

**UNIVERSIDADE MUNICIPAL DE SÃO CAETANO DO SUL
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO
MESTRADO PROFISSIONAL**

André Calixto Gonçalves

**A INFLUÊNCIA DO NÚMERO DE ALUNOS EM SALA DE AULA NO
DESEMPENHO ACADÊMICO DOS ESTUDANTES BRASILEIROS**

**São Caetano do Sul
2021**

ANDRÉ CALIXTO GONÇALVES

**A INFLUÊNCIA DO NÚMERO DE ALUNOS EM SALA DE AULA NO
DESEMPENHO ACADÊMICO DOS ESTUDANTES BRASILEIROS**

**Trabalho Final de Curso apresentado ao
Programa de Pós-Graduação em Educação
– Mestrado Profissional – Universidade
Municipal de São Caetano do Sul como
requisito parcial para a obtenção do título
de Mestre em Educação.**

**Área de concentração: Formação de
Professores e Gestores**

Orientador: Prof. Dr. Leandro Campi Prearo

**São Caetano do Sul
2021**

FICHA CATALOGRÁFICA

GONCALVES, André Calixto.

A Influência do número de alunos em sala de aula no desempenho acadêmico dos estudantes brasileiros / André Calixto Goncalves – São Caetano do Sul: USCS, 2021.

162p.

Orientador: Prof. Dr. Leandro Campi Prearo.

Dissertação (Mestrado) – USCS. Universidade de São Caetano do Sul. Programa de Pós-Graduação em Educação. Mestrado Profissional, 2021.

1. Educação básica. 2. Avaliação em larga escala. 3. Políticas Públicas de educação. 4. Despenho estudantil. 5. Propense Score Matching. Título II. Universidade Municipal de São Caetano do Sul.

**Reitor da Universidade Municipal de São Caetano do Sul
Prof. Dr. Leandro Campi Prearo**

**Pró-reitora de Pós-Graduação e Pesquisa
Prof.^a Dr.^a Maria do Carmo Romeiro**

**Gestão do Programa de Pós-Graduação em Educação
Prof. Dr. Nonato Assis de Miranda
Prof.^a Dr.^a Ana Sílvia Moço Aparício**

Trabalho Final de Curso defendido e aprovado em 12/08/2021 pela Banca Examinadora constituída pelos (as) professores (as):

Prof.Dr. Leandro Campi Prearo (USCS)

Prof.Dr. Nonato Assis de Miranda (USCS)

Prof.Dr. Estefano Vizconde Veraszto (UFSCar)

AGRADECIMENTOS

Gostaria de agradecer a Deus, sobretudo por mais esta conquista. À minha esposa, pelo incentivo e paciência, e à minha família, por sempre estar presente.

Agradeço principalmente ao meu orientador, o Prof. Dr. Leandro Campi Prearo, que não mediu esforços para me orientar desde o início do desenvolvimento desta pesquisa, pela confiança depositada em mim e, especialmente, por me ajudar nessa nova conquista.

Também não poderia deixar de agradecer ao Prof. Dr. Nonato de Assis Miranda, por estar sempre disposto a me ajudar.

Quero agradecer, ainda, à Universidade Municipal de São Caetano (USCS), pela concessão de bolsa.

Por fim, gostaria de dedicar este trabalho à memória do meu falecido colega mestrando, mais uma vítima do coronavírus, André Bueno Antonachi, que, desde o início das aulas, chamando-me pelo apelido carinhoso de “Xará”, foi uma pessoa muito presente e um amigo com quem sempre pude contar.

RESUMO

O presente estudo teve o objetivo de verificar se o número de alunos por turma influencia o seu desempenho escolar, tanto em relação aos Anos Iniciais quanto aos Anos Finais do Ensino Fundamental. Para o desenvolvimento do trabalho, utilizaram-se os microdados da Prova Brasil, referente ao ano de 2017. Assim, a partir de dados de pesquisas anteriormente realizadas sobre o tema, verificaram-se diversos caminhos a seguir, a fim de obter o resultado desejado. Optou-se, então, pela utilização do modelo estatístico de *Propense Score Matching*, com o estimador *Nearest Neighbor Matching*. Dessa forma, foi possível agrupar as turmas em *clusters* por grau de semelhança. A métrica usada para auxiliar no agrupamento dos *clusters* foi a distância euclidiana. Para verificar a aderência à distribuição normal dos dados, utilizou-se o teste Kolmogorov Smirnov e, para testar a hipótese nula das medianas, utilizou-se o teste de Wilcoxon. A fim de auxiliar a execução da análise estatística, optou-se pelo uso do software IBM SPSS *Statistic*. Os resultados, em geral, mostram que o número de alunos por turma tem pouca influência no desempenho escolar e, quando essa influência existe, as turmas grandes apresentam melhor desempenho em relação às turmas pequenas. Com base nos resultados, como produto final, foi proposto o desenvolvimento de um plano de ação para gestores públicos, notadamente diretores de escola, com foco na gestão e implementação da melhoria da qualidade da educação. Essa proposição está pautada no entendimento de que o presente trabalho servirá de parâmetro para subsidiar eventuais políticas públicas em educação que visam a controlar o número de alunos por turma, como metodologia de aumento de desempenho.

Palavras-chave: Educação básica. Avaliação em larga escala. Políticas Públicas de educação. Desempenho estudantil. *Propense Score Matching*

ABSTRACT

The present study aimed to verify whether the number of students per class influences the school performance of elementary school students, both in relation to the initial and final years of this level of education. For the development of the work, microdata from Prova Brasil for the year 2017 were used. Thus, based on data from previous research on the subject, it was possible to verify several paths to be followed to obtain the desired result. It was then decided to use the statistical model of Propense Score Matching, with the Nearest Neighbor Matching estimator. Thus, it was possible to group the classes into Clusters by degree of similarity, the metric used to assist in the clustering of clusters was the Euclidean Distance. To verify adherence to the normal distribution of data, the Kolmogorov Smirnov test was used and to test the null hypothesis of medians, the Wilcoxon test was used. To help perform the statistical analysis, we chose to use the IBM SPSS Statistic software. The results, in general, showed that the number of students per class has little influence on school performance, and when this influence exists, the results showed that large classes obtain better performances in relation to small classes. Based on the results, as a final product, it was proposed the development of an action plan for public managers, especially school directors, with a focus on the management and implementation of improving the quality of education. This proposition is based on the understanding that this work will serve as a parameter to support any public policies in education that aim to control the number of students per class as a methodology to increase performance.

Keywords: Basic education. Large scale assessment. Public Education Policies. Student performance. Propense Score Matching.

LISTA DE FIGURAS

Figura 1: Desempenho Escolar	25
Figura 2: Valor médio por aluno ao ano	28
Figura 3: Influência do Tamanho da classe na nota de Matemática	33
Figura 4: Desempenho Matemática PISA – 2018 – América Latina.....	41
Figura 5: Desempenho Leitura Pisa – 2018 – América Latina	41
Figura 6: Desempenho Ciências Pisa – 2018 – América Latina.....	42

LISTA DE TABELAS

Tabela 1: Média do número de alunos por turma em 2020 no Brasil	35
Tabela 2: Média do número de alunos por turma em 2020 por região	35
Tabela 3: Médias IDEB de 2015, 2017, 2019 e 2021.	56
Tabela 4: Relação de autores sobre os determinantes de desempenho educacional	94
Tabela 5: Alunos – Variáveis escolhidas	95
Tabela 6: Professor – Variáveis escolhidas.....	96
Tabela 7: Diretor – Variáveis escolhidas.....	98
Tabela 8: Escola – Variáveis escolhidas.....	98
Tabela 9: Diretor – Questão 39 – Critério de formação de turmas	104
Tabela 10: Pontuação atribuída a cada variável, pelo Critério Brasil.....	105
Tabela 11: Pontuação atribuída a cada variável, pelo Critério Brasil.....	105
Tabela 12: Pontuação atribuída a cada classe econômica, pelo Critério Brasil.....	106
Tabela 13: Teste Kolmogorov Smirnov referente aos dados do 5º ano.....	109
Tabela 14: Teste Kolmogorov Smirnov referente aos dados do 9º ano.....	109
Tabela 15: Pequena Comparação de Turmas – Língua Portuguesa – 9º Ano	111
Tabela 16: Comparações por Série – Seção 1 – Língua Portuguesa - 9º Ano	112
Tabela 17: Comparações por Série – Seção 2 – Língua Portuguesa - 9º Ano	113
Tabela 18: Comparação detalhada – Seção 1 – Língua Portuguesa - 9º Ano.....	114
Tabela 19: Comparação detalhada – Seção 2 – Língua Portuguesa - 9º Ano.....	115
Tabela 20: Comparação detalhada – Seção 3 – Língua Portuguesa - 9º Ano.....	116
Tabela 21: Comparação detalhada – Seção 4 – Língua Portuguesa - 9º Ano.....	117
Tabela 22: Pequeno número de comparações – Matemática - 9º Ano.....	118
Tabela 23: Comparações por Séries – Seção 1 – Matemática - 9º Ano	119
Tabela 24: Comparações por Séries – Seção 2 – Matemática - 9º Ano	120

Tabela 25: Comparação detalhada – Seção 1 – Matemática - 9º Ano.....	121
Tabela 26: Comparação detalhada – Seção 2 – Matemática - 9º Ano.....	122
Tabela 27: Comparação detalhada – Seção 3 – Matemática - 9º Ano.....	123
Tabela 28: Comparação detalhada – Seção 4 – Matemática – 9º Ano.....	123
Tabela 29: Pequeno número de comparações – Língua Portuguesa – 5º Ano.....	125
Tabela 30: Comparações por Séries – Seção 1 – Língua Portuguesa - 5º Ano.....	126
Tabela 31: Comparações Séries – Seção 2 – Língua Portuguesa - 5º Ano.....	127
Tabela 32: Comparação detalhada – Seção 1 – Língua Portuguesa - 5º Ano.....	127
Tabela 33: Comparação detalhada – Seção 2 – Língua Portuguesa - 5º Ano.....	128
Tabela 34: Comparação detalhada – Seção 3 – Língua Portuguesa - 5º Ano.....	129
Tabela 35: Comparação detalhada – Seção 4 – Língua Portuguesa - 5º Ano.....	130
Tabela 36: Pequeno número de comparações – Matemática - 5º Ano.....	131
Tabela 37: Comparações por Séries – Seção 1 – Matemática - 5º Ano.....	132
Tabela 38: Comparações por Séries – Seção 2 – Matemática - 5º Ano.....	132
Tabela 39: Comparação detalhada – Seção 1 – Matemática - 5º Ano.....	133
Tabela 40: Comparação detalhada – Seção 2 – Matemática - 5º Ano.....	134
Tabela 41: Comparação detalhada – Seção 3 – Matemática - 5º Ano.....	135
Tabela 42: Comparação detalhada – Seção 4 – Matemática - 5º Ano.....	136

LISTA DE ABREVIATURAS E SIGLAS

AACC - Avaliação do Ambiente de Aprendizagem Construtivista
ABEP - Associação Brasileira de Empresas de Pesquisa
ANEB - Avaliação Nacional da Educação Básica
ANRESC - Avaliação Nacional de Rendimento Escolar
CSR - Class Size Reduction
CSV - Comma Separated Values
ECA - Estatuto da Criança e Adolescente
EUA - Estados Unidos da América
IBM - Internacional Business Machines
IDEB - Índice de Desenvolvimento da Educação Básica
INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
K3 – Kindergarten
LDB - Lei de Diretrizes Básicas da Educação
MEC – Ministério da Educação
NNS - Nearest Neighbor Matching
OCDE - Organização para a Cooperação e Desenvolvimento Econômico
PIB - Produto Interno Bruto
PISA - Programme for International Student Assessment
PNE - Plano Nacional de Educação
PSM - Propensity Score Matching
SAEB - Sistema de Avaliação da Educação Básica
SAT - Stanford Achievement Test
SPSS - Statistical Package for the social Sciences
STAR - Student Teacher Achievement Ratio

SUMÁRIO

1 INTRODUÇÃO	15
1.1 Objetivos Geral e Específicos	17
1.1.1 Geral.....	17
1.1.2 Específicos	18
1.2 Abrangência do estudo	18
1.3 Justificativa do estudo	18
1.4 Organização da dissertação.....	19
2 REFERENCIALTEÓRICO	20
2.1 Determinantes de desempenho	20
2.2 Determinantes da qualidade educacional.....	22
2.3 Avaliações de Larga Escala	36
2.3.1 PISA - <i>Programme for International Student Assessment</i>	39
2.3.2 Prova Brasil	43
2.4 Turmas	46
2.5 Legislação	48
2.5.1 Lei 9.394 de 1996 – Lei de Diretrizes Básicas da Educação.....	48
2.5.2 Constituição Federal de 1988.....	52
2.5.3 Lei 8.069 de 1990 – Estatuto da Criança e do Adolescente.....	54
2.5.4 Lei 13.005 de 2014 – Plano Nacional de Educação.....	55
2.6 Panorama de pesquisas.....	58
2.6.1 América do Norte	59
2.6.2 Europa.....	75
2.6.3 América do Sul	79
2.6.4 África	84
2.6.5 Oriente Médio.....	88
3 MÉTODO	91
3.1 Tipo de pesquisa	91
3.2 Abrangência de dados.....	93
3.3 Período de Referência.....	93
3.4 Método de Preparação e Análise dos dados.....	93
3.5 Tamanho final da amostra.....	102

3.6 Preparação dos Dados	103
4 RESULTADOS	108
4.1 Apresentação dos resultados	110
4.1.1 Anos Finais ou 9ºano	111
4.1.2 Anos Iniciais ou 5º ano	125
5 PRODUTO FINAL: PLANO DE AÇÃO COM FOCO NA IMPLEMENTAÇÃO DE POLÍTICAS PÚBLICAS	137
6 CONCLUSÃO	138
REFERÊNCIAS	142

1 INTRODUÇÃO

O Brasil é um país continental que, em virtude de suas muitas diferenças regionais e sociais, apresenta diversos desafios na efetivação de uma educação pública de qualidade para sua população. Esses desafios estão relacionados à desigualdade de oportunidades, de aprendizagem, de acesso ao ambiente escolar, entre outros fatores. Diante disso, há o desafio de melhorar a estrutura do seu sistema educacional.

De acordo com dados do INEP, por meio dos microdados do Censo Escolar para o ano de 2020, o Brasil tem cerca de 47,3 milhões de alunos, divididos entre aproximadamente 179,5 mil escolas. Com uma taxa de 96,4% de crianças na escola, o maior desafio educacional brasileiro está na qualidade do processo.

Segundo Menezes-Filho (2007), desde a década de 1990, o país tem aumentado significativamente a frequência escolar. O autor ressalta que o problema está em proporcionar a melhoria da qualidade da educação, principalmente na rede pública de ensino.

Menezes-Filho (2015) destaca, ainda, que programas de transferência direta de renda, como o Bolsa Família – que condiciona o recebimento do benefício pelo interessado à frequência dos filhos na escola –, é uma das medidas que mais insere alunos nas escolas. No entanto, o autor adverte que somente esse tipo de política não é suficiente para garantir a melhoria da qualidade educacional. Nesse sentido, considera que uma maneira de proporcionar melhoria seria a diminuição de uma parcela de recursos destinados à educação superior, redirecionando-a ao ensino infantil. Com mais recursos nesse segmento, é possível chegar ao objetivo em questão, por meio de novas políticas públicas.

Em uma visão global, de acordo com Hanushek (2019), a combinação entre o capital humano, o desempenho escolar e a qualidade – em conjunto com outros fatores, incluindo educação na família, experiência no mercado de trabalho, por exemplo – é determinante para o sucesso educacional de um país.

O pesquisador também salienta que países com mais capital humano têm a tendência de aumentar sua produtividade. Em sua opinião, o determinante mais importante para o crescimento econômico é o capital do conhecimento.

Hanushek (2019) afirma que as diferenças de desempenho encontradas entre estudantes com qualidade de ensino diferentes permite investigar qual política pública é, de fato, a mais eficaz na promoção da qualidade do ensino. Assim, explica que qualquer melhoria demanda tempo para ser realizada, e seus impactos demoram ainda mais tempo para serem percebidos. Por isso, é necessário prestar atenção na dinâmica da melhoria do capital humano e na trajetória temporal dos impactos econômicos.

Segundo Crato (2010), diversos estudos revelam que os alunos de baixo desempenho que ingressam no Ensino Básico mantêm as mesmas médias ao longo dos anos. Por essa razão, é necessária a implementação de políticas públicas inclusivas desde cedo. Os estudos internacionais comprovam que as dificuldades no ensino podem ser superadas por meio de um ensino sistemático, coerente e atento aos progressos dos alunos. Sob essa perspectiva, Hanushek (2019) salienta que o valor econômico da melhoria da qualidade educacional está na combinação entre resultados educacionais e crescimento econômico de longo prazo.

No sentido de contribuir para a qualidade do ensino público no Brasil, Menezes-Filho (2007) acredita que o tempo que o aluno permanece na escola tem impacto relevante no desempenho acadêmico, ao passo que o número de alunos por classe parece não ter influência sobre o desempenho do estudante em qualquer série.

Em contrapartida, Glass e Smith (1979) ponderam que a pesquisa acerca do tamanho da classe é algo particularmente complexo. Para encontrar resultados favoráveis, os autores afirmam ser necessária a implementação de uma análise sofisticada e ressaltam que análises ingênuas podem prejudicar a percepção sobre os resultados. Além disso, informam ter encontrado uma forte relação entre o tamanho da classe e o desempenho, sobretudo no que se refere às notas dos alunos do ensino secundário.

