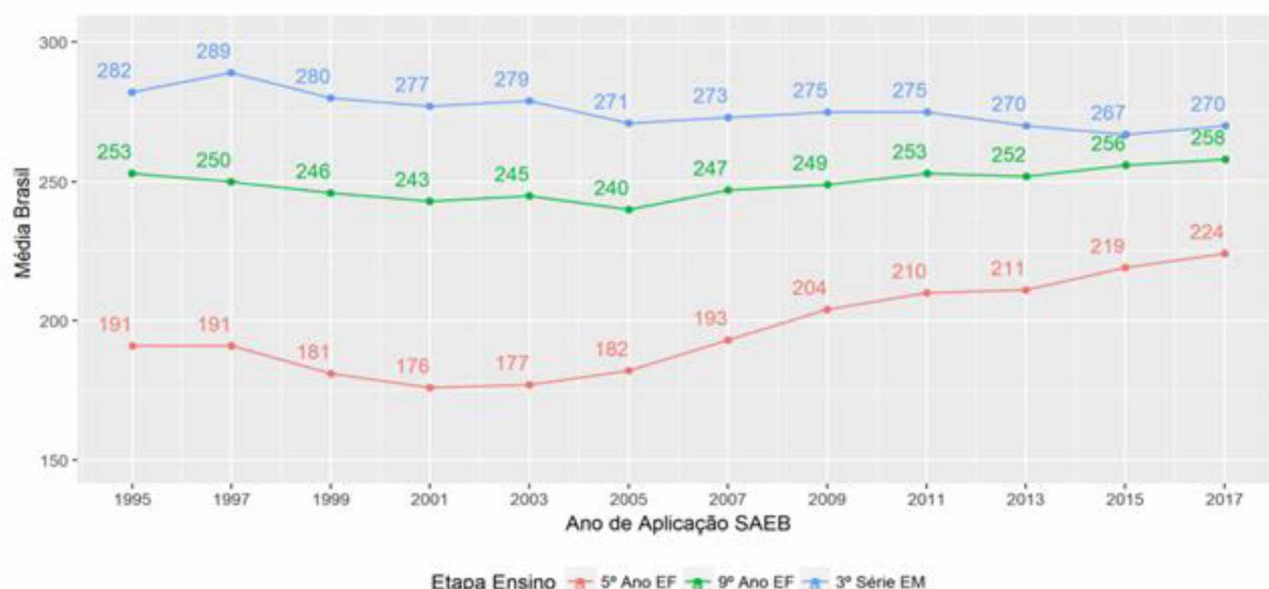


Fonte: INEP, 2018.

A partir do gráfico 5, é possível perceber um aumento contínuo na proficiência de Língua Portuguesa no 5º ano EF desde 2001, ano em que a proficiência foi de 165 e aumentou 50 pontos ao longo dos dezesseis anos, passando para 215. Isso representa a diferença de dois níveis na escala, do insuficiente para o básico - maior pontuação da série observada. Observa-se uma piora nos resultados de todas as etapas até 1999. No caso do 9º ano EF, o desempenho fica estagnado no mesmo nível, com 233 pontos em 1999, até 2007 com 235 pontos, até apresentar melhoras em 2009 (um aumento de nove pontos, passando para 244). Desde então apresenta trajetória de crescimento em todos os anos avaliados, ainda que não tão acentuado quanto no 5º ano EF, com um aumento de 14 pontos nesses oito anos.

A proficiência média de 2017 do 9º ano também foi a maior do período para essa etapa, mas ainda se enquadra no nível insuficiente. O 3º ano EM, no entanto, ainda não superou a queda de desempenho em 1997 e 1999. Em 1995 foi observada a maior nota, 290 pontos que representa o nível insuficiente, até 1999 a proficiência caiu 33 pontos, o que corresponde a mais que um nível da escala, e passou a ser 257. A partir dessa avaliação, o desempenho médio do 3º ano não teve grandes alterações e, até 2017, aumentou 11 pontos, chegando a 268 pontos, o que representa uma diferença negativa de 22 pontos em relação à primeira proficiência observada em 1995, 22 anos antes. Além deste ano e de 1997, as maiores notas do 3º ano EM foram de 2009 e 2011, na marca dos 269 pontos, ainda insuficiente.

Gráfico 6 – Evolução das proficiências médias demonstradas pelos estudantes brasileiros em Matemática 1995-2017.



Fonte: INEP, 2018.

A evolução da proficiência em Matemática é muito semelhante à de Língua Portuguesa, conforme gráfico 2. Em geral as notas de Língua Portuguesa são maiores que as de Matemática em todas as etapas em questão, com exceção da edição de 1995 no 9º ano EF e no 3º ano EM. As diferenças entre as notas das duas disciplinas são maiores para o 5º ano EF, (em média 11 pontos) seguido do 3º ano EM (média de 8 pontos) e do 9º ano EF (6 pontos em média de diferença). Em 2017 o 5º ano EF estava no nível básico enquanto o 9º ano EF e o 3º ano EM estavam no nível insuficiente de proficiência em Matemática.

Apesar das tendências similares, algumas distinções se destacam. Não houve uma queda tão acentuada na proficiência de Matemática em 1997 e 1999 no 9º ano EF como em Língua Portuguesa, mas a proficiência reduziu desde a primeira edição até 2005, quando registrou 240 pontos. A partir de 2007 a pontuação foi gradualmente aumentando até que, em 2017, com 258 pontos, finalmente superou-se a nota de 1995. No 3º ano EM, desde 1997, o desempenho de Matemática foi diminuindo a cada edição e nos anos que não houve queda da proficiência a variação em relação ao ano anterior foi muito pequena. Em 2017, o desempenho de Matemática no 3º ano EM foi de 270 pontos, uma das menores notas do período.

Para além do SAEB existem outras importantes avaliações tanto no âmbito estadual, quanto municipal. Entre elas o Sistema Mineiro de Avaliação da Educação Pública (SIMAVE), instituído pela Secretaria de Estado de Educação (SEE) em 2000. Um dos seus instrumentos de avaliação é o Programa de Avaliação da Rede Pública de Educação Básica (PROEB) que avalia o Ensino Fundamental e o Médio. O PROEB utiliza uma metodologia semelhante ao SAEB, o que garante uma comparabilidade do resultado dos dois exames. Os resultados educacionais, tais como o desempenho cognitivo dos alunos, dependem das políticas educacionais que variam a cada estado. Por isso, optou-se por realizar uma análise específica da rede pública mineira.

É necessário considerar ainda que a proficiência dos alunos não se distribui de forma equitativa, grande parte dessas diferenças estão associadas a características adscritas dos alunos, como cor, sexo, idade, e outros atributos sociais, como nível socioeconômico e cultural, além de diferenças regionais. Nesse sentido, definiu-se que a qualidade da educação também será analisada pela perspectiva da equidade. Um sistema de ensino equitativo é aquele que entende que todos os alunos têm direito à aprendizagem, independentemente de suas características pessoais, que reconhece a forte influência dos fatores extraescolares, principalmente atributos dos alunos e de suas famílias, e atua de forma a mitigar essas desigualdades. Sabe-se que tais desigualdades decorrem de diversos fatores, mas esse trabalho analisará principalmente o impacto do nível socioeconômico cultural.

Á vista desse quadro, entende-se que a escola é essencial para a melhora da qualidade da educação já que “é o cenário das relações interpessoais e profissionais, que influenciam de fato o que ocorre em sala de aula, é nela que se planta e que se colhem os frutos dos planos de melhoramento” (BROOKE; SOARES, 2008, p 334). Qual então o papel das escolas mineiras para promoção de qualidade

e equidade educacional? Esse é o principal ponto de análise dessa monografia, o capítulo seguinte apresentará os dados e os métodos utilizados para responder essa questão.

3 O MODELO DE ANÁLISE DOS DETERMINANTES INTRAESCOLARES DA EFICÁCIA ESCOLAR EM MINAS GERAIS

Como mencionado na Introdução, este trabalho analisa a relação entre processos escolares, a qualidade da educação e a equidade intraescolar na rede pública de Minas Gerais. Utilizou-se dados educacionais de estudantes do 3º ano do Ensino Médio de 2019, os mais recentes sobre a situação da educação no estado, conforme será descrito neste capítulo, que tece as considerações metodológicas do estudo. Primeiramente, são apresentados os dados e as técnicas de sistematização utilizadas. Em seguida, descreve-se as variáveis consideradas e suas formas de obtenção e por fim, apresenta-se os métodos de ajuste do modelo de regressão hierárquica que será utilizado neste estudo. Vale destacar que os métodos aqui adotados foram inspirados em análises anteriores feitas por Soares (2004a) e Franco et al. (2007).

3.1 Origem e técnicas de sistematização dos dados

O objetivo desse estudo é avaliar a eficácia das escolas públicas mineiras de ensino médio, estaduais e municipais, identificando que fatores intraescolares contribuem para ela. De início, cabe recordar, conforme apresentado na Introdução, que a eficácia escolar será analisada pela perspectiva da qualidade e da equidade. A qualidade da educação será mensurada pela proficiência dos alunos na prova de Matemática e Língua Portuguesa, que dizem respeito aos resultados cognitivos dos construtos que foram mensurados por meio das avaliações em larga escala. A qualidade sozinha, no entanto, não é o suficiente para verificar a eficácia escolar, dada a existência de atributos dos estudantes que favorecem academicamente uns em detrimento de outros. Assim, a análise das desigualdades educacionais considera as diferenças de desempenho entre grupos de alunos. Entende-se que a escola eficaz não apenas promove melhores desempenhos como viabiliza que todos os estudantes tenham resultados satisfatórios independente de suas características adscritas. Dessa forma, o conceito de equidade, por sua vez, foi operacionalizado pela influência das características socioeconômicas do estudante em seu desempenho nas avaliações, alunos de escolas equitativas sofrem menor influência desse fator.

Serão utilizados dados do Programa de Avaliação da Rede Pública de Educação Básica (PROEB) fornecidos pela Secretaria de Estado de Educação. O PROEB é uma avaliação externa censitária do Estado de Minas Gerais e mede a

proficiência dos estudantes a partir dos resultados de provas de múltipla escolha de Língua Portuguesa e Matemática. Utiliza os métodos da Teoria Clássica de Testes (TCT)³ e da Teoria de Resposta ao Item (TRI), explicitada adiante.

Além do teste, estudantes, professores das séries avaliadas e diretores escolares respondem um questionário contextual que visa analisar fatores extra e intraescolares do processo educativo, tais como nível socioeconômico do estudante, gestão e administração da escola, práticas pedagógicas, características individuais dos atores, qualificação dos profissionais, dentre outros fatores. São utilizados os resultados de proficiência dos alunos no 3º ano e dados dos questionários contextuais aplicados no ano de 2019. Ademais, uma das variáveis explicativas, conforme será descrito adiante, foi obtida a partir dos dados das escolas do Censo 2019, disponível na página dos microdados do INEP.

Vale ressaltar que no banco de dados utilizado estão ausentes alguns dados, tanto no nível do aluno, especificamente os que não participaram das provas ou não preencheram o questionário, quanto no nível da escola, há alunos e professores que não responderam o instrumento, ou casos de alguns itens sem resposta. Pela distribuição e natureza dos dados ausentes, eles foram desconsiderados. Não se encontrou evidências de que a ausência viesasse os resultados, principalmente por não haver correlação entre os valores das variáveis e a distribuição dos dados ausentes e não haver diferenças significativas entre as médias das variáveis antes e depois da exclusão. Optou-se por não utilizar os itens do questionário do diretor por conta do alto número de questionários sem resposta, o que prejudicaria a análise. Mais informações sobre os dados ausentes, como a comparação das médias do universo e da amostra e proporção dos dados utilizados constam no apêndice B.

Neste estudo, tem-se a intenção de avaliar determinados constructos teóricos, variáveis latentes (que não podem ser diretamente observadas), como é o caso da habilidade cognitiva em matemática dos alunos e o clima escolar. Por não poderem ser diretamente mensurados, esses constructos são avaliados segundo uma escala de valores que é produzida a partir de um conjunto de itens relacionados diretamente a eles. Por meio de diversas técnicas é possível obter-se, então, uma medida para um constructo (M. SOARES, 2005). A proficiência em matemática, o

³ A TCT faz parte do campo dos testes educacionais, mas não será utilizada neste estudo. Maiores detalhes sobre a TCT podem ser vistos em LORD (1980) e em WEISS & YOES (1991).

índice socioeconômico dos alunos e o indicador de complexidade da gestão foram obtidos pela Teoria de Resposta ao Item.

Em síntese, os escores calculados pela TRI são obtidos pelo produto da probabilidade de um indivíduo escolher certa alternativa de um item (questão), dado o nível de seu constructo, então a proficiência, por exemplo, é calculada pela probabilidade de o estudante acertar o item dada sua habilidade. Uma diferença entre o cálculo da proficiência e do índice socioeconômico, é que este é obtido por Teoria de Resposta ao Item não paramétrica, ou seja, não impõe uma forma algébrica explícita para a função resposta do item. A proficiência, por sua vez, leva em conta três parâmetros: a dificuldade do item, a discriminação⁴ do item e a chance de acerto ao acaso (ou chute) (PEREIRA, 2006).

Outra técnica utilizada foi a análise fatorial, ela permite reduzir um grande número de itens (variáveis) em uma variável sintética, um fator. O método utilizado foi de análise dos componentes principais, assim, pressupõe-se que:

“a) as variáveis mais correlacionadas combinam-se dentro de um mesmo fator; b) as variáveis que compõem um fator são estatisticamente não relacionadas com aquelas que compõem outros fatores (os fatores não são correlacionados entre si); e c) a derivação dos fatores processa-se visando maximizar a percentagem da variância total associada a cada fator consecutivo, de modo que o primeiro fator extraído sempre terá o maior poder de explicação das variáveis originais, e assim consecutivamente”. (DINIZ; CORRAR, 2006; FERREIRA JR.; FAHEL; HORTA, 2018 apud BERTÚ, 2020, p. 100 e 101)

Um dos passos da análise é a determinação do número de fatores a serem extraídos de acordo com diversos critérios, como a proporção da variância explicada pelo fator ou sua relevância teórica. Recomenda-se que os fatores extraídos tenham eigenvalues⁵ maiores que 1 e que o total da variância explicada pelos fatores seja de pelo menos 50%. Assim, os escores são obtidos pelo somatório de cada variável ponderado por sua carga fatorial, quanto maior a carga fatorial mais correlacionada a variável está ao fator. Dessa forma, pressupõe-se que exista correlação entre as variáveis, ou seja que elas sejam relacionadas (na medida em que elas correspondem

⁴ A discriminação de um item representa a capacidade que ele tem de diferenciar pessoas com diferentes habilidades.

⁵ Corresponde ao percentual da variância total do conjunto de variáveis que é explicado pelos fatores extraídos na análise.

a um mesmo fator). Para verificar este pressuposto recomenda-se os testes de Kaiser-Meyer-Olkin (KMO) e de esfericidade de Bartlett.

A estatística KMO avalia a adequabilidade da análise fatorial a partir dos coeficientes de correlação entre os pares de variáveis e dos coeficientes de correlação e de correlação parcial, que é a medida de associação entre duas variáveis controlado por um conjunto de outras variáveis. O KMO varia de 0 a 1 e quanto maior o valor, maior a dependência entre os valores. O teste de esfericidade de Bartlett testa a hipótese de que as variáveis não sejam correlacionadas na população. A hipótese nula é a matriz de correlação, formada pela correlação entre cada variável, ser uma matriz identidade, isto é, avalia se os componentes fora da diagonal principal são zero. Se os componentes fora da diagonal forem zero, indica que não há correlação entre as variáveis, pois a correlação é uma medida que varia de -1 a 1, sendo os valores positivos representando uma relação direta e os negativos, inversa; quanto mais distante de zero maior a associação entre as variáveis (BERTÙ, 2020).

Para construção dos fatores utilizou-se apenas variáveis com KMO maior que 0,7 e o módulo da carga fatorial maior que 0,3, com a exceção da variável “relação com a comunidade”. Considerou-se para cada caso que a extração de apenas um fator era satisfatória, portanto, fez-se essa opção pela facilidade da análise teórica. A construção dessas variáveis foi feita no software R pelo pacote psych, as cargas fatoriais e a extração dos fatores foram executadas pela função “principal” (REVELLE, 2020).

Tendo explicitado a origem e as técnicas de sistematização do banco de dados, apresenta-se a seguir o modelo (variáveis e níveis) que serão utilizadas neste estudo.

3.2 Dados e modelo de análise

As variáveis consideradas para analisar a relação entre fatores intraescolares, qualidade e equidade educacional são de dois níveis de agregação: dos alunos e o das escolas. A proficiência dos alunos em Matemática e Língua Portuguesa no PROEB 2019 serão as variáveis dependentes. O índice socioeconômico, o sexo, a raça do aluno e se ele é ou não defasado representam as características pessoais dos alunos e atributos de capital econômico e cultural de suas famílias, além disso o índice socioeconômico agregado por escola também foi acrescentado para um controle adicional. A importância da incorporação dessas

variáveis é reconhecida na literatura, conforme mencionado no capítulo 2, visto a influência das características pessoais do aluno em seu desempenho escolar. Sem o devido controle a análise tende a produzir resultados “falsos positivos” devido à alta correlação entre esses fatores (FRANCO et al, 2007).

A proficiência dos alunos em Matemática e Língua Portuguesa foi fornecida pela SEE e calculada por meio da TRI. O índice socioeconômico dos alunos também foi fornecido pela SEE e calculado por meio de TRI não paramétrica, a partir de itens ordinais do questionário dos alunos (anexo A). Tais itens dizem respeito a condições materiais e culturais da família do estudante, como a escolaridade dos pais, ou responsáveis e a posse de bens de consumo. O índice socioeconômico médio das escolas foi obtido por agregação da variável do índice socioeconômico dos alunos. As outras variáveis do aluno – sexo, raça e defasagem – foram obtidas por meio do questionário contextual dos alunos. A variável defasagem foi calculada a partir do item que perguntava a idade dos alunos, sendo considerados defasados os discentes com mais de 18 anos de idade.

O indicador de complexidade da gestão escolar e o indicador de regularidade docente são calculados a partir de dados do Censo Escolar e disponíveis na página de indicadores educacionais do INEP. No caso do primeiro, agrega-se por meio de TRI, informações sobre o porte da escola, o número de etapas ofertadas, a complexidade das etapas (entendendo que etapas de alunos mais velhos apresentam maior complexidade pelo maior desafio que apresenta) e a quantidade de turnos de funcionamento. Quanto ao segundo, são analisados dados dos professores e das escolas nos cinco anos anteriores e é atribuída uma pontuação pela presença do docente no quadro da escola e pela regularidade, ou seja, pela permanência em uma mesma escola em anos consecutivos (INEP, 2014, 2015). Essas variáveis foram acrescentadas para fornecer um controle, respectivamente, da situação contextual da escola e de seus insumos, de acordo com a classificação de Scheerens (1990).

As sete demais variáveis do modelo buscam representar o que autor chamou de processos escolares. As variáveis de gestão da matéria e clima escolar foram calculadas por análise fatorial a partir de itens ordinais do questionário dos alunos, enquanto efetividade da implementação do PPP e gestão do diretor foram calculadas por análise fatorial a partir de itens ordinais do questionário dos professores e relação com a comunidade por meio de itens dicotômicos do Censo Escolar 2019. Os itens dos questionários, bem como os resultados da análise fatorial,

foram discriminados no anexo A e B e no apêndice A, respectivamente. Para as variáveis da base do PROEB, os dados foram agregados pela proporção de alunos ou professores que responderam o questionário com “concordam” ou “concordam muito” em cada item.

A Gestão da matéria trata das estratégias didático pedagógicas que o professor utiliza em suas aulas. O clima escolar diz respeito ao relacionamento entre os atores escolares e à ênfase acadêmica da escola. A gestão do diretor, por sua vez, aborda a gestão democrática e pedagógica do Diretor, sobre o suporte e acompanhamento dele aos alunos e professores e quanto sua gestão é aberta e participativa. A relação com a comunidade trata da existência de meios para a participação da comunidade nas atividades da escola. A efetividade da implementação do PPP envolve a efetividade da participação dos atores escolares, elementos dificultadores na implementação e execução dele e a vinculação dos demais instrumentos de planejamento pedagógico (como planos de aula) a ele. Não se encontrou, na literatura analisada, estudos semelhantes que analisassem a relação entre a efetividade na implementação do PPP e o desempenho dos alunos. Ainda assim, acredita-se que seja uma análise relevante, na medida em que o PPP faz parte de uma dimensão muito importante da gestão escolar e é reconhecida a importância de uma gestão participativa e efetiva para a qualidade da educação. Assim, busca-se verificar se esse âmbito da gestão também é significativo.

O comprometimento do professor foi calculado pela proporção de alunos em uma escola que “discordavam” ou “discordavam muito” que seus professores faltavam muito. Por fim, o percentual do programa concluído representa a média desse percentual por escola, informação extraída do questionário dos professores. O quadro 1 sumaria os tipos, a descrição e a natureza das variáveis que compõem o modelo aplicado neste estudo, bem com as classifica de acordo com o modelo de Soares (2004a). A seção seguinte apresentará a regressão hierárquica linear, técnica que será utilizada para operacionalizar a análise proposta.

Quadro 1 – Variáveis utilizadas na análise

(Continua)

Variável	Tipo de Variável e Codificação	Descrição	Classificação
Nível 1: Aluno			
Matemática	Contínua	Proficiência em Matemática no PROEB 2019 calculada por TRI	Resultado escolar
Língua Portuguesa	Contínua	Proficiência em Língua Portuguesa no PROEB 2019 calculada por TRI	Resultado escolar
ISE	Contínua	Índice do socioeconômico do aluno. Obtida por TRI não paramétrica a partir de itens ordinais do questionário do aluno.	Recurso Familiar
Defasagem	Dicotômica (o aluno é defasado = 1)	Defasagem escolar. Calculado a partir de resposta do questionário do aluno sobre sua idade.	Característica Pessoal do aluno
Sexo	Dicotômica (masculino = 1)	Obtido a partir de resposta do questionário do aluno sobre seu sexo.	Característica Pessoal do aluno
Raça	Dicotômica (o aluno é preto ou pardo = 1)	Obtido a partir de resposta do questionário do aluno sobre sua raça.	Característica Pessoal do aluno

Fonte: Elaboração própria, 2020

Quadro 1 – Variáveis utilizadas na análise

(Continuação)

Variável	Tipo de Variável e Codificação	Descrição	Classificação
Nível 2: Escola			
ISE Escola	Contínua	Índice do nível socioeconômico médio dos alunos que estudam na mesma escola. Obtido via agregação da variável ISE	Controle
ICG	Contínua	Indicador de complexidade da gestão. Calculado pelo INEP a partir de dados do Censo Escolar.	Controle
IRD	Contínua	Indicador de regularidade docente. Calculado pelo INEP a partir de dados do Censo Escolar.	Controle
Clima Escolar	Contínua	Clima Escolar. Obtida por análise fatorial a partir da proporção de alternativas de itens ordinais do questionário do estudante.	Relações sociais na escola
Gestão	Contínua	Gestão do Diretor. Obtida por análise fatorial a partir da proporção de alternativas de itens ordinais do questionário do professor.	Direção e Gestão
PPP	Contínua	Efetividade da construção do PPP. Obtida por análise fatorial a partir da proporção de alternativas de itens ordinais do questionário do aluno e do professor.	Direção e Gestão
Comprometimento	Contínua	Comprometimento do professor. Obtida a partir da proporção de alternativas do item do questionário do aluno sobre absentismo do professor.	Comprometimento
Matéria	Contínua	Gestão da matéria. Obtida por análise fatorial a partir da proporção de alternativas de itens ordinais do questionário do estudante.	Gestão da Matéria
Programa	Contínua	Percentual do programa concluído pelo professor. Obtida a partir da proporção de alternativas do item do questionário do professor percentual do programa concluído pelo professor	Gestão da Matéria
Comunidade	Contínua	Relação com a comunidade. Obtida por análise fatorial a partir de itens dicotômicos do Censo Escolar.	Relação com a Comunidade Escolar

Fonte: Elaboração própria, 2020.

3.3 Abordagem analítica

Como a distribuição da proficiência dos alunos depende de sua distribuição por escola e se busca analisar a natureza da interação das variáveis dentro de cada escola, este estudo utiliza o modelo hierárquico linear. Isso pois, esse modelo permite a análise por dois níveis de agregação - aluno e escola - permite a interação entre variáveis de diferentes níveis e considera que a variação da variável dependente, proficiência, depende do agrupamento dos alunos em escolas. Os detalhes técnicos desses modelos estão descritos com mais detalhes em Bryk e Raudenbush (1992) e Goldstein (1995) e apresentados a seguir.

Na análise hierárquica é indicado a construção do modelo "de baixo para cima", ou seja, partir de um modelo incondicional ou nulo, sem variáveis explicativas, acrescentar as variáveis explicativas do nível 1, aquelas relacionadas ao aluno e em seguida acrescentar as variáveis explicativas do nível 2, associadas à escola. O modelo incondicional permite calcular quanto da variância da proficiência dos alunos ocorre pelo agrupamento em escolas, ou seja, identifica a proporção da variação da variável dependente (proficiência) entre as escolas e dentro das escolas. Na formalização dos modelos hierárquicos cada nível é representado por meio de um conjunto de equações. A equação do nível 1 do modelo incondicional é descrita por:

$$y_{ij} = \beta_{0j} + e_{ij}, i = 1, \dots, n_j, j = 1, \dots, J \quad (3.1)$$

E a equação do nível 2 como:

$$\beta_{0j} = \gamma_{00} + u_{0j}, i = 1, \dots, n_j, j = 1, \dots, J \quad (3.2)$$

Dessa forma tem-se a equação geral:

$$y_{ij} = \gamma_{00} + u_{0j} + e_{ij}, i = 1, \dots, n_j, j = 1, \dots, J \quad (3.3)$$

Sendo y_{ij} o valor de y para o i -ésimo indivíduo do j -ésimo grupo, ou seja, y_{ij} é o valor da proficiência do aluno i da escola j , n_j é o total de alunos e J o total de escolas. β_{0j} corresponde ao intercepto no eixo y e no caso desse modelo representa a proficiência média da j -ésima escola, enquanto γ_{00} é a média do resultado para todos os indivíduos. u_{0j} e e_{ij} são os valores do erro, ambos com distribuição normal, média igual a zero e variâncias constantes $\sigma_{u_0}^2$ e σ_e^2 , respectivamente, u_{0j} o erro, ou o efeito aleatório, no nível das escolas e e_{ij} no nível

dos alunos. Já σ_{u0}^2 e σ_e^2 representam respectivamente a variabilidade dentro da escola e entre as escolas, sendo a variância total de y_{ij} dada por:

$$\text{var}(y_{ij}) = \sigma_{u0}^2 + \sigma_e^2 \quad (3.4)$$

A partir desses valores é possível mensurar a proporção da variância que é devida ao nível da escola, a correlação intraclasse que é dada por:

$$\rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_e^2} \quad (3.5)$$

Em geral, pode-se pensar nos modelos hierárquicos a partir de dois componentes: uma parte fixa que especifica o relacionamento entre a média da variável dependente e as variáveis explicativas, e uma parte aleatória que consiste nos erros do nível 1 e 2. O efeito aleatório do nível 1 corresponde a e_{ij} que é a parcela da variabilidade do nível do indivíduo que não é explicada pelo modelo. Outro componente aleatório é u_{0j} que é o erro do intercepto, ou seja, o valor do intercepto β_{0j} da escola j é definido por γ_{00} mais o erro u_{0j} dessa escola. No nível 2, é possível ter outro componente aleatório, o do coeficiente, conforme será explicado. Como o modelo incondicional não tem variáveis explicativas, a parte fixa é definida apenas por γ_{00} .