Na visão de Hoxby (1998), reduções de 15 a 30 alunos não tiveram variações no desempenho discente por ela analisado. Já para Hanushek (1998), as reduções de estudantes por sala de aula têm um apelo popular considerável, pois trata-se de uma política pública de fácil implantação. Contudo, as evidências experimentais oferecem pouco apoio sobre a real eficácia dessa medida.

Em conformidade com Hanushek (1998), Chingos (2010) igualmente percebe a existência de algumas falhas nas pesquisas, prejudicando o resultado. Entre elas,

mencionam-se as diferenças entre as salas de aula. A seu ver, estas últimas devem variar bastante para que os resultados sejam confiáveis.

Woesmann e West (2002), ao analisar o efeito do tamanho da classe em 18 países, encontraram nas estimativas convencionais, um alto índice de pesquisas tendenciosas, pois contavam com a colocação não aleatória de alunos em classes menores. Por meio de seu estudo, os autores apontam efeitos benéficos de classes menores na Grécia e na Islândia. Em contrapartida, nos demais países, não encontraram evidências de efeitos positivos dessa política pública.

Partindo do baixo grau de consolidação do tema na literatura, pretende-se, na presente pesquisa, avaliar o cenário brasileiro com relação ao impacto que a redução na quantidade de alunos por turma tem no desempenho escolar. É importante ressaltar a observação de variáveis regionais, orçamentárias e outras que serão identificadas e avaliadas, a fim de que não gerem resultados inconsistentes e distorcidos da realidade.

Assim, este trabalho visa a proporcionar resultados que comprovem que a diminuição do número de alunos por sala de aula é ou não uma política pública eficaz, com vistas a melhorar a qualidade do ensino brasileiro. Além disso, espera-se que tais resultados figurem como uma maneira de dirimir eventuais dúvidas sobre a aplicação dessa política pública, com dados quantitativos capazes de oferecer aos gestores ferramentas para novas políticas públicas.

Dado o contexto apresentado, emerge o seguinte problema de pesquisa: qual a relação entre o número de estudantes por sala de aula e seu desempenho escolar?

1.1 Objetivos Geral e Específicos

Os objetivos foram separados para a melhor distribuição das metas a serem alcançadas neste trabalho.

1.1.1 Geral

O objetivo geral desta pesquisa é verificar a relação entre o número de estudantes por sala de aula e seu desempenho escolar.

1.1.2 Específicos

- a) identificar o panorama de número de alunos por sala, tanto para os anos iniciais, quanto para os anos finais no Brasil;
- b) identificar o panorama de número de alunos por sala, por macrorregião do país;
- c) verificar a relação entre o número de estudantes por sala de aula e seu desempenho escolar, por nível de ensino e por macrorregião do país.

1.2 Abrangência do estudo

A amplitude do estudo será nacional, e o nível de ensino abordado será o fundamental, tanto nos anos iniciais quanto nos finais. Por meio do uso dos microdados do INEP referente à Prova Brasil para o ano de 2017, a abrangência tende a ser sistêmica, proporcionando resultados amplos.

1.3 Justificativa do estudo

A literatura internacional sobre o número de alunos por turma como forma de proporcionar a melhora do desempenho escolar é ampla, e seus resultados são os mais diversos. Enquanto estudos como o STAR, do estado do Tennessee, encontraram resultados positivos acerca dessa política pública, outros estudos, como o CSR, do estado da Califórnia, também nos EUA, não encontraram evidências significativas dessa relação. Ao avaliar a literatura nacional, encontram-se poucos trabalhos acerca do tema. Assim, esse cenário proporciona a oportunidade de explorar tal lacuna e contribuir de forma satisfatória para a promoção da qualidade da educação brasileira.

Para fornecer ferramentas ao gestor escolar ou ao formador de políticas públicas, o presente estudo se faz muito importante, pois seu resultado pode tanto contribuir para solucionar problemas de desempenho, como também para mostrar que essa política pública é ineficaz. Ambos os resultados são relevantes a fim de garantir a eficiência da aplicação dos recursos educacionais.

1.4 Organização da dissertação

Esta dissertação está organizada em seções, e cada uma delas contribui de forma única para a contextualização do estudo, sendo divididos da seguinte maneira:

Nesta primeira seção, que figura como introdução, apresentam-se os objetivos do trabalho, a abrangência do estudo, sua justificativa e a organização da dissertação.

Na seção 2, expõe-se o referencial teórico, no que concerne ao desempenho escolar e aos determinantes que proporcionam a qualidade educacional. Também se discute o conceito de avaliação de larga escala, com destaque para as principais avaliações que impactam o cenário brasileiro, o PISA e a Prova Brasil.

Como o foco do trabalho é verificar se o tamanho das turmas influencia no desempenho acadêmico, o referencial teórico conta com uma subseção dedicada às turmas, e outra dedicada às legislações que influenciam diretamente tanto no desempenho escolar, quanto nas turmas.

Ainda no referencial teórico, são elencados os mais importantes estudos já realizados no mundo com relação ao tema aqui debatido, divididos por continentes e pelos países que os compõem.

Na seção 3, expõe-se, detalhadamente, a metodologia de execução deste trabalho, demonstrando como os dados são abordados e também a forma com que as variáveis intervenientes são identificadas e controladas.

Por fim, na seção 4, apresentam-se os resultados encontrados na análise estatística. Por intermédio de tabelas, demonstra-se como cada resultado impacta decisões futuras sobre políticas públicas em educação.

Na sequência, encontram-se a conclusão e a apresentação do produto a ser gerado a partir desta pesquisa.

2 REFERENCIAL TEÓRICO

Busca-se, nesta seção, apresentar os elementos teóricos que deram sustentação ao desenvolvimento da presente pesquisa.

2.1 Determinantes de desempenho

Quando se fala de desempenho escolar no Brasil, a palavra qualidade é constantemente citada. Para Siqueira e Gurgel-Giannetti (2011), com a gratuidade da educação no Brasil – um direito conquistado na Constituição –, as crianças brasileiras, em sua maioria, tiveram o direito de estudar. Porém, as autoras afirmam que essa medida não garantiu a qualidade do processo educacional.

Para Jorge (2014), a coleta de dados sobre o desempenho escolar dos alunos, bem como as pesquisas sobre avaliação de desempenho tiveram início no país a partir de 1990, com a criação do Sistema de Avaliação da Educação Básica – SAEB. Nessa perspectiva, o autor expõe que a ênfase dada à eficiência do sistema educacional se deve principalmente às políticas públicas desenvolvidas nos últimos dez anos, com foco nos exames de avaliação de desempenho.

Segundo o pesquisador, os sistemas de avaliação de desempenho são medidores importantes da qualidade escolar. A seu ver, as avaliações introduzem fatores de mercado na escola e, por meio de *rankings*, é possível criar a responsabilização tanto positiva quanto negativa dos agentes do processo sobre os desempenhos alcançados, além de estimular os estados e municípios brasileiros a criarem seus próprios sistemas de verificação de desempenho. Isso facilita o reconhecimento dos bons profissionais e, dessa forma, estimula melhores desempenhos.

De acordo com Caprara (2013), o desempenho escolar tem sido observado de forma mais detalhada e demonstra a existência de inúmeros fatores que influenciam no desempenho. Destaca que compreender quais fatores de fato influenciam no desempenho escolar é uma tarefa complexa, afirmando que fatores imponderáveis podem definir se um estudante vai ou não ter um desempenho satisfatório.

Sob essa ótica, Andrade e Laros (2007) sinalizam a importância de identificar as estruturas educacionais que influenciam diretamente no desempenho do aluno, sejam elas suas famílias, escolas, bairros ou até grupos de amigos. Menezes-Filho (2007) destaca, além das características da família, a educação da mãe e o número de livros, a existência de um computador em casa, entre outros. No mesmo sentido, Nascimento (2007) ressalta que os fatores determinantes no desempenho escolar, discutidos pela literatura educacional, podem ser divididos em três grandes categorias: as características socioeconômicas, as habilidades anteriormente adquiridas pelos estudantes e os recursos destinados à educação.

Para Castro (2007), somente uma avaliação externa de larga escala é capaz de mensurar a qualidade do sistema escolar, e os resultados educacionais não são facilmente observáveis. A autora afirma que é ideal o cruzamento dos dados no nível escolar com os dados da avaliação de desempenho. Dessa forma, possibilita-se a verificação de quais fatores têm, de fato, influência sobre o desempenho escolar.

Em contrapartida, Manfredi, Schneck e Zuanetti (2008) demonstram que o bom desempenho escolar pode até favorecer o enfrentamento de problemas acadêmicos, mas as dificuldades encontradas no caminho podem piorar problemas tanto emocionais como educacionais.

Conforme assinalam Correa e Trojan (2015), o fato de o estudante ter um bom desempenho em um exame não garante a qualidade do sistema escolar. Contudo, os autores afirmam que o mau desempenho indica a existência de deficiências no processo de aprendizagem, que devem ser investigadas. Eles afirmam que as avaliações de desempenho devem servir para moldar políticas públicas, e não para forçar escolas, professores e alunos a aceitarem um modelo com a tendência de responsabilizar individualmente certas pessoas, desconsiderando a realidade de muitas instituições escolares brasileiras em termos de infraestrutura.

Cabe destacar a afirmação de Esteban (2009), ao afirmar que, nos últimos anos, as avaliações externas têm obtido grande destaque como forma de medir a qualidade da educação. A autora ressalta que a interpretação dos resultados dos testes de desempenho é, em sua maioria, técnica e despreza a complexidade da escola no dia a dia. Segundo a pesquisadora, os níveis de desempenho aferidos expõem o aluno com dificuldades acadêmicas e fazem com que tenha ainda mais

dificuldades. Ela ainda tece uma crítica aos padrões de desempenho utilizados pelas avaliações externas para determinar quais metas devem ser alcançadas. Na sua opinião, ao focalizar esses objetivos, não se debate um modelo de escola hegemônico e dinâmico, prejudicando principalmente os estudantes de nível socioeconômico mais baixo.

Corroborando o ponto de vista de Esteban (2009), Stein, Knijnik e Giacomoni (2013) ressaltam ser necessário o aprimoramento e o desenvolvimento de melhores mecanismos de verificação de desempenho para auxiliar na melhora da qualidade tanto dos alunos quanto dos professores.

De acordo com Felício e Fernandes (2005), no caso da hipótese de que realmente as avaliações de desempenho sejam importantes indicadores para a qualidade do sistema educacional, a tarefa principal dos pesquisadores, em geral, é identificar quais determinantes de desempenho são de fundamental importância na criação de novas políticas públicas.

O desempenho escolar, por si só, é difícil de ser mensurado, pois são muitos os determinantes que influenciam seu resultado. Portanto, é importante entender quais são os principais determinantes e como são avaliados. Eles são apresentados na próxima subseção, na qual também registramos as opiniões dos principais autores sobre o tema.

2.2 Determinantes da qualidade educacional

Partindo do princípio de que a educação é o elemento fundamental do desenvolvimento e da integração social, Menezes-Filho (2007) pontua que a literatura educacional sinaliza alguns benefícios do aumento da escolaridade. Entre eles, destacam-se o aumento dos salários das pessoas, a diminuição dos índices de criminalidade, a melhora da saúde, além do aumento da probabilidade de encontrar um emprego.

Segundo Oliveira (2008), os governos de diversos países têm, como um de seus objetivos, promover o acesso dos cidadãos à educação. Nesse sentido, Sousa (2018) sinaliza a existência de uma relação intrínseca entre os aspectos socioeconômicos e a qualidade educacional, tanto das famílias como dos alunos.

Em um de seus estudos, Menezes-Filho (2007) afirma que, por meio da econometria, verificou-se que algumas variáveis têm influência sobre o desempenho

acadêmico, quais sejam: os atributos tanto dos alunos como das famílias, o nível de escolaridade da mãe, a cor da pele, o trabalho fora de casa, o atraso escolar e a reprovação prévia, além do número de livros e a presença de computadores em casa. Sobre isso, Sousa (2018) assevera que o fator de maior influência no desempenho escolar é o nível de formação acadêmica dos pais. Além disso, o autor salienta que o trabalho durante a fase escolar é uma variável que propicia a redução do desempenho.

Ao analisar as pesquisas mais recentes sobre os determinantes da qualidade na educação, constata-se que os autores utilizam metodologias diferentes para verificar se algumas variáveis são mais influentes que outras no desempenho escolar. Nesse sentido, o intuito desta subseção é discorrer sobre os achados desses pesquisadores.

Para tanto, parte-se de uma apresentação em formato de funil. Inicialmente, alude-se aos determinantes de qualidade de forma geral. Posteriormente, focalizam-se as variáveis que cada pesquisador considerou mais influentes em seus estudos. Por último, apresentam-se os pesquisadores que tratam da questão do número de estudantes por turma – tema deste trabalho –, a fim de elucidar a importância desse determinante para qualidade da educação.

Este roteiro visa a posicionar o leitor no cenário existente, bem como a direcioná-lo ao objetivo principal do presente estudo, que pretende verificar, por meio da análise dos microdados do INEP referente à Prova Brasil de 2017, se o número de alunos por sala de aula tem impacto no desempenho acadêmico.

Gramani (2017) relata, por meio de pesquisa realizada no estado do Ceará que o setor da educação teve mudanças significativas desde o ano de 2007. As melhorias identificadas pela autora estão principalmente nos anos iniciais do Ensino Fundamental.

Pass (2016) constata que, no estado do Paraná, há disparidades sociais entre as regiões do estado, o que contribui para diferentes desempenhos no Ensino Fundamental.

Com base na análise de desempenho de diferentes municípios do interior paulista, Oliveira (2015) ressalta que as escolas com alto desempenho têm elementos comuns em relação à preocupação com as avaliações externas e a gestão escolar.

Por meio da metodologia de programação linear, Gramani (2017) avaliou os dados do IDEB para medir a eficiência relativa de unidades produtivas semelhantes. Pass (2016), por sua vez, optou por utilizar a pesquisa quantitativa de coleta e apreciação de dados, valendo-se também do IDEB como a principal fonte de dados.

Já Oliveira (2015) optou pela análise documental e pela análise de conteúdo, combinadas com o processo de observação sistemática. Além disso, utilizou entrevistas semiestruturadas com os diferentes integrantes das escolas pesquisadas.

Isso posto, nota-se que os autores lançaram mão de metodologias diferentes a fim de avaliar o desempenho acadêmico. Pass (2016) constatou que o investimento de recursos na educação obteve baixa significância no desempenho escolar e concluiu que isso possivelmente aconteceu por ineficiência dos gastos pelos gestores educacionais. Ademais, destacou que a baixa significância do investimento em educação relacionada ao desempenho pode estar ligada ao fato de os recursos terem sido destinados a aumentar o desempenho em regiões com resultados mais baixos. Dessa forma, deixar de alocar recursos em outras regiões culminou no leve aumento no índice verificado.

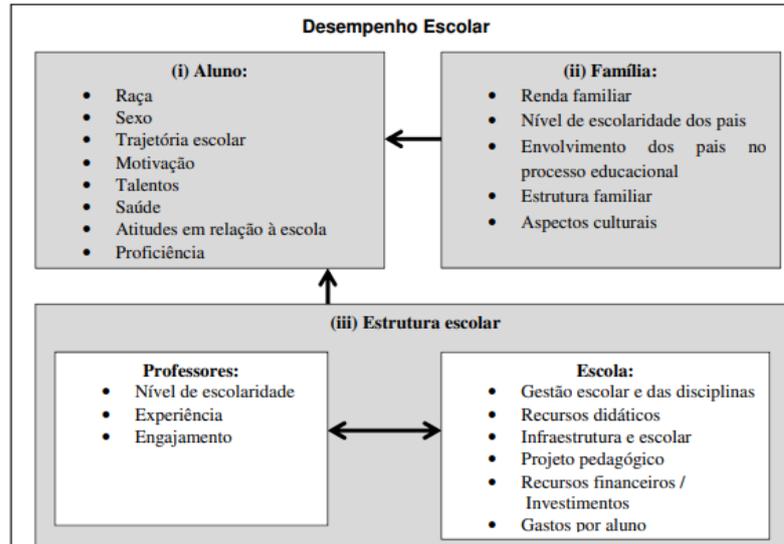
Oliveira (2015) afirmou que as escolas do interior de São Paulo apresentaram diferenças significativas no desempenho escolar, ao analisar a adesão ou não dos gestores escolares à política de avaliação externa. Os gestores com maior adesão tiveram resultados melhores em suas escolas, comparados com os que não aderiram.

Segundo Gramani (2017), no estado do Ceará, os municípios com melhores desempenhos escolares têm as seguintes variáveis em comum: gasto em educação, nível de escolaridade das mães, além do fator renda como determinantes de eficiência.

Observa-se que os autores buscaram mostrar quais são os determinantes da qualidade do ensino e do desempenho escolar. Para tanto, cada um deles utilizou-se de metodologias diferentes. Os resultados encontrados variam de um pesquisador para o outro, indicando que cada região têm sua particularidade.

A fim de ilustrar com maior clareza como o desempenho escolar é influenciado pelas variáveis determinantes, Oliveira (2016) adaptou o modelo apresentado por Soares (2005), conforme a figura 1, reproduzida a seguir:

Figura 1: Desempenho Escolar



Fonte: Oliveira (2016), com base em Soares (2005).

Outros autores também fizeram estudos sobre desempenho escolar, tais como Moraes e Becker (2015), Soares e De Farias (2018) e Dellazem e Coelho (2019). Por meio de suas pesquisas, os pesquisadores verificaram qual a influência do professor no desempenho escolar, assunto bastante debatido no ambiente educacional.

Com o intuito de obter melhores resultados em pesquisas neste sentido, Dellazem e Coelho (2019) afirmaram ser necessária a utilização de metodologias de fundamentação bibliográfica, de fundamentação documental e de pesquisa de campo. Desse modo, o estudo realizado torna-se mais completo, pois será composto tanto por análises qualitativas quanto quantitativas.

Soares e Farias (2018) utilizaram entrevistas semiestruturadas com os professores para identificar semelhanças discursivas que possibilitassem medir a qualidade, sob a ótica docente. Moraes e Becker (2015) optaram pela utilização de variáveis explicativas e variáveis determinantes para determinar proficiência dos alunos na Prova Brasil.

Conforme analisado por Soares e Farias (2018), os docentes entrevistados, em sua maioria, afirmaram que a indisciplina é consequência da relação do aluno em sua conjuntura familiar e, seguindo essa mesma linha, Dellazem e Coelho (2019) descobriram que as escolas com baixo desempenho acadêmico apresentaram 100% do índice de indisciplina.

Sob outra perspectiva, Soares e Farias (2018) citam o exemplo dos docentes que, em sua maioria, consideram a afetividade como principal ferramenta de aproximação com os alunos para o aumento do desempenho.

Ainda com relação à variável influência do docente no desempenho, Moraes e Becker (2015) encontraram alguns resultados intrigantes. Por exemplo, estudantes com professores do sexo masculino apresentaram desempenho menor em Português e Matemática.

Ao focar especificamente no modo de ensinar, Dellazem e Coelho (2019) consideram que, enquanto o docente garante apresentar seus conteúdos de forma significativa e dinâmica, as entrevistas dos pais e alunos seguiram o caminho contrário, como se não houvesse indicativo algum do afirmado pelos docentes.

A complexidade da atuação docente no desempenho acadêmico é evidente. Pode-se observar que as pesquisas citadas abordaram aspectos diferentes sobre o mesmo tema, demonstrando o quão difícil é dar uma resposta para esse tipo de variável.

Outra variável que também chama atenção nos determinantes da qualidade é a influência dos pais no aumento do desempenho escolar. Os autores Lima e Santos (2016), França e Gonçalves (2008) e Pozzobon, Falcke e Marin (2018) pesquisaram sobre o tema e contribuíram para a presente análise.

Segundo França e Gonçalves (2008), os pais são a principal referência para os filhos em termos de conduta na escola. Desse modo, os pais com capital humano elevado têm filhos com nível educacional semelhante ou superior. Lima e Santos (2016) atestam que a ausência dos pais no desenvolvimento educacional dos filhos resulta em impactos negativos no desempenho escolar.

Para a busca de respostas, Lima e Santos (2016) utilizaram os microdados da Prova Brasil referentes aos anos de 2007, 2009, 2011 e 2013. França e Gonçalves (2008) optaram pelo uso dos dados do SAEB referentes a 2003, ao passo que Pozzobon, Falcke e Marin (2018) valeram-se da metodologia de abordagem mista com delineamento pré-experimental, com coleta de dados via questionário. Essa

abordagem sinalizou que as conquistas acadêmicas podem estar relacionadas às características específicas das famílias.

Com base na abordagem multinível, França e Gonçalves (2008) verificaram que o efeito da variável família no desempenho escolar obteve resultados diferentes em cada município avaliado. Sob outra ótica, Lima e Santos (2016), por meio da metodologia de análise transversal, verificaram que o impacto negativo no desempenho deriva da vivência do aluno com alguma mudança na estrutura familiar, e não na influência direta do pai sobre o filho.

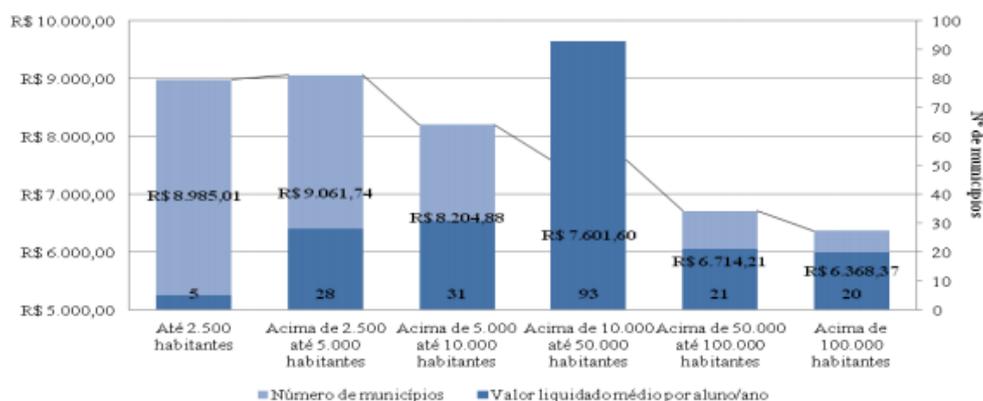
A maioria das políticas públicas que visam a melhorar a qualidade na educação necessita de recursos financeiros. Alguns autores como Oliveira (2016) e Panassol (2018) elaboraram estudos para verificar se o gasto educacional é um determinante da qualidade educacional.

Oliveira (2016), por exemplo, optou pelo uso dos microdados da Prova Brasil, por meio da metodologia de regressão quantílica. A autora usou esse modelo econométrico a fim de mensurar o efeito das variáveis explicativas. Ainda nessa linha, Panassol (2018), além de usar a regressão quantílica, adicionou a ela o método de mínimos quadrados ordinários.

Visando à aplicação de recursos públicos e observando os princípios da transparência e da eficiência, Oliveira (2016) avaliou as notas médias da Prova Brasil nos municípios goianos, nas disciplinas de Português e Matemática, e percebeu que houve uma melhora no desempenho. A autora justifica que o resultado encontrado deve-se ao uso do modelo quantílico estimado, pois foi por meio dele que se perceberam os impactos positivos e significativos nas notas médias.

Panassol (2018), após estimar cinco modelos por mínimos quadrados ordinários, concluiu que o desempenho acadêmico só é influenciado pelos gastos educacionais quando há a combinação de variáveis explicativas com o uso desses recursos. O gráfico a seguir demonstra o que o autor encontrou a partir dos dados colhidos em sua pesquisa sobre o valor médio por aluno ao ano, nos municípios do estado do Rio Grande do Sul:

Figura 2: Valor médio por aluno ao ano



Fonte: Panassol (2018).

Hoxby (1998) afirma que políticas que oferecem apenas recursos financeiros para a melhoria da educação tendem a ser menos eficazes do aquelas que atrelam recursos ao desempenho. Para Oliveira (2016), nos municípios de Goiás, apenas a combinação de gasto em educação com variáveis como infraestrutura, despesas correntes e melhoria na remuneração dos professores é suficiente para mostrar o aumento de desempenho.

Segundo Oliveira (2008), os efeitos dos recursos escolares é um tema controverso. Enquanto alguns pesquisadores concluem que não existe relação entre aumento dos gastos em educação com desempenho escolar, outros afirmam que essa relação existe e é relevante o suficiente para ser considerada. Para Camargo (2014), existe grande controvérsia na literatura sobre quais variáveis, de fato, influenciam na qualidade do ensino.

Considerando o objetivo deste trabalho, apresentam-se, na sequência, alguns pesquisadores que tratam da temática relacionada à influência do número de alunos por classe no desempenho acadêmico.

Segundo Hanushek (1998), políticas públicas para reduzir o tamanho da classe apresentam um apelo popular considerável. Chingos (2010) afirma que há muitos argumentos contrários à eficácia dessa medida.

Ainda para Hanushek (1998), nenhum assunto sobre educação recebe tanta atenção pública como o interesse em reduzir o número de alunos por turma. Hoxby(1998), por sua vez, expõe que as flutuações no número de estudantes em cada classe têm alta relevância para os debates políticos atuais.

Em contrapartida, Woesmann e West (2002) afirmam que, apesar de os EUA terem muitas evidências sobre a quantidade de alunos por turma, não há razões para generalizar as respostas encontradas, já que cada país tem um sistema escolar distinto.

Para Glass e Smith (1979), é nítido que as questões sobre a quantidade de alunos por sala são complexas. Elas preocupam os professores que querem menos alunos por turma, bem como os administradores que querem mais alunos para aumentar o lucro.

À luz da questão brasileira sobre a redução de alunos por classe, Matavelli (2018) afirma que o assunto tem grande relevância no debate sobre políticas públicas educacionais. Entretanto, Oliveira (2008) deixa claro que não há consenso sobre a eficácia das políticas de redução de turma na literatura mundial.

Nesse sentido, Hanushek (1998) afirma que o fato mais surpreendente sobre os levantamentos dos estudos acerca do número de alunos por classe é que, apesar da enorme quantidade de pesquisas sobre o tema, nada muito convincente foi encontrado acerca da influência dessa redução no desempenho dos estudantes.

Segundo o autor, o grande interesse em relação às políticas de redução do tamanho da turma ocorre por se tratar de um instrumento político conveniente, ou seja, é algo que pode ser facilmente feito. Nessa mesma linha, Woesmann e West (2002) observam que as variações do tamanho de classe podem ser as mais acessíveis em restrições orçamentárias.

Para Hanushek (1998), o caminho para entender o interesse político em torno das expectativas sobre desempenho está justamente na visão de que a educação é um processo cumulativo, baseado na realização do passado. Em sua opinião, é importante e necessário avaliar com atenção se essa medida não impacta fortemente no orçamento. O autor demonstra, a partir de sua análise, que os gastos reais por aluno nos EUA, no período de 1890 a 1990, tiveram uma média de crescimento de 3,5% ao ano.

Por meio de sua análise, Oliveira (2008) conclui que os custos da redução do número de alunos por classe envolvem professores e salas de aulas adicionais, além de aumentar o custo por estudante.

Já Woessmann e West (2002) salientam que os benefícios do controle do número de alunos por classe superam marginalmente os custos dessa medida, demonstrando, assim, o porquê das dúvidas sobre os custos de tal política pública.

Para Camargo (2014), o tamanho da classe está no centro do debate sobre quais variáveis influenciam na qualidade da educação, pois, segundo a autora, essa variável é de fácil aplicação. Hoxby (1998) afirma que, quando os formuladores de políticas públicas querem melhorar o desempenho dos estudantes, a redução da classe é uma das ferramentas mais propensas de serem usadas.

Hanushek (1998) levanta uma questão importante sobre o custo de implantação de classes menores, ao destacar que esses custos normalmente são altos. Caso seja uma política ineficaz, o pesquisador alerta que os custos de projetar, implementar e avaliar um projeto sem eficiência faz com que uma série de políticas alternativas deixem de ser realizadas.

De acordo com Chingos (2010), uma pesquisa eficaz seria direcionar recursos para a redução do tamanho da classe em alguns sistemas e direcionar a mesma quantidade de recursos a outros, a fim de que esses distritos gastem como quiserem. Segundo o autor, esse tipo de estudo elucidaria as dúvidas que assolam a literatura educacional acerca da redução de classe. Em contrapartida, Travitzki e Cássio (2017) textualizam que uma política pública efetiva voltada para a melhoria do desempenho acadêmico necessita superar a visão da sociedade como um bloco homogêneo.

Sob essa perspectiva, Chingos (2010) pondera que, apesar de uma vasta pesquisa que dá suporte ao efeito da redução do tamanho da classe no desempenho escolar, argumentos sobre a eficácia dessa medida persistem. Hanushek (1998) alega que, para alterar a situação e os resultados educacionais, basta que todos tivessem incentivos para usar os orçamentos de forma estratégica e, assim, melhorar o desempenho dos alunos.

Para Hoxby (1998), uma das principais razões pela qual o controle do número de alunos por classe é considerado controverso seria a evidência empírica contraditória. A autora argumenta que a controvérsia empírica é essencialmente simples. Os estudos usam diferentes fontes de variação no tamanho e composição de classe, e quase todas essas fontes são contaminadas pela correlação com outros determinantes de desempenho escolar.

Nesse sentido, Travitzki e Cássio (2017) afirmam que o único consenso existente na literatura sobre o tamanho das classes é a dificuldade de estudar essa variável. Os autores afirmam que isolar uma única variável para medir o desempenho é uma tarefa árdua. Para Oliveira (2008), ao avaliar as políticas de redução do tamanho da classe, faz-se necessário compreender as ferramentas pelas quais essa variável pode afetar a proficiência.

Segundo Hanushek (1998), há poucas razões para acreditar que turmas menores têm impacto no desempenho. Da mesma forma, Hoxby (1998) afirma, a partir de sua pesquisa, que não há evidências da eficácia da redução de alunos por classe nas séries iniciais, principalmente em escolas que contenham concentração alta de alunos de baixa renda ou em turmas inicialmente maiores. Vale destacar também que Camargo (2014), após avaliar os microdados da Prova Brasil referente ao Rio Grande do Sul, não encontrou nenhuma evidência significativa de que o tamanho da classe tem influência em seu desempenho.

Em contrapartida, Oliveira (2008) encontrou, em seu estudo sobre o Ensino Fundamental brasileiro, que, em classes com mais de 33 alunos e menos de 41 alunos, a diminuição dos estudantes por turma se mostra como uma política pública eficaz em termos de desempenho escolar.

Do mesmo modo, Chingos (2010) afirma que, desde 1978 ou antes, têm sido encontrados resultados positivos em pequenas classes, principalmente nas primeiras séries do jardim de infância e, segundo Hanushek (1998), as classes pequenas, caso ocorram no início do processo de escolaridade, têm um efeito único no desempenho discente.

Para Hoxby (1998), as reduções no tamanho da turma podem ser mais eficazes nas escolas nas quais as crianças são de famílias menos favorecidas, mesmo ainda sem evidências para isso. Sobre isso, Hanushek (1998) considera que é provável haver situações com professores específicos, grupos específicos de alunos e assuntos específicos, em que pequenas turmas podem ser muito benéficas para o desempenho dos estudantes.

Com o objetivo de encontrar resultados confiáveis sobre o impacto do número de alunos por classe, alguns pesquisadores defendem métodos diferentes para o alcance dessa resposta. Segundo Woessmann e West (2002), o método padrão – ou aquele mais utilizado para avaliar a relação entre o número de alunos por turma e o

desempenho escolar – é o de regressão de mínimos quadrados. Para Hoxby (1998), o emprego da metodologia de variação natural tem a vantagem de os participantes não saberem que estão sendo avaliados. Dessa forma, é possível imitar uma política pública real.

Outras metodologias para a busca dessa relação entre número de alunos por classe e desempenho escolar também foram utilizadas. É o caso de Camargo (2014), que utilizou o método de variáveis instrumentais juntamente com o desenho de regressão com descontinuidade. Já Travitzki e Cássio (2017) optaram pelo método do pareamento com escore de propensão generalizado.

A metodologia de *Propensity Score Matching* com estimador *Nearest Neighbor* foi utilizada por Oliveira (2008) para verificar a influência da quantidade de alunos por sala no desempenho acadêmico. No entanto, Matavelli (2018) optou por utilizar o método de desenhos de regressão descontínua por achar ser mais transparente.

Para possibilitar a melhor interpretação sobre os resultados, Woessmann e West (2002), afirmam ser necessário notar que há muitos aspectos de nível e de qualidade de recursos que influenciam no desempenho escolar, entre os quais o tamanho da classe é apenas um aspecto. Nesse sentido, Oliveira (2008) diz que é necessário identificar a relação do desempenho escolar com os efeitos dos insumos escolares, o que economizaria muitos recursos, principalmente em períodos de austeridade.

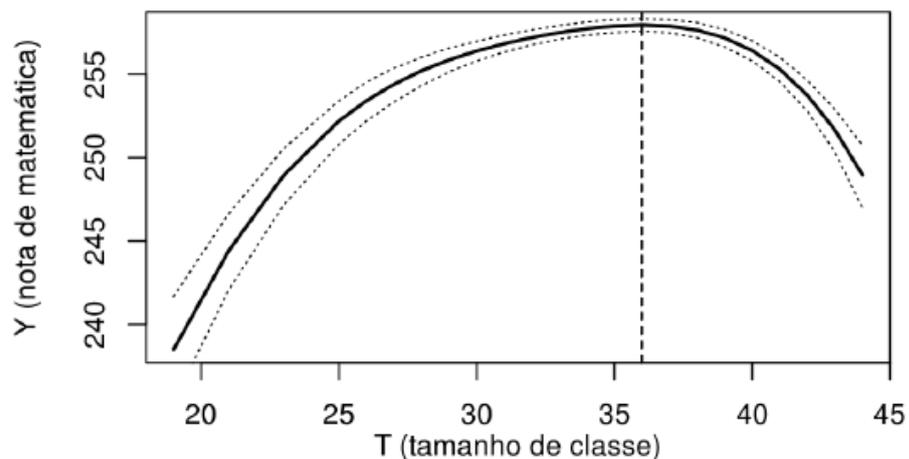
Ao focar a situação brasileira, a autora afirma que a média da quantidade de alunos por turma vem diminuindo, principalmente no Ensino Fundamental, nas séries de primeira à quarta em escolas municipais de ensino. A partir de seus resultados, ela descobriu que as escolas da rede municipal de ensino do Norte e Nordeste do Brasil são as que apresentam a maior média de alunos por turma entre os anos de 1997 e 2006. A pesquisadora ainda destaca que, apesar da diminuição do número de estudantes por turma no período de 1997 a 2006, segundo dados colhidos na OCDE, o Brasil ainda apresenta uma das maiores razões aluno-professor. Estudos brasileiros sobre o tema encontraram números ideais para sala de aula, e a quantidade nos anos iniciais difere da quantidade em anos finais.