A partir do modelo nulo, pode-se acrescentar variáveis explicativas do primeiro nível. A equação do nível 1 com a variável explicativa X_{1ij} é descrita por:

$$y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + e_{ij} \quad (3.6)$$

Quanto as equações do nível 2:

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (3.7)$$

$$\beta_{1j} = \gamma_{10} \quad (3.8)$$

Assim a equação geral é descrita por:

$$y_{ij} = \gamma_{00} + \gamma_{10}X_{1ij} + u_{0j} + e_{ij} \quad (3.9)$$

O que difere a equação 3.9 da equação 3.3 é que se acrescentou a variável do nível 1 X_{1ij} que tem o efeito fixo igual a γ_{10} na média das escolas. É possível permitir que esse efeito varie a cada escola, assim, a relação entre cada variável explicativa e a proficiência dos alunos é diferente para cada escola. Para tanto,

acrescenta-se um componente aleatório u_{1j} a β_{1j} e as equações de nível 2 passam a ser:

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (3.10)$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \quad (3.11)$$

Assim a equação geral fica:

$$y_{ij} = \gamma_{00} + \gamma_{10}X_{1ij} + u_{1j}X_{1j} + u_{0j} + e_{ij} \quad (3.12)$$

O modelo passa a ter, além dos efeitos aleatórios u_{0j} e e_{ij} na média respectivamente do nível dos alunos e das escolas, também $u_{1j}X_{1j}$, o efeito aleatório na inclinação. Ressalta-se que se pressupõe ainda que os efeitos aleatórios possuem distribuição normal com média zero e variâncias constantes. Em um modelo com inclinação aleatória a variância entre escolas deve estar em função da variável com coeficiente aleatório:

$$\begin{aligned} var(u_{0j} + u_{1j}X_{ij}) &= var(u_{0j}) + 2X_{ij}cov(u_{0j}, u_{1j}) + var(u_{1j}) = \\ &= \sigma_{u_0}^2 + 2\sigma_{u_0u_1}X_{ij} + \sigma_{u_1}^2X_{ij}^2 \end{aligned} \quad (3.13)$$

A parte fixa do modelo pode ser expandida acrescentando mais preditores, enquanto a parte aleatória pode ser expandida permitindo que o efeito de um ou mais preditores varie entre grupos. Pode-se então, acrescentar as variáveis explicativas do nível 2 W_j , as equações, considerando efeito fixo das variáveis do nível 1 são:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + u_{0j} \quad (3.14)$$

$$\beta_{1j} = \gamma_{10} \quad (3.15)$$

Dessa forma, tem-se a equação geral:

$$y_{ij} = \gamma_{00} + \gamma_{10}X_{1ij} + \gamma_{1j}W_j + u_{0j} + e_{ij} \quad (3.16)$$

Assim, tem-se γ_{10} como o efeito de X_{1ij} dentro das escolas e $\gamma_{00} + \gamma_{10}$ o efeito de X_{1ij} entre as escolas. Não obstante, é possível permitir a interação entre as variáveis explicativas, como X_{1ij} do nível dos alunos e as variáveis explicativas W_j do nível das escolas. Assim, tem-se as seguintes equações do nível 2, considerando efeito aleatório para a variável de nível 1:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + u_{0j} \quad (3.17)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_j + u_{1j} \quad (3.18)$$

Com a equação geral expressa por:

$$y_{ij} = \gamma_{00} + \gamma_{10}X_{1ij} + u_{1j}X_{1ij} + \gamma_{1j}W_jX_{1ij} + \gamma_{01}W_j + u_{0j} + e_{ij} \quad (3.19)$$

Nessa equação, γ_{10} pode ser interpretado como o efeito de X_{1ij} na média das escolas quando W_j é zero. Dessa forma, para cada unidade que W_j variar, o efeito de X_{1ij} varia o correspondente a γ_{1j} , pode-se entender que o coeficiente fixo de X_{1ij} passa a ser $(\gamma_{10} + \gamma_{1j}W_j)$.

Com isso, serão construídos quatro modelos, conforme será descrito na seção 4.4, para tanto, utilizou-se a função `lmer` do pacote `lme4` no software R (BATES et al, 2015). Inicia-se com o modelo nulo, sem variáveis independentes, acrescenta-se então as variáveis do primeiro nível com inclinação aleatória para o índice socioeconômico e em seguida as do segundo. Por fim, no modelo seguinte incluiu-se a interação entre o nível socioeconômico do aluno e as demais variáveis. A partir dessas regressões é possível extrair evidências que contribuem para responder às questões de interesse deste estudo.

4 AS ESCOLAS MINEIRAS FAZEM DIFERENÇA NO APRENDIZADO DE ESTUDANTES DO ENSINO MÉDIO?

Este capítulo tem o objetivo de, a partir do referencial teórico discutido no capítulo 2 e dos dados e métodos apresentados no capítulo 3, responder à questão de pesquisa: As escolas públicas mineiras foram eficazes em 2019 para os alunos do 3º ano do EM? Dessa forma, inicialmente serão apresentados os dados da amostra do PROEB 2019 utilizada neste estudo. A partir disso, averiguar-se-á as desigualdades educacionais enfrentadas por esses estudantes. Então, serão apresentados os resultados dos modelos analisando o papel da escola e seus processos na promoção de qualidade e equidade da educação. Concluir-se-á com a síntese dos resultados.

4.1 Análise descritiva dos dados

Nessa seção, serão apresentados os dados e as variáveis utilizadas no estudo. A base de dados do PROEB 2019 contava com 177208 alunos de 2363 escolas. Após a exclusão dos alunos e escolas que não possuíam todas as informações utilizadas na análise, ficou-se com 114325 alunos para Língua Portuguesa e 114319 alunos para Matemática (alguns alunos fizeram uma prova mas não fizeram a outra) de 2027 escolas, que são a unidade de análise que será tratada adiante. Em primeiro lugar, serão apresentadas as estatísticas descritivas das variáveis, em seguida, a distribuição das variáveis explicativas do nível da escola e por fim, as proficiências dos alunos em Língua Portuguesa e Matemática.

A tabela 1 apresenta os valores mínimos, máximos, média e desvio padrão das variáveis. As médias de proficiência em Língua Portuguesa e Matemática são bem próximas, 267,33 e 270,62, respectivamente. Essa nota corresponde ao nível baixo de proficiência, de acordo com a escala utilizada pelo PROEB que será apresentada a seguir. Os desvios padrões também são bem próximos, aproximadamente 51, isso indica que os dados dispersam em torno de 50 pontos da média, quanto menor o desvio padrão, maior a homogeneidade da distribuição dos dados. Os valores mínimos e principalmente os máximos, já são diferentes, a menor nota de Língua Portuguesa foi 127,78 (nível baixo) enquanto de Matemática foi 158,66 (também nível baixo), 30,88 pontos a mais. A maior nota de Língua Portuguesa foi 393,98 (nível avançado), já Matemática (nível avançado), 59,37 pontos a mais.

As variáveis “sexo”, “raça” e “defasagem escolar” são dicotômicas, ou seja, são variáveis que assumem valor igual a zero ou a um, são indicativas de um atributo, como descrito no capítulo anterior. “Sexo” indica que o aluno é do sexo masculino, “raça” indica que o aluno é preto ou pardo e “defasagem escolar” indica que o aluno não cursa o 3º ano na idade adequada. Por serem dicotômicas, a média dessas variáveis indica sua proporção. Assim, 44% dos estudantes do 3º ano do Ensino Médio que participaram da avaliação são do sexo masculino e 56% do feminino. A maioria são pretos ou pardos (69%), seguidos de 31% de discentes que se declaram como brancos, amarelos ou indígenas. Nota-se que poucos estudantes estão defasados (11%), o que aponta que a maioria (89%) tem trajetória regular.

Outro ponto importante a salientar é que as variáveis “índice socioeconômico”, “índice socioeconômico da escola”, “efetividade do PPP”, “gestão do diretor”, “clima escolar”, “gestão da matéria”, “indicador de regularidade docente” e “indicador de complexidade da gestão” foram padronizadas e centralizadas na média para facilitar a interpretação, de modo que cada valor foi subtraído da média e dividido pelo desvio padrão. Por isso, todas têm média igual a zero e desvio padrão igual a um. Essas variáveis se referem a construtos que não podem ser diretamente observados, portanto, foram definidos escores para as variáveis e os valores de máximo e mínimo apresentados são a escala em que essas variáveis serão analisadas.

Os escores permitem uma comparação ordinal entre as escolas, por exemplo, uma escola com um clima escolar igual a 2 tem um clima melhor que outra escola onde o fator é igual a 1. No entanto, neste estudo, não foi definida uma escala que atribuísse sentido a cada valor da variável, como existe para proficiência, que também representa um construto latente. Se um estudante do 3º ano EM tem uma proficiência em Língua Portuguesa igual a 350 pontos no SAEB é possível interpretar que esse resultado evidencia um aprendizado de nível básico e pode-se até investigar quais são as prováveis habilidades que esse estudante foi capaz de consolidar. Quanto as demais variáveis serão analisadas em termos ordinais, seu valor em si não evidencia um sentido, pois o que importa para esse estudo é sua distribuição

Por fim, as variáveis “comprometimento do professor” e “percentual do programa concluído” representam proporções. Observa-se que há escolas em que os professores não faltam muito, sendo que em outras até 50% dos professores faltam muito, a média de absenteísmo nas escolas foi 12%. Quanto ao programa, em média,

os professores trabalharam 84% do conteúdo programado, sendo que todos concluíram até 25% e há aqueles que abordaram todo o conteúdo.

Tabela 1 – Estatísticas descritivas

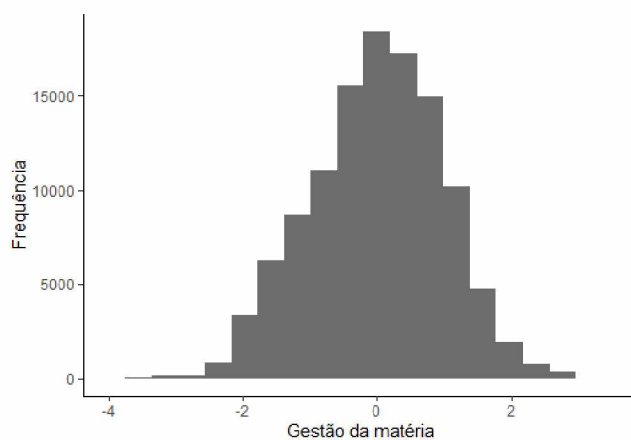
Variável	Média	Desvio		
		Padrão	Mínimo	Máximo
PROFICIENCIA_LP	267,33	51,41	127,78	393,98
PROFICIENCIA_MT	270,62	51,50	158,66	453,35
SEXO	0,445	0,5	0	1
RACA	0,687	0,46	0	1
DEFAS	0,117	0,32	0	1
ISE	0	1	-3,50	3,16
ISE_ESCOLA	0	1	-4,24	2,85
ICG	0	1	-3,98	1,27
IRD	0	1	-3,51	3,51
GESTAO	0	1	-0,81	6,42
PPP	0	1	-6,66	1,45
MATERIA	0	1	-7,38	3,01
CLIMA_ESCOLAR	0	1	-4,59	3,50
COMUNIDADE	0	1	-3,10	1,38
COMPROMETIMENTO	0,876	0,07	0,50	1
PROGRAMA	0,840	0,08	0,25	1

Fonte: Dados: SEE/MG, 2019 e INEP, 2019.

Elaboração própria, 2020.

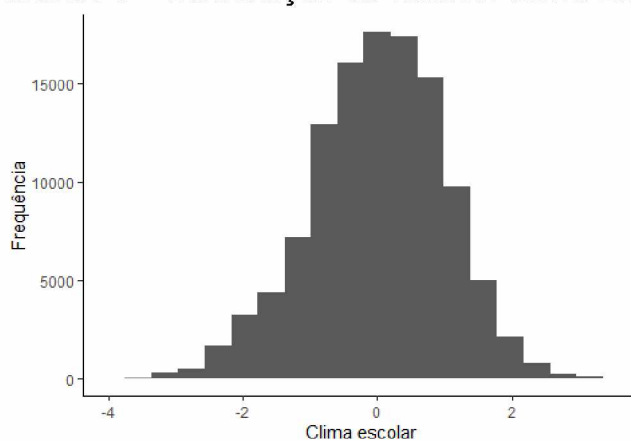
Os gráficos 7 a 17 apresentam a distribuição de frequência das variáveis do nível da escola. Os gráficos 7 e 8 apresentam, respectivamente, o “clima escolar” e a “gestão da matéria” do professor. Observa-se que ambas têm uma distribuição muito semelhante, próxima à de sino, a maior parte dos valores é igual ou próxima a zero que é a média, progressivamente, diminui ou aumenta de forma bem simétrica. Verifica-se que o “clima escolar” e a “gestão da matéria” são muito correlacionadas ($r = 0,66$) o que significa que nas escolas que tem bom clima escolar, os professores também conseguem fazer uma boa gestão da matéria. Já sobre o “índice socioeconômico”, apresentado no gráfico 5, embora não esteja correlacionado ao “clima escolar” ($r = -0,03$) nem à “gestão da matéria” ($r = -0,15$), também apresenta distribuição semelhante – em formato de sino. O “índice socioeconômico das escolas”, gráfico 10, por ser a média do primeiro NSE também tem uma distribuição de frequência muito similar. Outras duas variáveis que têm maiores frequências quanto mais próximo da média são o indicador de regularidade docente e relação com a comunidade, observa-se ainda que este não é tão simétrico, há mais observações acima da média.

Gráfico 7 – Distribuição da variável Gestão da Matéria, 3º ano EM, Minas Gerais, 2019.



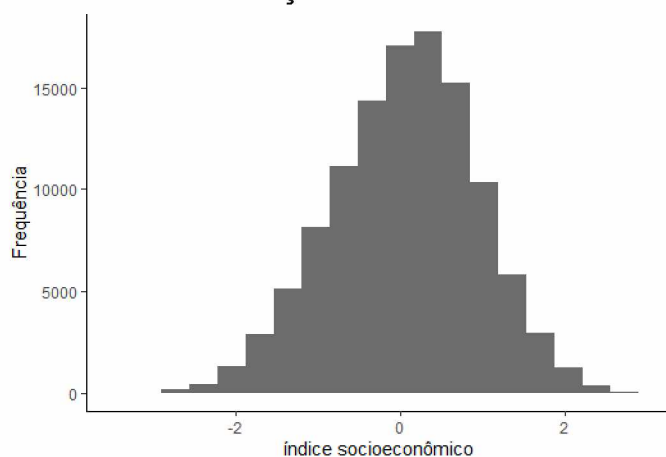
Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Gráfico 8 – Distribuição da variável Clima Escolar, 3º ano EM, Minas Gerais, 2019.



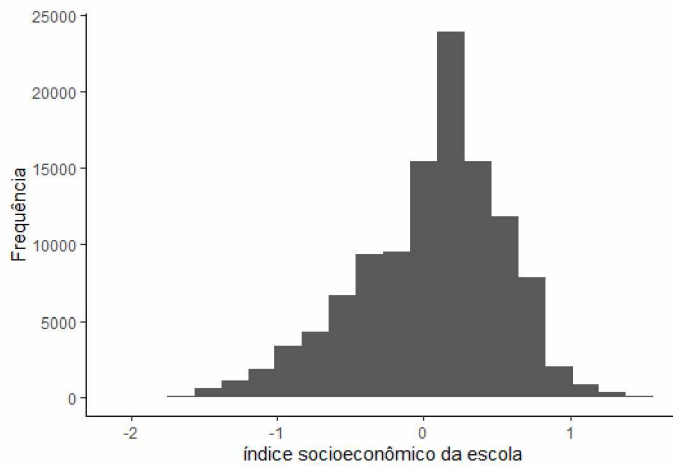
Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Gráfico 9 – Distribuição da variável ISE dos alunos, 3º ano EM, Minas Gerais, 2019.



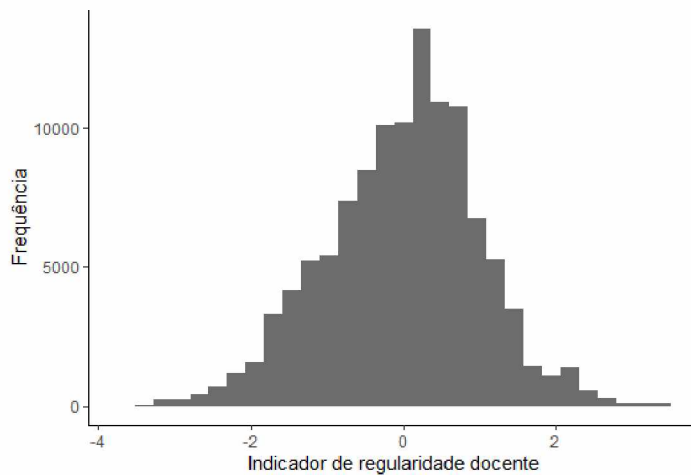
Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Gráfico 10 – Distribuição da variável ISE Escola, 3º ano EM, Minas Gerais, 2019.



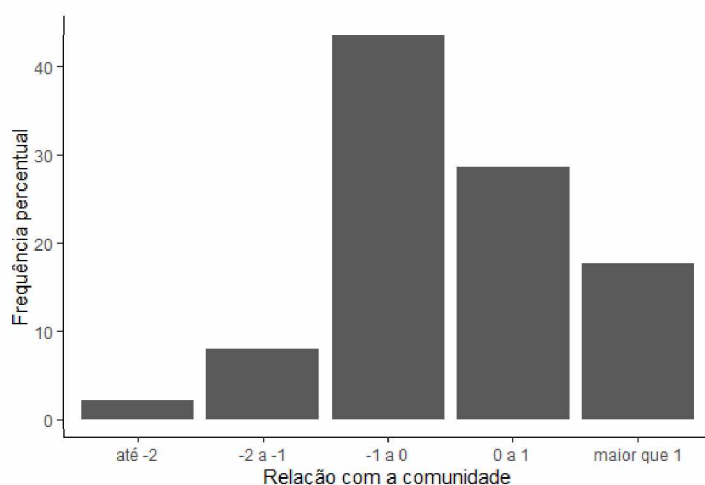
Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Gráfico 11 – Distribuição da variável Indicador de Regularidade Docente, 3º ano EM, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Gráfico 12 – Distribuição da variável Relação com a Comunidade, 3º ano EM, Minas Gerais, 2019.

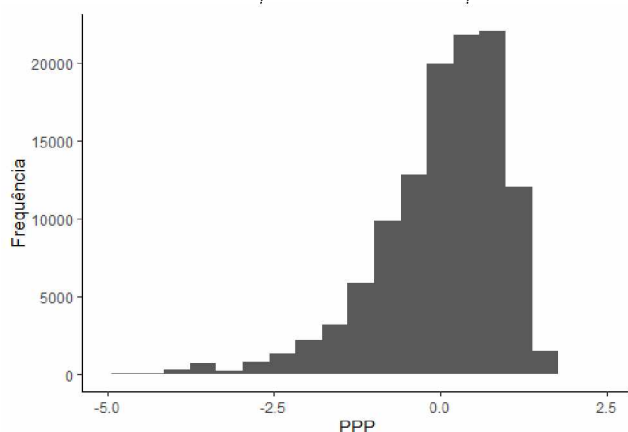


Fonte: Dados: SEE/MG, 2019.

Elaboração própria, 2020.

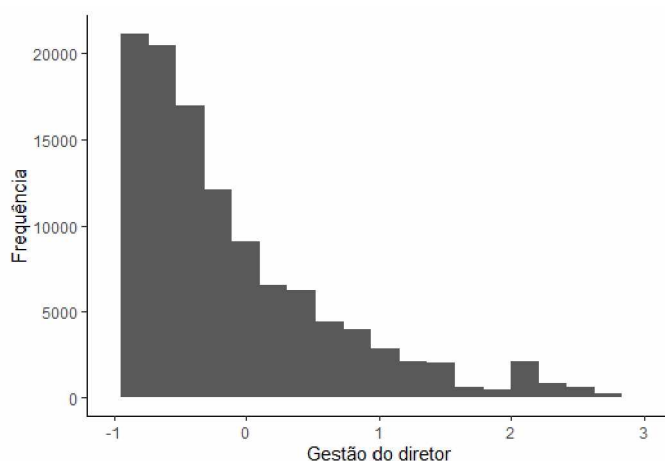
As variáveis “efetividade do PPP” e “gestão do diretor” são apresentadas nos gráficos 7 e 8. Ao contrário das anteriores, os valores com maior ocorrência estão nos extremos da distribuição. A primeira distribui-se mais à direita, têm mais valores acima da média (0) enquanto a segunda tende à esquerda, tem mais valores abaixo da média. É curioso notar que esses dois fatores têm uma relação inversa ($r = -0,56$), ou seja, escolas com uma gestão escolar melhor que a média tendem a ter uma implementação do PPP pior que a média, sendo o inverso igualmente válido. Esse resultado é curioso, pois esperava-se que uma boa gestão do diretor induziria a uma maior efetividade na implementação do PPP. As duas variáveis, inclusive, averiguam alguns aspectos muito próximos, tais como o nível participativo da gestão do professor e a efetividade da participação de diversos atores educacionais na implementação do PPP. Talvez, essa relação deva ao fato de o PPP ainda não ter sido incorporado como instrumento de gestão pedagógica pelos atores escolares, assim provavelmente, apesar do que se esperava, esse fator não terá relação significativa com o desempenho dos alunos. Dessa forma, sugere-se que estudos futuros investiguem essa relação contra intuitiva, a fim de um melhor entendimento sobre ela.

Gráfico 13 – Distribuição da variável Efetividade da implementação do PPP, 3º ano EM, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Gráfico 14 – Distribuição da variável Gestão do Diretor, 3º ano EM, Minas Gerais, 2019.

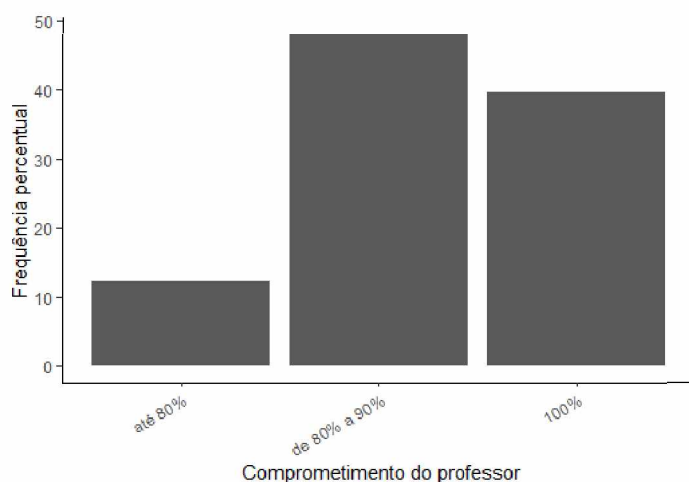


Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Os gráficos 9 e 10 mostram a distribuição das variáveis “comprometimento do professor”, “percentual do programa concluído” e “indicador de complexidade da gestão” por categorias. Verifica-se assim que em 41,01% das escolas até 10% dos professores faltam muito, em 47,74% delas, de 10% a 20% dos professores faltam muito e em 11,24% das instituições mais de 20% dos professores faltam muito. Quanto a conclusão do programa, observa-se que metade dos professores (50,09%) concluiu até 90% na data da aplicação do questionário (o PROEB é aplicado sempre no final do ano letivo). Da outra metade, 21,65% concluíram todo o programa, 23,41% completaram até 80% e 4,84% concluíram menos que esse percentual. A maioria das escolas têm nível de complexidade a partir de 4 (94,7%), quanto maior o nível significa uma maior complexidade. Isso se justifica pelo fato de a etapa do 3º ano do Ensino

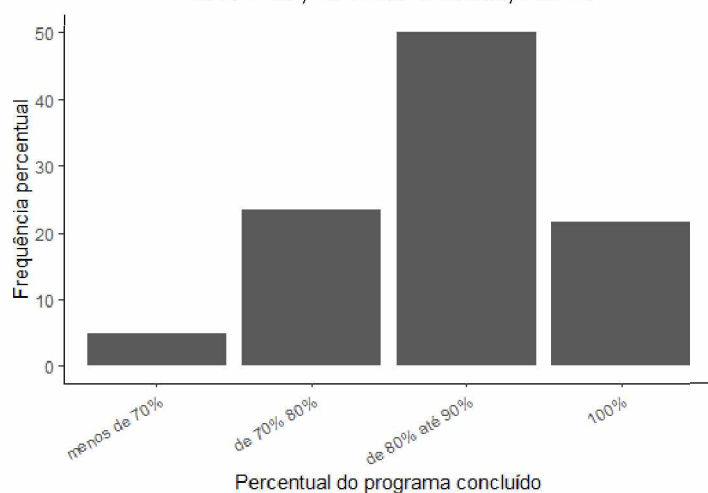
Médio ser considerada a mais complexa, ou seja, demandar uma maior articulação de professores e do currículo e apresentar maiores desafios, como a manutenção do estudante na escola. Dentre os três maiores níveis 36,7% das escolas está no nível 4, 30% no nível 5 e 27,9% no nível 6.

Gráfico 15 – Frequência percentual da variável Comprometimento do professor, 3º ano EM, Minas Gerais, 2019.



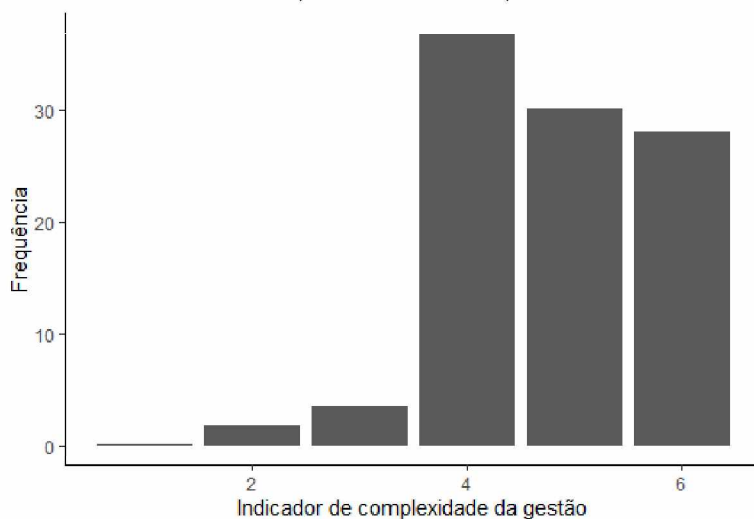
Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Gráfico 16 – Frequência percentual da variável Percentual do programa concluído, 3º ano EM, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Gráfico 17 – Frequência percentual da variável Percentual do programa concluído, 3º ano EM, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Por fim, nos gráficos 18 e 19 tem-se a distribuição percentual das proficiências dos alunos do 3º ano em Língua Portuguesa e Matemática por nível. Para auxiliar na interpretação da proficiência, o quadro 3 apresenta os níveis de proficiência conforme as notas alcançadas no PROEB. Os resultados seguem a escala do SAEB, ou seja, 0 a 500 pontos, que representam níveis de desempenho diferentes, do baixo ao avançado. Espera-se que o estudante esteja pelo menos no nível recomendado.

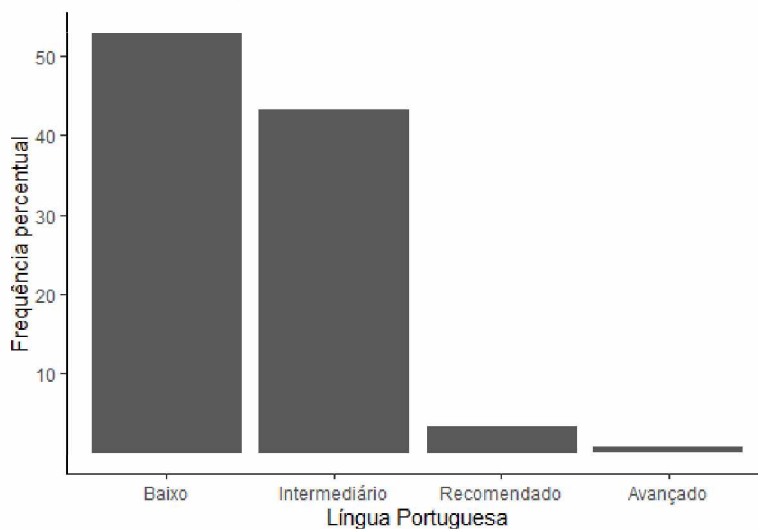
Observa-se que 52,75% dos estudantes têm uma proficiência baixa em Língua Portuguesa, 43,17% intermediária, 3,27% recomendada e apenas 0,80% avançada. Esses resultados demonstram que um baixo nível de qualidade em que somente 4,08% dentre todos os alunos analisados nesse estudo têm níveis satisfatórios de proficiência em Língua Portuguesa. Os dados da prova de Matemática não diferem muito, mas são um pouco mais otimistas: 56,98% com proficiência baixa, 34,96% intermediária, 4,56% recomendada e 3,49% avançada. Assim, 91,94% dos alunos da amostra não têm o mínimo da proficiência esperada em Matemática e apenas 8,06% dos estudantes atingiram esse patamar em 2019.

Quadro 3 - Escala de proficiência do 3º ano EM - PROEB

Faixa de proficiência	Níveis
0 a 275	Baixo
275 a 350	Intermediário
350 a 375	Recomendado
375 a 500	Avançado

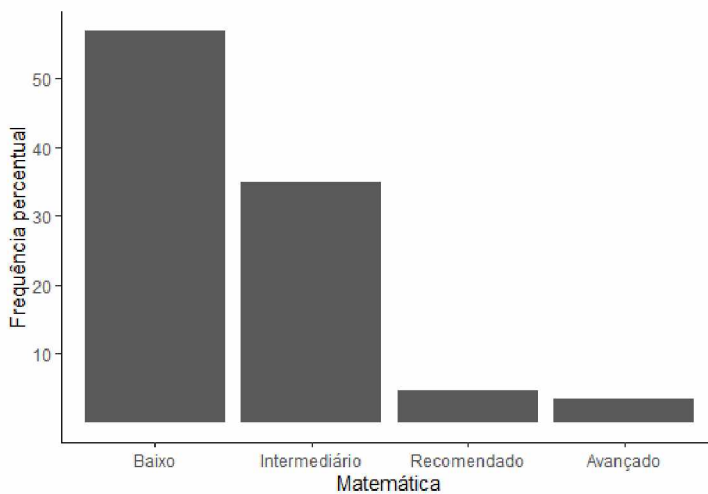
Fonte: SEE/MG, 2019.