Para Travitzki e Cássio (2017), a partir de seu estudo sobre a influência do tamanho das turmas na disciplina de Matemática, dos anos finais do ensino fundamental no estado de São Paulo, o tamanho ideal de classe encontrado seria o

de 34 alunos. Em contrapartida, Oliveira (2008), quando estudou os anos iniciais do Ensino Fundamental brasileiro chegou ao tamanho ideal de classe de 30 alunos.

Segundo Travitzki e Cássio (2017), as políticas educacionais paulistas têm fundamentação consistente acerca do tamanho ideal da classe quando avaliada pela média.

Figura 3: Influência do Tamanho da classe na nota de Matemática



Fonte: Travitzki e Cássio (2017).

O uso de médias, no entanto, é criticado por Hanushek (1998). O autor descobriu que, ao utilizar a estratégia de avaliar a proporção aluno por professor para descobrir o tamanho das salas, às vezes, cometem-se erros, pois esse número se refere à quantidade de alunos dividida pelo número de professores e não diz respeito ao número de estudantes por sala. Em concordância, Chingos (2010) afirma que o uso de médias nos estudos pode muito bem ofuscar a eficácia da política pública de redução de alunos por turma.

Para Woessmann e West (2002), o principal problema da avaliação dos impactos do tamanho da classe no desempenho escolar é justamente o fato de os pais colocarem seus filhos em escolas com classes maiores ou menores de acordo com o seu desempenho. Isso obscurece a relação de causa e efeito entre desempenho e quantidade de estudantes por sala. Nesse mesmo sentido, Hoxby (1998) considera que as reduções de tamanho de classe parecerão mais eficazes do

que realmente são, se os pais que se preocupam mais com a educação dos filhos escolherem escolas que oferecem classes menores.

Woesmann e West (2002) afirmam também que as escolas podem colocar alunos em salas menores ou maiores, de acordo com seu comportamento ou seu desempenho escolar. Além disso, os autores pontuam que uma análise ingênua pode comprometer a eficácia de pesquisas sobre o tema. Os pesquisadores também alertam que muitas pesquisas podem sofrer com o “Efeito *Hawthorne*”, pois os participantes sabem que estão sendo estudados e, por isso, aumentam seu esforço, o que em condições normais não aconteceria.

Ao aplicarem uma estratégia de identificação que explica os efeitos de classificação, os estudiosos demonstraram que os efeitos de tamanho de classe podem ser identificados, evitando-se resultados ingênuos. Eles afirmam que o fato de alguns países terem resultados positivos e outros negativos com relação ao controle da quantidade de alunos por sala ocorre de acordo com seu nível de desenvolvimento. No entanto, alertam que não há um padrão claro entre o Produto Interno Bruto – PIB – e os benefícios desse controle.

Travitzki e Cássio (2017) afirmam que a maioria dos estudos sobre o controle do tamanho da classe se restringe a ser apenas observacional. Segundo os autores, isso acontece pela dificuldade de realizar experimentos sociais randômicos e com controle de variáveis de forma adequada. Para Woesmann e West (2002), é o potencial do efeito do tamanho da classe que, ao final, valerá seus custos.

De acordo com Oliveira (2008), uma educação de melhor qualidade é importante tanto no nível individual quanto no nível geral, sendo um benefício fundamental para o crescimento econômico da sociedade. Sobre isso, Woesmann e West (2002) afirmam que os efeitos de tamanho de classe encontrados nos estudos não devem ser interpretados como resposta geral para todos os sistemas escolares.

Para ilustrar a situação brasileira sobre a quantidade de alunos por turma, os dados dos indicadores educacionais do INEP mostram, por meio de médias, um panorama da situação em 2020.

Tabela 1: Média do número de alunos por turma em 2020 no Brasil

Média do número de alunos por turma – 2020				
Unidade e Geográfica	Localização	Dependência Administrativa	Ensino Fundamental - Anos Iniciais	Ensino Fundamental - Anos Finais
Brasil	Urbana	Público	25,3	28,8
		Privado	18,3	23,2

Fonte: Elaboração própria.

Observa-se que, em questão de média, a iniciativa privada apresenta quantidades menores de alunos por sala, quando comparada à iniciativa pública. Um maior detalhamento permite visualizar que cada região do país tem sua particularidade.

Tabela 2: Média do número de alunos por turma em 2020 por região

Média do número de alunos por turma – 2020				
Unidade Geográfica	Localização	Dependência Administrativa	Ensino Fundamental - Anos Iniciais	Ensino Fundamental - Anos Finais
Norte	Urbana	Público	26,9	30,4
		Privado	18	24,1
Nordeste	Urbana	Público	25,3	29,3
		Privado	17,2	22,6
Sudeste	Urbana	Público	25,8	29,6
		Privado	18,7	22,9
Sul	Urbana	Público	22,9	25,2
		Privado	20,4	25,1
Centro-Oeste	Urbana	Público	25,7	29,4
		Privado	18,3	24,4

Fonte: Elaboração própria.

Na região Sul, nos anos finais do Ensino Fundamental, é possível perceber que as médias de alunos por turma, tanto da iniciativa pública quanto da iniciativa privada, são aproximadamente as mesmas. Esta situação pode fornecer um importante comparativo de desempenho.

A fim de visualizar de forma mais clara os impactos do número de alunos por classe no desempenho escolar, faz-se necessário mostrar como esse desempenho pode ser avaliado. Há diversas formas para isso, e uma delas é avaliação de larga escala. Na próxima subseção, discute-se como essas avaliações funcionam.

2.3 Avaliações de Larga Escala

Segundo Menegão (2016), o objetivo da avaliação de larga escala, de modo geral, é medir a qualidade do sistema educacional. A autora ressalta que, para isso ocorrer, é preciso definir, inicialmente, qual o conceito de qualidade se busca alcançar e, posteriormente, estabelecer quais serão as disciplinas avaliadas.

Para Vianna (2001), não é por improvisado que se cria uma avaliação de larga escala. É necessária toda uma estrutura de funcionamento, monitoramento e gerenciamento. De acordo com o autor, qualquer avaliação precisa de instrumentos anteriormente testados, e isso normalmente exige uma amostra de grande representatividade.

Klein e Fontanive (1995) salientam que uma avaliação de larga escala é composta de um conjunto de informações que objetivam fornecer um diagnóstico da educação de um país. Os autores afirmam que, para ela ser concebida, é necessário um monitoramento constante, a fim de identificar os aspectos positivos e negativos que interferem no processo.

De acordo com Castro, Lopes e Silva (2016), a avaliação em larga escala é uma ferramenta para gerar melhorias na qualidade educacional e, ao possibilitar o recolhimento de informações, é possível ter base para criação de novas políticas públicas.

Na opinião de Hypolito e Jorge (2020), essas avaliações são normalmente criações de governos e entidades tanto nacionais quanto internacionais do campo educacional. Sua finalidade é identificar os ganhos e os resultados acadêmicos e, a partir disso, formular novas políticas educacionais.

Para Cária e Oliveira (2015), os sistemas de avaliação em larga escala estão sendo amplamente discutidos nos últimos dez anos, especialmente, com relação à formulação de políticas públicas em educação. As pesquisadoras afirmam que, a partir de 1990, como resultado da regulamentação do estado brasileiro como condutor de políticas públicas educacionais, se iniciou uma nova perspectiva de

busca pela melhoria do sistema educacional. Segundo as autoras, foi com base em modelos internacionais que se montou um modelo capaz de verificar os resultados a partir de critérios de eficiência e eficácia.

Nesse sentido, Silva, Gimenes e Lozano (2013) afirmam que, por meio da interlocução entre os entes federativos na educação básica, é notável o aumento da importância de sistemas de avaliações de desempenho no Brasil. Contudo, ressaltam que a utilização dos resultados como meio de mudanças no sistema educacional ainda é um assunto a ser mais explorado.

Cária e Oliveira (2015) mostram que, a partir da criação do Sistema Nacional de Avaliação Básica (SAEB), começou a concretização das avaliações de larga escala no Brasil. Para as autoras, foi com o SAEB que a Prova Brasil ganhou um papel de destaque nos últimos anos. Por esse ângulo, Schneider (2013) ressalta que o SAEB e a Prova Brasil foram criados com base no modelo de avaliação desenvolvido pelo PISA.

De acordo com Menezes (2014), atualmente, muitos sistemas escolares brasileiros centralizaram seu foco nas avaliações de larga escala para definir suas políticas públicas em educação.

Sobre isso, Gatti (2012) ressalta que o foco principal em avaliações como forma de fazer política está justamente em âmbito federal, indicando um aspecto produtivista da educação ligado a pressões de financiadores econômicos internacionais. No mesmo sentido, Camba e Tavares (2017) afirmam que essa mudança na forma de medir a qualidade do ensino no Brasil decorreu de uma reforma educacional, consequência de pressões internacionais sobre contrapartidas para disponibilização de créditos na economia.

Antunes (2018) demonstra que a avaliação de larga escala brasileira está em expansão e vem se fortalecendo nos últimos anos. Em contrapartida, a autora constata a existência de contradições sobre os entendimentos teóricos relacionados ao tema.

Sob esta perspectiva, Schneider (2013) expõe que o fato de a avaliação em escala ser formada tanto de benefícios como de contradições faz com que seja necessário conhecer suas possibilidades e seus limites.

Lindoso (2019) enfatiza que a avaliação externa, por estar em foco tanto em políticas nacionais como em políticas internacionais, mesmo apresentando um

caráter complexo e contraditório, é uma ferramenta importante para a transformação educacional.

De acordo com Cária e Oliveira (2015), há certos pesquisadores que criticam o processo de avaliação em larga escala por considerá-lo alheio ao ambiente escolar, uma vez que é realizado por pessoas de fora da escola.

Menegão (2016) evidencia que as avaliações de larga escala brasileiras contribuem para manter os professores vinculados a um modelo complexo de responsabilização e meritocracia, que prejudica a prática pedagógica. Para a autora, se, antigamente, o bom professor era aquele que reprovava a maioria de seus alunos, atualmente, o bom educador é aquele que conquista bons resultados em avaliações de larga escala. Segundo a pesquisadora, isso explica a atual dificuldade de os docentes adotarem ações mais reflexivas.

Cária e Oliveira (2015) destacam, ainda, que a redução de currículos para focar nessas avaliações tem submetido os estudantes a incessantes práticas de classificação escolar, por meio de *rankings* de desempenho que geram diversos problemas psicológicos nos alunos.

Segundo Rocha e Ferreira (2017), a avaliação em larga escala é sempre um tema de grande discussão, sendo reconhecida na literatura a dificuldade de fazer com que esse processo traga, de fato, melhores benefícios na melhoria da qualidade educacional.

Para Cária e Oliveira (2015), suas características positivas conduziram, nos últimos anos, as avaliações em larga escala à importância que hoje elas têm sobre as políticas educacionais.

É justo dizer, de acordo com Becker (2010), que as avaliações necessitam de aprimoramentos para que seus resultados possam contribuir de forma mais efetiva para qualidade educacional. A autora argumenta que é necessário construir, no Brasil, novos processos para que gestores e professores utilizem, de forma mais efetiva, os dados obtidos.

Sob essa ótica, Santos e Carbonera (2010) destacam que as avaliações em larga escala são capazes de proporcionar tanto reflexões positivas, como novas ferramentas de controle da eficiência do sistema. Becker (2010) evidencia que a avaliação não é o fim de um processo, e sim uma ferramenta a ser utilizada para corrigir os rumos e planejar o futuro.

Ao focar no âmbito global, a avaliação de larga escala que mais se destaca entre os países é o *Programme for International Student Assessment* (PISA), ao passo que, no âmbito local ou brasileiro, tem-se a Prova Brasil como principal instrumento.

Para demonstrar como essas avaliações ocorrem, as próximas duas subseções tratam das características do PISA e da Prova Brasil, juntamente com o ponto de vista dos autores, acerca do cenário educacional sobre elas.

2.3.1 PISA - *Programme for International Student Assessment*

Segundo dados colhidos na Organização para Cooperação e Desenvolvimento Econômico (OCDE), até o ano de 1990, as comparações de desempenho dos alunos eram baseadas nos anos de escolaridade. A OCDE considerava tal metodologia falha, pois esse indicador não demonstrava realmente a capacidade dos estudantes.

Foi por meio da criação do *Programme for International Student Assessment* (PISA), que a OCDE tentou mudar essa realidade. A ideia da organização foi criar uma avaliação que pudesse testar as habilidades dos estudantes mediante métricas combinadas internacionalmente.

Conforme explica a OCDE, a missão do PISA é ajudar tanto as escolas quanto os formuladores de políticas públicas a verificar quais fatores podem contribuir para o aumento do desempenho dos alunos de seus países. Para que tal medida ocorra, comparam-se sistemas escolares de diversos lugares do mundo.

Segundo Villar e Zoido (2016), o PISA foi iniciado a partir do ano 2000 e repetido a cada dois anos. Os autores demonstram que a avaliação é capaz de fornecer o mais completo agrupamento de dados sobre as competências e as habilidades dos alunos de 15 anos de idade.

Para os pesquisadores, o PISA é capaz de proporcionar a obtenção de dados que vão desde o histórico pessoal dos alunos até seus hábitos de aprendizagem, bem como as características de sua escola e de sua família. Villar e Zoido (2016) ainda expõem que a avaliação tem por objetivo medir o desempenho dos estudantes nas disciplinas de leitura, Matemática e Ciências. Ademais, afirmam que o PISA oferece ferramentas que podem gerar a melhoria da qualidade educacional de cada país

participante. A avaliação, segundo os autores, traz dados sobre os alunos, seus desempenhos, sua escola, sua família, seus hábitos de aprendizagem, suas atitudes, seu engajamento e sua motivação. Por meio desses dados, é possível correlacionar o desempenho estudantil com as variáveis que influenciam nos resultados. A seu ver, a avaliação estabelece seis níveis na escala de proficiência dos participantes, e esses níveis são parametrizados pelas pontuações que os alunos atingem em cada disciplina avaliada.

Para OCDE, medir o desempenho dos estudantes com 15 anos de idade é importante, já que se trata de uma idade chave na transição entre o desenvolvimento físico e emocional dos jovens. Para falar sobre o desempenho do Brasil no PISA, é necessário contextualizar a situação em que se encontra a América Latina, pois somente assim é possível verificar a amplitude dos resultados brasileiros.

Segundo Villar e Zoido (2016), ao analisar os resultados do PISA para América Latina, percebe-se a necessidade de esses países melhorarem o desempenho dos discentes socioeconomicamente mais vulneráveis. De acordo com os autores, os estudantes latinos mais desfavorecidos têm maiores chances de apresentarem desempenhos baixos.

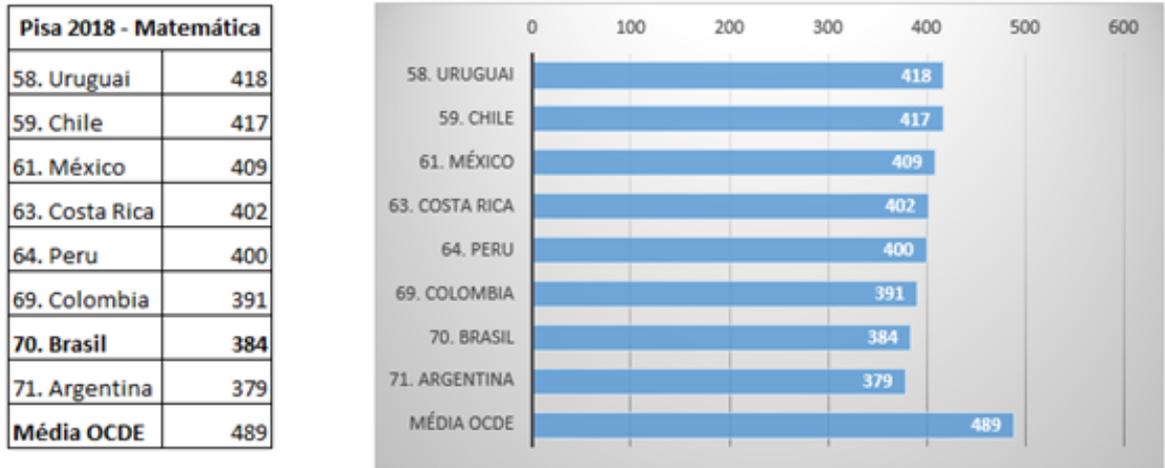
Sobre isso, Monteiro (2015) demonstra que o Brasil experimentou ganhos consideráveis em escolaridade da sua população nos últimos anos. Segundo a pesquisadora, dados recentes indicam que o país tem 93% de suas crianças com idades entre 6 e 14 anos matriculadas no ensino fundamental.

Em contrapartida Villar e Zoido (2016) textualizam que, embora os países da América Latina – incluindo o Brasil – tenham feito progressos significativos em acesso à escolaridade, a qualidade da educação ainda carece de melhorias. Na visão dos pesquisadores, os sistemas educacionais desses países precisam ter como foco o ganho de capital próprio e a melhoria da qualidade. Monteiro (2015) corrobora essa visão, assinalando que muitos problemas persistem, principalmente com relação à qualidade do ensino brasileiro.