Gráfico 18 – Frequência percentual da proficiência em Língua Portuguesa do 3º ano EM, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Gráfico 19 – Frequência percentual da proficiência em Matemática do 3º ano EM, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Dessa forma, percebe-se que os resultados educacionais do 3º ano do EM, em termos de desempenho cognitivo, não foram satisfatórios no ano de 2019. Tendo vista a distribuição dos fatores escolares, resta-se compreender, quem são esses alunos, como a proficiência está distribuída entre eles e como suas características influenciam em seu desempenho. A seção seguinte busca justamente lançar luz a essas questões, para, em seguida, analisar as escolas que esses alunos estudaram e avaliar a relação entre os seus processos internos e sua eficácia.

4.2 A relação entre origem social e qualidade no ensino médio

Nessa seção serão apresentados os dados da amostra de alunos analisados nesse estudo. Espera-se com isso identificar quem são esses estudantes e verificar como suas características pessoais influenciam o seu desempenho. Primeiramente, serão abordados os impactos das diferenças entre as escolas e no desempenho. Em seguida, será feita uma análise da relação entre nível socioeconômico e proficiência. Por último, discute-se a relação entre características adscritas (sexo, raça), trajetória escolar dos estudantes e desempenho, pelo o que Alves e Soares (2003) e Soares (2004b) denominam de gradiente socioeconômico.

A tabela 2 apresenta as diferenças de desempenho médio entre grupos de discentes, bem como sua distribuição por grupo. É possível, assim, compreender com mais clareza os perfis dos estudantes que compõem a amostra. Inicialmente, observa-se que cada nível socioeconômico tem 25% dos alunos, isso pois, esses níveis foram definidos por quartil da distribuição da variável. A maioria dos alunos (55,5%) é do sexo feminino, isso pode ser indicativo do que já foi verificado por outros estudos, as meninas enfrentam menos barreiras no acesso e progressão no Ensino Médio (ASSIS, 2012). Nota-se também que a maioria dos alunos é preto ou pardo (68,7%) em comparação aos brancos, amarelos ou indígenas (31,3%) e que apenas 11,7% dos alunos é defasado.

Essa tabela evidencia claramente as desigualdades educacionais, não é justificável que existam diferenças de desempenho dada as características dos alunos, por isso a necessidade de que hajam medidas compensativas que tornem os resultados mais equitativos. Quanto ao nível socioeconômico do aluno, apenas os alunos com nível alto têm proficiência média considerada intermediária, enquanto os demais grupos estão no nível baixo para os dois componentes curriculares. A diferença de proficiência dos grupos dos dois extremos sociais é de 26,75 pontos para

Língua Portuguesa e 29,36 para Matemática, isso significa uma diferença de cerca de 9 pontos para cada nível social. Também há diferenças por sexo, raça e trajetória escolar. As meninas têm uma proficiência média bem maior em Língua Portuguesa (17,45 pontos) e um pouco menor em Matemática (-3,85 pontos). As diferenças por raça e trajetória escolar independem do componente curricular, alunos pretos e pardos têm uma média cerca de 15 pontos menor, enquanto alunos defasados têm uma média em torno de 30 pontos menor.

Tabela 2 – Caracterização dos alunos por ISE, sexo, raça e trajetória escolar

Variável	Categoria	Número de alunos	Percentual de alunos	Proficiência média	
				Língua Portuguesa	Matemática
Nível socioeconômico	Baixo	28586	25,0%	253,81	256,75
	Médio Baixo	28585	25,0%	263,29	265,41
	Médio Alto	28581	25,0%	271,63	274,21
	Alto	28582	25,0%	280,56	286,11
Sexo	Feminino	63505	55,5%	275,08	268,91
	Masculino	50829	44,5%	257,63	272,76
Raça	Branco, amarelo ou indígena	35780	31,3%	277,65	281,46
	Preto ou pardo	78554	68,7%	262,62	265,68
Trajetória escolar	Defasado	13337	11,7%	239,16	244,22
	Regular	100997	88,3%	271,04	274,11

Fonte: Dados: SEE/MG, 2019
Elaboração própria.

Para além disso, há de se considerar que os alunos não são apenas uma raça ou um gênero, Alves e Soares (2016) verificaram que quando os grupos são definidos por múltiplos fatores as diferenças são ainda mais acentuadas. Visto isso, a tabela 3 apresenta a distribuição das características dos alunos por grupos. É possível verificar que a condição socioeconômica não varia muito por gênero, mas que os baixos níveis são bem mais frequentes entre os alunos pretos e pardos e entre os alunos defasados. Recordar-se aqui que o índice socioeconômico foi centralizado e possui média igual a zero, assim, valores positivos são acima da média e negativos abaixo. Além disso, alunos defasados tendem a ser pretos ou pardos e ser do sexo masculino.

Tabela 3 – Caracterização dos alunos por ISE, sexo, raça e trajetória escolar

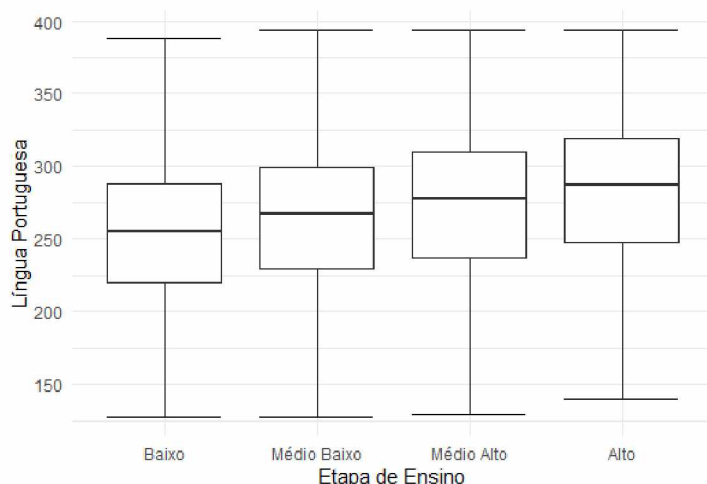
Variável	Categoria	Índice socioeconômico médio	Proporção de alunos do sexo masculino	Proporção de alunos pretos ou pardos	Proporção de alunos defasados
Nível socioeconômico	Baixo	-1,315	0,392	0,797	0,183
	Médio Baixo	-0,286	0,414	0,723	0,119
	Médio Alto	0,368	0,449	0,659	0,092
	Alto	1,234	0,523	0,569	0,073
Sexo	Feminino	-0,086	0	0,685	0,097
	Masculino	0,108	1	0,690	0,142
Raça	Branco, amarelo ou indígena	0,274	0,441	0	0,090
	Preto ou pardo	-0,125	0,446	1	0,129
Trajetória escolar	Defasado	-0,350	0,540	0,759	1
	Regular	0,046	0,432	0,678	0

Fonte: Dados: SEE/MG, 2019
Elaboração própria.

Os gráficos 20 e 21 apresentam as proficiências em Língua Portuguesa e Matemática por nível socioeconômico. Mais uma vez, entende-se que cada nível corresponde a um quartil, ou seja, 25% da distribuição dos dados, conforme metodologia adotada pelo PROEB. Tratam-se de gráficos de caixa, utilizados para representar a variação dos dados, ele é composto dos limites inferiores e superiores, que correspondem aos valores máximos e mínimos. As extremidades da “caixa” representam o primeiro e o terceiro quartil, quanto a reta central corresponde a mediana. É possível avaliar o grau de dispersão de cada subamostra pela distância entre cada valor, quanto mais distante, mais disperso.

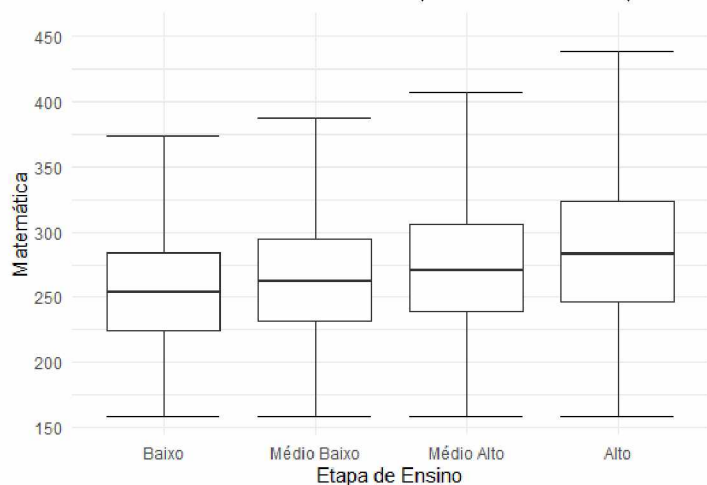
A partir disso, é possível perceber que os valores mínimos de proficiência, independem do nível socioeconômico nos dois componentes curriculares. Pontua-se apenas que o valor mínimo em Língua Portuguesa é um pouco maior para o nível do ISE alto. Neste mesmo componente, os valores máximos também são os mesmos para todos os níveis. Ou seja, a amplitude das notas é a mesma, no entanto, quanto maiores os níveis sociais, maiores tendem ser as notas médias, observa-se, nos dois componentes curriculares, que as caixas representam quartis cada vez maiores, demonstrando desigualdades existentes entre esses alunos. No caso de Matemática, verifica-se ainda essa mesma progressão das notas, não apenas quanto ao primeiro e terceiro quartis, mas também em relação aos valores máximos.

Gráfico 20 – Distribuição da proficiência em Língua Portuguesa do 3º ano do EM por nível socioeconômico, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Gráfico 21 – Distribuição da proficiência em Matemática do 3º ano do EM por nível socioeconômico, Minas Gerais, 2019.



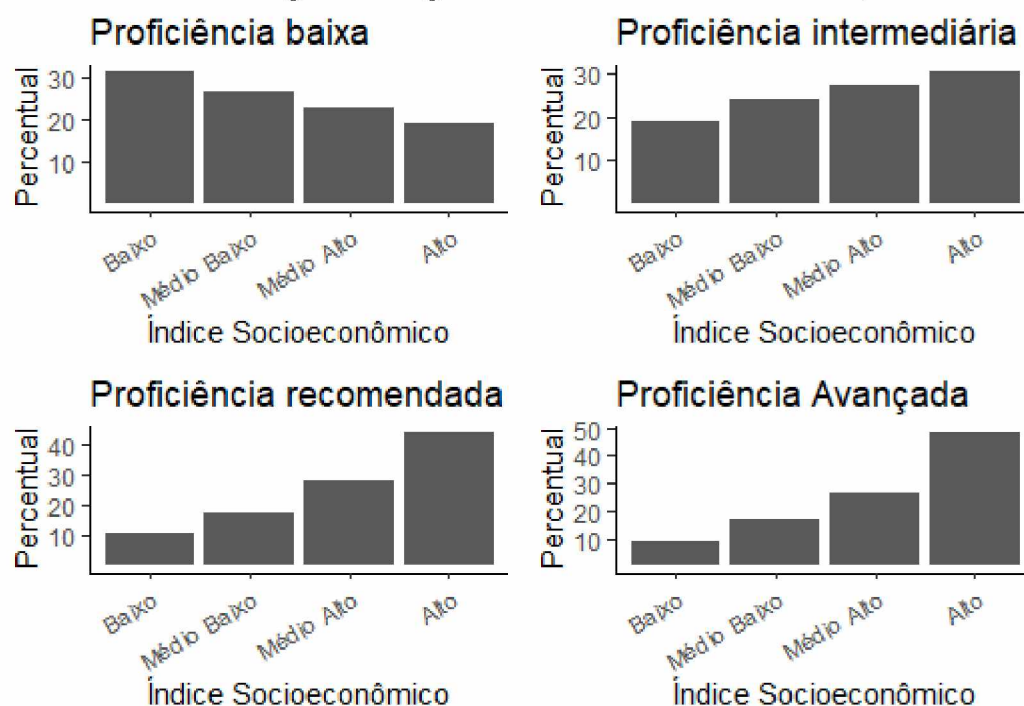
Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Os gráficos 22 e 23 trazem os dados da comparação entre a composição social dos alunos em cada nível de proficiência em Língua Portuguesa e Matemática, respectivamente. Aqui tem-se mais uma evidência de que o nível socioeconômico desses alunos é determinante no seu desempenho. Para o nível de proficiência baixa, nos dois componentes curriculares, nota-se uma distribuição de frequência decrescente para os níveis socioeconômicos, havendo, deste modo, mais alunos com piores condições sociais. Já no nível de proficiência intermediária a distribuição inverte

e quanto maior o nível de proficiência (recomendado e avançado) mais discrepante passa a ser a distribuição dos alunos por estrato socioeconômico.

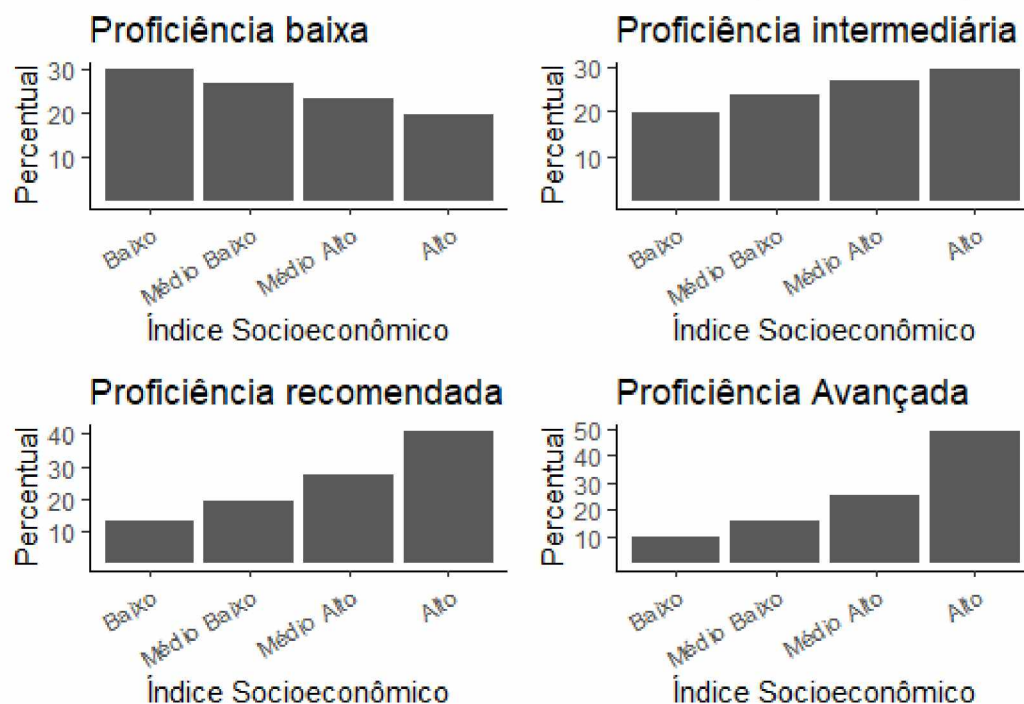
Enquanto na proficiência baixa, cerca de 30% dos alunos têm nível socioeconômico baixo e um pouco mais de 20% alto, na proficiência avançada são cerca de 10% com nível socioeconômico baixo em comparação aos 50% de alunos com nível socioeconômico alto. Assim percebe-se que o nível socioeconômico é mais determinante para o desempenho nos maiores níveis de proficiência. Em outras palavras, ter nível socioeconômico alto foi mais determinante para atingir melhores notas do que ter nível socioeconômico baixo foi para piores desempenhos.

Gráfico 22 – Distribuição percentual do nível socioeconômico por nível de proficiência em Língua Portuguesa 3º ano do Ensino Médio, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Gráfico 23 – Distribuição percentual do nível socioeconômico por nível de proficiência em Matemática 3º ano do Ensino Médio, Minas Gerais, 2019.

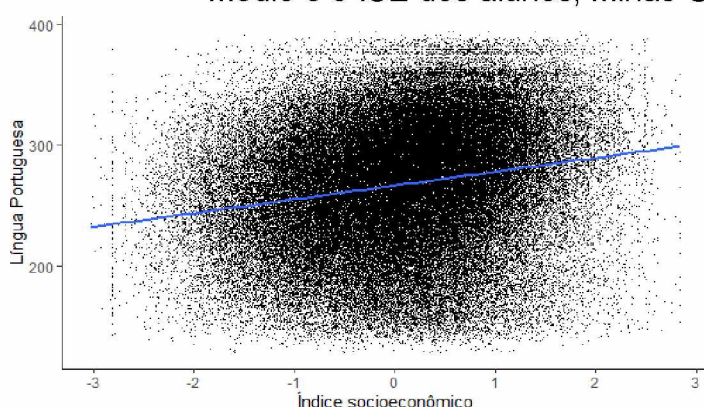


Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Os gráficos 24 a 31 a seguir mostram a relação entre proficiência e índice socioeconômico. Cada ponto no gráfico representa uma observação, um aluno, quanto mais alto o ponto, maior a proficiência, quanto mais a direita maior o índice socioeconômico. A reta representa a regressão linear simples dos dados, ou seja, indica o valor esperado da proficiência dado o índice socioeconômico. Assume-se que uma diferença no índice socioeconômico está associada com a mesma diferença de proficiência ao longo de toda a distribuição, ou seja, o benefício marginal da vantagem socioeconômica é o mesmo em qualquer ponto da posição social.

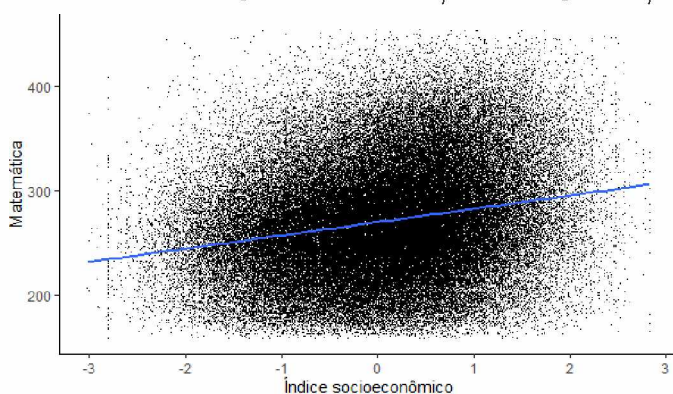
Os gráficos 24 e 25 apresentam o gradiente socioeconômico. Mais uma vez fica clara a forte influência da condição social no desempenho nos dois componentes curriculares, uma vez que a reta apresenta inclinação positiva, quanto maior o índice, maior a proficiência esperada. Já nos gráficos 26 a 31 discriminam esses resultados por sexo, raça e trajetória escolar. Neles, os dados foram divididos de acordo com cada categoria das variáveis, sexo feminino ou masculino, por exemplo, cada uma representada por uma reta. Com elas, é possível comparar os valores de proficiência de alunos com um mesmo índice socioeconômico, sendo que quanto maior a distância entre as retas, maior a diferença entre os grupos de estudantes.

Gráfico 24 – Relação entre a proficiência em Língua Portuguesa do 3º ano do Ensino Médio e o ISE dos alunos, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Gráfico 25 – Relação entre a proficiência em Matemática do 3º ano do Ensino Médio e o ISE dos alunos, Minas Gerais, 2019.

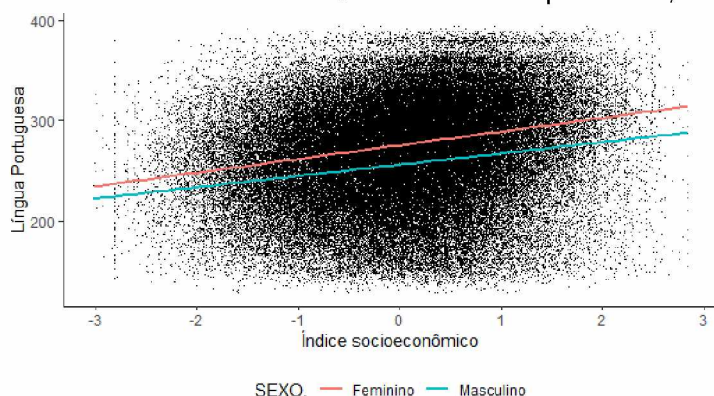


Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Primeiramente, aponta-se para o maior desempenho das meninas, em todos os níveis socioeconômicos em Língua Portuguesa, sendo que a diferença entre sexos é maior para alunos com maiores níveis socioeconômicos. Em Matemática tem-se uma situação totalmente inversa, os meninos tendem a ter maiores notas e as discrepâncias são suavizadas com o aumento do índice socioeconômico, até as retas se encontrarem, sinalizando que as diferenças por gênero se anulam nos estratos socioeconômicos mais elevados. Além disso, mesmo entre os alunos com piores condições sociais, a diferença é muito menor do que a observada em Língua Portuguesa. Conclui-se, então, que existem desigualdades de gênero entre os alunos e que essa desigualdade varia de acordo com o componente curricular e o nível socioeconômico do aluno, sendo que é mais expressiva em Português, nas

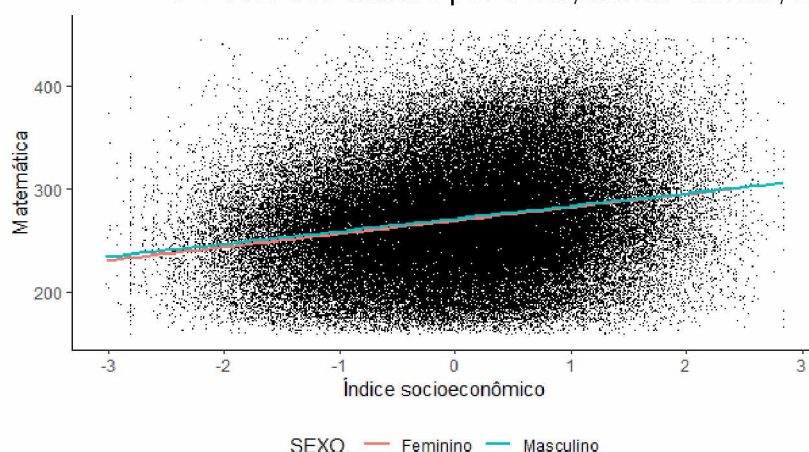
proficiências mais altas e entre alunos com condições socioeconômicas mais favorecidas.

Gráfico 26 – Relação entre a proficiência em Língua Portuguesa do 3º ano do Ensino Médio e o ISE dos alunos por sexo, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

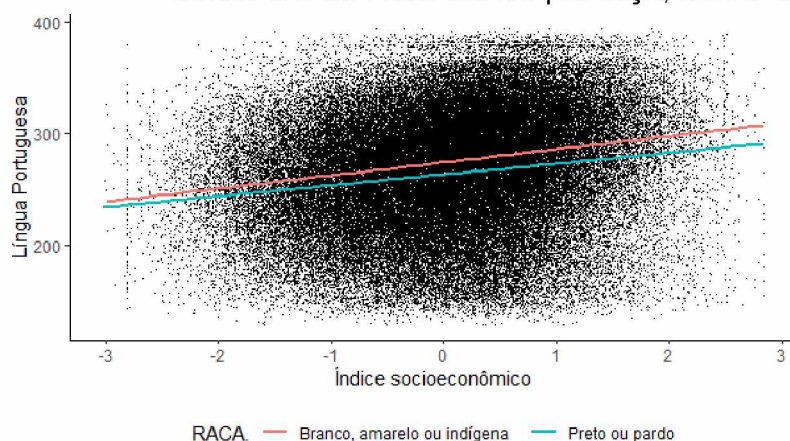
Gráfico 27 – Relação entre a proficiência em Matemática do 3º ano do Ensino Médio e o ISE dos alunos por sexo, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

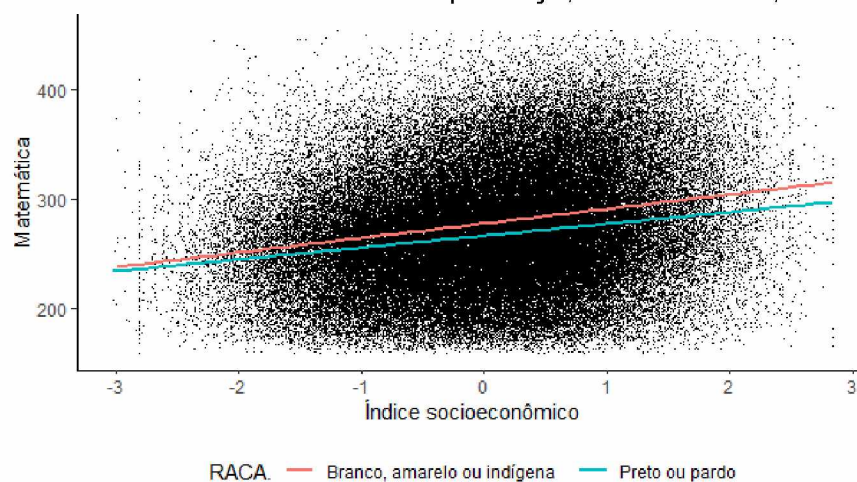
Já as diferenças raciais são muito semelhantes nos dois componentes curriculares. Alunos brancos, amarelos ou indígenas tendem a ter notas maiores independente de seu nível socioeconômico. Percebe-se, no entanto, que a raça do aluno faz mais diferença em seu desempenho entre aqueles alunos que tem maior nível socioeconômico. Portanto, há também desigualdades raciais entre esses alunos e ela se expressa menos pelo componente curricular avaliado e mais pelo nível socioeconômico dos alunos, sendo maior nos estratos mais abastados.

Gráfico 28 – Relação entre a proficiência em Língua Portuguesa do 3º ano do Ensino Médio e o ISE dos alunos por raça, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

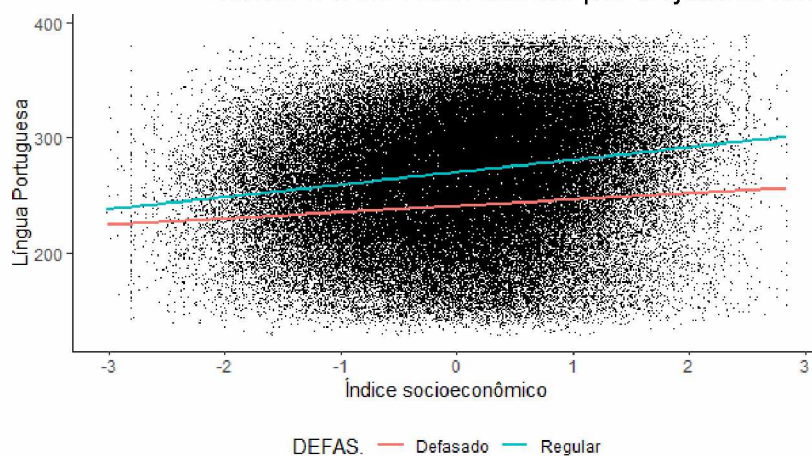
Gráfico 29 – Relação entre a proficiência em Matemática do 3º ano do Ensino Médio e o ISE dos alunos por raça, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

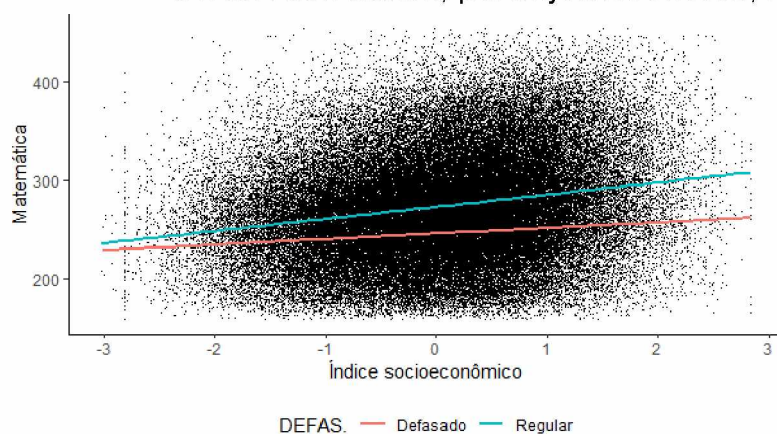
Por último, quanto a trajetória escolar, tem-se um cenário muito similar ao analisado quanto a raça. Não há diferenças entre os componentes curriculares na relação entre trajetória escolar e proficiência, alunos defasados sempre têm menores desempenhos. Percebe-se ainda que na medida em que o índice socioeconômico cresce, as diferenças entre os dois grupos de alunos também se acirram, ainda mais do que observado em relação a raça. Dessa forma, para além da defasagem em si já demonstrar a diferença entre os alunos, o desempenho deles também tende a ser menor quando está em situação de trajetória irregular.

Gráfico 30 – Relação entre a proficiência em Língua Portuguesa do 3º ano do Ensino Médio e o ISE dos alunos por trajetória escolar, Minas Gerais, 2019



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Gráfico 31 – Relação entre a proficiência em Matemática do 3º ano do Ensino Médio e o ISE dos alunos, por trajetória escolar, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Assim, as desigualdades educacionais são complexas, existem maiores desigualdades de gênero em relação à Língua Portuguesa e de raça e de trajetória escolar entre alunos com melhores condições sociais.

A análise dos dados do PROEB 2019 evidencia um cenário preocupante quanto a qualidade educacional, já que menos de 10% dos alunos da amostra apresentam desempenho satisfatório. A desigualdade educacional também se mostra evidente, ao se observar as diferenças de desempenho determinadas por nível socioeconômico, sexo, raça e trajetória escolar. Evidentemente, trata-se de cenário muito desafiador para a gestão da política educacional no estado, diante do papel

central da escola para superação desse quadro. A seção seguinte avaliará a relação entre processos escolares, qualidade e equidade no ensino médio de Minas Gerais.

4.3 A diferença da escola na proficiência do ensino médio e na promoção da equidade

Os estudos revisados no capítulo 2 demonstraram que as escolas podem fazer muita diferença no desempenho dos alunos. Essa seção pretende verificar a influência de fatores escolares na qualidade e equidade das escolas municipais e estaduais de EM de Minas Gerais. Para isso serão analisados dados de 114325 alunos para Língua Portuguesa e 114319 alunos para Matemática de 2027 escolas por meio de modelos de regressão hierárquica linear. A qualidade será mensurada pelo desempenho cognitivo dos alunos (proficiência nas provas de Matemática e Língua Portuguesa), enquanto a equidade será medida pelo efeito das características pessoais dos alunos em seu desempenho. É importante esclarecer que a seção anterior buscou averiguar se existiam desigualdades educacionais e como elas são distribuídas, o objetivo agora passa a ser verificar como as escolas influenciam nessas desigualdades, especificamente as geradas pelo ISE, e se as escolas são capazes de mitigá-las.

Nesse ponto fica claro que a proficiência é determinada por uma série de fatores. Dessa forma, cada um tem uma trajetória escolar única que resulta em uma grande variabilidade das proficiências. Pode-se dizer então que o clima escolar, por exemplo, explica a variabilidade entre as proficiências, assim, quando agrupamos as escolas por clima escolar, a variabilidade dentro de cada grupo é menor. Ao selecionarmos aleatoriamente um par de alunos, se eles estudarem em escolas com o mesmo clima, as chances de os desempenhos deles serem mais parecidos é maior.

O cálculo de uma regressão serve justamente para identificar fatores, como o clima escolar, que explicam a variabilidade de resultados. Os coeficientes da regressão indicam se e em que medida um fator determina a proficiência. A partir da equação da regressão é possível identificar o valor dela, dado o clima da escola em que um aluno estuda, por exemplo. O coeficiente diz, então, o quanto o valor médio| varia, na medida em que os outros valores variam. Por se tratar de uma regressão linear, assume-se que essa relação é a mesma independentemente do valor da variável. Há de se considerar que não é interessante nem viável acrescentar todos os preditores da proficiência em um modelo. Dessa forma, parte da variação da

proficiência não é explicada pelos fatores analisados e sempre haverá uma diferença entre o valor esperado e o valor real, o que se chama de erro.

Um ponto particularmente interessante da regressão hierárquica é que é possível calcular o valor da proficiência esperada de um aluno considerando a escola que ele estuda. O chamado de intercepto aleatório consiste em um valor de proficiência médio para cada escola. O intercepto é o ponto da reta de regressão que cruza o eixo y quando as variáveis explicativas são iguais a zero. Assim, o intercepto da escola é calculado pelo intercepto global mais o erro do intercepto daquela escola. Como é possível separar o erro por nível de agregação, pode-se calcular o quanto da variação decorre da escola. O erro é uma medida da variação não explicada pelo modelo. Assim, pode-se calcular a proporção da variação que ocorre dentro da escola (entre os alunos) e entre as escolas.

Em seguida, serão apresentados os resultados dos modelos. A proficiência em Matemática e em Língua Portuguesa foram utilizadas como variáveis dependentes, além de quatro variáveis do nível do aluno que representam as características pessoais do aluno e de dez variáveis do nível da escola, o “índice socioeconômico médio”, o “indicador de regularidade docente” e o “indicador de complexidade da gestão” acrescentados como controle e as variáveis que representam os processos escolares, conforme apresentado no Capítulo 3. Para cada componente curricular, foram construídos quatro modelos. O primeiro é o modelo nulo, sem qualquer variável. Este modelo servirá de parâmetro para os demais, além disso traz evidências interessantes quanto à proporção da variância entre escolas.

No segundo modelo, foram adicionadas as variáveis de nível do aluno, sendo que o índice socioeconômico também tem efeito aleatório. Este efeito corresponde a parte do coeficiente que varia em cada escola, de forma que o impacto dessa variável é diferente para cada uma. Os coeficientes fixos apontam para a influência das características dos alunos em sua proficiência, coeficientes positivos são promotores de qualidade, pois aumentam a proficiência esperada e os coeficientes aleatórios revelam como estes variam por escola, ou seja, demonstram como, em algumas escolas, esses efeitos são menos expressivos do que em outras. Uma das razões dessa diferença, pode ser que uma escola é mais equitativa que outra.

A partir disso, acrescentou-se as variáveis do nível da escola (modelo 3). Os coeficientes fixos desse modelo trazem evidências quanto a importância dos processos escolares para melhores desempenhos. Justifica-se manter a variável do

índice socioeconômico com efeito aleatório para garantir um melhor ajuste dos dados. Por fim, foi construído um quarto modelo acrescentando as interações entre as variáveis explicativas dos modelos e o índice socioeconômico, o que busca identificar o quanto uma afeta na outra. Dessa forma, os fatores escolares que promovem equidade são aqueles que diminuem o efeito das características dos alunos sobre a proficiência.

Os resultados dos modelos 1, 2 e 3 estão na tabela 4 para proficiência em Língua Portuguesa e 5 para Matemática. Consta dizer que, para o teste da razão de verossimilhança, o agrupamento hierárquico em escolas foi significativo nos dois modelos nulos, é dizer, constatou-se que há, de fato, diferenças de desempenho médio entre as escolas e que essa diferença não é aleatória. Sem qualquer controle, a proporção da variância entre escolas no primeiro modelo é de 10,17% e 12,54% no segundo. Como discutido, esse valor é calculado verificando quanto do erro, variação não explicada pelo modelo, reside no nível da escola. Como no modelo nulo não há preditores, esse valor é interpretado como uma medida de quanto a escolha da escola importa para o desempenho dos alunos.

Tabela 4 - Resultados dos modelos hierárquicos lineares 1 a 3 de Língua Portuguesa (continua)

Variáveis independentes	Modelo 1	Modelo 2	Modelo 3
Efeitos Fixos	Coefficientes		
Intercepto	264.15	282.27	237.24
Índice socioeconômico		7.26*	6.94*
Sexo		-17.13*	-17.08*
Raça		-8.45*	-8.17*
Defasagem		-24.79*	-24.66*

Obs.: (*) $p < 0,001$; (**) $p \leq 0,05$; (***) $p \leq 0,1$

Fonte: Dados: SEE/MG, 2019 e INEP, 2019.

Elaboração própria, 2020.

Tabela 4 - Resultados dos modelos hierárquicos lineares 1 a 3 de Língua Portuguesa
(continuação)

Variáveis independentes	Modelo 1	Modelo 2	Modelo 3
Efeitos Fixos			
Coeficientes			
PPP			-
Gestão do diretor			-
ISE Escola			2.82*
IRD			2.49*
ICG			-1.01*
Comprometimento			36.21*
Clima Escolar			4.18*
Matéria			0.65***
Programa			16.81*
Comunidade			0.49*
Efeitos Aleatórios			
Variância			
Intercepto	263.8*	177.62*	108.38*
Inclinação do índice socioeconômico		12.09*	12.12*
% da variância entre escolas	9,97%	7,49%	

Obs.: (*) $p < 0,001$; (**) $p \leq 0,05$; (***) $p \leq 0,1$

Fonte: Dados: SEE/MG, 2019 e INEP, 2019.

Elaboração própria, 2020.

Tabela 5 - Resultados dos modelos hierárquicos lineares 1 a 3 de Matemática
(continua)

Variáveis independentes	Modelo 1	Modelo 2	Modelo 3
Efeitos Fixos			
Coeficientes			
Intercepto	267.36	275.72	228.67
Índice socioeconômico		7.01*	6.82*
Sexo		4.20*	4.24*
Raça		-8.38*	-8.15*
Defasagem		-24.59*	-24.50*

Obs.: (*) $p < 0,001$; (**) $p \leq 0,05$; (***) $p \leq 0,1$

Fonte: Dados: SEE/MG, 2019 e INEP, 2019.

Elaboração própria, 2020.

Tabela 5 - Resultados dos modelos hierárquicos lineares 1 a 3 de Matemática
(continuação)

Variáveis independentes	Modelo 1	Modelo 2	Modelo 3
Efeitos Fixos		Coefficientes	
PPP			-
Gestão do diretor			0.69***
ISE Escola			2.55*
IRD			3.72*
ICG			-1.86*
Comprometimento			30.66*
Clima Escolar			4.58*
Matéria			-
Programa			25.36*
Comunidade			0.62***
Efeitos Aleatórios		Variância	
Intercepto	325.4*	239.53*	142.00*
Inclinação do índice socioeconômico		17.48*	16.41*
% da variância entre escolas	12,22%	9,75%	

Obs.: (*) $p < 0,001$; (**) $p \leq 0,05$; (***) $p \leq 0,1$

Fonte: Dados: SEE/MG, 2019 e INEP, 2019.

Elaboração própria, 2020.

Uma medida importante usada nos estudos de eficácia escolar para justamente verificar a contribuição das escolas é o efeito-escola (PERERIRA, 2006). Ressalta-se, porém, que essa variância não será considerada como o efeito-escola, na medida em que não foi controlada pela proficiência prévia dos alunos, ou seja, não é possível de fato distinguir quanto a escola agregou ao aluno e quanto ele já sabia. Para um cálculo fidedigno, deve-se ser utilizar dados longitudinais, uma série temporal de observações sobre a proficiência do mesmo estudante. Só assim poder-se-ia discriminar quanto da proficiência é resultante do trabalho da escola e quanto já havia sido desenvolvido ou deriva de características pessoais do próprio aluno.

Ainda assim, fica claro que as escolas têm uma contribuição importante na determinação do desempenho de um aluno, há escolas mais eficazes que outras e essas diferenças foram responsáveis por pelo menos 10% da variação da proficiência dos alunos do 3º ano EM em 2019. A proporção da variância entre escolas é de 7,59% em Língua Portuguesa e 9,93% em Matemática, quando controladas pelas características pessoais dos alunos. Assim, a escola teve um papel mais determinante

para proficiência em Matemática do que de Língua Portuguesa, de fato há um entendimento de que a aprendizagem da Matemática depende mais da escola, enquanto outros componentes curriculares também são desenvolvidos em outros espaço que a criança frequenta, por isso em geral, prefere-se analisar a proficiência nessa disciplina (FRANCO et al, 2007).

Dessa forma, mesmo com a adição das características pessoais dos alunos, a maior parte da variância da proficiência ocorre dentro das escolas, depende dos alunos. Isso não altera, entretanto, a conclusão de que é possível melhorar o desempenho por ações internas a escola. Esses valores são inferiores àqueles observados no modelo nulo de Franco et. al (2007) e no modelo com controle pelo nível socioeconômico de Soares (2004a), 39% e 12%, respectivamente, para dados do SAEB de 2001 de Matemática. Isso deve, em parte, pelo fato de os dados daqueles estudos incluírem escolas particulares o que, naturalmente, introduz maior variabilidade entre as escolas.

Antes da análise dos coeficientes, é importante esclarecer que as estimativas da influência de cada fator refletem as variáveis incluídas no modelo, de forma que não é possível obter uma estimativa completamente não enviesada da contribuição de um fator escolar específico para o desempenho do aluno (ALVES; SOARES, 2003). Assim, ainda que a análise da magnitude do efeito seja importante, ela é apenas um indicativo, sendo importante considerar, principalmente, a significância dos coeficientes. Dessa forma, o efeito de uma variável é considerado, controlando-se o efeito das outras. Um exemplo disso é encontrado no modelo 2 da tabela 4: o aumento de uma unidade do índice socioeconômico gera um aumento de 7,3 unidades na proficiência esperada, considerando que as demais características do aluno, defasagem, raça e sexo, mantenham-se as mesmas.

Isso posto, verifica-se pela comparação da tabela 4 e 5, que a diferença entre os modelos 2 de cada componente é principalmente o efeito do sexo do aluno, relação verificada na seção anterior, bem como em outros estudos (ANDRADE et al, 2003). No PROEB 2019, meninos tiveram uma nota esperada 4,2 pontos maior em Matemática que meninas com as mesmas características de defasagem, raça e índice socioeconômico. Em Língua Portuguesa as meninas que têm mais vantagem, com uma diferença ainda maior de 17,17 pontos. Historicamente, as mulheres têm resultados educacionais piores, o que vem sendo observado é a inversão desse quadro, por exemplo, pela maior proporção de mulheres no 3º ano, como visto na

seção anterior, e não só pelos maiores desempenhos em Língua Portuguesa, como também pela progressiva redução das diferenças em Matemática. Ainda assim, elas tendem a ocupar carreiras menos relacionadas às ciências exatas e as de menor prestígio (ALVES; SOARES, 2016; SILVA; HASENBALG, 2000). Cabe ressaltar ainda, que em ambos os modelos, tanto sexo quanto as demais variáveis foram significativas.

Outra diferença é quanto o intercepto, mais uma vez, valor da proficiência esperada quando todas as variáveis equivalem a zero. No caso do modelo 2, é a proficiência esperada de um estudante do sexo feminino, branco, amarelo ou indígena, com trajetória regular e índice socioeconômico igual a média. Língua Portuguesa tem intercepto 6,5 pontos maior que Matemática (282,27, 275,72, respectivamente). Em relação à média da amostra, a proficiência de Matemática é cerca de 3 pontos maior. Essa diferença provavelmente é efeito das relações de gênero, que são inversas em cada componente curricular e muito acentuadas em favor das meninas em Língua Portuguesa.

Controlado pelos demais fatores, a “defasagem escolar” é a que mais impacta a proficiência, independente do componente avaliado, uma diferença de cerca de 25 pontos em relação a alunos com trajetória regular. A mesma situação é observada para o “índice socioeconômico” e para “raça”. Alunos pretos ou pardos têm uma proficiência 8,4 pontos menor que brancos, amarelos ou indígenas. O aumento de uma unidade do índice acarreta um aumento de cerca de 7 pontos na proficiência esperada, 7,3 no caso da proficiência em Língua Portuguesa. Salienta-se aqui que o modelo 2 também tem efeito aleatório para essa variável, conforme será discutido em seguida. Esses resultados refletem toda a discussão consolidada desde o Relatório Coleman (1966) sobre o quanto as características dos alunos são determinantes de seus resultados educacionais.

Pela proporção da variância entre escolas dos modelos 1 e 2 foi possível identificar que houve escolas mais eficazes em promover melhores resultados aprendizado do que outras o que condiz com a literatura apresentada no Capítulo 2. Então, o modelo 3 permite investigar, quais são as características das escolas que contribuem para a eficácia. Inicialmente destaca-se que a variável “efetividade da implementação do PPP” não foi significativa em nenhum modelo, conforme hipótese levantada na seção 4.1, o PPP provavelmente ainda não foi incorporado pelos atores escolares como instrumento de gestão pedagógica, por isso não apresenta efeito na

proficiência média da amostra. Por sua vez, “gestão do diretor” foi significativa apenas para o modelo de Matemática, enquanto “gestão da matéria” apenas para o modelo de Língua Portuguesa.

Isso, no entanto, não é suficiente para concluir que tais fatores não são relevantes de forma geral para educação, nem sequer que eles não tenham sido relevantes para estes alunos. Este estudo traz a análise de apenas uma turma e uma edição do PROEB, é importante que todos os resultados aqui encontrados sejam interpretados em vista da teoria consolidada até então, como discutido no Capítulo 2. Além disso, achados fora do esperado precisam ser investigados por estudos mais profundos, o que já se apresenta como um desdobramento dessa monografia.

De toda forma, as demais variáveis foram significativas e apresentaram o sinal esperado. A diferença entre os interceptos de cada componente curricular é 8,6 pontos a mais em Língua Portuguesa (com coeficiente igual a 237,24) do que em Matemática (228,67). Este é mais um indicativo de que as escolas são mais determinantes para os resultados dos alunos em Matemática, os fatores escolares são responsáveis por um aumento do desempenho médio esperado maior do que em Língua Portuguesa. Ainda que haja outras diferenças entre os modelos, nenhuma tão expressiva.

Em relação ao modelo de Língua Portuguesa, o fator que mais influencia a proficiência é o “comprometimento do professor” ($\gamma = 36,21$). Representa a proporção de professores que não faltam muito em cada escola, em escolas que os professores faltam menos a proficiência esperada é significativamente maior. Outra variável com coeficiente muito alto em comparação aos demais é o “percentual do programa concluído pelo professor”, que equivale a 16,81. Nas escolas em que foi concluído 25% do intercepto (o mínimo) a nota esperada é 241,44, enquanto dentre aquelas que concluíram todo o programa é 12,61 pontos maior (254,05). No entanto, a “gestão da matéria”, variável construída por análise fatorial de itens do questionário do aluno que perguntavam sobre os métodos de ensino do professor, tem o menor coeficiente igual a 0,65. Ainda que não tão expressivamente, de certo a gestão da matéria está relacionada ao desempenho. Além disso, o “indicador de regularidade docente”, que diz respeito a permanência de um professor em uma mesma escola, uma medida de seus insumos, impacta 2,49 pontos no desempenho esperado dos alunos.

Observa-se assim, a relevância de fatores escolares, em especial os relacionados ao professor, no desempenho dos alunos, mesmo após o devido controle

pelas suas características pessoais. Resultados semelhantes, que demonstram a relevância da atuação do professor para a proficiência dos alunos, foram encontrados por Machado Soares (2003, 2005), Lee, Franco e Albernaz (2004) e Soares (2004a), conforme revisado na seção 2.3.

Vale destacar ainda o “clima escolar”, com coeficiente igual a 4,18, mensurado por análise fatorial dos itens do questionário dos alunos sobre as relações entre os atores escolares, a ênfase acadêmica e um ambiente disciplinado, entende-se assim que esses são aspectos determinantes da proficiência. A relação com a comunidade, ainda que não tão expressiva, também é determinante ($\gamma = 0,49$). Quanto ao “indicador de complexidade da gestão” que fornece um controle dos fatores contextuais teve um coeficiente negativo igual a -1,01, o que significa que quanto mais complexa a gestão da escola, menor tende a ser seus resultados.

Enquanto a variável ISE diz respeito aos recursos econômicos e culturais da família do aluno, o ISE Escola mensura a “bagagem” cultural dos alunos de uma escola, o impacto positivo de estudar em uma escola com alto ISE médio é chamado de efeito dos pares. A interação que existe entre alunos com melhores condições é benéfica ao desempenho e esse benefício se estende a todos alunos da escola, mesmo os com ISE baixo (ALVES; SOARES, 2003; RODRIGUES; RIOS-NETO; PINTO, 2011). O coeficiente do ISE Escola foi igual a 2,82, observa-se que o coeficiente do ISE diminui quando acrescenta-se as variáveis da escola, assim o impacto do índice socioeconômico passa a ser o efeito do capital econômico e cultural das famílias (6,94) mais o impacto da interação do grupo (2,82), o equivalente a 9,76, maior que o efeito do ISE no segundo modelo (sem variáveis da escola, 7,26). Ressalta-se ainda que o efeito dos pares encontrado é menor do que observado em outros estudos, provavelmente por não incluir escolas privadas, as quais além de terem alunos com as melhores condições sociais, ainda possuem políticas de seleção (SOARES, 2004).

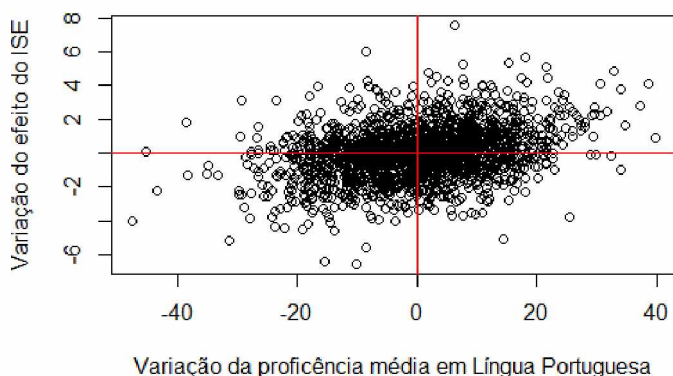
As conclusões para o modelo de Matemática são muito similares. O “comprometimento do professor” também é o fator que mais determina o aprendizado desse componente, mas com um coeficiente menor ($\gamma = 30,66$). A conclusão do programa, no entanto, é quase 10 pontos mais importante que para Língua Portuguesa ($\gamma = 25,36$). As variáveis “índice socioeconômico da escola” ($\gamma = 2,55$), “indicador de regularidade docente” ($\gamma = 3,72$), “indicador de complexidade da gestão” ($\gamma = -1,86$), “clima escolar” ($\gamma = 4,58$) e “relação com a comunidade” ($\gamma = 0,62$) também

são determinantes do desempenho em Matemática, têm coeficientes bem semelhantes, mas um pouco maiores para esse modelo. Uma diferença importante é que a “gestão do diretor” é significativa nesse modelo, ainda que seu coeficiente, 0,62, não seja tão alto, é um fator importante para o desempenho desse componente curricular. Os modelos 3 deixam claro a importância da escola para a promoção da qualidade educacional do 3º ano em 2019.

Para além da qualidade, as escolas também podem ser eficazes por promoverem equidade. Esta é observada quando os fatores intraescolares diminuem o impacto no índice socioeconômico no desempenho dos estudantes, relação que pode ser verificada nos modelos 2 e 3 pelo coeficiente aleatório do índice socioeconômico. Estes são uma parcela do efeito de um fator que varia para cada escola, esse efeito é dado pelo coeficiente fixo mais o erro do coeficiente da escola, ou efeito aleatório. O efeito aleatório foi significativo para o índice socioeconômico nos dois modelos. Isso significa que o efeito das condições sociais dos alunos não é o mesmo para todos os alunos, ele varia de acordo com a escola, com seus processos internos, indicando que há escolas mais equitativas que outras dentro da amostra analisada.

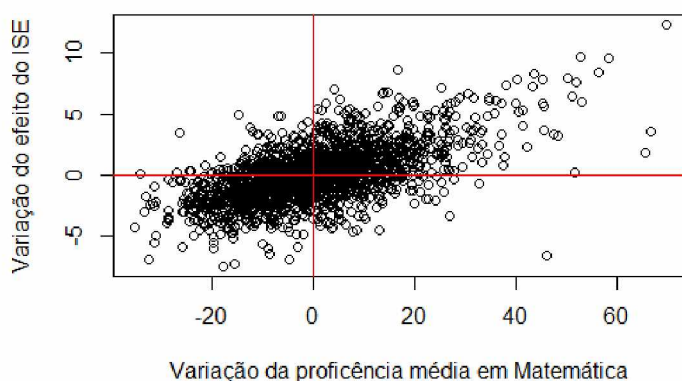
Os gráficos 25 e 26 demonstram a relação entre os dois componentes aleatórios dos modelos, o erro do intercepto e o erro do coeficiente do índice socioeconômico. Mais uma vez, o primeiro refere-se a distância da proficiência média de uma escola das demais, já o segundo, a quanto o efeito do índice socioeconômico varia para cada escola. Entende-se que quanto mais a direita está uma escola (representada pelos círculos pretos) maior é sua proficiência média e quanto mais abaixo, menor o efeito do nível socioeconômico. Assim, a melhor situação é das escolas do quadrante inferior direito, maior qualidade e equidade. No entanto, quanto maior a proficiência média das escolas maior tende a ser o efeito do índice socioeconômico no desempenho médio dos alunos, principalmente para proficiência em Matemática, como discutido na seção anterior.

Gráfico 32 – Relação entre o efeito aleatório da proficiência em Língua Portuguesa do 3º ano do Ensino Médio e o efeito aleatório do ISE dos alunos, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Gráfico 33 – Relação entre o efeito aleatório da proficiência em Matemática do 3º ano do Ensino Médio e o efeito aleatório do ISE dos alunos, Minas Gerais, 2019.



Fonte: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Para uma melhor compreensão de como as escolas se distribuem por critérios de qualidade e equidade, dividiu-se os dados para cada quartil, as médias das variáveis, bem como a proporção de escolas por quartil são apresentados nas tabelas 6 e 7. Há de se reforçar mais uma vez a relação inversa entre qualidade e equidade, principalmente para Matemática. A maioria das escolas estão no segundo quartil (qualidade a cima da média e equidade abaixo da média) ou no terceiro quartil (qualidade abaixo da média e equidade a cima da média), para Língua Portuguesa, respectivamente 29,6% e 28,7% das escolas, quanto para Matemática, mais escolas ainda estão nesses quartis 32,1% e 39,7%. O primeiro quartil são das escolas com os dois parâmetros abaixo da média, existem uma proporção maior de escolas nesse quartil em Língua Portuguesa (20,7%) do que em Matemática (13,4%), ainda assim,

a situação é inversa no quarto quartil (em que os dois parâmetros são a cima da média) 21% em Língua Portuguesa e 14,8% em Matemática.

Como esperado, dentre as escolas de maior qualidade (quartis 2 e 4) a proficiência média é maior para aquelas com menor equidade, o mesmo vale para o grupo de escolas com qualidade abaixo da média (1 e 3). Pontua-se apenas que não há muita diferença na proficiência em Matemática entre as escolas abaixo da média de qualidade em Língua Portuguesa (tabela 6). A proporção de meninos por quartil é a mesma independente do componente curricular e varia muito pouco por quartil, o primeiro com 46% e o quarto com 43%. A distribuição de alunos pretos ou pardos e defasados, por sua vez, também independe o componente curricular, quanto menor a proficiência média maior a proporção desses alunos. O inverso para o “ISE”, quanto maior, maior tendem a ser as proficiências. Ressalta-se aqui que o segundo quartil é o único em que a média dessa variável é maior que a média da amostra de todos os alunos analisados.

Tabela 6 – Análise descritiva por quartil do gráfico 32

	1° quartil	2° quartil	3° quartil	4° quartil
Qualidade	-	+	-	+
Equidade	-	-	+	+
Número de Escolas	420	600	581	426
Percentual de Escolas	20,7%	29,6%	28,7%	21,0%
PROFICIENCIA_LP	253,27	281,12	251,92	275,83
PROFICIENCIA_MT	257,49	283,49	257,14	277,46
SEXO	0,46	0,44	0,45	0,43
RACA	0,73	0,63	0,73	0,69
DEFAS	0,13	0,10	0,13	0,12
ISE	-0,16	0,23	-0,14	-0,07

Fonte: Dados: SEE/MG, 2019.

Elaboração própria, 2020.

Tabela 7 - Análise descritiva por quartil do gráfico 33

	1° quartil	2° quartil	3° quartil	4° quartil
Qualidade	-	+	-	+
Equidade	-	-	+	+
Número de Escolas	271	651	805	300
Percentual de Escolas	13,4%	32,1%	39,7%	14,8%
PROFICIENCIA_LP	258,00	279,50	253,92	275,44
PROFICIENCIA_MT	259,64	285,95	254,22	279,17
SEXO	0,46	0,44	0,45	0,43
RACA	0,72	0,63	0,74	0,67
DEFAS	0,12	0,10	0,14	0,11
ISE	-0,09	0,19	-0,15	-0,03

Fonte: Dados: SEE/MG, 2019.

Elaboração própria, 2020.

Rodando o modelo 3 com os grupos de escolas de cada quartil, tem-se os resultados apresentados nas tabelas 8 e 9. A primeira coluna apresenta os resultados com todos os dados, enquanto as subsequentes apenas as escolas referentes a cada quartil. Observa-se que os interceptos têm uma distribuição semelhante às proficiências médias das tabelas 6 e 7. Quanto ao “ISE”, verifica-se que seu coeficiente é menor nos modelos das escolas mais equitativas, chegando a ser negativo (-0,57) no modelo do quarto quartil de Matemática, o que mostra que há escolas que de fato são mais equitativas, inclusive dentre aquelas com qualidade acima da média. Enfatiza-se, que esse coeficiente pode indicar uma homogeneidade do grupo, ou seja, como todos os alunos são igualmente pobres o ISE não tem efeito, no entanto, o desvio padrão de cada quartil é muito semelhante ao da amostra dos dados, ou seja, a dispersão em cada grupo e entre todos os dados é bem próxima.