Segundo dados colhidos na OCDE, referentes ao ano de 2018, é possível observar, nas seguintes figuras, o desempenho do Brasil em comparação com outros países da América Latina:

Figura 4: Desempenho Matemática PISA – 2018 – América Latina

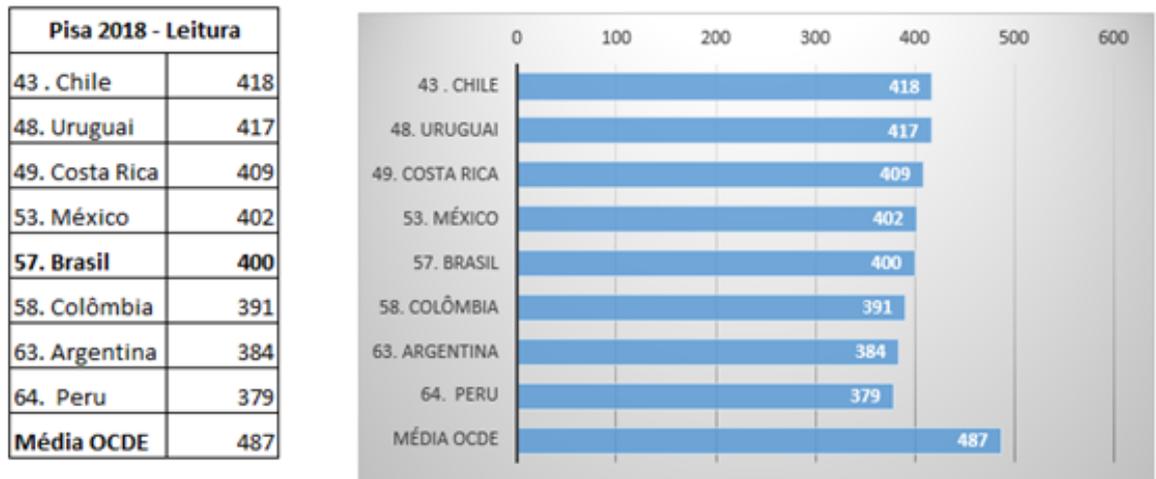
Desempenho Matemática Pisa - 2018 – Latin América



Fonte: Elaboração própria.

Figura 5: Desempenho Leitura Pisa – 2018 – América Latina

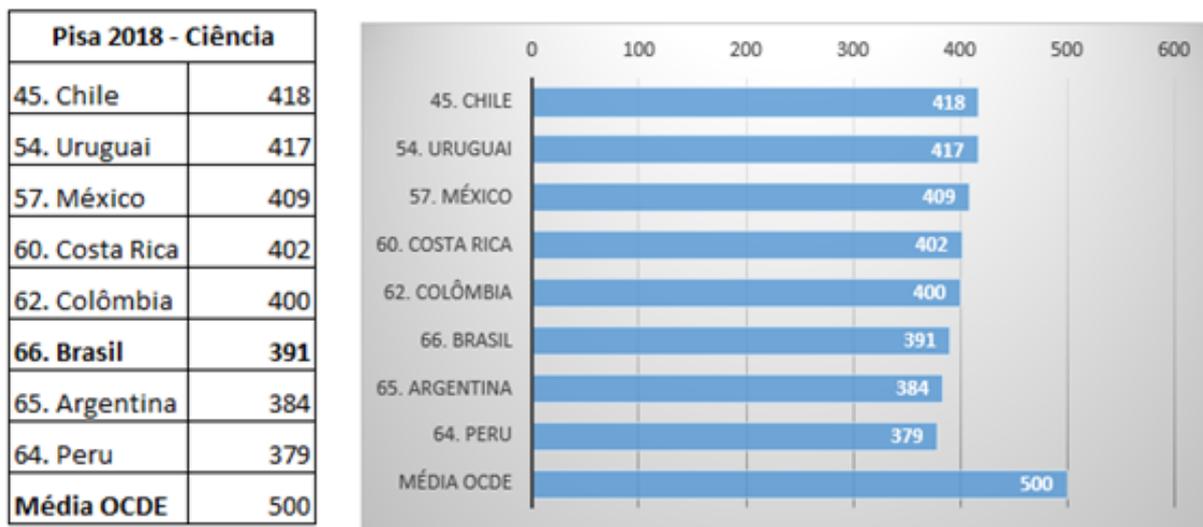
Desempenho Leitura Pisa - 2018 – Latin América



Fonte: Elaboração própria.

Figura 6: Desempenho Ciências Pisa – 2018 – América Latina

Desempenho Pisa Ciência - 2018 – Latin América



Fonte: Elaboração própria.

Para Araújo (2012), a participação brasileira no PISA ocorre em meio a mudanças políticas no sistema educacional. A autora atribui os resultados do Brasil à metodologia utilizada pela avaliação, com relação à participação de alunos com 15 anos de idade. Afirma que, em outros países, essa idade representa o fim do ciclo educacional, mas, no Brasil, isso nem sempre acontece. Ao avaliar os resultados do PISA, a pesquisadora chegou à conclusão de que os estudantes brasileiros, mesmo após frequentarem as escolas por um período grande de tempo, são considerados mal alfabetizados pela avaliação internacional.

Sasaki, Di Pietra, Menezes-Filho e Komatsu (2018) destacam que os alunos brasileiros apresentam desempenho baixo no PISA, especialmente pelo fato de os participantes não conseguirem terminar a prova.

Schneider (2013) pondera que, apesar de o Brasil não apresentar um desempenho satisfatório no PISA, essa avaliação de larga escala internacional serviu como base para formulação da Prova Brasil, principal instrumento avaliativo do país. Para facilitar a compreensão do que se trata a Prova Brasil, a ela dedica-se a próxima subseção.

2.3.2 Prova Brasil

Segundo dados do Ministério da Educação (MEC), a Prova Brasil se constitui de uma avaliação de larga escala com o objetivo de fazer o diagnóstico da situação educacional do país. Ela abrange escolas públicas e privadas, nos âmbitos federal, estadual e municipal, tanto da parte urbana quanto da parte rural.

Criada a partir da Portaria 931 de 21 de março de 2005, a Avaliação Nacional de Rendimento Escolar – ANRESC –, popularmente conhecida como Prova Brasil, é uma avaliação aplicada a cada dois anos. Além disso, compõe o Sistema de Avaliação da Educação Básica – SAEB, em conjunto com a Avaliação Nacional da Educação Básica – ANEB (BRASIL, 2005).

Para Oliveira (2011), a Portaria 931 define que as avaliações têm o papel de fomentar uma ampla mobilização dos envolvidos, além de conscientizá-los da necessidade de proporcionar a melhoria da realidade diagnosticada.

O documento legal afirma que os dados produzidos pelas avaliações de larga escala serão capazes de oferecer subsídios para criação de novas políticas públicas promovendo a melhoria educacional. Segundo Oliveira (2011), se todos os atores envolvidos na educação buscassem a melhoria da qualidade do sistema, existiria uma maior participação da sociedade sobre os resultados produzidos.

Nesse sentido, Szatkoski (2014) destaca que, a partir de 2005, por meio da criação da Prova Brasil, foi possível comparar as unidades de ensino brasileiras, proporcionando um melhor entendimento dos pontos positivos e negativos de cada uma delas.

Na visão da autora, é a partir desse momento que a Prova Brasil passa a ser importante não somente na gestão municipal, mas também na gestão escolar, na figura do diretor, algo que o SAEB, em sua configuração original, não proporcionava. Dessa forma, Skatkoski (2014) ressalta que a nova configuração, instituída pela Portaria 931, oferece ao diretor maior potencial para realizar alterações no âmbito escolar, com o objetivo de melhorar a qualidade escolar.

Os dados apresentados pelo Ministério da Educação demonstram que as avaliações, em geral, têm por configuração serem aplicadas tanto nos anos iniciais como nos anos finais do Ensino Fundamental. As disciplinas avaliadas, conforme explica o ministério, são: Língua Portuguesa, com foco na leitura, e Matemática,

com foco na resolução de problemas. Além disso, há a aplicação de questionários socioeconômicos capazes de fornecer um diagnóstico sobre os fatores de contexto que podem ser diretamente influentes no desempenho acadêmico dos avaliados (BRASIL, 2020).

Segundo Parandekar, Amorim e Welsh (2008), a Prova Brasil é uma base importante de medição de resultados na aprendizagem, mas a sua criação e implementação criaram um desafio urgente. Os autores acreditam que esse desafio está justamente ligado à forma como os dados colhidos serão efetivamente aplicados a fim de resultar num real aumento da qualidade educacional brasileira.

A Portaria 931 procurou, ainda, por meio de seu texto, garantir que as informações produzidas pelas avaliações não fossem utilizadas com o intuito de identificar escolas, turmas, alunos, professores e diretores. Essa medida favoreceu o uso de dados apenas para diagnosticar a situação educacional brasileira, sem expor um ou outro ator no processo.

De acordo com Oliveira (2011), avaliações externas cuja base seja uma eficiente disseminação de dados são capazes de proporcionar uma verdadeira articulação entre as esferas federal, estadual e municipal. A autora ressalta, ainda, que a adoção de um mecanismo de avaliação externo à escola, como ferramenta de regulação da qualidade no aprendizado e nos resultados, é um ganho para os gestores públicos, principalmente nos processos de formulação, implantação e acompanhamento dos resultados das políticas públicas ora implementadas, com bases nos dados colhidos.

Por meio de procedimentos metodológicos formais e científicos, a Portaria 931 garante a coleta e a sistematização dos dados produzidos gerando informações sobre o desempenho escolar, assim como sobre condições intra e extraescolares que tenham influência no processo de ensino e aprendizado.

Soares (2013) lembra que, ao longo de todos esses anos, desde a criação da Prova Brasil, é possível verificar um amadurecimento dos processos de avaliação e coleta dos dados, ampliando seus objetivos, sem desprezar, entretanto, as críticas a ela direcionadas.

Dourado (2018) expõe que, desde a primeira edição das avaliações de larga escala no Brasil, os parâmetros de aferição vêm sendo criticados por repercutirem diretamente em políticas públicas. Para a autora, a Prova Brasil não mede a

aprendizagem em sentido amplo, mas sim a competência dos estudantes em disciplinas como leitura e Matemática.

Segundo Gimenes, Silva, Príncipe, Louzano e Moriconi (2013), as avaliações de larga escala no Brasil não tem seus resultados aplicados para melhorar o cotidiano escolar. Nesse sentido, uma crítica comum é feita com relação à demora na divulgação dos resultados. Os pesquisadores também demonstram ter um consenso entre os gestores sobre a necessidade de as avaliações de larga escala brasileiras serem realizadas anualmente, em vez de bienalmente. Isso oportunizaria a utilização das informações coletadas, permitindo a melhoria da qualidade no cotidiano escolar. Para os autores, essa melhoria pode fornecer, além do diagnóstico da qualidade educacional brasileira, ferramentas para proporcionar a melhoria das escolas, de seus profissionais e até mesmo do currículo efetivamente ensinado.

Para Machado (2017), além da demora na divulgação dos resultados da Prova Brasil, o grande distanciamento entre o professor e o aluno nas avaliações gera um déficit nas políticas de avaliação e nos processos didáticos e pedagógicos. A pesquisadora também tece outra crítica importante, relacionada ao envolvimento das escolas na criação de políticas públicas. De acordo com a autora, essas políticas são formuladas por pessoas fora do contexto escolar, e isso faz com que contribuições significativas de quem tem o conhecimento do processo no dia a dia fiquem esquecidas.

Ainda sobre isso, Dourado (2018) sinaliza que o fato de os professores terem de planejar toda uma estratégia de currículo a ser ensinado é um problema para as escolas, sem contar a necessidade de , e que ainda atenda aos requisitos da avaliação externa, para demonstrar a qualidade da sua escola.

Para Soares (2013), as críticas direcionadas às avaliações de larga escala brasileiras não anulam o potencial que o acúmulo de dados relacionados à realidade educacional tem sobre a formulação de novas políticas públicas. Na sua opinião, a qualidade diagnosticada pelas avaliações é capaz de proporcionar mudanças, por exemplo, na infraestrutura física da escola, facilitando o aprimoramento de diversas práticas pedagógicas. Além disso, a autora ressalta se tratar de possíveis mudanças nas condições de trabalho dos profissionais da educação, na gestão escolar e em outros processos que vão além do desempenho do aluno. Desse

modo, as avaliações de larga escala brasileiras compõem um projeto institucional de educação ampla e, ao mesmo tempo, servem de base para formulação de políticas públicas e de mecanismo de regulação do estado.

Para Parandekar, Amorim e Welsh (2008), o fato de algumas escolas produzirem resultados escolares melhores que outras – mesmo apresentando contexto socioeconômico semelhante – demonstra ser possível gerar a melhoria da qualidade, ainda que com recursos limitados. Nesse sentido, Dourado (2018) salienta a importância das avaliações externas como sinalizadores das condições de aprendizado, proporcionando o aumento das habilidades e competências necessárias para formar bons alunos.

A fim de analisar os dados referentes à Prova Brasil, é necessário descrever as características das turmas no contexto brasileiro, suas variações, formas de montagem, e sua contribuição para o desempenho discente. A seguir, desenvolve-se o conceito de turma.

2.4 Turmas

Alves e Soares (2007) são enfáticos ao mencionar a importância de discutir os impactos que as turmas têm no desempenho escolar dos estudantes. Os autores demonstram que há muitas pesquisas sobre esse tema, pois, em geral, há enormes variações na montagem de turmas no contexto brasileiro.

Sobre isso, Marassi (2010) afirma não haver legislação federal referente ao modo como as turmas devem ser compostas, ficando essa tarefa a cargo do diretor da escola. Nos microdados do INEP, é possível encontrar uma questão, contida no questionário do diretor, que corrobora a visão dessa autora. Essa questão trata da divisão de turmas utilizada pelo diretor naquele ano específico.

Como não há legislação específica para tal divisão, Marassi (2010) aponta a existência de uma maior incidência de turmas homogêneas, já que essa corrente se difundiu desde o século XX. De acordo com a pesquisadora, a maior defesa dessa corrente é o fato de ela possibilitar ao docente o desenvolvimento de um trabalho uniforme com atividades iguais para todos os estudantes. No entanto, a autora destaca que as críticas a esse modelo estão justamente na impossibilidade de fazer uma divisão homogênea justa, já que há uma enorme diversidade psicológica,

cognitiva e social dos estudantes. Diante disso, utilizar o modelo heterogêneo estimularia a socialização dos alunos.

Canário (2005) considera que a homogeneidade da turma representa um diferencial para a escola, pois, ao separar os alunos, seja em termos etários ou em termos de conhecimento, é possível melhorar a administração do espaço, do tempo e do conteúdo ensinado.

Para Bernado (2013), a divisão dos discentes, seja pelo modelo homogêneo, seja pelo modelo heterogêneo ou por nenhum deles, é uma medida de grandes consequências para escola. A autora enfatiza que a divisão de turmas é um dos fatores que impactam diretamente no ambiente de aprendizagem, na instrução por parte do professor, na disseminação do conteúdo, no clima escolar e até na prática pedagógica.

Nessa perspectiva, Caldeira (2018), expõe que a divisão de alunos por turma, também chamada “enturmação”, baseia-se em diversos critérios como idade, desempenho, gênero, entre outros. Ela acredita ser importante que o diretor responsável por essa divisão compreenda os aspectos que melhor se insiram na realidade de sua escola.

De acordo com Oliveira e Bueno (2014), a “enturmação” e seus critérios de divisão são dispositivos muito relevantes na estruturação e configuração escolar. Os autores ressaltam que esse processo deve acontecer da melhor forma possível, pois é na sala de aula que os alunos passam mais tempo.

Marassi (2010) observa que o debate sobre a “enturmação” é vasto, por não haver uma orientação específica para tal divisão. A pesquisadora afirma que a formação das turmas é importante no desempenho escolar e, por isso, deve obedecer a critérios abrangentes.

O decreto-lei número 75 de 2008 define, no artigo 20, que compete ao diretor escolar comandar a formação das turmas. Além disso, assegura a autonomia do diretor na criação de normas próprias para a divisão de turmas (BRASIL, 2008).

Sobre isso Bernado (2013) demonstra que o mais comum nas escolas é que os professores mais qualificados e experientes sejam alocados em turmas com melhores desempenhos. Em contrapartida, Alves e Soares (2007) alertam que a criação de turmas homogêneas por nível de desempenho do estudante, bem como a

alocação dos melhores professores nessas turmas não favorecem os estudantes com maior dificuldade.

Os autores afirmam que, geralmente, o nível de habilidade do aluno coincide com o seu nível socioeconômico. Na opinião dos pesquisadores, o critério mal utilizado para “enturmação” faz com que pequenas diferenças entre grupos se transformem em grandes diferenças no futuro.

Como se pode observar, a falta de legislação específica sobre o tema cria diversas divisões de sala dentro de um mesmo sistema educacional. Apesar de não haver especificamente nenhuma lei sobre turmas, alguns aparatos legais são utilizados para normatizar o processo de forma mais ampla. Na subseção seguinte, apresentam-se essas legislações.

2.5 Legislação

Com o objetivo de contextualizar o cenário legal brasileiro na área da educação, faz-se necessário apresentar as legislações que, de fato, norteiam o entendimento e o desenvolvimento da educação básica.