No entanto, há outros fatores de desigualdade que precisam ser considerados, em relação ao “sexo”, a diferença é maior em favor dos meninos entre o grupo de escola com mais qualidade para os dois componentes curriculares. Os coeficientes de “raça” variaram muito pouco nos modelos de Língua Portuguesa. Já para Matemática, o efeito é maior para as escolas com qualidade acima da média, assim como “defasagem escolar” (em ambos componentes). Esses resultados reforçam as evidências da seção 4.2 de que as desigualdades por gênero, raça e trajetória escolar são acentuadas quanto maior o ISE.

Quanto as variáveis da escola, destaca-se que houveram bem mais variáveis não significativas. Isso se deve tanto pela redução da amostra, que impacta no grau de significância principalmente para o primeiro e quarto quartis, quanto pela

redução da variabilidade entre as escolas, ou seja, é possível que uma variável tenha efeito para um grupo de escolas, mas como todas as escolas do grupo são muito parecidas, esse efeito não é captado pelo modelo. De fato, para além dos coeficientes que não foram significativos, os que foram em geral têm coeficientes menores do que os do modelo com todos os dados. O “ISE Escola” foi significativo para os dois componentes curriculares no terceiro quartil, aquele com a menor média do ISE e em que a equidade é acima da média, justamente porque essa variável indica o efeito dos pares. Por outro lado, teve um coeficiente negativo (-1,29) no primeiro quartil de Língua Portuguesa, provavelmente por conta do alto valor do coeficiente do ISE (13,79).

As variáveis “efetividade do PPP” e “gestão do diretor” não foram significativas em nenhum modelo. Já “indicador de regularidade docente”, “comprometimento do professor”, “clima escolar” e “percentual do programa concluído” são variáveis que quando foram significativas tiveram coeficientes menores. O “indicador de regularidade docente” e o “clima escolar” foram as duas variáveis significativas para os quatro quartis dos modelos de Língua Portuguesa, para Matemática foi o “comprometimento do professor.” Quanto a “gestão da matéria” e a “relação com a comunidade escolar”, quando foram significativas tiveram um coeficiente maior. A “gestão da matéria”, inclusive, não foi significativa para o modelo de Matemática com todos os dados, mas o foi para as escolas do segundo quartil, provavelmente porque considerando todas as escolas, o efeito dessa variável é muito associado ao efeito de outra variável.

Tabela 8 - Resultados das interações significativas dos modelos hierárquicos lineares 4 de Língua Portuguesa e Matemática

(continua)

Qualidade		-	+	-	+
Equidade		-	-	+	+
Variáveis independentes	Modelo 3	1° quartil	2° quartil	3° quartil	4° quartil
Efeitos Fixos		Coefficientes			
Intercepto	237,24	264.68	274.48	250.70	272.37
Índice socioeconômico	6.94*	13.79*	11.42*	1.021**	0.44*
Sexo	-17.08*	-19.36*	-15.29*	-19.83*	-14.3*
Raça	-8.17*	-8.06*	-8.02*	-8.29*	-7.85*
Defasagem PPP	-24.66*	-22.68*	-27.61*	-22.31*	-25.35*
Gestão do diretor	-	-	-	-	-
ISE Escola	2.82*	-1.26**	-	3.13*	-
IRD	2.49*	1.53*	0.86**	1.61*	0.31*
ICG	-1.01*	-	-1.59*	-	-
Comprometimento	36.21*	-	12.08**	13.97**	-
Clima Escolar	4.18*	2.56*	2.12*	0.92***	1.65*
Matéria	0.65***	1.26***	-	1.33**	-
Programa	16.81*	-	8.29***	10.86**	-
Comunidade	0.49*	-	1.32**	0.98**	-

Tabela 9 - Resultados das interações significativas dos modelos hierárquicos lineares 4 de Língua Portuguesa e Matemática

(continua)

Qualidade		-	+	-	+
Equidade		-	-	+	+
Variáveis independentes	Modelo 3	1° quartil	2° quartil	3° quartil	4° quartil
Efeitos Fixos		Coefficientes			
Intercepto	228,67	255.41	255.83	243.87	269.61
Índice socioeconômico	6.82*	13.84*	11.56*	2.51*	-0.57*
Sexo	4.24*	2.23**	6.08*	2.22*	6.66**
Raça	-8.15*	-7.70*	-9.26*	-6.17*	-8.47*
Defasagem	-24.50*	-20.73*	-29.76*	-19.35*	-29.82*
PPP	-	-	-	-	-
Gestão do diretor	0.69***	-	-	-	-
ISE Escola	2.55*	-	-	2.67*	-
IRD	3.72*	-	1.09**	2.08*	-
ICG	-1.86*	-	-2.09*	-	-
Comprometimento	30.66*	10.40***	15.23***	8.01***	8.57***
Clima Escolar	4.58*	-	2.57*	1.95*	-
Matéria	-	-	2.41*	-	-
Programa	25.36*	-	23.12*	15.05	-
Comunidade	0.62***	-	0.97***	-	-

Tendo visto que existem escolas em que o efeito do ISE é bem mais baixo, vale verificar o quanto os processos escolares promovem equidade. Ou seja, investigar se os processos escolares incluídos no modelo, além de terem contribuído para a qualidade educacional também foram promotores de equidade intraescolar. Os modelos 4 possuem interação das variáveis já apresentadas com o “índice socioeconômico”. Em uma regressão é possível permitir que o efeito de uma variável explicativa dependa do valor de outra variável explicativa. Especificamente na regressão hierárquica, é possível que a interação seja entre variáveis de diferentes níveis. Dessa forma, o efeito de uma variável explicativa é dado pelo seu coeficiente mais o coeficiente de interação multiplicado pelo valor da outra variável.

A tabela 10 apresenta os resultados das interações significativas dos modelos 4. A análise deve levar em consideração que o coeficiente do “índice

socioeconômico” é positivo, uma redução em seu efeito seria causada por variáveis com interações negativas. Nos dois modelos observa-se as mesmas variáveis com coeficiente significativos e com valores não muito distintos. As variáveis com coeficientes de interação negativos são “raça”, “defasagem”, “gestão da matéria” e “relação com a comunidade”. Assim, ser preto ou pardo reduz o efeito da condição socioeconômica em 1,38 pontos para Língua Portuguesa e 1,52 para Matemática. Na mesma direção, ser defasado reduz em, respectivamente, 3,18 pontos e 4,22 pontos o impacto do ISE.

A “gestão da matéria” e a “relação com a comunidade” são os únicos fatores escolares com potencial de gerar equidade, pois a primeira reduz o efeito das condições socioeconômicas em 0,44 em Português e 0,54 em Matemática, já o segundo, respectivamente 0,22 e 0,36. Por outro lado, as variáveis “clima escolar” e “regularidade docente” acirram as desigualdades, principalmente para o aprendizado de Matemática. Todos os alunos se beneficiam com o efeito dessas variáveis, no entanto, aqueles com maiores níveis socioeconômicos também se beneficiam pelo efeito adicional da interação. Para a proficiência em Língua Portuguesa o aumento é de 0,43 em relação ao “clima escolar” e 0,33 quanto a “regularidade docente”. Já para Matemática os coeficientes das interações equivalem a 0,59 e 0,89.

Tabela 10 - Resultados das interações significativas dos modelos hierárquicos lineares 4 de Língua Portuguesa e Matemática

	Raça	Defasagem	Gestão da matéria	Relação com a comunidade	Regularidade docente	Clima Escolar
Língua Portuguesa	-1.38*	-3.18*	-0.44**	-0.22***	0.33*	0.43**
Matemática	-1.52*	-4.22*	-0.54**	-0.36***	0.89*	0.59**

Obs.: (*) $p < 0,001$; (**) $p \leq 0,05$; (***) $p \leq 0,1$

Fonte: Dados: SEE/MG, 2019 e INEP, 2019.

Elaboração própria, 2020.

Percebe-se assim, que as escolas respondem por parcela significativa da variabilidade de notas. Não obstante, muitos dos fatores intraescolares analisados são promotores de qualidade, ou seja, estão relacionados a uma melhor proficiência. Ainda assim, fica claro que a relação entre qualidade e equidade é muito complexa, pois, desses fatores, apenas “gestão da matéria” e “relação com a comunidade” reduziram o impacto do nível socioeconômico, sendo que “clima escolar” e “indicador

de regularidade docente”, por outro lado, aumentam o efeito dessa variável. Dessa forma, práticas que aumentam a qualidade da educação não necessariamente reduzem as desigualdades, podendo inclusive ampliá-las. Como então, melhorar os baixos níveis de desempenho do ensino médio mineiro sem desconsiderar as fortes desigualdades existentes? Esta monografia, buscou trazer evidências que contribuíssem para o enfrentamento desse desafio, a síntese desses resultados é apresentada na seção seguinte.

4.4 A (in)eficácia das escolas mineiras de ensino médio

Essa monografia faz uma análise dos resultados das escolas públicas mineiras no PROEB 2019, nessa seção será feita uma síntese dos resultados encontrados. Primeiramente, destaca-se um quadro de baixa qualidade da educação, a média de proficiência em Língua Portuguesa e em Matemática é considerada baixa, mais de 90% dos alunos têm proficiências abaixo do recomendado, sendo que destes, a maioria possui nível baixo.

Em segundo lugar, foram evidenciadas desigualdades educacionais existentes entre os alunos. Há um aumento das médias de proficiência nos níveis socioeconômicos mais elevados, sendo que as maiores disparidades estão entre os alunos com maiores proficiências, indicando que o impacto da condição social do aluno pode variar de acordo com sua proficiência. Verificou-se também o efeito de outras características pessoais dos alunos no desempenho nas avaliações, como sexo, raça e trajetória escolar. Meninas tendem a ter menores notas em matemática e meninos notas menores ainda em Língua Portuguesa. Ser preto ou pardo e especialmente, ser defasado, impacta negativamente a proficiência. Essas diferenças se mantêm mesmo considerando o efeito do índice socioeconômico. Na verdade, quanto melhores as condições sociais dos alunos, mais significativo é o efeito da raça e da defasagem escolar.

A análise dos modelos de regressão linear hierárquica indica que, de fato, as características dos alunos importam muito para o aprendizado, mas que se somam à diferença que as escolas fazem, principalmente para as notas de Matemática. Destaca-se a relevância de fatores relacionados aos professores: “comprometimento do professor”, “percentual do programa concluído” e menos expressivo, mas ainda relevante para Língua Portuguesa, “a gestão da matéria”. Outros fatores como o “clima escolar”, “relação com a comunidade”, “regularidade docente” e, no caso de

Matemática, a “gestão do diretor” também foram determinantes. Observa-se também que a complexidade da gestão da escola tem efeito negativo no desempenho. No entanto, a “efetividade da implementação do PPP” não foi significativa.

Verificou-se também que o impacto do “índice socioeconômico” difere a cada escola o que pode indicar que umas escolas são mais equitativas que outras, principalmente aquelas que tem menores proficiências médias. Dentre as escolas com maior equidade, o modelo por quartis mostrou que o impacto das condições socioeconômicas dos alunos é bem menor, mas persistem outras desigualdades. Para mais, não foi tão evidente a influência dos fatores escolares na redução dos efeitos das condições sociais sobre as proficiências. Apenas a “gestão da matéria” e “relação com a comunidade” geram um efeito pró equidade, enquanto “clima escolar” e “indicador de regularidade docente” aumentavam as desigualdades sociais. Os fatores dos alunos, “raça” e “defasagem” também diminuíram o efeito do “índice socioeconômico” sobre o aprendizado em Português e Matemática.

A tabela 11 sumaria os resultados dos modelos quanto a qualidade e equidade, indica se cada variável teve efeito positivo (“+”) ou negativo (“-“), tanto na proficiência de Língua Portuguesa e Matemática, quanto no efeito do “índice socioeconômico” nessa proficiência. O cenário mineiro é de qualidade educacional baixa, mas a escola é capaz de contribuir para o aumento do desempenho dos alunos, ainda assim, existem desigualdades e a relação dos fatores que as promovem é muito complexa. A escola contribui para diminuir as desigualdades, mas muito pouco, em geral, na verdade, a melhora do desempenho é feita às custas do aumento das desigualdades.

Tabela 11 - Resultados das interações significativas dos modelos hierárquicos lineares 4 de Língua Portuguesa e Matemática

Variáveis	Impacta o desempenho?		Impacta o efeito do ISE?
	Língua Portuguesa	Matemática	
Índice socioeconômico	+	+	
Sexo (masculino)	-	+	
Raça (preto ou pardo)	-	-	-
Defasagem Escolar	-	-	-
Efetividade da implementação do PPP			
Gestão do diretor		+	
Índice socioeconômico da escola	+	+	
Indicador de regularidade docente	+	+	+
Indicador de complexidade da gestão	-	-	
Comprometimento do professor	+	+	
Clima Escolar	+	+	+
Gestão da matéria	+		-
Percentual do programa concluído	+	+	
Relação com a comunidade	+	+	-

Fonte: Dados: SEE/MG, 2019 e INEP, 2019.
Elaboração própria, 2020.

5 CONSIDERAÇÕES FINAIS

O objetivo desse trabalho foi analisar a eficácia das escolas públicas do EM de Minas Gerais em 2019. Assim, verificou-se quais os processos escolares contribuíram para a qualidade da educação, mensurada pela proficiência em Língua Portuguesa e Matemática no PROEB, como as características dos alunos geravam desigualdades nestes resultados, medidas pelas diferenças de desempenho médio entre grupos de alunos e, dentre processos escolares, quais deles também contribuem para a equidade intraescolar ou igualdade de capacidades, ou seja, para a mitigação dos efeitos do nível socioeconômico no aprendizado.

Para tanto, valeu-se dos modelos conceituais de Scheerens (1990) e Soares (2004a) apresentados no capítulo 2 e utilizou-se de modelos hierárquicos lineares, descritos no capítulo 3. Já o capítulo 4 apresenta evidências de desigualdades educacionais por índice socioeconômico, gênero, raça e trajetória escolar dentre os alunos analisados. Constata-se ainda que os efeitos dessas desigualdades não são os mesmos entre todos os alunos. Maiores índices socioeconômicos acirram as diferenças por gênero (no caso de Língua Portuguesa), raça e trajetória escolar. Além disso, as condições socioeconômicas parecem mais determinantes entre alunos com maiores níveis.

Não obstante, verifica-se que o quadro histórico de baixos desempenhos a nível nacional, principalmente no ensino médio, também é evidente no contexto estadual: a maior parte dos alunos analisados tiveram proficiências baixas. Ainda assim, há de salientar que as escolas foram sim importantes para os resultados obtidos e a maioria dos processos intraescolares analisados se mostram importantes para o desempenho dos alunos, em especial os relacionados ao professor. Ainda que o efeito do nível socioeconômico varie para cada escola, mostrando que é possível que haja escolas mais equitativas que outras, os únicos processos escolares que reduzirem o efeito do nível socioeconômico na proficiência são a “gestão da matéria” e a “relação com a comunidade”, enquanto o “clima escolar” e o “ indicador de regularidade docente” agravavam tais efeitos.

Os estudos de eficácia escolar surgiram para evidenciar que as escolas importam para o desempenho dos alunos, tal como foi evidenciado por essa monografia também. Não se pode deixar de ressaltar que o efeito das escolas não é irrestrito e que as características pessoais dos alunos impactam muito. Para além

disso, é preciso ter em conta que mudanças marginais no modelo da escola têm efeito limitado, principalmente quando se constata uma estrutura escolar tão descolada da realidade das crianças e jovens (TARTUCE et al., 2018). Por isso, mais que uma melhora dos fatores escolares é necessário um avanço de sua estrutura, de forma que se adeque a como a sociedade se organiza, a nova forma de aprendizado e ao contexto sociocultural em que os alunos se inserem.

Esse trabalho, de certo, contribui para a compreensão da realidade das escolas públicas do ensino médio de Minas Gerais. Ainda assim, é preciso considerar que se trata apenas de um conjunto de dados específicos e, como explicitado anteriormente, as variáveis utilizadas e a forma que são agregadas, a escolha e modelagem da regressão influenciam muito os resultados encontrados. Dessa forma, espera-se que as conclusões apresentadas sejam interpretadas a luz da literatura referenciada, brevemente discutida no capítulo 2.

As evidências mostram que as escolas fazem a diferença no aprendizado. Isso significa dizer que há, dentro delas, espaço significativo para mudanças que podem contribuir de forma efetiva para promoção de qualidade de aprendizado. Mas ainda persiste a relação entre origem e alcance educacional, ou seja, as características dos alunos e de suas famílias têm peso na proficiência. Dessa forma, as políticas educacionais de promoção de qualidade devem sempre estar atreladas a políticas de promoção de equidade, pois as evidências mostram que essa relação é complexa e muitas vezes fatores que aumentam o desempenho médio também acirram as desigualdades entre diferentes grupos de estudantes.

A melhora dos resultados educacionais é um processo multiforme, na medida em que os fatores escolares são construídos a partir das relações sociais na escola e na comunidade escolar, dos valores que orientam essas relações e como isso se materializa em ações concretas. Assim, os fatores aqui estudados são interdependentes, e bons resultados educacionais dependem de um conjunto de bons processos escolares. É necessário ainda, que as políticas e práticas pedagógicas sejam orientadas de forma a compensar pelas desigualdades de resultado.

Entender essas relações e como os processos escolares impactam nos resultados é, então, fundamental. Ainda assim, deve-se ter em conta as especificidades de cada realidade e para que as mudanças sejam efetivas é importante que elas partam da escola, o que realmente são grandes desafios, principalmente considerando o cenário premente evidenciado pelos dados do PROEB

analisados nesta monografia. Dessa forma, o objetivo não é ser prescritivo, nem providenciar “receitas” de quais fatores são importantes. Afinal, não apenas o “o que”, ou quais fatores são importantes, devem variar de acordo com o contexto, mas também o “como”, ou seja, a forma de promover a melhora desses fatores, de maneira geral, a forma de promover um avanço educacional. Este é o grande desafio que se coloca para gestão educacional, como garantir uma melhora dos fatores escolares dada as múltiplas realidades e em especial como fazer com que o aumento da qualidade seja acompanhado de um aumento da equidade.

Vale reiterar que uma análise adequada dos efeitos das escolas deve ser feita a partir de dados longitudinais. Em primeiro lugar, pela questão evidente de que se os resultados educacionais dependem tanto dos alunos, não seria adequado comparar escolas através do desempenho de alunos diferentes. Mas também porque o resultado educacional não é único, ele é construído a partir da trajetória do aluno. Já existem pesquisas brasileiras que utilizam de dados longitudinais (“projeto GERES”, “Avaliação do desempenho: fatores associados”(LAROS; KARINO, 2017 e PEREIRA,2006)), mas pelos altos custos desse tipo de pesquisa e pelas dificuldades operacionais da coleta desses dados, eles não são tão utilizados.

A alta proporção de alunos com baixa proficiência desperta a noção de que tão importante quanto analisar os fatores correlatos de eficácia seria analisar àqueles correlatos à ineficácia. Sammons (1999) levanta o ponto de que não é possível assumir que a forma que as “boas escolas” alcançam e mantenham bons resultados seja a mesma que uma “escola ineficaz” possa melhorá-los. Principalmente em um contexto de altas desigualdades, como no Brasil, é preciso então, buscar entender também, por que algumas escolas não agregam tanto aos seus alunos. Além disso, é interessante investigar melhor o fenômeno da equidade, como promover medidas compensativas que compensem as desigualdades para além de um aumento ou uma melhor distribuição de recursos.

Ademais, verifica-se uma limitação dessa pesquisa por apenas analisar o efeito dos fatores dos alunos e das escolas na média da escola. No entanto, encontrou-se evidências de que, ao menos as características dos alunos, não têm o mesmo efeito em todos os alunos da amostra. O índice socioeconômico, por exemplo, parece ser mais determinante de altos desempenhos que de baixos desempenhos. Assim, é preciso compreender não apenas se é eficaz, mas para quem, ou seja, é

interessante que se investigue como os fatores escolares afetam diferentes grupos de alunos.

A partir disso, espera-se que os gestores das políticas educacionais orientem suas ações cientes da centralidade escola para o alcance de bons resultados. Ressalta-se também seu papel não apenas em induzir as mudanças nas escolas, como também estarem aptos a prestar o suporte necessário a elas. É importante sempre evidenciar as desigualdades de resultados existentes entre os estudantes, promover, e provocar os atores escolares a promoverem, políticas e práticas educacionais que diminuam essas disparidades.

REFERÊNCIAS

- ALBERNAZ, A.; FERREIRA, F.; FRANCO, C. **Qualidade e equidade no ensino fundamental brasileiro**. Pesquisa e Planejamento Econômico, Rio de Janeiro, v. 32, n. 3, 2002.
- ALVES, Maria Teresa Gonzaga; SOARES, José Francisco. **Efeito-escola e estratificação escolar: o impacto da composição de turmas por nível de habilidade dos alunos**. Educação em Revista, v. 45, p. 25-58, jun. 2007.
- ALVES, Maria Teresa Gonzaga; SOARES, José Francisco; XAVIER, Flavia Pereira. **Desigualdades educacionais no ensino fundamental de 2005 a 2013: hiato entre grupos sociais**. Belo Horizonte. 2016.
- ALVES, Maria Teresa Gonzaga. **Efeito-Escola e Fatores Associados ao Progresso Acadêmico dos Alunos entre o Início da 5ª Série e o fim da 6ª Série do Ensino Fundamental: Um Estudo Longitudinal em Escolas Públicas no Município de Belo Horizonte – MG**. Orientador: José Francisco Soares. Belo Horizonte, 2006. Tese de Doutorado Faculdade de Educação, Universidade Federal de Minas Gerais
- ANDRADE, M., FRANCO, C.; CARVALHO, J. B. P. **Gênero e desempenho em matemática ao final do ensino médio: quais as relações?** Estudos em Avaliação Educacional, n. 27, p. 77-95, jan./jun. 2003.
- ASSIS, Marcos Arcanjo. **Educação e desigualdade em Minas Gerais: um diagnóstico a partir das transições educacionais**. 2012. Tese (Mestrado em Administração Pública).- Fundação João Pinheiro, Belo Horizonte, 2012.
- BARBOZA, Eleuza Maria Rodrigues. **A Composição das Turmas e o Desempenho Escolar na Rede Pública de Ensino Minas Gerais**. Orientador: Alicia Catalano Bonamino. Rio de Janeiro, 2006. 100 p. Tese de Doutorado – Departamento de Educação, Pontifícia Universidade Católica do Rio de Janeiro.
- BATES D., MÄCHLER M., BOLKER B., WALKER S. **Fitting Linear Mixed-Effects Models Using lme4**. Journal of Statistical Software, 67(1), 1–48. 2015. doi: 10.18637/jss.v067.i01.
- BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Brasília, DF: Centro Gráfico, 1988.
- BROOKE, N. Eficácia escolar. In: OLIVEIRA, D.A.; DUARTE, A.M.C.; VIEIRA, L.M.F. **DICIONÁRIO: trabalho, profissão e condição docente**. Belo Horizonte: UFMG/Faculdade de Educação, 2010. CDROM
- BROOKE, Nigel; SOARES, José Francisco (orgs). **Pesquisa em Eficácia Escolar: origem e trajetória**. Tradução: Viamundi Idiomas e Traduções; Cleusa Aguiar Brooke; Rômulo MonteAlto. Belo Horizonte: Editora UFMG, 2008.

BROOKOVER, W. B. **School Social Systems and Student Achievement: Schools Can Make a Difference**. New York: Praeger Publishers, 1979.

BRYK, A.; RAUDENBUSH, S. **Hierarchical Linear Models: applications and data analysis methods**. London: Sage Publications, 1992.

CACE - CENTRAL ADVISORY COUNCIL FOR EDUCATION. **Children and their Primary Schools** (The Plowden Report). London, 1967.

CARVALLO-PONTÓN, M. **Eficacia escolar: antecedentes, hallazgos y futuro**. Revista Internacional de Investigación en Educación, v. 3, n. 5, p. 199-214, jul./dez. 2010.

COLEMAN, J. S.; CAMPBELL, E.; HOBSON, C.; MCPARTLAND, J.; MOOD, A.; WEINFELD, R.; YORK, R. **Equality of Educational Opportunity**. Washington, DC: US Department of Health, Education & Welfare, 1966.

EDMONDS, R.R. **Effective School for the Urban Poor**. Educational Leadership, v37, n. 1 p. 15-27, 1979.

FIELD A. **Descobrimo a estatística usando o SPSS**. 2nd ed. Porto Alegre: Artmed; 2009.

FRANCO, Creso et al. **Qualidade e equidade em educação: reconsiderando o significado de "fatores intra-escolares"**. Ensaio: aval.pol. públ.Educ., Rio de Janeiro, v. 15, n. 55, p. 277-298, June 2007.

GOLDSTEIN, H. **Multilevel statistical models**. 2.ed. Londres: Edward Arnold, 1995.

HOX, J. J. **Applied multilevel analysis**. Amsterdam: TT-Publicities, 1995. Caps. 2, 4.

INEP, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Censo da Educação Básica: Sinopse Estatística da Educação Básica – 2019**. Brasília, 2019. Disponível em: < <http://www.inep.gov.br/superior/censosuperior/sinopse/default.asp>>. Acesso em jul. 2020.

INEP, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Evidências da Edição 2017**. Brasília, 2018. Disponível em: < http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=94161-saeb-2017-versao-ministro-revfinal&category_slug=agosto-2018-pdf&Itemid=30192>. Acesso em jul. 2020.

INEP, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Nota Técnica CGCQTI/DEED/INEP nº 11/2015**. Brasília, 2015.

INEP, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Nota Técnica Nº 040/2014**. Brasília, 2014.

INEP, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Sistema de Avaliação da Educação Básica (Saeb)**. Brasília, 2020. Disponível em: <<https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/saeb>>. Acesso em jul. 2020.

JANUARIO, André Alves. **Pesquisa em eficácia escolar no Brasil: um estudo sobre a produção científica brasileira com foco nos fatores relacionados à eficácia escolar (2005- 2017)**. 2019. 139f. Dissertação (Mestrado em Educação) Pontifícia Universidade Católica de Campinas, Centro de Ciências Humanas e Sociais Aplicadas, Programa de Pós-Graduação em Educação, Campinas, 2019.

JENCKS, C. S; BROWN, M. D. **Effects of High Schools on their Students**. Havard Center for Educational Policy Research, 1972.

KARINO, Camila A.; LAROS, Jaap A. **Estudos brasileiros sobre eficácia escolar: uma revisão de literatura**. Revista Examen, v. 1, n. 1, p. 95-126, 2017.

LEE, V. E.; FRANCO, C.; ALBERNAZ, A. **Quality and equality in brazilian secondary schools: a multilevel cross-national school effects study**. International Review of Contemporary Sociology, 2007.

LEE, V.E.; BRYK, A.S.; SMITH, J. **The organization of effective secondary schools**. En: Darling-Hammond, L. (Org). Review of Research in Education. Washington, American Educational Research Association, n. 19, pp. 171-268, 1993.

LORD, F. M. **Applications of item response theory to practical testing problems**. Hillsdale: Lawrence Erlbaum, New York, 1980.

MADAUS, G.F.; AIRASIAN, Peter W.; KELLAGHAN, Thomas. **School Effectiveness: A Reassessment of the Evidence**. New York; MacGraw-Hill Book Company, 1980.

MORTIMORE, P.; SAMMONS, P.; STOLL, L.; LEWIS, D.; ECOB, R. **School Matters: The Junior Years**. Shepton Mallett: Open Books, 1988.

MOSTELLER, Frederick, MOYNIHAN, Daniel P. (Ed.). **On Equality of Educational Opportunity** (Sobre a igualdade de oportunidades educacionais). New York: Vintage Books, 1972.

MOURA, Nathalia Bertú. **Análise do desempenho municipal na educação em Minas Gerais: uma proposta metodológica a partir do indicador do tribunal de contas do estado de minas gerais**. 2020. 191 f. Dissertação (Mestrado) - Curso de Administração Pública, Fundação João Pinheiro, Escola de Governo Professor Paulo Neves de Carvalho, Belo Horizonte, 2020.

OLIVEIRA, Romualdo Portela de; ARAUJO, Gilda Cardoso de. **Qualidade do ensino: uma nova dimensão da luta pelo direito à educação**. Rev. Bras. Educ., Rio de Janeiro, n. 28, p. 5-23, Apr. 2005. Available from <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-24782005000100002&lng=en&nrm=iso>. access on 02 Feb. 2021.