2.5.1 Lei 9.394 de 1996 – Lei de Diretrizes Básicas da Educação

Segundo Souza (2007), a Lei de Diretrizes Básicas da Educação – LDB – é muito relevante para o debate em torno da educação, e sua importância é inquestionável para o país. O autor argumenta que a lei propicia um meio de fazer importantes mudanças no cenário escolar e, assim, legitimar as melhorias outrora conquistadas.

No artigo primeiro da referida lei lê-se o seguinte: “A educação abrange os processos formativos que se desenvolvem na vida familiar, na convivência humana, no trabalho, nas instituições de ensino e pesquisa, nos movimentos sociais e organizações da sociedade civil e nas manifestações culturais”(BRASIL, 1996), Para Cury (2006), esse artigo é de suma importância para definir a educação brasileira. O pesquisador destaca o quão importante é definir os processos formativos que de fato influenciam na educação, sendo possível verificar a preocupação do legislador em atrelar esses processos à educação.

No entanto, Machado (2014) afirma que, ao diminuir o conceito da educação da forma, como foi feito no primeiro artigo, ocorre a redução da responsabilidade do Estado, bem como o desgaste da relação entre a educação escolar e o ensino com as outras perspectivas educacionais.

Na sequência, o artigo segundo diz o seguinte: “A educação é dever da família e do Estado, inspirada nos princípios de liberdade e nos ideais de solidariedade humana, tem por finalidade o pleno desenvolvimento do educando, seu preparo para o exercício da cidadania e sua qualificação para o trabalho”(BRASIL, 1996). Segundo Utzgi, Costa e Bueno (2015), esse texto é capaz de apontar quais as finalidades da educação, ressaltando a importância da promoção do desenvolvimento educacional e também do fomento à preparação do aluno para o exercício da sua cidadania.

Já o artigo terceiro, em alguns incisos, tem como foco os princípios que deverão nortear o ensino brasileiro:

I - igualdade de condições para o acesso e permanência na escola; II - liberdade de aprender, ensinar, pesquisar e divulgar a cultura, o pensamento, a arte e o saber; III - pluralismo de ideias e de concepções pedagógicas; VII - valorização do profissional da educação escolar; VIII - gestão democrática do ensino público, na forma desta Lei e da legislação dos sistemas de ensino; IX - garantia de padrão de qualidade; XIII - garantia do direito à educação e à aprendizagem ao longo da vida (BRASIL, 1996).

Na visão de Cury (2010), o artigo terceiro norteia a proposta pedagógica, ao expressar o pluralismo de ideias e a concepção do ensino. Ademais, nota-se que o texto da lei é capaz de criar uma autoconsciência que reflete na educação brasileira.

Para Oliveira e Romano (2015), o artigo terceiro propõe a gestão democrática na escola e faz com que a participação dos professores na criação de um projeto pedagógico em conjunto esteja assegurada, o que mostra a importância da participação de todos os agentes na organização escolar.

O artigo quatro, no inciso nove, preconiza que o dever do Estado com educação escolar pública será efetivado mediante a garantia de: “IX - padrões mínimos de qualidade de ensino, definidos como a variedade e quantidade mínimas, por aluno, de insumos indispensáveis ao desenvolvimento do processo de ensino e aprendizagem”(BRASIL, 1996). Para Schneider e Gouveia (2011), a garantia do padrão de qualidade na educação já é um princípio da Constituição Federal. Ao ser

novamente inserido no artigo quarto da LDB, observa-se a preocupação do Estado em reforçar esse objetivo. Sena (2015) ressalta que esse artigo favorece ações capazes de corrigir, de forma progressiva, as disparidades relacionadas ao acesso à educação e de garantir um padrão mínimo de qualidade para educação brasileira.

No artigo nono, inciso quarto, a LDB alude ao papel da União de se encarregar de:

IV - estabelecer, em colaboração com os Estados, o Distrito Federal e os Municípios, competências e diretrizes para a educação infantil, o ensino fundamental e o ensino médio, que nortearão os currículos e seus conteúdos mínimos, de modo a assegurar formação básica comum". (BRASIL, 1996)

De acordo com Pinto e Gomes (2018), o inciso quatro do artigo nono determina que a União seja responsável pelas avaliações de desempenho escolar em todos os níveis de ensino. Além disso, obriga a colaboração de outros entes da federação, a fim de garantir as diretrizes de uma formação básica e de um currículo com conteúdo suficiente.

A educação básica tem sua finalidade definida no artigo 22 da LDB: "A educação básica tem por finalidades desenvolver o educando, assegurar-lhe a formação comum indispensável para o exercício da cidadania e fornecer-lhe meios para progredir no trabalho e em estudos posteriores". (BRASIL, 1996). Já no artigo 23, pode-se observar a forma como a lei dita a organização da educação básica:

A educação básica poderá organizar-se em séries anuais, períodos semestrais, ciclos, alternância regular de períodos de estudos, grupos não seriados, com base na idade, na competência e em outros critérios, ou por forma diversa de organização, sempre que o interesse do processo de aprendizagem assim o recomendar" (BRASIL, 1996).

Para Vasconcelos (2008), é no artigo 23 que a organização pretendida pelo Estado fica explicitamente definida. Entretanto, Ramal (1997) afirma que esse artigo, no trecho "sempre que o interesse do processo de aprendizagem assim o recomendar" (BRASIL, 1996), abre margem para que qualquer critério que pareça conveniente seja utilizado.

A divisão das turmas e os critérios de "enturmação" na educação básica, em conjunto com o rendimento escolar, estão contemplados no artigo 24, Inciso IV e V da LDB, que diz o seguinte:

A educação básica, nos níveis fundamental e médio, será organizada de acordo com as seguintes regras comuns: IV - poderão organizar-se classes, ou turmas, com alunos de séries distintas, com níveis equivalentes de adiantamento na matéria, para o ensino de línguas estrangeiras, artes, ou outros componentes curriculares e V - a verificação do rendimento escolar observará os seguintes critérios: a) avaliação contínua e cumulativa do desempenho do aluno, com prevalência dos aspectos qualitativos sobre os quantitativos e dos resultados ao longo do período sobre os de eventuais provas finais; b) possibilidade de aceleração de estudos para alunos com atraso escolar; c) possibilidade de avanço nos cursos e nas séries mediante verificação do aprendizado; d) aproveitamento de estudos concluídos com êxito; e) obrigatoriedade de estudos de recuperação, de preferência paralelos ao período letivo, para os casos de baixo rendimento escolar, a serem disciplinados pelas instituições de ensino em seus regimentos (BRASIL, 1996).

Segundo Cury (2006), o artigo 24 trata basicamente da organização, dos objetivos a serem alcançados pela educação e das regras comuns do processo. Gadotti (2020) ressalta que é no inciso V que se verifica a preocupação do legislador com os alunos que apresentam baixo desempenho.

No artigo 32 da LDB, nota-se um importante marco na definição da estrutura do ensino. Nesse artigo, lê-se o seguinte:

O ensino fundamental é obrigatório, com duração de 9 anos, gratuito na escola pública, iniciando-se aos 6 anos de idade, terá por objetivo a formação básica do cidadão, mediante: I - o desenvolvimento da capacidade de aprender, tendo como meios básicos o pleno domínio da leitura, da escrita e do cálculo; II - a compreensão do ambiente natural e social, do sistema político, da tecnologia, das artes e dos valores em que se fundamenta a sociedade; III - o desenvolvimento da capacidade de aprendizagem, tendo em vista a aquisição de conhecimentos e habilidades e a formação de atitudes e valores; IV - o fortalecimento dos vínculos de família, dos laços de solidariedade humana e de tolerância recíproca em que se assenta a vida social (BRASIL, 1996)

Para Reis, Souza, Bello e Andrade (2019), o artigo 32, já em seus primeiros parágrafos, focaliza a formação básica dos estudantes, ressaltando quais são os direitos básicos dos alunos em geral.

No artigo 61, a Lei de Diretrizes Básicas da Educação contempla os docentes, tal como se observa no seguinte trecho:

Consideram-se profissionais da educação escolar básica os que, nela estando em efetivo exercício e tendo sido formados em cursos reconhecidos, são: I – professores habilitados em nível médio ou superior para a docência na educação infantil e nos ensinos fundamental e médio; II – trabalhadores em educação portadores de diploma de pedagogia, com habilitação em administração, planejamento, supervisão, inspeção e

orientação educacional, bem como com títulos de mestrado ou doutorado nas mesmas áreas (BRASIL, 1996).

Ademais, o artigo 67, incisos II, IV e VI, a LDB textualiza a valorização docente:

Os sistemas de ensino promoverão a valorização dos profissionais da educação, assegurando-lhes, inclusive nos termos dos estatutos e dos planos de carreira do magistério público: II - aperfeiçoamento profissional continuado, inclusive com licenciamento periódico remunerado para esse fim; IV - progressão funcional baseada na titulação ou habilitação, e na avaliação do desempenho; VI - condições adequadas de trabalho” (BRASIL, 1996).

Segundo Graciano (2007), é nos artigos 61 e 67 que a legislação alerta para a necessidade de um aprimoramento constante dos docentes, tanto em questão de formação continuada, como em questão salarial. Nesse sentido, Cury (2010) considera que, nesses artigos, nota-se o apreço do legislador pela aprendizagem dos estudantes e pela garantia das condições de trabalho. Grazziotini, Waimannii e Mendézii (2010) ressaltam, ainda, que o artigo 67, com seus incisos, ampara legalmente a formação continuada do professor, criando a obrigação de o sistema educacional fomentar o desenvolvimento do docente.

2.5.2 Constituição Federal de 1988

A Constituição Federal brasileira, em alguns artigos, trata da garantia do direito à educação. Logo no artigo sexto, é possível verificar essa preocupação: “São direitos sociais a educação, a saúde, a alimentação, o trabalho, a moradia, o transporte, o lazer, a segurança, a previdência social, a proteção à maternidade e à infância, a assistência aos desamparados, na forma desta Constituição”. (BRASIL, 1998)

Moreira, Rezende e Rangel (2019) destacam que esse artigo reconhece, de maneira clara, o direito à educação como um direito social, condicionando o Estado à garantia de seu cumprimento. Diante disso, Melo (2017) resalta que a educação não é um favor do Estado à sociedade, mas sim um direito a ser garantido, e seu desrespeito deve ser punido na forma da lei.

Já no artigo 22, inciso XXIV, a constituição disciplina a competência privativa da União em legislar sobre: “diretrizes e bases da educação nacional” (BRASIL, 1988). De acordo com Travincas (2016), esse artigo assegura à União a atribuição

de editar as diretrizes e as bases da educação brasileira. Sobre isso, Kim e Bolzam (2015) salientam que a repartição de competências é uma técnica comum da Constituição e, por meio do princípio da predominância do interesse público, a União legisla sobre as questões de importância nacional.

A Constituição, em seu artigo 23 inciso V, reforça as competências da União e apresenta as que são comuns aos outros entes: “É competência comum da União, dos Estados, do Distrito Federal e dos Municípios: V - proporcionar os meios de acesso à cultura, à educação, à ciência, à tecnologia, à pesquisa e à inovação”(BRASIL, 1988).

De acordo com Travincas (2016), a competência prevista no artigo 23 é de natureza administrativa, com foco em um tipo de federalismo cooperativo. Sob outro ponto de vista, Andrade (2013) faz uma crítica à ausência de uma lei complementar para reforçar o descrito nesse artigo e acredita que tal omissão gera limitações quanto à uniformidade do atendimento às demandas educacionais. O artigo 205 da Carta Magna preconiza:

A educação, direito de todos e dever do Estado e da família, será promovida e incentivada com a colaboração da sociedade, visando ao pleno desenvolvimento da pessoa, seu preparo para o exercício da cidadania e sua qualificação para o trabalho (BRASIL, 1988).

Para Melo (2017), esse artigo define o dever familiar de matricular o filho na escola, zelar por sua aprendizagem e, ainda, de participar de reuniões pedagógicas. A autora ressalta a importância tanto do Estado como da família de fiscalizar a continuidade desses processos.

O dever do Estado com relação à educação básica é melhor especificado no artigo 208, incisos I e VII:

O dever do Estado com a educação será efetivado mediante a garantia de: I - educação básica obrigatória e gratuita dos 4 aos 17 anos de idade, assegurada inclusive sua oferta gratuita para todos os que a ela não tiveram acesso na idade própria; VII - atendimento ao educando, em todas as etapas da educação básica, por meio de programas suplementares de material didático escolar, transporte, alimentação e assistência à saúde (BRASIL, 1988).

Saveli e Tenreiro (2012) consideram que, nesse artigo, a Constituição determina a gratuidade de ensino a todos. Para os autores, é imperativo dizer que o texto garante o direito subjetivo de todos aqueles que queiram estudar. Nesse

sentido, Cury (2004) ressalta que o Estado tem a obrigação de atender à demanda educacional de qualquer cidadão, por mérito do direito subjetivo, assegurado no artigo 208, sob pena de punições legais.

2.5.3 Lei 8.069 de 1990 – Estatuto da Criança e do Adolescente

Segundo Ferreira (2004), o Estatuto da Criança e do Adolescente – ECA – representou uma mudança de padrão, reconhecendo as crianças e os adolescentes como dignos de direitos, assegurando-lhes principalmente o acesso à educação. Em seu artigo quarto, diz o ECA:

É dever da família, da comunidade, da sociedade em geral e do poder público assegurar, com absoluta prioridade, a efetivação dos direitos referentes à vida, à saúde, à alimentação, à educação, ao esporte, ao lazer, à profissionalização, à cultura, à dignidade, ao respeito, à liberdade e à convivência familiar e comunitária (BRASIL, 1990).

Andrade (2012) afirma ser perceptível, nesse texto, a complementação da Constituição, no sentido de assegurar os direitos das crianças e jovens brasileiros. Contudo, é no artigo 53 que o detalhamento desses direitos é melhor apresentado:

A criança e o adolescente têm direito à educação, visando ao pleno desenvolvimento de sua pessoa, preparo para o exercício da cidadania e qualificação para o trabalho, assegurando-lhes: I - igualdade de condições para o acesso e permanência na escola; II - direito de ser respeitado por seus educadores; III - direito de contestar critérios avaliativos, podendo recorrer às instâncias escolares superiores; IV - direito de organização e participação em entidades estudantis; V - acesso à escola pública e gratuita próxima de sua residência. VI - Acesso à escola pública e gratuita, próxima de sua residência, garantindo-se vagas no mesmo estabelecimento a irmãos que frequentem a mesma etapa ou ciclo de ensino da educação básica. Parágrafo único. É direito dos pais ou responsáveis ter ciência do processo pedagógico, bem como participar da definição das propostas educacionais (BRASIL, 1990).

Para Cury (2005), o inciso V do artigo 53 é um complemento bastante eficaz da Constituição Federal e da Lei de Diretrizes Básicas da Educação, pois é capaz de garantir o acesso à escola pública e gratuita na vizinhança. Sob essa ótica, Ferreira (2009) salienta, ainda, que esse artigo estabelece a educação como base de preparo à cidadania.

Para Antunes, Caroline, Lara e Soffa (2017), cabe aos órgãos públicos e aos entes da Federação garantir o efetivo cumprimento dos dispositivos legais e

promover a efetiva adequação das instituições para atendimento das crianças e adolescentes.

2.5.4 Lei 13.005 de 2014 – Plano Nacional de Educação

A Lei 13.005 de 25 de junho de 2014 se baseia nas premissas do artigo 214 da Carta Magna para instituir o Plano Nacional da Educação – PNE – detalhando todos os seus objetivos e metas para o período de 10 anos (BRASIL, 2014). Mais especificamente,

A lei estabelecerá o Plano Nacional de Educação, de duração decenal, com o objetivo de articular o sistema nacional de educação em regime de colaboração e definir diretrizes, objetivos, metas e estratégias de implementação para assegurar a manutenção e desenvolvimento do ensino em seus diversos níveis, etapas e modalidades por meio de ações integradas dos poderes públicos das diferentes esferas federativas que conduzam a: I - erradicação do analfabetismo; II - universalização do atendimento escolar; III - melhoria da qualidade do ensino; IV - formação para o trabalho; V - Promoção humanística, científica e tecnológica do País. VI - Estabelecimento de meta de aplicação de recursos públicos em educação como proporção do produto interno bruto (BRASIL, 1988).

Pereira e Martins (2017) destacam que a educação básica, por meio do PNE, é capaz de gerar políticas públicas de inclusão que possibilitem à lei alcançar todos os estudantes, considerando todas as minorias e outras especificidades.

No artigo segundo do PNE, lê-se:

São diretrizes do PNE: I - erradicação do analfabetismo; II - universalização do atendimento escolar; III - superação das desigualdades educacionais, com ênfase na promoção da cidadania e na erradicação de todas as formas de discriminação; IV - melhoria da qualidade da educação; V - formação para o trabalho e para a cidadania, com ênfase nos valores morais e éticos em que se fundamenta a sociedade; VI - promoção do princípio da gestão democrática da educação pública; VII - promoção humanística, científica, cultural e tecnológica do País; VIII - estabelecimento de meta de aplicação de recursos públicos em educação como proporção do Produto Interno Bruto - PIB, que assegure atendimento às necessidades de expansão, com padrão de qualidade e equidade; IX - valorização dos (as) profissionais da educação; X - promoção dos princípios do respeito aos direitos humanos, à diversidade e à sustentabilidade socioambiental (BRASIL, 2014).