PEREIRA, Danielle Ramos de Miranda. **FATORES ASSOCIADOS AO DESEMPENHO ESCOLAR NAS DISCIPLINAS DE MATEMÁTICA E DE PORTUGUÊS NO ENSINO FUNDAMENTAL**: uma perspectiva longitudinal. 2006. 118 f. Tese (Doutorado) - Curso de Demografia, Centro de Desenvolvimento e Planejamento Regional Faculdade de Ciências Econômicas – Ufmg 2006, Belo Horizonte, 2006.

REVELLE, W. (2020). **psych: Procedures for Psychological, Psychometric, and Personality Research**. Northwestern University, Evanston, Illinois, 2020. R package version 2.0.12, <https://CRAN.R-project.org/package=psych>.

RODRIGUES, Clarissa Guimarães; RIOS-NETO, Eduardo Luiz Gonçalves; PINTO, Cristine Campos de Xavier. **Diferenças intertemporais na média e distribuição do desempenho escolar no Brasil: o papel do nível socioeconômico, 1997 a 2005**. Rev. bras. estud. popul., São Paulo , v. 28, n. 1, p. 5-36, June 2011 . Available from <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-30982011000100002&lng=en&nrm=iso>. access on 02 Feb. 2021.

RUTTER, M.; MAUGHAN, B.; MORTIMORE, P.; OUSTON, J.; SMITH, A. **Conclusões, especulações e implicações**. In: BROOKE, N.; SOARES, J. F. (eds.). Pesquisa em eficácia escolar: origem e trajetórias. Belo Horizonte: Editora UFMG, 1979. p. 225-251.

SAMMONS, Pam. **As Características-chave das Escolas Eficazes (1999)** In: BROOKE, N., SOARES, J. F. Pesquisa em Eficácia Escolar: origem e trajetórias. Belo Horizonte: Editora UFMG, 2008

SAMMONS, P; HILLMAN, J; MORTIMORE, P. **Key Characteristics of Effective Schools: A Review of School Effectiveness Research**. London: Office for Standards in Education (OFSTED), 1995.

SCHEERENS, J. **School effectiveness research and the development of process indicators of school functioning**. School Effectiveness and School Improvement, v. 1, n. 1, p. 61-80, 1990.

SEE-MG, Secretaria de Estado de Educação de Minas Gerais. **Avaliação da Rede Pública de Educação Básica – PROEB**. Minas Gerais, 2008. Disponível em: <<http://www2.educacao.mg.gov.br/aspectos-legais-e-responsabilidades/page/297-proeb>> . Acesso: abr. 2020.

SEE-MG, Secretaria de Estado de Educação de Minas Gerais. **Padrões de desempenho**. Minas Gerais, 2008. Disponível em: <<http://simave.educacao.mg.gov.br/#!/programa>> . Acesso: abr. 2020.

SILVA, Nelson do Valle; HASENBALG, Carlos. **Tendências da desigualdade educacional no Brasil**. Dados, Rio de Janeiro , v. 43, n. 3, p. 423-445, 2000 . Available from <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0011-52582000000300001&lng=en&nrm=iso>. Acesso: 02 Feb. 2021.

SOARES, J. F. **O Efeito da Escola no Desempenho Cognitivo dos seus alunos.** Revista Eletrônica Iberoamericana sobre Calidad, Eficácia y Cambio em Educacion. Vol.02, nº2, 2004.

SOARES, J. F. **Qualidade e equidade na educação básica brasileira: a evidência do SAEB-2001.** Archivos Analíticos de Políticas Educativas, v.12, n. 38, 2004.
SOARES, J. F.; ALVES, M. T. Desigualdades raciais no Sistema Brasileiro de Educação Básica. Educação e Pesquisa, São Paulo, v. 29, p. 147-165, 2003.

SOARES, T. M. **"Utilização da teoria da resposta ao item na produção de indicadores sócio-econômicos"**. Pesquisa Operacional, Rio de Janeiro, v. 25, n. 1, 2005 .

SOARES, T. M. **Influência do professor e do ambiente em sala de aula sobre a proficiência alcançada pelos alunos avaliados no Simave 2002.** Estudos em Avaliação Educacional, São Paulo, v. 28, p. 103-123, 2003.

SOARES, T. M. **Modelo de 3 níveis hierárquicos para a proficiência dos alunos de 4a série avaliados no teste de língua portuguesa do SIMAVE/PROEB-2002.** Revista Brasileira de Educação. Revista Brasileira de Educação, São Paulo, v. 29, p. 73-87, 2005.

SOARES, T. M., MENDONÇA, M. C. M., (2003). **Construção de um modelo de regressão hierárquico para os dados do SIMAVE-2000.** Pesquisa Operacional, Rio de Janeiro, nº 3, v. 23, p. 421-442

TARTUCE, Gisela Lobo B. P. et al . **Desafios do ensino médio no Brasil: iniciativas das secretarias de educação.** Cad. Pesqui., São Paulo , v. 48, n. 168, p. 478-504, June 2018 .
Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-15742018000200478&lng=en&nrm=iso>. access on 02 Feb. 2021.

VAN DEN EEDEN, P.; HOX, J.; HAUER, J. **Theory and model in multilevel research: Convergence or divergence?** Amsterdam: SISWO, 1990.

WEISS, D.J. & YOES, M. E. **Item Response Theory.** In: Hambleton, R. K. & Zaal, J. N. Advances in Educational and Psychological Testing: Theory and Applications. Boston: Kluwer Academic Publishers, 1991.

WILLMS, J. D. **Monitoring school performance.** Washington, D.C.: The Falmer Press, 1992.

WILLMS, J. D.; RAUDENBUSH, S.W. **A Longitudinal Hierarchical Linear Model for Estimating School Effects and their Stability.** Journal of Educational Measurement, 26(3), p 209-232, 1989.

APÊNDICE A – RESULTADOS DA ANÁLISE FATORIAL

Tabela 12 – Resultados gerais das análises fatoriais

	Teste de esfericidade de Bartlett			KMO	Proporção da variância explicada pelo fator
	Estatística de teste qui-quadrado	Valor p	Graus de liberdade		
Gestão da Matéria	2989,813	0	10	0,81	53,0%
Clima Escolar	8300,02	0	28	0,9	54,0%
Efetividade da implementação do PPP	44382,71	0	378	0,9	32,2%
Relação com a comunidade	649,1668	5,077	10	0,58	31,6%
Gestão do Diretor	21979,96	0	91	0,95	54,9%

Fonte: Dados: SEE/MG, 2019 e INEP, 2019.
Elaboração própria, 2020.

Tabela 13 – Resultados da análise fatorial por item – Gestão da Matéria

Item	Carga Fatorial	KMO
RQ_041	0,633	0,86
RQ_042	0,792	0,8
RQ_043	0,643	0,84
RQ_044	0,826	0,77
RQ_045	0,726	0,83

Fontes: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Tabela 14 – Resultados da análise fatorial por item – Clima Escolar

Item	Carga Fatorial	KMO
RQ_033	0,634	0,92
RQ_034	0,713	0,9
RQ_035	0,83	0,9
RQ_036	0,746	0,91
RQ_037	0,669	0,93
RQ_038	0,832	0,91
RQ_039	0,849	0,89
RQ_040	-0,555	0,91

Fontes: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Tabela 15 – Resultados da análise fatorial por item – Efetividade da implementação do PPP

Item	Carga Fatorial	KMO
RQ_030	-0,448	0,73
RQ_031	-0,462	0,74
RQ_032	-0,561	0,88
RQ_033	-0,55	0,88
RQ_034	-0,557	0,93
RQ_035	-0,56	0,93
RQ_036	-0,544	0,95
RQ_037	-0,536	0,78
RQ_038	-0,526	0,79
RQ_039	0,407	0,87
RQ_040	0,489	0,86
RQ_041	0,563	0,91
RQ_042	0,521	0,94
RQ_043	0,587	0,93
RQ_044	0,68	0,91
RQ_045	0,695	0,89
RQ_046	0,676	0,93
RQ_047	0,64	0,97
RQ_048	0,698	0,95
RQ_049	0,642	0,93
RQ_050	0,591	0,95
RQ_051	0,604	0,93
RQ_052	0,601	0,94
RQ_053	-0,529	0,91
RQ_054	-0,492	0,91
RQ_055	-0,528	0,93
RQ_056	-0,529	0,91
RQ_057	-0,54	0,9

Fontes: Dados: SEE/MG, 2019.
Elaboração própria, 2020.

Tabela 16 – Resultados da análise fatorial por item – Gestão do Diretor

Item	Carga	
	Fatorial	KMO
RQ_096	0,642	0,97
RQ_097	0,711	0,94
RQ_098	0,762	0,89
RQ_099	0,779	0,93
RQ_100	0,677	0,96
RQ_101	0,753	0,96
RQ_102	0,799	0,95
RQ_106	0,793	0,94
RQ_107	0,754	0,95
RQ_108	0,691	0,97
RQ_109	0,691	0,96
RQ_110	0,775	0,96
RQ_111	0,694	0,97
RQ_112	0,83	0,93

Fontes: Dados: SEE/MG, 2019.

Elaboração própria, 2020.

Tabela 17 – Resultados da análise fatorial por item – Gestão do Diretor

Item	Carga	
	Fatorial	KMO
IN_REDES_SOCIAIS	0,544	0,63
IN_ESPACO_ATIVIDADE	0,677	0,57
IN_ESPACO_EQUIPAMENTO	0,628	0,6
IN_ORGAO_CONSELHO_ESCOLAR	0,477	0,54
IN_ORGAO_NENHUM	-0,453	0,54

Fontes: Dados: INEP, 2019.

Elaboração própria, 2020.

APÊNDICE B – ANÁLISE DOS DADOS AUSENTES

Tabela 18 – Análise descritiva da amostra de dados utilizada

Variável	Média	Desvio Padrão	Mínimo	Máximo
VL_PROFICIENCIA_LP	267,32	51,41	127,78	393,98
VL_PROFICIENCIA_MT	270,62	51,50	158,66	453,35
SEXO	0,445	0,50	0	1
RACA	0,687	0,46	0	1
DEFAS	0,117	0,32	0	1
ISE	0,058	0,88	-3,02	2,84
ISE_ESCOLA	0,051	0,49	-2,04	1,46
ICG	4,789	0,95	1,00	6,00
IRD	2,788	0,41	1,36	4,21
GESTAO	0,028	1,03	-0,81	6,68
PPP	-0,059	1,02	-6,83	1,41
MATERIA	-0,133	0,85	-6,40	2,43
CLIMA_ESCOLAR	-0,013	0,84	-3,89	2,94
COMUNIDADE	0,011	1,00	-3,09	1,40
COMPROMETIMENTO	0,876	0,07	0,50	1
PROGRAMA	0,840	0,08	0,25	1

Fonte: Dados: SEE/MG, 2019 e INEP, 2019.

Elaboração própria, 2020.

Tabela 19 – Análise descritiva da base de dados completa

Variável	Média	Desvio Padrão	Mínimo	Máximo
VL_PROFICIENCIA_LP	265,29	51,92	127,75	393,98
VL_PROFICIENCIA_MT	268,73	51,49	158,65	454,74
SEXO	0,443	0,50	0	1
RACA	0,693	0,46	0	1
DEFAS	0,118	0,32	0	1
ISE	0,059	0,88	-3,02	2,84
ISE_ESCOLA	0,066	0,49	-2,04	1,46
ICG	4,802	0,96	1,00	6,00
IRD	2,765	0,41	1,36	4,21
GESTAO	0,059	1,04	-0,81	6,68
PPP	-0,093	1,03	-6,83	1,41
MATERIA	-0,213	0,88	-7,13	2,43
CLIMA_ESCOLAR	-0,107	0,87	-4,00	2,94
COMUNIDADE	-0,029	1,00	-3,09	1,40
COMPROMETIMENTO	0,872	0,07	0,50	1
PROGRAMA	0,838	0,08	0,25	1

Fonte: Dados: SEE/MG, 2019 e INEP, 2019.

Elaboração própria, 2020.

Tabela 20 – Diferença das médias do universo e da amostra

Variável	Diferença entre as médias
VL_PROFICIENCIA_LP	2,037
VL_PROFICIENCIA_MT	1,896
SEXO	0,002
RACA	-0,005
DEFAS	-0,002
ISE	-0,001
ISE_ESCOLA	-0,016
ICG	-0,014
IRD	0,023
GESTAO	-0,031
PPP	0,034
MATERIA	0,080
CLIMA_ESCOLAR	0,095
COMUNIDADE	0,040
COMPROMETIMENTO	0,004
PROGRAMA	0,002

Fonte: Dados: SEE/MG, 2019 e INEP, 2019.
Elaboração própria, 2020.

Tabela 21 – Quantidade de dados ausentes por variável

Variável	Número de observações	Proporção de dados completos
NU_SEQUENCIAL	177208	100%
CD_ESCOLA	177208	100%
ISE	144788	81,7%
VL_PROFICIENCIA_MT	150475	84,9%
SEXO	144016	81,3%
RACA	136970	77,3%
DEFAS	141511	79,9%
VL_PROFICIENCIA_LP	150493	84,9%
GESTAO	159867	90,2%
PPP	159880	90,2%
MATERIA	176686	99,7%
CLIMA_ESCOLAR	176686	99,7%
COMUNIDADE	176722	99,7%
ISE_ESCOLA	176686	99,7%
COMPROMETIMENTO	176686	99,7%
PROGRAMA	151535	85,5%
ICG	177208	100,0%
IRD	176481	99,6%

Fonte: Dados: SEE/MG, 2019 e INEP, 2019.
Elaboração própria, 2020.

Tabela 22 – Quantidade e proporção dos dados da amostra e do universo

	Dados disponíveis	Dados utilizados	Proporção dos dados utilizados
Alunos	177208	114334	64,5%
Escolas	2363	2027	85,8%

Fonte: Dados: SEE/MG, 2019 e INEP, 2019.

Elaboração própria, 2020.

APÊNDICE C – SINTAXES

Sintaxes para replicação utilizadas no software livre R para toda a análise realizada

```
rm(list = ls(all = TRUE))
```

```
#carregar pacotes
```

```
library(tidyverse)
```

```
library(mice)
```

```
library(ggplot2)
```

```
library(corrplot)
```

```
library(psych)
```

```
library(lme4)
```

```
library(lattice)
```

```
library(data.table)
```

```
library(gridExtra)
```

```
library(readxl)
```

```
#####
```

```
#carregar base de dados
```

```
estudante <- read_delim("SIMAVE 2019 MICRODADOS QUESTIONARIO  
ESTUDANTE.csv",
```

```
  ";", escape_double = FALSE,
```

```
  locale = locale(decimal_mark = ","), trim_ws = TRUE)
```

```
professor <- read_delim("SIMAVE 2019 MICRODADOS QUESTIONARIO  
PROFESSOR.csv",
```

```
  ";", escape_double = FALSE, trim_ws = TRUE)
```

```
censo <- read_delim("ESCOLAS.CSV",
```

```
  "|", escape_double = FALSE, trim_ws = TRUE)
```

```
ICG <- read_excel("ICG_ESCOLAS_2019.xlsx")
```

```
IRD <- read_excel("IRD_ESCOLAS_2019.xlsx")
```

```
#####
```

```
#preparar base de dados
```

```
censo[censo[,] == 9] = NA
```

```
censo <- censo %>%
```

```
  mutate(CD_ESCOLA = censo$CO_ENTIDADE)
```

```
colnames(ICG) <- c("CD_ESCOLA", "ICG")
```

```
colnames(IRD) <- c("CD_ESCOLA", "IRD")
```

```
estudante <- estudante %>% filter(ID_ETAPA_CENSO == "27")
```

```

professor <- professor %>%
  mutate(RQ_058 = case_when(RQ_058 == 4 ~ 0, RQ_058 < 4 ~ 3),
         RQ_020 = case_when(RQ_020 == 1 ~ 0.25, RQ_020 == 2 ~ 0.5,
                             RQ_020 == 3 ~ 0.6, RQ_020 == 4 ~ 0.7,
                             RQ_020 == 5 ~ 0.8, RQ_020 == 6 ~ 0.9,
                             RQ_020 == 7 ~ 1),
         RQ_022 = case_when(RQ_022 < 3 ~ 3, RQ_022 > 2 ~ 0))

estudante <- estudante %>%
  mutate(RQ_001 = case_when(RQ_001 == 1 ~ 1, RQ_001 == 2 ~ 0),
         RQ_002 = case_when(RQ_002 == 2 | RQ_002 == 3 ~ 1, RQ_002 == 1 | RQ_002
< 3 ~ 0),
         RQ_003 = case_when(RQ_003 > 12 ~ 1, RQ_003 < 13 ~ 0),
         RQ_046 = case_when(RQ_046 == 1 ~ 3, RQ_046 == 2 ~ 3,
                             RQ_046 == 3 ~ 1, RQ_046 == 4 ~ 1))

escolas <- professor %>%
  select(CD_ESCOLA, RQ_016, RQ_020)

estudante_agr <- estudante

estudante_agr <- estudante_agr %>%
  select(-NM_TURMA)

estudante_agr[estudante_agr[,] == 1] = 0
estudante_agr[estudante_agr[,] == 2] = 0
estudante_agr[estudante_agr[,] == 3] = 1
estudante_agr[estudante_agr[,] == 4] = 1

professor[professor[,] == 1] = 0
professor[professor[,] == 2] = 0
professor[professor[,] == 3] = 1
professor[professor[,] == 4] = 1

perguntas_p <- colnames(professor[,33:144])

professor <- professor %>% select(CD_ESCOLA, all_of(perguntas_p)) %>%
  pivot_longer(cols = all_of(perguntas_p), names_to = "pergunta") %>%
  group_by(CD_ESCOLA, pergunta) %>%
  summarise(prop = sum(1 - value, na.rm = TRUE)/length(na.omit(value)))

professor <- professor %>% select(CD_ESCOLA, pergunta, prop) %>%
  pivot_wider(names_from = pergunta, values_from = prop)

```

```

perguntas_e <- colnames(estudante_agr[,33:82])

estudante_agr <- estudante_agr %>%
  pivot_longer(cols = all_of(perguntas_e), names_to = "pergunta") %>%
  group_by(CD_ESCOLA, pergunta) %>%
  summarise(prop = sum(value, na.rm = TRUE)/length(na.omit(value)))

estudante_agr <- estudante_agr %>%
  select(CD_ESCOLA, pergunta, prop) %>%
  pivot_wider(names_from = pergunta, values_from = prop)

estudante_professor <- left_join(estudante_agr, professor, by = "CD_ESCOLA") #x
aluno, y professor

materia <- estudante_professor %>% select(paste('RQ_0', c(41:45), '.x', sep = ''))
clima_escolar <- estudante_professor %>% select(paste('RQ_0', c(33:40), '.x', sep =
''))
PPP <- estudante_professor %>% select(c(paste('RQ_0', 30:50, '.y', sep = ''),
paste('RQ_0', 51:58, sep = '')))
gestao <- estudante_professor %>% select(c(paste('RQ_0', c(96:99), sep = ''),
paste('RQ_', c(100:102, 106:112), sep = '')))

#####
# calculando as variáveis

cd_escola <- estudante_professor[,1]

comprometimento <- estudante_agr %>% select(RQ_046)

escolas <- escolas %>%
  group_by(CD_ESCOLA) %>%
  summarise(programa = mean(RQ_020))
escolas <- left_join(cd_escola, escolas, by = "CD_ESCOLA")
escolas <- escolas %>%
  left_join(ICG, by = "CD_ESCOLA") %>%
  left_join(IRD, by = "CD_ESCOLA") %>%
  left_join(censo, by = "CD_ESCOLA")

ISE_ESCOLA <- estudante %>%
  group_by(CD_ESCOLA) %>%
  summarise(ISE_ESCOLA = mean(VL_ISE, na.rm = T))

comunidade <- left_join(cd_escola, censo, by = "CD_ESCOLA") %>%
  select(IN_REDES_SOCIAIS,

```

```
    IN_ESPACO_ATIVIDADE,
    IN_ESPACO_EQUIPAMENTO,
    IN_ORGAO_CONSELHO_ESCOLAR,
    IN_ORGAO_NENHUM)
#####
# análise fatorial
###clima_escolar

matcor_clima_escolar <- cor(clima_escolar[,2:9], use = "complete.obs")
corrplot(matcor_clima_escolar, method = "circle")

(bartlett_clima_escolar <- cortest.bartlett(clima_escolar[,2:9]))
(kmo_clima_escolar <- KMO(clima_escolar[,2:9]))

CP_clima_escolar <- princomp(na.omit(clima_escolar[,2:9]), cor = TRUE)
summary(CP_clima_escolar)

screplot(CP_clima_escolar)
plot(CP_clima_escolar, type = "lines")

PCA_clima_escolar <- principal(clima_escolar[,2:9], nfactors = 1,
                               n.obs = 8,
                               scores = TRUE)

PCA_clima_escolar

PCA_clima_escolar$values

PCA_clima_escolar$loadings

biplot(PCA_clima_escolar)

###PPP

matcor_PPP <- cor(PPP[,2:30], use = "complete.obs")
corrplot(matcor_PPP, method = "circle")

(bartlett_PPP <- cortest.bartlett(PPP[,2:29]))
(kmo_PPP <- KMO(PPP[,2:29]))

CP_PPP <- princomp(na.omit(PPP[,2:30]), cor = TRUE)
summary(CP_PPP)

screplot(CP_PPP)
```

```
plot(CP_PPP, type = "lines")

PCA_PPP <- principal(PPP[,2:29], nfactors = 1,
                    n.obs = 28,
                    scores = TRUE)

PCA_PPP

PCA_PPP$values

PCA_PPP$loadings

biplot(PCA_PPP)

###matéria

matcor_materia <- cor(materia[,2:6], use = "complete.obs")
corrplot(matcor_materia, method = "circle")

(bartlett_materia <- cortest.bartlett(materia[,2:6]))
(kmo_materia <- KMO(materia[,2:6]))

CP_materia <- princomp(na.omit(materia[,2:6]), cor = TRUE)
summary(CP_materia)

screplot(CP_materia)
plot(CP_materia, type = "lines")

PCA_materia <- principal(materia[,2:6], nfactors = 1,
                        n.obs = 5,
                        scores = TRUE)

PCA_materia

PCA_materia$values

PCA_materia$loadings

biplot(PCA_materia)

###gestão

matcor_gestao <- cor(gestao[,2:15], use = "complete.obs")
corrplot(matcor_gestao, method = "circle")
```

```
(bartlett_gestao <- cortest.bartlett(gestao[,2:15]))
(kmo_gestao <- KMO(gestao[,2:15]))

CP_gestao <- princomp(na.omit(gestao[,2:15]), cor = TRUE)
summary(CP_gestao)

screplot(CP_gestao)
plot(CP_gestao, type = "lines")

PCA_gestao <- principal(gestao[,2:15], nfactors = 1,
                       n.obs = 14,
                       scores = TRUE)

PCA_gestao

PCA_gestao$values

PCA_gestao$loadings

biplot(PCA_gestao)

### comunidade

(matcor_comunidade <- cor(comunidade[,2:6], use = "complete.obs"))
corrplot(matcor_comunidade, method = "circle")

(bartlett_comunidade <- cortest.bartlett(comunidade[,2:6]))
(kmo_comunidade <- KMO(comunidade[,2:6]))

CP_comunidade <- princomp(na.omit(comunidade), cor = TRUE)
summary(CP_comunidade)

screplot(CP_comunidade)
plot(CP_comunidade, type = "lines")

PCA_comunidade <- principal(comunidade[,2:6], nfactors = 1,
                           n.obs = 5,
                           scores = TRUE)

PCA_comunidade

PCA_comunidade$values
```

```
PCA_comunidade$loadings
```

```
biplot(PCA_comunidade)
```

```
#####
```

```
# juntando as variáveis
```

```
estudante_professor["GESTAO"] <- as.vector(PCA_gestao$scores)
estudante_professor["PPP"] <- as.vector(PCA_PPP$scores)
estudante_professor["MATERIA"] <- as.vector(PCA_materia$scores)
estudante_professor["CLIMA_ESCOLAR"] <- as.vector(PCA_clima_escolar$scores)
estudante_professor["COMUNIDADE"] <- as.vector(PCA_comunidade$scores)
estudante_professor["ISE_ESCOLA"] <- ISE_ESCOLA$ISE_ESCOLA
estudante_professor["COMPROMETIMENTO"] <- comprometimento$RQ_046
estudante_professor["PROGRAMA"] <- escolas$programa
estudante_professor["ICG"] <- escolas$ICG
estudante_professor["ICG."] <- escolas$ICG
estudante_professor["IRD"] <- escolas$IRD
```

```
estudante_professor <- estudante_professor %>%
  select(GESTAO, PPP, MATERIA, CLIMA_ESCOLAR, COMUNIDADE,
         ISE_ESCOLA,
         COMPROMETIMENTO, PROGRAMA, DEP_ADMINISTRATIVA,
         LOCALIZACAO, ICG., ICG, IRD)
```

```
estudante <- estudante %>%
  mutate(SEXO = RQ_001, RACA = RQ_002, DEFAS = RQ_003, ISE = VL_ISE) %>%
  select(c(NU_SEQUENCIAL, CD_ESCOLA, ISE, VL_PROFICIENCIA_MT,
         SEXO, RACA, DEFAS, VL_PROFICIENCIA_LP))
```

```
estudante_mt <- estudante %>%
  select(c(NU_SEQUENCIAL, CD_ESCOLA, ISE, VL_PROFICIENCIA_MT,
         SEXO, RACA, DEFAS))
```

```
estudante_lp <- estudante %>%
  select(c(NU_SEQUENCIAL, CD_ESCOLA, ISE,
         SEXO, RACA, DEFAS, VL_PROFICIENCIA_LP))
```

```
df1 <- left_join(estudante, estudante_professor,
                by = "CD_ESCOLA") %>%
  mutate(CD_ESCOLA = as.integer(CD_ESCOLA))
```

```
df1_mt <- left_join(estudante_mt, estudante_professor,
                   by = "CD_ESCOLA") %>%
```

```

mutate(CD_ESCOLA = as.integer(CD_ESCOLA))
df1_mt <- na.omit(df1_mt)

df1_lp <- left_join(estudante_lp, estudante_professor,
                  by = "CD_ESCOLA") %>%
  mutate(CD_ESCOLA = as.integer(CD_ESCOLA))
df1_lp <- na.omit(df1_lp)

alunos <- full_join(df1_mt, df1_lp, by = "NU_SEQUENCIAL") %>%
  select(NU_SEQUENCIAL)

df <- left_join(alunos, df1)

#####
# análise dos dados ausentes

md.pattern(df1)

na_df <- as.data.frame(abs(is.na(df1)))
na_df[na_df[,] == 1] = 2
na_df[na_df[,] == 0] = 1
na_df[na_df[,] == 2] = 0

qnt_na <- data.frame(`Variável` = colnames(na_df),
                   `Número de observações` = map_dbl(na_df, sum),
                   `Proporção de dados completos` = map_dbl(na_df, mean))
fwrite(qnt_na, "qnt_na.csv")

(matcor_na_df <- cor(df1[,3:21]))

(cor_na <- cor(df1[3:21], na_df[3:21], use="pairwise.complete.obs"))

var_resumo_na <- data.frame(`Variável` = colnames(df1[,3:20]),
                          `Média` = round(map_dbl(df1[,3:20], mean, na.rm = T), 4),
                          `Desvio Padrão` = round(map_dbl(df1[,3:20], sd, na.rm = T), 2),
                          `Mínimo` = round(map_dbl(df1[,3:20], min, na.rm = T), 2),
                          `Máximo` = round(map_dbl(df1[,3:20], max, na.rm = T), 2))
fwrite(var_resumo_na, "var_resumo_na.csv")

var_resumo1 <- data.frame(`Variável` = colnames(df[,3:20]),
                        `Média` = round(map_dbl(df[,3:20], mean, na.rm = T), 4),
                        `Desvio Padrão` = round(map_dbl(df[,3:20], sd, na.rm = T), 2),
                        `Mínimo` = round(map_dbl(df[,3:20], min, na.rm = T), 2),
                        `Máximo` = round(map_dbl(df[,3:20], max, na.rm = T), 2))