De acordo com Yaegashi, Oliveira, Ruiz e Souza (2020), uma das estratégias do PNE é a estruturação dos projetos pedagógicos voltados à alfabetização. Para tal, os autores acreditam que há a necessidade da valorização docente, algo que está presente no inciso IX do artigo segundo.

Com o intuito de entender como o PNE norteia o desenvolvimento da educação básica, apresentam-se aqui algumas metas, a fim de elucidar a sua importância no processo educacional brasileiro. As metas apresentadas são a 2, 5, 6, 7, 15, 16 e 17.

A meta número 2 do PNE busca: “universalizar o ensino fundamental, num total de nove anos para toda a população de 6 a 14 anos e garantir que pelo menos 95% dos alunos concluam essa etapa na idade recomendada, até o último ano de vigência deste PNE”(BRASIL, 2014). De acordo com Oliveira, Gouveia e Araújo (2018), em conjunto com outras, essa meta depende da realização de outras questões presentes no PNE, por exemplo, a valorização docente. Os autores ainda alertam a necessidade de participação das secretarias de educação, das escolas, e dos professores para efetivação desta meta.

A fim de garantir a alfabetização dos estudantes brasileiros, a meta 5 objetiva: “alfabetizar todas as crianças, no máximo, até o final do 3º ano do ensino fundamental” (BRASIL 2014). Segundo Ferreira (2016), para a efetivação dessa meta, é necessário implantar mecanismos de avaliação para cada ano de ensino, que tenham como foco aferir a alfabetização. Dessa forma, será possível estimular os sistemas de ensino a corrigir determinados déficits.

Com objetivo de oferecer a educação em tempo integral, a meta 6 propõe: “oferecer educação em tempo integral em, no mínimo, 50% das escolas públicas, de forma a atender, pelo menos, 25% dos alunos da educação básica” (BRASIL, 2014).

Segundo Amaral (2014), ela carece de muitas discussões sobre as condições para sua realização e depende de um volume grande de recursos financeiros.

Visando a atingir melhores médias de aprendizagem, a meta 7 procura: “fomentar a qualidade da educação básica em todas as etapas e modalidades, com a melhoria do fluxo escolar e da aprendizagem de modo a atingir as seguintes médias nacionais para o IDEB” (BRASIL, 2014).

Tabela 3: Médias IDEB de 2015, 2017, 2019 e 2021.

IDEB	2015	2017	2019	2021
Anos Iniciais do Ensino Fundamental - 5º Ano	5,2	5,5	5,7	6,0
Anos Finais do Ensino Fundamental - 9º Ano	4,7	5,0	5,2	5,5
Ensino Médio	4,3	4,7	5,0	5,2

Fonte: Elaboração própria.

Oliveira, Gouveia e Araújo (2018) textualizam que, para o alcance das médias desejadas na meta 7, as ações do governo atual necessitam de articulação com as propostas do Congresso, a fim de ampliarem as avaliações de larga escala brasileiras. Nesse sentido, Ferreira (2016) deixa claro que o IDEB, por ter um caráter censitário, é o maior aliado para busca da qualidade educacional refletida nos aumentos das médias. A autora afirma ainda que a meta 7 do PNE é a que mais favorece o sistema brasileiro de avaliações de larga escala, pois, apesar de mencionar aumento de médias ou da qualidade, a referida meta tem como objetivo principal a melhoria dos fluxos escolares.

Para garantir que os professores da educação básica apresentem uma melhor qualidade em termos de currículo, a meta 15 busca

Garantir, em regime de colaboração entre a União, os Estados, o Distrito Federal e os Municípios, no prazo de um ano de vigência deste PNE, política nacional de formação dos profissionais da educação, assegurado que todos os professores e as professoras da educação básica possuam formação específica de nível superior, obtida em curso de licenciatura na área de conhecimento em que atuam (BRASIL, 2014).

Seguindo essa mesma linha, a meta 16 tenciona

[...] formar, em nível de pós-graduação, 50% dos professores da educação básica, até o último ano de vigência deste PNE, e garantir a todos os profissionais da educação básica formação continuada em sua área de atuação, considerando as necessidades, demandas e contextualizações dos sistemas de ensino (BRASIL, 2014).

Oliveira, Gouveia e Araújo (2018) tecem uma crítica contundente a essas metas, em virtude da visão tecnicista do Estado com relação à formação docente. Para os autores, elas representam um grave retrocesso e desvalorizam inúmeros professores já atuantes no sistema educacional do país.

Por fim, a meta 17 trata de: “valorizar os profissionais do magistério das redes públicas de educação básica de forma a equiparar seu rendimento médio ao dos demais profissionais com escolaridade equivalente, até o final do sexto ano de vigência deste PNE”. (BRASIL, 2014). Gomez e Ramos (2018) ressaltam que ela foi responsável por evidenciar, de forma categórica, a urgência de valorizar os profissionais da educação básica em termos salariais. Nesse ponto, Oliveira, Gouveia e Araújo (2018) destacam que a valorização docente é o tema mais

importante no momento para educação brasileira. Ademais, a realização da meta 17 é capaz de assegurar ao professor melhores condições de trabalho.

2.6 Panorama de pesquisas

Apenas com o conhecimento do que foi exposto até o presente momento, não é possível estar plenamente inserido no contexto educacional relacionado ao tema. Para tanto, faz-se necessário conhecer as pesquisas já realizadas acerca do objeto aqui estudado. Por isso, nesta subseção, focalizam-se os estudos realizados tanto no Brasil, como no exterior.

Segundo Franco (2009), a qualidade do ensino não é uma questão que interessa apenas a países subdesenvolvidos com altos níveis de analfabetismo. Os países desenvolvidos também investem muito nesse setor, proporcionando a melhoria da qualidade de vida e da educação de sua população. Para Menezes-Filho (2007), no Brasil, são diversas evidências que apontam para a importância da educação em setores econômicos e sociais.

Considerando o objetivo deste trabalho, ao analisar o cenário internacional, percebe-se que nem todos os problemas relacionados ao número de alunos em sala de aula foram estudados no montante necessário. Desse modo, é possível encontrar apenas algumas pesquisas sobre a manipulação do tamanho das turmas.

Apesar disso, Hanushek (1998) admite que nenhum tema na educação é tão discutido no nível escolar como o número de alunos por turma, tanto no nível público como no privado. Chingos (2010) afirma que há um grande debate sobre a real eficácia dessa medida para o desempenho escolar.

O desafio que se apresenta, ao analisar estas pesquisas, está justamente em encontrar quais delas contribuíram de forma satisfatória para elucidação do problema de pesquisa aqui proposto.

Constata-se, então, que alguns experimentos são considerados pelos pesquisadores da área mais significativos que outros. Portanto, após análise de diversos artigos científicos, selecionaram-se todos os trabalhos encontrados, inclusive os mais significativos. As pesquisas são divididas por área geográfica, a fim de organizar a apresentação.

2.6.1 América do Norte

Nesta subseção, podem-se observar todos os trabalhos encontrados desenvolvidos na América do Norte.

2.6.1.1 The Tennessee STAR Project

Acerca do desenvolvimento dessa pesquisa, Word (1990) expõe que os relatórios técnicos do projeto STAR apresentaram os dados referentes aos anos de 1985 a 1990. Esses relatórios foram preparados pelo departamento de educação do estado do Tennessee e redigidos por Elizabeth Word.

De acordo com grande parte da literatura sobre o tema, o projeto STAR do Tennessee foi um estudo que marcou época e, por isso, faz parte do referencial teórico de inúmeros trabalhos. Chingos (2010) considera que o Projeto STAR foi uma pesquisa de grandes proporções, que aplicou o método randomizado dentro da escola, o “padrão ouro” em pesquisas empíricas. No entanto, Hanushek (1998) afirma que o projeto foi uma pesquisa extensa e complicada.

Word (1990) informa que o projeto STAR foi fruto de um projeto experimental para verificar o efeito do tamanho da classe na aprendizagem dos alunos. A autora explica que pesquisa foi realizada por meio da Universidade Estadual do Tennessee, principal universidade do estado. Para Hanushek (1998), o *design* do Projeto STAR foi muito influenciado pelo resumo inicial do trabalho realizado por Glass e Smith, em 1979.

Sob essa perspectiva, Word (1990) destaca que a legislação vigente à época do estudo determinava que, para um professor, era aceitável o máximo de 25 alunos para as turmas primárias. A autora informa que o tamanho médio das turmas do Tennessee, nos anos de 1985 e 1986, era de 22,3 alunos no jardim de infância e 23,5 alunos no primeiro ano. Esses números, portanto, estavam abaixo do que a legislação considerava aceitável.

Segundo ela, de acordo com os relatórios do Projeto STAR, a legislação ainda estabeleceu que 13 seria o mínimo de alunos por turma. A pesquisadora também lembra que, para o desenvolvimento do projeto, houve um investimento de três milhões de dólares por ano, em um período de quatro anos.

Para Word (1990), a vantagem principal do projeto STAR foi sua rápida implantação, e isso ocorreu devido a um esforço cooperativo dos envolvidos, visto que, em agosto de 1985, as escolas já estavam iniciando o projeto. Naquele momento, a autora percebeu que, dos 3 milhões de dólares disponibilizados por ano, cerca de 2,5 milhões de dólares eram destinados exclusivamente ao custeio de professores e auxiliares do projeto. Diante disso, Hoxby (1998) ressalta que o Projeto STAR se destacou no cenário educacional, pois conseguiu alcançar os resultados em apenas um ano de projeto.

Word (1990) relata que todos os professores participantes do projeto foram certificados de acordo com nível de ensino em que atuavam. Ademais, recorda que à época, era necessário que as escolas concordassem com a atribuição aleatória de professores e alunos a diferentes condições de ensino. Segundo a pesquisadora, inicialmente, 180 escolas pertencentes a 50 dos 141 sistemas escolares do estado manifestaram interesse em participar da pesquisa.

A autora informa que os dados colhidos no relatório do projeto STAR demonstraram que o estudo teve a participação de 128 turmas pequenas, com cerca de 1.900 alunos, 101 turmas regulares, com cerca de 2.300 alunos, e 99 classes regulares com auxiliares de professores, com cerca de 2.200 alunos.

Para contornar o problema das variáveis intervenientes que poderiam influenciar no resultado, a pesquisadora optou por criar um projeto dentro de cada escola, com o intuito de reduzir as principais fontes de variações possíveis. Isso porque tais variáveis poderiam impactar no desempenho dos alunos.

Por meio de seu relatório, ela registra que algumas escolas do estado do Tennessee apresentavam um número de matrículas superior às demais e, para que isso não influenciasse no resultado, aulas adicionais foram criadas. Dessa forma, garantiu-se que todos os grupos de alunos comparados tivessem as mesmas características, como currículos, direção, política, agenda, despesas, entre outras.

Word (1990) afirma, ainda, que, por meio da legislação do Tennessee, se determinou o teste de alunos uma vez por ano, em datas prefixadas. Para isso, definiu-se o nível adequado do *Stanford Achievement Test* – SAT – que deveria ser alcançado por cada turma.

A autora também informa que, por meio de uma análise transversal dos dados, avaliando os alunos participantes em cada série e analisando regionalmente

os impactos do estudo, se determinou, por meio de procedimentos de variância, a divisão da pesquisa em:

1. Tipo de Classe - Dimensão fixa; Diferença média entre os tipos de classes compreendidas no objetivo do estudo;
2. Tipo de escola – Dimensão fixa; Tipo de classe Cruzada;
3. Escolas – Dimensão Aleatória; Alinhada em Locais; cruzada ao tipo de Classe (três tipos de turmas presentes em cada escola);
4. Classes – Dimensão Aleatória (quando há mais de um tipo de classe dentro de alguma escola em particular);
5. Alunos – Amostra Aleatória (alinhada a cada aula)

Ao avaliar o Projeto Star, Woessman e West (2002) salientam que o uso de experimentos randomizados no estudo apresentou problemas intrínsecos. Sobre isso, Hanushek (1998) sinaliza a existência de efeitos de realização limitados e específicos no STAR.

Word (1990), por sua vez, considera que, por meio do resultado do projeto, foi possível observar quatro análises transversais, feitas ao final de cada ano, compreendendo desde o jardim de infância até a terceira série. Sendo assim, a autora percebeu que o primeiro ano de projeto teve como foco a busca por amostras que identificaram as diferenças existentes entre cada escola e cada sistema.

Hanushek (1998) afirma que o Projeto STAR tem efeito único e duradouro nas reduções de tamanho de classe no jardim de infância.

Para Word (1990), a análise dos resultados mostra que as classes menores tiveram um desempenho significativo em algumas pontuações. A pesquisadora percebeu também que as pontuações sobre a motivação dos alunos permaneceram iguais, tanto em sala pequena quanto em sala regular. Além disso, a autora identifica diferenças significativas por localização ou região no desempenho escolar. Entretanto, quando as variáveis intervenientes foram controladas, as distorções desapareceram.

Em seu relatório, Word (1990) expõe que o efeito entre o desempenho do professor e a sala de aula com um número de alunos reduzido apresentou-se de forma positiva, sugerindo que salas de aula com poucos estudantes, além de melhorar o desempenho escolar, também beneficiavam os docentes.

A partir da análise dos resultados do referido projeto, Hanushek (1998) considera que tanto para Leitura como para Matemática, os alunos de classes pequenas têm um desempenho significativamente maior no fim do jardim de infância. Nesse sentido, Hoxby (1998) afirma que o Projeto STAR testemunhou uma melhora única na realização e, apesar de o projeto contar com reduções pequenas no tamanho da turma, há resultados estatisticamente significantes.

Segundo Word (1990), quando a equipe de pesquisadores direcionou a análise ao jardim de infância, percebeu-se um resultado definitivamente superior, ao comparar salas pequenas com salas regulares. Entretanto, ao analisar o impacto do uso de um auxiliar de classe, não foi encontrada nenhuma vantagem significativa.

A partir de sua análise, Hanushek (1998) constata que, no STAR, classes menores têm efeito significativo no jardim infância, mas não há evidências de que sejam eficazes posteriormente. Nesse sentido, Chingos (2010) salienta que os alunos que ingressaram em salas menores nos anos iniciais do jardim de infância apresentaram resultados mais duradouros que os ingressantes na segunda e terceira séries.

Para Hanushek (1998), uma maneira de melhorar os resultados do STAR seria colocar os mesmos alunos em classes pequenas nos anos iniciais e, em determinado momento, em classes grandes.

Hoxby (1998) ressalta que o projeto STAR foi um experimento político. Por essa razão, alunos e professores sabiam que estavam sendo avaliados, e isso influenciou negativamente o resultado.

Woessman e West (2002) afirmam que a validade do Projeto STAR foi prejudicada por decisões específicas de colocação descentralizada. Justamente nesse sentido, Hanushek (1998) considera que inúmeros fatores, tanto de *design* como naturais, criaram incertezas na análise e interpretação dos dados no Projeto STAR. O autor também pontua que o projeto nunca foi replicado, o que representa dificuldades para sanar as dúvidas surgidas com o experimento.

2.6.1.2 Wisconsin's SAGE Program

Com base na análise dos dados publicados em dezembro de 1997, por Peter Maier, Alex Molnar, Stephen Percy, Philip Smith e John Zahorik, foi possível verificar a complexidade do projeto SAGE, que surgiu por meio da parceria entre o

departamento de instrução pública e o Centro de Iniciativas Urbanas e Pesquisas da Universidade de Wisconsin, Milwaukee.

Segundo Graue, Hatch, Rao e Oen (2007), o SAGE é a maior engrenagem na máquina da escola. Para os autores, se a maior máquina não se aperfeiçoa, o restante fica comprometido.

Maier, Molnar, Percy, Smith e Zahorik (1997) informam que o projeto teve, como objetivo verificar qual a eficácia na melhoria do desempenho escolar de alunos de baixa renda do terceiro ano do jardim de infância com a aplicação da diminuição de estudantes por turma. O projeto foi promulgado pelo legislativo de Wisconsin em 1995.

Graue, Hatch, Rao e Oen (2007) afirmam que o SAGE se configurou como uma reforma legislativa de caráter multifacetado, cujo intuito era reduzir o impacto da falta de recursos familiares no desempenho escolar. Conforme apontam Maier, Molnar, Percy, Smith e Zahorik (1997), por intermédio do Estatuto do programa SAGE, os objetivos principais do projeto foram divididos na seguinte ordem: redução do tamanho da turma para 15 alunos no jardim de infância e primeira série, em 1996 e 1997; redução do tamanho da turma para 14 alunos no jardim de infância e primeira série, em 1997 e 1998; redução do tamanho da turma para 13 alunos no jardim de infância e primeira série, em 1998 a 2001. De acordo com Graue, Hatch, Rao e Oen (2007), a lei do SAGE definiu o número máximo de 15 alunos por classe.

Maier, Molnar, Percy, Smith e Zahorik (1997) mencionam a existência de outros objetivos no projeto, tais como o de manter as escolas abertas desde o início da manhã até o final do dia e colaborar com as organizações comunitárias. Com isso, visava-se a fornecer serviços educacionais, recreativos, comunitários e sociais. Além disso, tencionava-se disponibilizar um currículo acadêmico rigoroso para melhorar a realização do projeto e estabelecer mecanismos de desenvolvimento e responsabilização de pessoal. Os autores afirmam que o termo “Escola Iluminada” foi criado para caracterizar a escola que funcionava desde o início da manhã até o fim do dia e ainda colaborava com organizações comunitárias, com o intuito de fornecer os serviços já citados.