```

```

fwrite(var_resumo1, "var_resumo1.csv")

# centralizando os valores na média
df$MATERIA <- scale(df$MATERIA)
df$CLIMA_ESCOLAR <- scale(df$CLIMA_ESCOLAR)
df$ISE <- scale(df$ISE)
df$ISE_ESCOLA <- scale(df$ISE_ESCOLA)
df$PPP <- scale(df$PPP)
df$GESTAO <- scale(df$GESTAO)
df$COMUNIDADE <- scale(df$COMUNIDADE)
df$ICG <- scale(df$ICG)
df$IRD <- scale(df$IRD)

#####
# análise descritiva - transformando as variáveis em categóricas

df2 <- df %>%
  mutate(SEXO. = ifelse(SEXO == 0, "Feminino", "Masculino"),
         RACA. = ifelse(RACA == 0, "Branco, amarelo ou indígena", "Preto ou pardo"),
         DEFAS. = ifelse(DEFAS == 0, "Regular", "Defasado"),
         ISE. = ifelse(ISE <= quantile(df$ISE, na.rm = T)[2], "Baixo",
                      ifelse(ISE <= quantile(df$ISE, na.rm = T)[3] &
                             ISE > quantile(df$ISE, na.rm = T)[2], "Médio Baixo",
                             ifelse(ISE <= quantile(df$ISE, na.rm = T)[4] &
                                     ISE > quantile(df$ISE, na.rm = T)[3],
                                     "Médio Alto", "Alto"))),
         VL_PROFICIENCIA_MT. =
           ifelse(VL_PROFICIENCIA_MT <= 275, "Baixo",
                 ifelse(VL_PROFICIENCIA_MT <= 350 & VL_PROFICIENCIA_MT > 275,
                        "Intermediário", ifelse(VL_PROFICIENCIA_MT <= 375 &
                                                VL_PROFICIENCIA_MT > 350,
                                                "Recomendado", "Avançado"))),
         VL_PROFICIENCIA_LP. =
           ifelse(VL_PROFICIENCIA_LP <= 275, "Baixo",
                 ifelse(VL_PROFICIENCIA_LP <= 350 & VL_PROFICIENCIA_LP > 275,
                        "Intermediário", ifelse(VL_PROFICIENCIA_LP <= 375 &
                                                VL_PROFICIENCIA_LP > 350,
                                                "Recomendado", "Avançado"))),
         COMPROMETIMENTO. = ifelse(COMPROMETIMENTO <= 0.8, "até 80%",
                                    ifelse(COMPROMETIMENTO <= 0.9 & COMPROMETIMENTO
                                             > 0.8, "de 80% a 90%", "100%")),
         PROGRAMA. = ifelse(PROGRAMA <= 0.7 & PROGRAMA > 0.6, "menos de
70%",
                            ifelse(PROGRAMA <= 0.8 & PROGRAMA > 0.7, "de 70% 80%",

```

```

        ifelse(PROGRAMA <= 0.9 & PROGRAMA > 0.8, "de 80% até
90%", "100%")),
    COMUNIDADE. = ifelse(COMUNIDADE <= -2, "até -2", ifelse(COMUNIDADE >
-2 & COMUNIDADE <= -1, "-2 a -1",
        ifelse(COMUNIDADE <= 0 & COMUNIDADE > -1, "-1 a 0",
            ifelse(COMUNIDADE <= 1 & COMUNIDADE > 0, "0 a 1", "maior
que 1")))))

```

```

df2 <- df2 %>%
  mutate(ISE. = factor(ISE., levels = c("Baixo", "Médio Baixo", "Médio Alto", "Alto")),
    VL_PROFICIENCIA_MT. = factor(VL_PROFICIENCIA_MT., levels = c("Baixo",
"Intermediário", "Recomendado", "Avançado")),
    VL_PROFICIENCIA_LP. = factor(VL_PROFICIENCIA_LP., levels = c("Baixo",
"Intermediário", "Recomendado", "Avançado")),
    PROGRAMA. = factor(PROGRAMA., levels = c("menos de 70%", "de 70% 80%",
"de 80% até 90%", "100%")),
    COMPROMETIMENTO. = factor(COMPROMETIMENTO., levels = c("até 80%",
"de 80% a 90%", "100%")),
    COMUNIDADE. = factor(COMUNIDADE., levels = c("até -2", "-2 a -1", "-1 a 0",
"0 a 1", "maior que 1")))

```

```
summary(df2)
```

```
# estatísticas descritivas
```

```

var_resumo <- data.frame(`Variável` = colnames(df[,3:20]),
  `Média` = round(map_dbl(na.omit(df[,3:20]), mean), 4),
  `Desvio Padrão` = round(map_dbl(na.omit(df[,3:20]), sd), 2),
  `Mínimo` = round(map_dbl(na.omit(df[,3:20]), min), 2),
  `Máximo` = round(map_dbl(na.omit(df[,3:20]), max), 2))

```

```
fwrite(var_resumo, "var_resumo.csv")
```

```
# tabela resumo por variável de controle
```

```

ise_resumo <- df2 %>%
  group_by(ISE.) %>%
  summarise(`Proficiência média em Matemática` = mean(VL_PROFICIENCIA_MT,
na.rm = T),
    `Proficiência média em Língua Portuguesa` = mean(VL_PROFICIENCIA_LP,
na.rm = T),
    `Número de alunos` = n(),

```

```

`Percentual de alunos` = n()/nrow(df2),
`Índice socioeconômico médio` = mean(ISE),
`Proporção de alunos defasados` = mean(DEFAS),
`Proporção de alunos do sexo masculino` = mean(SEXO),
`Proporção de alunos pretos ou pardos` = mean(RACA))

fwrite(ise_resumo, "ise_resumo.csv")

defas_resumo <- df2 %>%
  group_by(DEFAS.) %>%
  summarise(`Proficiência média em Matemática` = mean(VL_PROFICIENCIA_MT,
na.rm = T),
            `Proficiência média em Língua Portuguesa` = mean(VL_PROFICIENCIA_LP,
na.rm = T),
            `Número de alunos` = n(),
            `Percentual de alunos` = n()/nrow(df2),
            `Índice socioeconômico médio` = mean(ISE),
            `Proporção de alunos defasados` = mean(DEFAS),
            `Proporção de alunos do sexo masculino` = mean(SEXO),
            `Proporção de alunos pretos ou pardos` = mean(RACA))

fwrite(defas_resumo, "defas_resumo.csv")

sexo_resumo <- df2 %>%
  group_by(SEXO.) %>%
  summarise(`Proficiência média em Matemática` = mean(VL_PROFICIENCIA_MT,
na.rm = T),
            `Proficiência média em Língua Portuguesa` = mean(VL_PROFICIENCIA_LP,
na.rm = T),
            `Número de alunos` = n(),
            `Percentual de alunos` = n()/nrow(df2),
            `Índice socioeconômico médio` = mean(ISE),
            `Proporção de alunos defasados` = mean(DEFAS),
            `Proporção de alunos do sexo masculino` = mean(SEXO),
            `Proporção de alunos pretos ou pardos` = mean(RACA))

fwrite(sexo_resumo, "sexo_resumo.csv")

raca_resumo <- df2 %>%
  group_by(RACA.) %>%
  summarise(`Proficiência média em Matemática` = mean(VL_PROFICIENCIA_MT,
na.rm = T),
            `Proficiência média em Língua Portuguesa` = mean(VL_PROFICIENCIA_LP,
na.rm = T),

```

```

`Número de alunos` = n(),
`Percentual de alunos` = n()/nrow(df2),
`Índice socioeconômico médio` = mean(ISE),
`Proporção de alunos defasados` = mean(DEFAS),
`Proporção de alunos do sexo masculino` = mean(SEXO),
`Proporção de alunos pretos ou pardos` = mean(RACA))

fwrite(raca_resumo, "raca_resumo.csv")

#matriz de correlações

mat_cor <- cor(na.omit(df))

fwrite(mat_cor, "mat_cor.csv")

# gráficos
## gráficos de distribuição de frequência
#Distribuição do fator: gestão da matéria
ggplot(df,
  aes(MATERIA)) +
  geom_histogram(bins = 20) +
  labs(x = "Gestão da matéria", y = "Frequência") +
  theme_classic() +
  scale_x_continuous(limits = c(-4,3.5))

#Distribuição do fator: clima escolar
ggplot(df,
  aes(CLIMA_ESCOLAR)) +
  labs(x = "Clima escolar", y = "Frequência") +
  geom_histogram(bins = 20) +
  theme_classic() +
  scale_x_continuous(limits = c(-4,3.5))

#Distribuição do fator: efetividade da construção e implementação do PPP
ggplot(df,
  aes(PPP)) +
  labs(x = "PPP", y = "Frequência") +
  geom_histogram(bins = 20) +
  theme_classic()+
  scale_x_continuous(limits = c(-5,2.5))

#Distribuição do fator: gestão do diretor
ggplot(df,
  aes(GESTAO)) +

```

```
labs(x = "Gestão do diretor", y = "Frequência") +
geom_histogram(bins = 20) +
theme_classic()+
scale_x_continuous(limits = c(-1,3))
```

```
#Distribuição do fator: relação com a comunidade
```

```
ggplot(df2,
  aes(COMUNIDADE., y = (..count..)*100/sum(..count..))) +
labs(x = "Relação com a comunidade", y = "Frequência percentual") +
geom_bar() +
theme_classic()
```

```
freq_comunidade <- df2 %>%
  group_by(COMUNIDADE.) %>%
  summarise(freq = n()*100/nrow(df2))
```

```
#Distribuição do fator: ird
```

```
ggplot(df,
  aes(IRD)) +
labs(x = "Indicador de regularidade docente", y = "Frequência") +
geom_histogram(bins = 30) +
theme_classic()
```

```
#Distribuição do fator: icg
```

```
ggplot(df,
  aes(ICG., y = (..count..)*100/sum(..count..))) +
labs(x = "Indicador de complexidade da gestão", y = "Frequência Percentual") +
geom_bar() +
theme_classic()
```

```
freq_icg <- df %>%
  group_by(ICG.) %>%
  summarise(freq = n()*100/nrow(df))
```

```
#Distribuição do fator: comprometimento do professor
```

```
ggplot(df2,
  aes(COMPROMETIMENTO. , y = (..count..)*100/sum(..count..))) +
labs(x = "Comprometimento do professor", y = "Frequência percentual") +
geom_bar() +
theme_classic() +
theme(axis.text.x = element_text(angle = 30, hjust = 1))
```

```
freq_comprometimento <- df2 %>%
  group_by(COMPROMETIMENTO.) %>%
```

```

summarise(freq = n()*100/nrow(df2))

#Distribuição do fator: frequência do programa concluída pelo professor
ggplot(df2,
  aes(PROGRAMA. , y = (..count..)*100/sum(..count..)) +
  labs(x = "Percentual do programa concluído", y = "Frequência percentual") +
  geom_bar() +
  theme_classic() +
  theme(axis.text.x = element_text(angle = 30, hjust = 1))

freq_programa <- df2 %>%
  group_by(PROGRAMA.) %>%
  summarise(freq = n()*100/nrow(df2))

#Distribuição do fator: índice socioeconômico da escola
ggplot(df,
  aes(ISE_ESCOLA)) +
  labs(x = "índice socioeconômico da escola", y = "Frequência") +
  geom_histogram(bins = 20) +
  theme_classic()

#Distribuição do fator: índice socioeconômico
ggplot(df,
  aes(ISE)) +
  labs(x = "índice socioeconômico", y = "Frequência") +
  geom_histogram(bins = 20) +
  theme_classic() +
  scale_x_continuous(limits = c(-3.5,3))

## gráficos de proficiência e ISE

#Distribuição percentual da proficiência em Matemática 3° ano do Ensino Médio por
nível
(x1 <- ggplot(data = df2,
  aes(x = VL_PROFICIENCIA_MT., y = (..count..)*100/sum(..count..)) +
  labs(x = "Matemática", y = "Frequência percentual") +
  geom_bar() +
  scale_y_continuous(breaks = c(10, 20, 30, 40, 50, 60)) +
  theme_classic())

freq_ma <- df2 %>%
  group_by(VL_PROFICIENCIA_MT.) %>%
  summarise(freq = n()*100/nrow(df2))

```

#Distribuição percentual da proficiência em Língua Portuguesa 3° ano do Ensino Médio por nível

```
(x2 <- ggplot(data = df2,
  aes(x = VL_PROFICIENCIA_LP., y = (..count..)*100/sum(..count..))) +
  labs(x = "Língua Portuguesa", y = "Frequência percentual") +
  geom_bar() +
  scale_y_continuous(breaks = c(10, 20, 30, 40, 50, 60)) +
  theme_classic())
```

```
freq_lp <- df2 %>%
  group_by(VL_PROFICIENCIA_LP.) %>%
  summarise(freq = n()*100/nrow(df2))
```

#Proficiência em Matemática 3° ano do Ensino Médio por nível socioeconômico

```
ggplot(df2,
  aes(ISE., y = VL_PROFICIENCIA_MT)) +
  stat_boxplot(geom = 'errorbar') +
  geom_boxplot(outlier.shape = NA, show.legend = FALSE) +
  labs(x = "Etapa de Ensino", y = "Matemática") +
  scale_y_continuous(breaks = c(100, 150, 200, 250, 300, 350, 400, 450)) +
  theme_minimal())
```

#Proficiência em Língua Portuguesa 3° ano do Ensino Médio por nível socioeconômico

```
ggplot(df2,
  aes(ISE., y = VL_PROFICIENCIA_LP)) +
  stat_boxplot(geom = 'errorbar') +
  geom_boxplot(outlier.shape = NA, show.legend = FALSE) +
  labs(x = "Etapa de Ensino", y = "Língua Portuguesa") +
  scale_y_continuous(breaks = c(100, 150, 200, 250, 300, 350, 400, 450)) +
  theme_minimal())
```

#Distribuição percentual do nível socioeconômico por nível de proficiência em Matemática 3° ano do Ensino Médio

```
df2_1 <- df2 %>%
  filter(VL_PROFICIENCIA_MT. == "Baixo")
df2_2 <- df2 %>%
  filter(VL_PROFICIENCIA_MT. == "Intermediário")
df2_3 <- df2 %>%
  filter(VL_PROFICIENCIA_MT. == "Recomendado")
df2_4 <- df2 %>%
```

```

filter(VL_PROFICIENCIA_MT. == "Avançado")

g1 <- ggplot(data = df2_1,
             aes(x = ISE., y = (..count..)*100/sum(..count..))) +
  labs(title = "Proficiência baixa", x = "Índice Socioeconômico", y = "Percentual") +
  geom_bar() +
  scale_y_continuous(breaks = c(10, 20, 30, 40, 50, 60)) +
  theme_classic() +
  theme(axis.text.x = element_text(angle = 30, hjust = 1))
g2 <- ggplot(data = df2_2,
             aes(x = ISE., y = (..count..)*100/sum(..count..))) +
  labs(title = "Proficiência intermediária", x = "Índice Socioeconômico", y =
"Percentual") +
  geom_bar() +
  scale_y_continuous(breaks = c(10, 20, 30, 40, 50, 60)) +
  theme_classic()+
  theme(axis.text.x = element_text(angle = 30, hjust = 1))
g3 <- ggplot(data = df2_3,
             aes(x = ISE., y = (..count..)*100/sum(..count..))) +
  labs(title = "Proficiência recomendada", x = "Índice Socioeconômico", y =
"Percentual") +
  geom_bar() +
  scale_y_continuous(breaks = c(10, 20, 30, 40, 50, 60)) +
  theme_classic()+
  theme(axis.text.x = element_text(angle = 30, hjust = 1))
g4 <- ggplot(data = df2_4,
             aes(x = ISE., y = (..count..)*100/sum(..count..))) +
  labs(title = "Proficiência Avançada", x = "Índice Socioeconômico", y = "Percentual") +
  geom_bar() +
  scale_y_continuous(breaks = c(10, 20, 30, 40, 50, 60)) +
  theme_classic()+
  theme(axis.text.x = element_text(angle = 30, hjust = 1))

grid.arrange(g1, g2, g3, g4,
             ncol=2, nrow=2)

#Distribuição percentual do nível socioecoômico por nível de Língua Portuguesa 3º
ano do Ensino Médio
df2_1 <- df2 %>%
  filter(VL_PROFICIENCIA_LP. == "Baixo")
df2_2 <- df2 %>%
  filter(VL_PROFICIENCIA_LP. == "Intermediário")
df2_3 <- df2 %>%
  filter(VL_PROFICIENCIA_LP. == "Recomendado")

```

```

df2_4 <- df2 %>%
  filter(VL_PROFICIENCIA_LP. == "Avançado")

g1 <- ggplot(data = df2_1,
  aes(x = ISE., y = (..count..)*100/sum(..count..))) +
  labs(title = "Proficiência baixa", x = "Índice Socioeconômico", y = "Percentual") +
  geom_bar() +
  scale_y_continuous(breaks = c(10, 20, 30, 40, 50, 60)) +
  theme_classic() +
  theme(axis.text.x = element_text(angle = 30, hjust = 1))
g2 <- ggplot(data = df2_2,
  aes(x = ISE., y = (..count..)*100/sum(..count..))) +
  labs(title = "Proficiência intermediária", x = "Índice Socioeconômico", y =
"Percentual") +
  geom_bar() +
  scale_y_continuous(breaks = c(10, 20, 30, 40, 50, 60)) +
  theme_classic()+
  theme(axis.text.x = element_text(angle = 30, hjust = 1))
g3 <- ggplot(data = df2_3,
  aes(x = ISE., y = (..count..)*100/sum(..count..))) +
  labs(title = "Proficiência recomendada", x = "Índice Socioeconômico", y =
"Percentual") +
  geom_bar() +
  scale_y_continuous(breaks = c(10, 20, 30, 40, 50, 60)) +
  theme_classic()+
  theme(axis.text.x = element_text(angle = 30, hjust = 1))
g4 <- ggplot(data = df2_4,
  aes(x = ISE., y = (..count..)*100/sum(..count..))) +
  labs(title = "Proficiência Avançada", x = "Índice Socioeconômico", y = "Percentual") +
  geom_bar() +
  scale_y_continuous(breaks = c(10, 20, 30, 40, 50, 60)) +
  theme_classic()+
  theme(axis.text.x = element_text(angle = 30, hjust = 1))

grid.arrange(g1, g2, g3, g4,
  ncol=2, nrow=2)

## gráficos de dispersão ISE por variável de controle

#Relação entre a proficiência em Matemática 3º ano do Ensino Médio e o índice
socioeconômico dos alunos
(z1 <- ggplot(df2,
  aes(x = ISE, y = VL_PROFICIENCIA_MT)) +
  geom_point(shape = '.') +

```

```
geom_smooth(se = FALSE, method = "lm") +
labs(x = "Índice socioeconômico", y = "Matemática") +
theme_classic())
```

```
(mod1 <- lm(VL_PROFICIENCIA_MT ~ ISE, df2))
```

#Relação entre a proficiência em Matemática 3° ano do Ensino Médio e o índice socioeconômico dos alunos por sexo

```
(z2 <- ggplot(df2,
  aes(x = ISE, y = VL_PROFICIENCIA_MT, color = SEXO.)) +
  geom_point(shape = '.', colour = "black") +
  geom_smooth(se = FALSE, method = "lm") +
  labs(x = "Índice socioeconômico", y = "Matemática") +
  theme_classic() +
  theme(legend.position = "bottom"))
```

```
df2a <- df2 %>%
```

```
  filter(SEXO == 1)
```

```
df2b <- df2 %>%
```

```
  filter(SEXO == 0)
```

```
(mod2a <- lm(VL_PROFICIENCIA_MT ~ ISE, df2a))
```

```
(mod2b <- lm(VL_PROFICIENCIA_MT ~ ISE, df2b))
```

#Relação entre a proficiência em Matemática 3° ano do Ensino Médio e o índice socioeconômico dos alunos por raça

```
(z3 <- ggplot(df2,
  aes(x = ISE, y = VL_PROFICIENCIA_MT, color = RACA.)) +
  geom_point(shape = '.', colour = "black") +
  geom_smooth(se = FALSE, method = "lm") +
  labs(x = "Índice socioeconômico", y = "Matemática") +
  theme_classic() +
  theme(legend.position = "bottom"))
```

```
df3a <- df2 %>%
```

```
  filter(RACA == 1)
```

```
df3b <- df2 %>%
```

```
  filter(RACA == 0)
```

```
(mod3a <- lm(VL_PROFICIENCIA_MT ~ ISE, df3a))
```

```
(mod3b <- lm(VL_PROFICIENCIA_MT ~ ISE, df3b))
```

#Relação entre a proficiência em Matemática 3° ano do Ensino Médio e o índice socioeconômico dos alunos por trajetória de ensino

```
(z4 <- ggplot(df2,
  aes(x = ISE, y = VL_PROFICIENCIA_MT, color = DEFAS.)) +
  geom_point(shape = '.', colour = "black") +
  geom_smooth(se = FALSE, method = "lm") +
  labs(x = "Índice socioeconômico", y = "Matemática") +
  theme_classic() +
  theme(legend.position = "bottom"))
```

```
df4a <- df2 %>%
  filter(RACA == 1)
df4b <- df2 %>%
  filter(RACA == 0)
```

```
(mod4a <- lm(VL_PROFICIENCIA_MT ~ ISE, df4a))
(mod4b <- lm(VL_PROFICIENCIA_MT ~ ISE, df4b))
```

#Relação entre a proficiência em Língua Portuguesa 3º ano do Ensino Médio e o índice socioeconômico dos alunos

```
(w1 <- ggplot(df2,
  aes(x = ISE, y = VL_PROFICIENCIA_LP)) +
  geom_point(shape = '.') +
  geom_smooth(se = FALSE, method = "lm") +
  labs(x = "Índice socioeconômico", y = "Língua Portuguesa") +
  theme_classic())
```

```
mod1. <- lm(VL_PROFICIENCIA_LP ~ ISE, df2)
summary(mod1)
```

#Relação entre a proficiência em Língua Portuguesa 3º ano do Ensino Médio e o índice socioeconômico dos alunos por sexo

```
(w2 <- ggplot(df2,
  aes(x = ISE, y = VL_PROFICIENCIA_LP, color = SEXO.)) +
  geom_point(shape = '.', colour = "black") +
  geom_smooth(se = FALSE, method = "lm") +
  labs(x = "Índice socioeconômico", y = "Língua Portuguesa") +
  theme_classic() +
  theme(legend.position = "bottom"))
```

```
df2a <- df2 %>%
  filter(SEXO == 1)
df2b <- df2 %>%
  filter(SEXO == 0)
```

```
(mod2a. <- lm(VL_PROFICIENCIA_LP ~ ISE, df2a))
```

```
(mod2b. <- lm(VL_PROFICIENCIA_LP ~ ISE, df2b))
```

#Relação entre a proficiência em Língua Portuguesa 3º ano do Ensino Médio e o índice socioeconômico dos alunos por raça

```
(w3 <- ggplot(df2,
  aes(x = ISE, y = VL_PROFICIENCIA_LP, color = RACA.)) +
  geom_point(shape = '.', colour = "black") +
  geom_smooth(se = FALSE, method = "lm") +
  labs(x = "Índice socioeconômico", y = "Língua Portuguesa") +
  theme_classic() +
  theme(legend.position = "bottom"))
```

```
df3a <- df2 %>%
```

```
  filter(RACA == 1)
```

```
df3b <- df2 %>%
```

```
  filter(RACA == 0)
```

```
(mod3a. <- lm(VL_PROFICIENCIA_LP ~ ISE, df3a))
```

```
(mod3b. <- lm(VL_PROFICIENCIA_LP ~ ISE, df3b))
```

#Relação entre a proficiência em Língua Portuguesa 3º ano do Ensino Médio e o índice socioeconômico dos alunos por trajetória de ensino

```
(w4 <- ggplot(df2,
  aes(x = ISE, y = VL_PROFICIENCIA_LP, color = DEFAS.)) +
  geom_point(shape = '.', colour = "black") +
  geom_smooth(se = FALSE, method = "lm") +
  labs(x = "Índice socioeconômico", y = "Língua Portuguesa") +
  theme_classic() +
  theme(legend.position = "bottom"))
```

```
df4a <- df2 %>%
```

```
  filter(DEFAS == 1)
```

```
df4b <- df2 %>%
```

```
  filter(DEFAS == 0)
```

```
(mod4a. <- lm(VL_PROFICIENCIA_LP ~ ISE, df4a))
```

```
(mod4b. <- lm(VL_PROFICIENCIA_LP ~ ISE, df4b))
```

```
#####
```

```
#modelos
```

```
#MT
```

```
##modelo nulo
```

```
fit <- lmer(VL_PROFICIENCIA_MT ~ (1 | CD_ESCOLA),
```

```

data = df, REML = FALSE)

summary(fit)
(VPC <- 325.4*100/(325.4 + 2337.1))

fita <- lm(VL_PROFICIENCIA_MT ~ 1, data = df)
(a <- anova(fit, fita))

## modelo de nível 1
fit1a <- lmer(VL_PROFICIENCIA_MT ~ ISE + SEXO + RACA + DEFAS +
             (1 | CD_ESCOLA), data = df,
             REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
(d1 <- drop1(fit1, test="Chisq"))

#testando efeitos aleatórios
#ISE
fit1 <- lmer(VL_PROFICIENCIA_MT ~ ISE + SEXO + RACA + DEFAS +
            (1 + ISE | CD_ESCOLA), data = df,
            REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
(i1 <- isSingular(fit1))
(a1 <- anova(fit1, fit1a))
(d1 <- drop1(fit1, test="Chisq"))

summary(fit1)

(VPC1 <- 239.53*100/(239.53 + 17.48 + 2199.16))

#análise dos efeitos aleatórios
# ISE
(ranef <- ranef(fit1, condVar = TRUE))

plot(ranef[[1]], xlab = "Variação da proficiência média em Matemática", ylab =
"Variação do efeito do ISE")
abline(h = 0, col = "red")
abline(v = 0, col = "red")

## modelo de nível 2
fit2 <- update(fit1, . ~ . + PPP + GESTAO + ISE_ESCOLA + COMPROMETIMENTO +
              CLIMA_ESCOLAR + MATERIA + PROGRAMA + ICG + IRD +
              COMUNIDADE)
(d2 <- drop1(fit2, test="Chisq"))
(i2 <- isSingular(fit2))
summary(fit2)

```

```
(a2 <- anova(fit2, fit1))
```

```
(VPC2 <- 142.00*100/(142.00 + 16.41 + 2199.59))
```

```
## permitindo interações
```

```
#ISE
```

```
fit3 <- lmer(VL_PROFICIENCIA_MT ~ SEXO*ISE + RACA*ISE + DEFAS*ISE +  
PPP*ISE +
```

```
    GESTAO*ISE + ISE_ESCOLA + COMPROMETIMENTO*ISE +
```

```
    CLIMA_ESCOLAR*ISE + MATERIA*ISE + PROGRAMA*ISE +
```

```
    ICG*ISE + IRD*ISE + COMUNIDADE*ISE +
```

```
    (1 + ISE | CD_ESCOLA), data = df,
```

```
    REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
```

```
(d3 <- drop1(fit3, test="Chisq"))
```

```
#(-): RACA, DEFAS, MATERIA, COMUNIDADE
```

```
#(+): CLIMA_ESCOLAR, IRD
```

```
summary(fit3)
```

```
(a3 <- anova(fit3, fit2))
```

```
#LP
```

```
##modelo nulo
```

```
fit. <- lmer(VL_PROFICIENCIA_LP ~ (1 | CD_ESCOLA),  
            data = df, REML = FALSE)
```

```
summary(fit.)
```

```
(VPC. <- 263.8*100/(263.8 + 2381.8))
```

```
fita. <- lm(VL_PROFICIENCIA_LP ~ 1, data = df)
```

```
(a. <- anova(fit., fita.))
```

```
## modelo de nível 1
```

```
fit1a. <- lmer(VL_PROFICIENCIA_LP ~ ISE + SEXO + RACA + DEFAS +  
              (1 | CD_ESCOLA), data = df,
```

```
              REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
```

```
(d1a. <- drop1(fit1., test="Chisq"))
```

```
#testando efeitos aleatórios
```

```
#ISE
```

```
fit1. <- lmer(VL_PROFICIENCIA_LP ~ ISE + SEXO + RACA + DEFAS +  
              (1 + ISE | CD_ESCOLA), data = df,
```

```
              REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
```

```

(i1. <- isSingular(fit1.))
(a1. <- anova(fit1., fit1a.))
(d1. <- drop1(fit1.,test="Chisq"))

summary(fit1.)

(VPC1. <- 177.62*100/(177.62 + 12.09 + 2181.05))

#análise dos efeitos aleatórios

# ISE
raneff. <- ranef(fit1., condVar = TRUE)

plot(ranef. [[1]], xlab = "Variação da proficiência média em Língua Portuguesa", ylab =
"Variação do efeito do ISE")
abline(h = 0, col = "red")
abline(v = 0, col = "red")

### modelo de nível 2
fit2. <- update(fit1., . ~ . + PPP + GESTAO + ISE_ESCOLA + COMPROMETIMENTO
+
          CLIMA_ESCOLAR + MATERIA + PROGRAMA + ICG + IRD +
          COMUNIDADE)
(d2. <- drop1(fit2., test="Chisq")) #PPP, GESTAO, RS
(i2. <- isSingular(fit2.))
summary(fit2.)

(a2. <- anova(fit2., fit1.))

(VPC2. <- 108.38*100/(108.38 + 12.12 + 2181.42))

### permitindo interações
#ISE
fit3. <- lmer(VL_PROFICIENCIA_LP ~ SEXO*ISE + RACA*ISE + DEFAS*ISE +
PPP*ISE +
          GESTAO*ISE + ISE_ESCOLA + COMPROMETIMENTO*ISE +
          CLIMA_ESCOLAR*ISE + MATERIA*ISE + PROGRAMA*ISE +
          ICG*ISE + IRD*ISE + COMUNIDADE*ISE +
          (1 + ISE | CD_ESCOLA), data = df,
          REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
(d3. <- drop1(fit3a,test="Chisq"))
#(-): RACA, DEFAS, MATERIA, , COMUNIDADE
#(+): CLIMA_ESCOLAR, IRD
summary(fit3.)

```

```
(a3. <- anova(fit3., fit2.))
```

```
#####
```

```
#analizando os grupos de escola
```

```
#preparando bases de dados
```

```
## MT
```

```
raneffq <- raneff$CD_ESCOLA
```

```
raneff1 <- raneffq %>% filter(`(Intercept)` < 0 & ISE > 0)
raneff1["CD_ESCOLA"] <- as.numeric(row.names(raneff1))
raneff1 <- raneff1 %>% select(CD_ESCOLA)
df_q1 <- left_join(raneff1, df, by = "CD_ESCOLA")
```

```
raneff2 <- raneffq %>% filter(`(Intercept)` > 0 & ISE > 0)
raneff2["CD_ESCOLA"] <- as.numeric(row.names(raneff2))
raneff2 <- raneff2 %>% select(CD_ESCOLA)
df_q2 <- left_join(raneff2, df, by = "CD_ESCOLA")
```

```
raneff3 <- raneffq %>% filter(`(Intercept)` < 0 & ISE < 0)
raneff3["CD_ESCOLA"] <- as.numeric(row.names(raneff3))
raneff3 <- raneff3 %>% select(CD_ESCOLA)
df_q3 <- left_join(raneff3, df, by = "CD_ESCOLA")
```

```
raneff4 <- raneffq %>% filter(`(Intercept)` > 0 & ISE < 0)
raneff4["CD_ESCOLA"] <- as.numeric(row.names(raneff4))
raneff4 <- raneff4 %>% select(CD_ESCOLA)
df_q4 <- left_join(raneff4, df, by = "CD_ESCOLA")
```

```
##LP
```

```
raneffq. <- raneff.$CD_ESCOLA
```

```
raneff1. <- raneffq. %>% filter(`(Intercept)` < 0 & ISE > 0)
raneff1.["CD_ESCOLA"] <- as.numeric(row.names(raneff1.))
raneff1. <- raneff1. %>% select(CD_ESCOLA)
df_q1. <- left_join(raneff1., df, by = "CD_ESCOLA")
```

```
raneff2. <- raneffq. %>% filter(`(Intercept)` > 0 & ISE > 0)
raneff2.["CD_ESCOLA"] <- as.numeric(row.names(raneff2.))
raneff2. <- raneff2. %>% select(CD_ESCOLA)
df_q2. <- left_join(raneff2., df, by = "CD_ESCOLA")
```

```
raneff3. <- raneffq. %>% filter(`(Intercept)` < 0 & ISE < 0)
raneff3.["CD_ESCOLA"] <- as.numeric(row.names(raneff3.))
raneff3. <- raneff3. %>% select(CD_ESCOLA)
```

```

df_q3. <- left_join(raneff3., df, by = "CD_ESCOLA")

raneff4. <- raneffq. %>% filter(`(Intercept)` > 0 & ISE < 0)
raneff4.["CD_ESCOLA"] <- as.numeric(row.names(raneff4.))
raneff4. <- raneff4. %>% select(CD_ESCOLA)
df_q4. <- left_join(raneff4., df, by = "CD_ESCOLA")

## análise descritiva dos dados
# MT
var_resumo1 <- data.frame(`Variável` = colnames(df_q1[,3:21]),
                        `Média1` = round(map_dbl(na.omit(df_q1[,3:21]), mean), 4))
var_resumo2 <- data.frame(`Média2` = round(map_dbl(na.omit(df_q2[,3:21]), mean),
4))
var_resumo3 <- data.frame(`Média3` = round(map_dbl(na.omit(df_q3[,3:21]), mean),
4))
var_resumo4 <- data.frame(`Média4` = round(map_dbl(na.omit(df_q4[,3:21]), mean),
4))
var_resumoq <- cbind(var_resumo1, var_resumo2, var_resumo3, var_resumo4)

fwrite(var_resumoq, "var_resumoq.csv")

nrow(raneff1)
nrow(raneff1)/length(unique(df$CD_ESCOLA))
nrow(raneff2)
nrow(raneff2)/length(unique(df$CD_ESCOLA))
nrow(raneff3)
nrow(raneff3)/length(unique(df$CD_ESCOLA))
nrow(raneff4)
nrow(raneff4)/length(unique(df$CD_ESCOLA))

# LP
var_resumo1. <- data.frame(`Variável` = colnames(df_q1.[,3:21]),
                        `Média1` = round(map_dbl(na.omit(df_q1.[,3:21]), mean), 4))
var_resumo2. <- data.frame(`Média2` = round(map_dbl(na.omit(df_q2.[,3:21]), mean),
4))
var_resumo3. <- data.frame(`Média3` = round(map_dbl(na.omit(df_q3.[,3:21]), mean),
4))
var_resumo4. <- data.frame(`Média4` = round(map_dbl(na.omit(df_q4.[,3:21]), mean),
4))
var_resumoq. <- cbind(var_resumo1., var_resumo2., var_resumo3., var_resumo4.)

fwrite(var_resumoq., "var_resumoq..csv")

nrow(raneff1.)

```

```

nrow(raneff1.)/length(unique(df$CD_ESCOLA))
nrow(raneff2.)
nrow(raneff2.)/length(unique(df$CD_ESCOLA))
nrow(raneff3.)
nrow(raneff3.)/length(unique(df$CD_ESCOLA))
nrow(raneff4.)
nrow(raneff4.)/length(unique(df$CD_ESCOLA))

## modelos por quartil
#MT
# 1° quartil
fit2q1 <- lmer(VL_PROFICIENCIA_MT ~ ISE + SEXO + RACA + DEFAS +
              PPP + GESTAO + ISE_ESCOLA + COMPROMETIMENTO +
              CLIMA_ESCOLAR + MATERIA + PROGRAMA + ICG + IRD +
              COMUNIDADE
              + (1 + ISE | CD_ESCOLA), data = df_q1,
              REML = FALSE, control = lmerControl(optimizer = "bobyqa"))

(d2q1 <- drop1(fit2q1, test="Chisq"))

summary(fit2q1)

(a2q1 <- anova(fit2q1, fit1a))

(VPC2q1 <- 162.78*100/(162.79 + 16.74 + 2199.47))

## permitindo interações
#ISE
fit3q1 <- lmer(VL_PROFICIENCIA_MT ~ SEXO*ISE + RACA*ISE + DEFAS*ISE +
              PPP*ISE +
              GESTAO*ISE + ISE_ESCOLA + COMPROMETIMENTO*ISE +
              CLIMA_ESCOLAR*ISE + MATERIA*ISE + PROGRAMA*ISE +
              ICG*ISE + IRD*ISE + COMUNIDADE*ISE +
              (1 + ISE | CD_ESCOLA), data = df_q1,
              REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
(d3q1 <- drop1(fit3q1, test="Chisq"))
#(-): RACA, DEFAS

summary(fit3q1)

(a3q1 <- anova(fit3q1, fit2q1))

# 2° quartil
fit2q2 <- lmer(VL_PROFICIENCIA_MT ~ ISE + SEXO + RACA + DEFAS +

```

```

      PPP + GESTAO + ISE_ESCOLA + COMPROMETIMENTO +
      CLIMA_ESCOLAR + MATERIA + PROGRAMA + ICG + IRD +
COMUNIDADE
      + (1 + ISE | CD_ESCOLA), data = df_q2,
      REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
(d2q2 <- drop1(fit2q2,test="Chisq"))

summary(fit2q2)

(a2q2 <- anova(fit2q2, fit1a))

(VPC2q2 <- 162.78*100/(162.79 + 16.74 + 2199.47))

### permitindo interações
#ISE
fit3q2 <- lmer(VL_PROFICIENCIA_MT ~ SEXO*ISE + RACA*ISE + DEFAS*ISE +
PPP*ISE +
      GESTAO*ISE + ISE_ESCOLA + COMPROMETIMENTO*ISE +
      CLIMA_ESCOLAR*ISE + MATERIA*ISE + PROGRAMA*ISE +
      ICG*ISE + IRD*ISE + COMUNIDADE*ISE +
      (1 + ISE | CD_ESCOLA), data = df_q2,
      REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
(d3q2 <- drop1(fit3q2,test="Chisq"))
#(-): RACA, DEFAS, PROGRAMA
#(+): MATERIA
summary(fit3q2)

(a3q2 <- anova(fit3q2, fit2q2))

# 3° quartil
fit2q3 <- lmer(VL_PROFICIENCIA_MT ~ ISE + SEXO + RACA + DEFAS +
      PPP + GESTAO + ISE_ESCOLA + COMPROMETIMENTO +
      CLIMA_ESCOLAR + MATERIA + PROGRAMA + ICG + IRD +
COMUNIDADE
      + (1 + ISE | CD_ESCOLA), data = df_q3,
      REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
(d2q3 <- drop1(fit2q3,test="Chisq"))

summary(fit2q3)

(a2q3 <- anova(fit2q3, fit1a))

(VPC2q3 <- 162.78*100/(162.79 + 16.74 + 2199.47))

```

```

## permitindo interações
#ISE
fit3q3 <- lmer(VL_PROFICIENCIA_MT ~ SEXO*ISE + RACA*ISE + DEFAS*ISE +
PPP*ISE +
      GESTAO*ISE + ISE_ESCOLA + COMPROMETIMENTO*ISE +
      CLIMA_ESCOLAR*ISE + MATERIA*ISE + PROGRAMA*ISE +
      ICG*ISE + IRD*ISE + COMUNIDADE*ISE +
      (1 + ISE | CD_ESCOLA), data = df_q3,
      REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
(d3q3 <- drop1(fit3q3, test="Chisq"))
#(-): RACA, DEFAS, MATERIA, , COMUNIDADE
#(+): CLIMA_ESCOLAR, IRD
summary(fit3q3)

(a3q3 <- anova(fit3q3, fit2q3))

# 4° quartil
fit2q4 <- lmer(VL_PROFICIENCIA_MT ~ ISE + SEXO + RACA + DEFAS +
      PPP + GESTAO + ISE_ESCOLA + COMPROMETIMENTO +
      CLIMA_ESCOLAR + MATERIA + PROGRAMA + ICG + IRD +
      COMUNIDADE
      + (1 + ISE | CD_ESCOLA), data = df_q4,
      REML = FALSE, control = lmerControl(optimizer = "bobyqa"))

(d2q4 <- drop1(fit2q1, test="Chisq"))
(i2q4 <- isSingular(fit2q1))
summary(fit2q4)

(a2q4 <- anova(fit2q4, fit1a))

(VPC2q4 <- 162.78*100/(162.79 + 16.74 + 2199.47))

## permitindo interações
#ISE
fit3q4 <- lmer(VL_PROFICIENCIA_MT ~ SEXO*ISE + RACA*ISE + DEFAS*ISE +
PPP*ISE +
      GESTAO*ISE + ISE_ESCOLA + COMPROMETIMENTO*ISE +
      CLIMA_ESCOLAR*ISE + MATERIA*ISE + PROGRAMA*ISE +
      ICG*ISE + IRD*ISE + COMUNIDADE*ISE +
      (1 + ISE | CD_ESCOLA), data = df_q4,
      REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
(d3q4 <- drop1(fit3q4, test="Chisq"))
#(-): RACA, DEFAS, MATERIA, , COMUNIDADE
#(+): CLIMA_ESCOLAR, IRD

```

```
summary(fit3q4)
```

```
(a3q4 <- anova(fit3q4, fit2q4))
```

```
#LP
```

```
# 1° quartil
```

```
fit2q1. <- lmer(VL_PROFICIENCIA_LP ~ ISE + SEXO + RACA + DEFAS +
  PPP + GESTAO + ISE_ESCOLA + COMPROMETIMENTO +
  CLIMA_ESCOLAR + MATERIA + PROGRAMA + ICG + IRD +
  COMUNIDADE
  + (1 + ISE | CD_ESCOLA), data = df_q1.,
  REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
```

```
(d2q1. <- drop1(fit2q1.,test="Chisq"))
```

```
(i2q1. <- isSingular(fit2q1.))
```

```
summary(fit2q1.)
```

```
(a2q1. <- anova(fit2q1., fit1a.))
```

```
(VPC2q1 <- 162.78*100/(162.79 + 16.74 + 2199.47))
```

```
## permitindo interações
```

```
#ISE
```

```
fit3q1. <- lmer(VL_PROFICIENCIA_LP ~ SEXO*ISE + RACA*ISE + DEFAS*ISE +
  PPP*ISE +
  GESTAO*ISE + ISE_ESCOLA + COMPROMETIMENTO*ISE +
  CLIMA_ESCOLAR*ISE + MATERIA*ISE + PROGRAMA*ISE +
  ICG*ISE + IRD*ISE + COMUNIDADE*ISE +
  (1 + ISE | CD_ESCOLA), data = df_q1.,
  REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
```

```
(d3q1. <- drop1(fit3q1.,test="Chisq"))
```

```
#(-): RACA, DEFAS, MATERIA, , COMUNIDADE
```

```
#(+): CLIMA_ESCOLAR, IRD
```

```
summary(fit3q1.)
```

```
(a3q1. <- anova(fit3q1., fit2q1.))
```

```
# 2° quartil
```

```
fit2q2. <- lmer(VL_PROFICIENCIA_LP ~ ISE + SEXO + RACA + DEFAS +
  PPP + GESTAO + ISE_ESCOLA + COMPROMETIMENTO +
  CLIMA_ESCOLAR + MATERIA + PROGRAMA + ICG + IRD +
  COMUNIDADE
```

```

+ (1 + ISE | CD_ESCOLA), data = df_q2.,
REML = FALSE, control = lmerControl(optimizer = "bobyqa"))

(d2q2. <- drop1(fit2q2.,test="Chisq"))
(i2q2. <- isSingular(fit2q2.))
summary(fit2q2.)

(a2q2. <- anova(fit2q2., fit1a.))

(VPC2q2. <- 162.78*100/(162.79 + 16.74 + 2199.47))

## permitindo interações
#ISE
fit3q2. <- lmer(VL_PROFICIENCIA_LP ~ SEXO*ISE + RACA*ISE + DEFAS*ISE +
PPP*ISE +
      GESTAO*ISE + ISE_ESCOLA + COMPROMETIMENTO*ISE +
      CLIMA_ESCOLAR*ISE + MATERIA*ISE + PROGRAMA*ISE +
      ICG*ISE + IRD*ISE + COMUNIDADE*ISE +
      (1 + ISE | CD_ESCOLA), data = df_q2.,
      REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
(d3q2. <- drop1(fit3q2.,test="Chisq"))
#(-): RACA, DEFAS, MATERIA, , COMUNIDADE
#(+): CLIMA_ESCOLAR, IRD
summary(fit3q2)

(a3q2. <- anova(fit3q2., fit2q2.))

# 3° quartil
fit2q3. <- lmer(VL_PROFICIENCIA_LP ~ ISE + SEXO + RACA + DEFAS +
      PPP + GESTAO + ISE_ESCOLA + COMPROMETIMENTO +
      CLIMA_ESCOLAR + MATERIA + PROGRAMA + ICG + IRD +
      COMUNIDADE
      + (1 + ISE | CD_ESCOLA), data = df_q3.,
      REML = FALSE, control = lmerControl(optimizer = "bobyqa"))

(d2q3. <- drop1(fit2q3.,test="Chisq"))
(i2q3. <- isSingular(fit2q3.))
summary(fit2q3.)

(a2q3. <- anova(fit2q3., fit1a.))

(VPC2q3. <- 162.78*100/(162.79 + 16.74 + 2199.47))

## permitindo interações

```

```

#ISE
fit3q3. <- lmer(VL_PROFICIENCIA_LP ~ SEXO*ISE + RACA*ISE + DEFAS*ISE +
PPP*ISE +
          GESTAO*ISE + ISE_ESCOLA + COMPROMETIMENTO*ISE +
          CLIMA_ESCOLAR*ISE + MATERIA*ISE + PROGRAMA*ISE +
          ICG*ISE + IRD*ISE + COMUNIDADE*ISE +
          (1 + ISE | CD_ESCOLA), data = df_q3.,
          REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
(d3q3. <- drop1(fit3q3.,test="Chisq"))
#(-): RACA, DEFAS, MATERIA, , COMUNIDADE
#(+): CLIMA_ESCOLAR, IRD
summary(fit3q3.)

(a3q3. <- anova(fit3q3., fit2q3.))

# 4° quartil
fit2q4. <- lmer(VL_PROFICIENCIA_LP ~ ISE + SEXO + RACA + DEFAS +
          PPP + GESTAO + ISE_ESCOLA + COMPROMETIMENTO +
          CLIMA_ESCOLAR + MATERIA + PROGRAMA + ICG + IRD +
          COMUNIDADE
          + (1 + ISE | CD_ESCOLA), data = df_q4.,
          REML = FALSE, control = lmerControl(optimizer = "bobyqa"))

(d2q4. <- drop1(fit2q1.,test="Chisq"))
(i2q4. <- isSingular(fit2q1.))
summary(fit2q4.)

(a2q4. <- anova(fit2q4., fit1a.))

(VPC2q4. <- 162.78*100/(162.79 + 16.74 + 2199.47))

### permitindo interações
#ISE
fit3q4. <- lmer(VL_PROFICIENCIA_LP ~ SEXO*ISE + RACA*ISE + DEFAS*ISE +
PPP*ISE +
          GESTAO*ISE + ISE_ESCOLA + COMPROMETIMENTO*ISE +
          CLIMA_ESCOLAR*ISE + MATERIA*ISE + PROGRAMA*ISE +
          ICG*ISE + IRD*ISE + COMUNIDADE*ISE +
          (1 + ISE | CD_ESCOLA), data = df_q4.,
          REML = FALSE, control = lmerControl(optimizer = "bobyqa"))
(d3q4. <- drop1(fit3q4.,test="Chisq"))
#(-): RACA, DEFAS, MATERIA, , COMUNIDADE
#(+): CLIMA_ESCOLAR, IRD
summary(fit3q4.)

```

```
(a3q4. <- anova(fit3q4., fit2q4.))
```

ANEXO A – QUESTIONÁRIO CONTEXTUAL PROEB

Figura 3 – Itens do questionário contextual do PROEB 2019

(continua)

Itens dos questionários do estudante e do professor do PROEB 2019 utilizados para consolidação das variáveis

Itens do questionário do estudante**Sexo**

1 Qual é o seu sexo?

Masculino.

Feminino.

Raça

2 De acordo com as categorias do IBGE, como você se considera?

Branco(a).

Pardo(a).

Preto(a).

Amarelo(a).

Indígena.

Defasagem Escolar

3 Qual é a sua idade?

7 anos ou menos.

8 anos.

9 anos.

10 anos.

11 anos.

12 anos.

13 anos.

14 anos.

15 anos.

16 anos.

17 anos.

18 anos.

19 anos.

20 anos ou mais.

Fonte: SEE/MG, 2019.

Figura 3 – Itens do questionário contextual do PROEB 2019

(continuação)

Índice Socioeconômico	
5	Até que série/ano sua mãe ou a responsável por você estudou? Nunca estudou ou não completou a 4a série/5° ano do ensino fundamental (antigo primário). Completou a 4a série/5° ano, mas não completou a 8a série/9° ano do ensino fundamental (antigo ginásio). Completou a 8a série/9° ano, mas não completou o ensino médio (antigo 2o grau). Completou o ensino médio, mas não completou a faculdade. Completou a faculdade. Não sei.
6	Até que série/ano seu pai ou o responsável por você estudou? Nunca estudou ou não completou a 4a série/5° ano do ensino fundamental (antigo primário). Completou a 4a série/5° ano, mas não completou a 8a série/9° ano do ensino fundamental (antigo ginásio). Completou a 8a série/9° ano, mas não completou o ensino médio (antigo 2o grau). Completou o ensino médio, mas não completou a faculdade. Completou a faculdade. Não sei.
7	A rua onde você mora é asfaltada ou tem calçamento? Sim. Não.
8	A rua onde você mora tem coleta de lixo? Sim. Não.
9	Alguém que mora com você recebe Bolsa Família? Sim. Não.
10	Você tem acesso à internet na sua casa? Sim. Não.
11	Você tem um quarto só para você na sua casa? Sim. Não.
12	Quantos banheiros existem na sua casa? Nenhum. Um. Dois ou mais.
15	Quantos micro-ondas existem na sua casa? Nenhum. Um. Dois ou mais.
16	Quantos ares-condicionados existem na sua casa? Nenhum. Um. Dois ou mais.
17	Quantos automóveis (carros) existem na sua casa? Nenhum. Um. Dois ou mais.
19	Quantos computadores (notebook etc.) existem na sua casa? Nenhum. Um. Dois ou mais.
20	Quantos videogames (Xbox, PS4 etc.) existem na sua casa? Nenhum. Um. Dois ou mais.
21	Sem considerar livros escolares, jornais e revistas, quantos livros existem na sua casa? Não existem livros na minha casa. O bastante para encher uma prateleira (1 a 10 livros). O bastante para encher duas prateleiras (11 a 20 livros). O bastante para encher uma estante (21 a 50 livros). O bastante para encher duas estantes (51 a 100 livros). O bastante para encher várias estantes (mais de 100 livros)

Fonte: SEE/MG, 2019.

Figura 3 – Itens do questionário contextual do PROEB 2019

(continuação)

<p>Comprometimento do professor Os estudantes tinham que escolher entre "Não concordo.", "Concordo pouco.", "Concordo." e "Concordo muito."</p> <p>46 Meus professores faltam muito às aulas.</p> <p>Gestão da Matéria Os estudantes tinham que escolher entre "Não concordo.", "Concordo pouco.", "Concordo." e "Concordo muito."</p> <p>Meus professores utilizam diferentes recursos para apresentar a matéria (vídeo, música, computador etc.).</p> <p>41 música, computador etc.).</p> <p>42 Meus professores respondem às dúvidas que aparecem durante as aulas.</p> <p>43 Meus professores utilizam o livro didático.</p> <p>44 Meus professores falam sobre a importância de estudar.</p> <p>45 Meus professores revisam a matéria para as provas.</p> <p>Clima Escolar Os estudantes tinham que escolher entre "Não concordo.", "Concordo pouco.", "Concordo." e "Concordo muito."</p> <p>Na minha escola, temos várias atividades interessantes ao longo do ano (gincanas, apresentações, esportes etc.).</p> <p>33 apresentações, esportes etc.).</p> <p>Na minha escola, o(a) diretor(a) e os(as) professores(as) deixam claro que estamos na escola para aprender.</p> <p>34 escola para aprender.</p> <p>Na minha escola, os estudantes se relacionam bem com o(a) diretor(a) e os(as) professores(as)</p> <p>35 professores(as)</p> <p>36 Na minha escola, os estudantes se relacionam bem uns com os outros.</p> <p>Na minha escola, quando há brigas, os funcionários/ servidores da escola interferem na situação.</p> <p>37 situação.</p> <p>38 Sinto-me seguro dentro da minha escola.</p> <p>Os estudantes obedecem quando o(a) diretor(a) ou os(as) professores(as) chamam a atenção</p> <p>39 atenção</p> <p>Na minha escola, muitos estudantes ficam do lado de fora da sala de aula fazendo barulho.</p> <p>40 barulho.</p>

Fonte: SEE/MG, 2019.

Figura 3 – Itens do questionário contextual do PROEB 2019

(continuação)

Itens do questionário do professor**Percentual do programa concluído**

Neste ano, nesta escola, nestas turmas, em média, qual foi a porcentagem do planejamento curricular que você conseguiu concluir?

- Até 25%.
- De 25% à 50%.
- De 50% à 60%.
- De 60% à 70%.
- De 70% à 80%.
- De 80% à 90%.
- De 90% à 100%.

Efetividade da implementação do PPP

Nessas questões os professores tinham que escolher entre "Não foi efetiva.", "Foi pouco efetiva.", "Foi efetiva." e "Foi muito efetiva."

Agora, diga sua opinião sobre a efetividade da participação dos seguintes atores educacionais na implementação do PPP na sua escola.

- 30 Secretaria Estadual de Educação.
- 31 . Superintendência Regional de Ensino.
- 32 Equipe gestora.
- 33 Diretor escolar.
- 34 Coordenador pedagógico.
- 35 Professores.
- 36 Estudantes.
- 37 Família dos estudantes.
- 38 Comunidade.

Nessas questões os professores tinham que escolher entre "Não dificultou.", "Dificultou pouco.", "Dificultou." e "Dificultou muito."

Em relação a implementação e execução do PPP na sua escola, informe o quanto cada elemento listado dificultou o processo.

- 39 Os recursos financeiros.
- 40 Os recursos pedagógicos.
- 41 A participação da comunidade.
- 42 A infraestrutura para o trabalho.
- 43 As ideias inovadoras da comunidade
- 44 A participação dos atores escolares.
- 45 A articulação entre os atores escolares.
- 46 . As ideias inovadoras dos atores escolares.
- 47 A comunicação com as instâncias superiores.
- 48 O planejamento para execução das atividades propostas.
- 49 O interesse da equipe em implementar algumas novidades.
- 50 Os prazos estabelecidos para execução das atividades propostas.
- 51 A assiduidade da equipe escolar nas reuniões de articulação do PPP.
- 52 A assiduidade da comunidade nas reuniões de articulação do PPP

Nessas questões os professores tinham que escolher entre "Em nenhuma medida.", "Em pouca medida.", "Em alguma medida." e "Em grande medida."

Em que medida os instrumentos abaixo foram construídos pela direção escolar, pela coordenação pedagógica e/ou pelos(as) professores(as) em consonância com PPP da escola?

- 53 Planos de ensino.
- 54 Planos de aula.
- 55 . Planos curriculares.
- 56 Regimento escolar
- 57 Projetos educacionais.

Fonte: SEE/MG, 2019.

Figura 3 – Itens do questionário contextual do PROEB 2019

(continuação)

Gestão do Diretor

Nessas questões os professores tinham que escolher entre "Não foi efetiva.", "Foi pouco efetiva.", "Foi efetiva." e "Foi muito efetiva.".

- 96 O(A) diretor(a) da minha escola acompanha com frequência a entrada e a saída dos estudantes.
- 97 O(A) diretor(a) da minha escola auxilia os(as) professores(as) na escolha ou troca do livro didático.
- 98 O(A) diretor(a) da minha escola orienta os(as) professores(as) na elaboração do plano de trabalho e outras produções.
- 99 O(A) diretor(a) da minha escola orienta os(as) professores(as) na elaboração de projetos didáticos diferenciados.
- 100 O(A) diretor(a) da minha escola promove reuniões pedagógicas e/ou grupos de estudos com os(as) professores(as).
- 101 O(A) diretor(a) da minha escola observa o trabalho nas turmas e orienta pedagogicamente os(as) professores(as).
- 102 O(A) diretor(a) da minha escola compromete-se com a gestão pedagógica.
- 106 O(A) diretor(a) da minha escola resolve coletivamente as questões relevantes.
- 107 O(A) diretor(a) da minha escola consulta boa parte dos(as) professores(as) nas decisões importantes.
- 108 O(A) diretor(a) da minha escola convoca várias reuniões com os pais dos estudantes ao longo do ano letivo.
- 109 O(A) diretor(a) da minha escola ouve os estudantes quando eles o procuram.
- 110 O(A) diretor(a) da minha escola presta conta de suas decisões.
- 111 O(A) diretor(a) da minha escola deixa claro para professores(as), funcionários(as), pais e estudantes quais são as regras a serem obedecidas na escola.
- 112 O(A) diretor(a) da minha escola tem práticas democráticas de gestão.

Fonte: SEE/MG, 2019.

ANEXO B – VARIÁVEIS DO CENSO 2019

Figura 4 – Itens do questionário do Censo Escolar 2019

Itens do Censo Escolar 2019 utilizados para consolidação das variáveis	
Relação com a comunidade	
1	A escola possui site ou blog ou página em redes sociais para comunicação institucional
	Sim
	Não
2	A escola compartilha espaços para atividades de integração escola-comunidade
	Sim
	Não
3	A escola usa espaços e equipamentos do entorno escolar para atividades regulares com os alunos
	Sim
	Não
4	Órgãos colegiados em funcionamento na escola - Conselho Escolar
	Sim
	Não
5	Órgãos colegiados em funcionamento na escola - Não há órgãos colegiados em funcionamento
	Sim
	Não

Fonte: SEE/MG, 2019.