Para o desenvolvimento do projeto, os pesquisadores relatam que houve a participação de 30 escolas do estado de Wisconsin, nos anos de 1996 e 1997. O

SAGE, segundo eles, contou com 3.614 estudantes, 220 professores e 190 salas de aula do jardim de infância e da primeira série.

Graue *et al.* (2007) textualizam que a implementação do SAGE considerou o contexto de diversos locais distintos, refletindo aspectos físicos, sociais e intelectuais. Nesse sentido, Ehrenberg *et al.* (2001) afirmam que o SAGE contou com uma avaliação bem projetada e foi um exemplo de programa altamente direcionado.

Para Maier *et al.* (1997), as salas de aula do projeto variavam entre 9 a 37 alunos. No entanto, algumas delas excederam a proporção de 15 alunos para um professor, apenas por poucos estudantes. A sala de aula média continha 17,4 estudantes.

Conforme esclarecem os autores, o objetivo principal do SAGE foi determinar se a proporção de 15 alunos por professor, a adoção de um currículo mais rigoroso a implantação do projeto escola iluminada e o desenvolvimento do pessoal, beneficiavam os estudantes na obtenção de melhores resultados acadêmicos.

Na visão de Ehrenber *et al.* (2001) o projeto comparou os alunos da primeira série das escolas participantes com um grupo de estudantes das escolas não participantes, que mantinham semelhanças em relação a renda familiar, desempenho, matrículas e composição racial.

Maier *et al.* (1997) afirmam que o projeto SAGE foi composto por uma avaliação formativa dividida em duas etapas que objetivavam determinar a eficácia do projeto. A primeira estava direcionada para a principal variável, ou seja, segundo eles a redução da proporção entre aluno e professor, utilizando para tanto, o método de mudança comparativa.

De acordo com os autores, a análise dos resultados entre as médias do grupo SAGE e do grupo de comparação, no primeiro momento, não encontrou qualquer diferença estatisticamente significativa para qualquer uma das variáveis aplicadas.

Finn e Pannozzo (2003) afirmam que os dados comportamentais dos estudantes do projeto sofreram com a ausência de uma comparação com classes maiores. Na busca dos resultados, Maier *et al.* (1997) utilizaram elementos diferentes da mudança comparativa, a fim de encontrar uma explicação para variação desempenho. Segundo os autores, o objetivo da pesquisa somente foi alcançado depois de medidas repetidas, bem como do desenvolvimento de controles reflexivos.

Os pesquisadores afirmam que a avaliação dos resultados do SAGE, realizada tanto no nível individual quanto no de classe, contou com vários métodos, utilizando-se de testes estatísticos bivariados e multivariados. Para obtenção de êxito, os autores observaram os ganhos de pontuação dos alunos participantes na comparação entre o pré-teste e o pós-teste para a disciplina de linguagens. Já na disciplina de leitura, observou-se o menor ganho relativo para os alunos participantes.

Segundo Milesi e Gamoran (2006), o projeto encontrou maiores benefícios das turmas pequenas para estudantes negros.

Maier *et al.* (1997) informam que o projeto demonstrou, em termos de renda familiar e de potenciais impedimentos ao ingresso na escola, certa vantagem dos alunos não participantes em relação aos que participaram do projeto. Mais especificamente, os autores perceberam que, para verificar essa diferenciação, foi necessário o controle estatístico de algumas variáveis intervenientes por meio do método de análises multivariadas, a fim de que tal discrepância não influenciasse o resultado.

Para testar os efeitos do SAGE no desempenho discente, os pesquisadores controlaram os fatores com base em modelos de regressão de mínimos quadrados ordinários, observando cada subteste e pontos totais na escala. De acordo com os autores, foi importante o destaque dos efeitos, pois eles variavam de uma sala de aula para outra. No entanto, os pesquisadores informam que todos os resultados dos subtestes e suas pontuações foram consistentes. Com o objetivo de encontrar um resultado que fosse além da manipulação dos alunos por salas, Maier *et al.* (1997) optaram por verificar se a combinação entre o tamanho da classe e a participação no SAGE apresentou efeitos significativamente satisfatórios.

Para Graue *et al.* (2007), a configuração do SAGE não permaneceu constante, e essa falta de conexão fez os autores adotarem uma dimensão longitudinal de testes.

Com o objetivo de buscar um resultado mais abrangente no projeto, Maier *et al.* (1997) utilizaram entrevistas com professores. Nelas, questionou-se aos docentes qual era a percepção de cada um deles com as mudanças na maneira de ensinar, já que, com turmas menores, foi possível maior interação com os alunos. Segundo os autores, do total de docentes entrevistados, somente duas equipes não observaram

mudanças significativas; os demais indicaram melhora com relação à disciplina, à instrução e ao entusiasmo pessoal dos estudantes.

De acordo com Graue *et al.* (2007), os professores do projeto tiveram de mudar suas práticas para que atuassem lado a lado com outro professor e, assim, aproveitar todos os espaços em sala de aula, proporcionando o melhor aprendizado dos discentes.

Maier *et al.* (1997) atestam que, com uma turma reduzida, os docentes perceberam que poderiam passar mais tempo com alunos individualmente e, desse modo, proporcionar um aprendizado diferenciado.

Ao direcionar o foco das entrevistas na proporção entre aluno e professor, os pesquisadores perceberam que, quando citada a efetividade do ensino, os professores mencionaram com mais frequência o termo “individualização”. Em sua totalidade, segundo os autores, os educadores entrevistados argumentaram que, no período em que participaram da pesquisa, foram capazes de se dedicar mais às necessidades de cada estudante. Além disso, puderam diagnosticar deficiências com maior antecedência.

Finn e Pannozzo (2003) informam que os professores atuantes no SAGE afirmaram entender melhor seus alunos do projeto, em comparação com os de salas maiores.

Maier *et al.* (1997) destacam a existência de argumentos no sentido de que, ao melhorar a capacidade de aprendizagem, se tornou possível a aproximação de cada aluno de forma mais pessoal, criando um vínculo de trabalho mais duradouro. Outro benefício percebido pelos professores do projeto SAGE foi com relação à locomoção dentro de sala de aula, já que, com menos alunos, havia mais espaço entre as carteiras.

Ao entrevistar os estudantes, os pesquisadores perceberam benefícios, tais como a vantagem de não ter de esperar pela atenção docente, proporcionando ainda mais facilidade de entendimento e disponibilização de conteúdos mais aprofundados.

Graue *et al.* (2007) afirmam que o SAGE, em particular, e as pesquisas de controle do tamanho da classe, em geral, permitiam que os professores possibilitassem oportunidades de aprendizados para os discentes. Os dados encontrados pela pesquisa não ficaram apenas na individualização da

aprendizagem; uma turma menor também apresentou maior tempo de compartilhamento de ideias, bem como maior interação entre professor e aluno.

Para Maier *et al.* (1997) algumas atividades foram recorrentes, por exemplo, a leitura em voz alta e atividades de perguntas e respostas. Nesse sentido, os pesquisadores ressaltam que uma das mudanças mais mencionadas pelos professores nas entrevistas foi o aumento do uso de atividades centradas no aluno. Dessa forma, geraram-se mais oportunidades de os estudantes resolverem problemas e se envolverem em mais atividades ligadas à criatividade.

Para os autores, os educadores ficaram mais entusiasmados em ensinar, em consequência da diminuição do número de alunos por turma. Esses profissionais passaram a ter uma atitude muito mais positiva em relação ao ensino, tendo muito mais energia e motivação. Assim, segundo os autores, os docentes experimentaram menos situações de estresse em geral.

Graue *et al.*(2007) afirmam que quase todos os professores relataram o problema do espaço em sala como um desafio nas escolas do projeto SAGE.

Já as entrevistas realizadas por Maier *et al.* (1997) no SAGE mostraram que os professores foram unânimes quanto ao crescimento acadêmico dos estudantes participantes, com relação à leitura, conhecimento em línguas e Matemática. Os autores informam que, depois de um ano de programa, os resultados comprovam a diminuição dos problemas de disciplina em sala de aula, melhorando a gestão da classe pelos docentes. Além disso, os ganhos acadêmicos dos alunos SAGE superaram as expectativas.

2.6.1.3 Indiana's Project Prime Time

Conforme Gilman (1998), o projeto Prime Time do estado de Indiana, nos EUA, teve como objetivo investigar o impacto de uma turma menor no desempenho das turmas de primeiro ano. Para isso, participaram do projeto quatro grupos de alunos, alocados em salas pequenas pelo período de um ano. Eles foram comparados a uma turma de alunos também de primeira série com instrução em turmas maiores.

Hunter (2000) informa que o programa Prime Time foi criado para incentivar escolas à redução da proporção de alunos e professores do jardim de infância à terceira série.

Segundo Englehart (2007), no ano de 1984, o estado de Indiana, a partir do projeto Prime Time, liderou as iniciativas de reduções de classe, incorporando uma agenda de pesquisa.

Gilman (1988) afirma que todos os participantes pertenciam ao sistema de educação de Indiana e, para a realização do estudo, aplicaram-se testes de Matemática e leitura de abrangência local, bem como testes de autoconceito e a mensuração da atitude em relação à escola. O autor acrescenta que todos os testes foram aplicados ao final dos anos de 1983 e 1984.

Com intuito de avaliar os resultados, o pesquisador informa que, quando comparadas as médias dos grupos participantes e dos grupos não participantes, os dados revelaram que houve diferenças significativas em todas as áreas que foram objeto de comparação, quais sejam: :

1. Leitura;
2. Matemática;
3. Autoconceito;
4. Atitude em relação à escola;
5. Afetiva Total.

De acordo com Englehart (2007), o projeto encontrou resultados impressionantes em locais específicos. Nesse sentido, Gilman (1988) constata diferenças positivas em grupos específicos pertencentes ao projeto. O autor salienta que foram identificadas notas mais altas nas disciplinas avaliadas, se comparadas ao ano anterior, quando os estudantes pertenciam a uma turma maior.

Para Gilman (1988), o principal problema de pesquisa do *Prime Time* foi verificar o que poderia ter sido feito para melhorar o desempenho e as atitudes dos alunos em sala de aula no estado de Indiana. Com base nos dados oficiais do projeto, o pesquisador menciona que a amostra trabalhada constituiu-se de 866 alunos de primeira série, pertencentes a três escolas diferentes, sendo que o projeto-piloto do Prime Time consistiu de 142 estudantes, com uma média de 17,5 alunos por sala para o ano letivo.

O autor ainda afirma que, para o alcance do resultado, compararam-se as pontuações dos 142 estudantes participantes com as pontuações dos 190 alunos

não participantes, ou seja, aqueles que pertenciam a classes maiores, de tamanho médio de 23,7 alunos, no ano letivo de 1983 e 1984. Com relação à consecução do objetivo do projeto, Gilman (1988) expõe que o estudo conseguiu comparar a porcentagem média com a pontuação bruta obtida no teste de habilidade em Matemática. Afirma também que, para encontrar o resultado, foi comparada a porcentagem média bruta no teste de habilidade em leitura. O critério utilizado para comparar as pontuações de cada aluno foi a média de 70% do total de pontuações, referente a um total de dez habilidades cursada por cada estudante.

Vandenberg (2012) afirma que a análise dos dados de desempenho dos alunos durante o ano em que o projeto Prime Time foi implantado também forneceu resultados positivos para o programa

Gilman (1988) destaca que o estudo comparou o número médio de habilidades em cada grupo participante. Além disso, ressalta que, depois de uma extensa pesquisa realizada pela equipe do projeto, os pesquisadores adotaram um teste afetivo de 36 itens de variedade Likert para medir atitudes e pontuações dos discentes. Assim, observa-se que resultados significativos foram encontrados, em relação à atitude em sala de aula e, conseqüentemente, na escola.

Para Ruggles (2003), os ganhos experimentados pelos alunos durante o último ano do estudo forneceram evidências de que o ajuste docente em relação a suas instruções os levou a cumprir seus objetivos instrucionais de forma mais eficaz.

A fim de realizar a reflexão sobre os ganhos acadêmicos dos estudantes no último ano de estudo, Gilman (1988) informa que foi utilizado um método de comparação que determinava se os resultados do último ano foram maiores ou menores do que os encontrados nos demais anos. Nesse sentido, o autor ressalta que os resultados do estudo demonstraram algo inesperado: encontraram-se ganhos significativos de desempenho no último ano do estudo. Notou-se que os alunos em ano escolar, de 1987 e 1988, obtiveram pontuações mais altas do que em todas as outras medidas de desempenho dos anos anteriores ao pesquisado. Com base nas análises realizadas no projeto, ela afirma que as melhorias detectadas no primeiro ano não somente se mantiveram, mas também se fortaleceram.

Em contrapartida, Ruggles (2003) salienta que o Prime Time foi elaborado como um projeto de demonstração e, por isso, não seguiu procedimentos rigorosos para a determinação de uma avaliação completa. Wilson (2007) lembra que infelizmente não havia recursos suficientes para alcançar a proporção desejada de 14 alunos por professor, mas as classes foram reduzidas para 18.

Para Hunter (2000), a implementação do projeto não foi rigorosamente controlada e, por essa razão, os resultados foram variados. Já Vandenberg (2012) ressalta que os resultados do projeto não foram muito bem aproveitados devido à grande quantidade de críticas sobre a maneira como ele foi realizado.

2.6.1.4 The California Class Size Reduction Program-CSR

Stecher e Bohrnstedt (2002) relatam que o projeto se iniciou em 1996, com aprovação da legislação de nome SB 1777. Os autores destacam que o intuito do estudo foi reduzir o tamanho das turmas nas séries iniciais.

Gilraine, Macartney e McMilan (2018) afirmam que o Califórnia CSR foi realmente muito grande, sendo a maior reforma educacional liderada pelo estado até hoje. Ademais, mencionam que o projeto foi inspirado no Projeto Star do Tennessee.

Segundo Stecher e Bohrnstedt (2002), as escolas da Califórnia continham 28 alunos em média, e o objetivo era reduzir essa quantidade para o número máximo de 20 alunos por turma. Para tanto, o CSR ofereceu aos distritos participantes aproximadamente 850 dólares para cada estudante do jardim de infância até a terceira série, matriculados em uma classe de 20 estudantes ou menos. Em suma, custou aproximadamente 1,5 bilhão dólares por ano e afetou quase 1,9 milhão de estudantes. Hanushek (1998) ressalta que o Califórnia CSR tinha programas nacionais que vão muito além da despesa anual de mais de 1 bilhão de dólares por ano.

Para Stecher e Bohrnstedt (2002), o projeto produziu resultados relativamente grandes em desempenho para todos os alunos participantes e obteve o dobro de ganhos acadêmicos para alunos de minorias sociais, em comparação aos demais alunos do CSR.

Os autores relatam que o projeto teve a preocupação de considerar não só o efeito da redução no tamanho da classe e o desempenho dos alunos, mas também o impacto da mudança em todos os aspectos do sistema educacional. Conforme apresentado pelos pesquisadores, a implantação quase total do Projeto ocorreu no seu quinto ano, entre os anos de 2000 e 2001. Nesse período, 97% dos alunos de nível *kindergarten* 3 (K-3) foram matriculados em turma de 20 ou menos alunos.

Stecher e Bohrnstedt (2002) ratificam que o projeto teve também o objetivo de verificar qual foi o real impacto da reforma no sistema educacional. A ideia, segundo eles, era separar o que de fato a reforma CSR conquistou em relação às demais reformas que já aconteciam ao mesmo tempo. Na visão dos autores, o projeto exigiu que os distritos e escolas participantes realocassem os fundos e remanejassem a infraestrutura para que o projeto pudesse ser desenvolvido. Diante disso, um problema prematuramente identificado foi a pouca qualificação dos professores disponíveis no mercado e que já atuavam no primário.

Com base em entrevistas com os docentes participantes, os pesquisadores expõem relatos de que, nas disciplinas de Matemática e Linguagens, o tempo de instrução foi maior que em turma normais. Isso ocorreu porque a participação dos alunos em sala de aula aumentou. Além disso, verificou-se, a partir dos depoimentos, a melhora no atendimento de alunos com maior dificuldade.

Além disso, Stecher e Bohrnstedt (2002) informam que, por meio dos depoimentos, foi possível perceber poucas diferenças em questão de currículo, independentemente da quantidade de alunos por turma. Eles afirmam que houve relatos de que, tanto em classes reduzidas como não reduzidas, foi possível cobrir aproximadamente o mesmo número de tópicos para cada disciplina estudada.

Ao longo do desenvolvimento do projeto, os pesquisadores saíram que o apoio público foi substancial, tendo sido capaz até de aumentar o potencial de aprendizagem e de realização dos discentes, além de proporcionar um financiamento adicional para os distritos escolares.

Para Gilraine, Macartney e McMilan (2018), os benefícios principais do California CSR foram: combinar as reduções no tamanho da turma com as mudanças na qualidade do professor e a reclassificar os alunos.

Stecher e Bohrnstedt (2002) ainda ressaltam que a pesquisa provou que os professores atuantes em classes menores eram significativamente mais propensos a se capacitar profissionalmente, se comparados àqueles que atuavam em classes maiores.

Jones (2015) admite que o Califórnia CSR deixou uma lição importante sobre as reduções do tamanho da turma e explica que essas reduções devem ser implantadas com professores preparados e apoiados. Segundo Stecher e Bohrnstedt (2002), a avaliação das demais reformas que aconteceram ao mesmo tempo possibilitou perceber a ligação entre o projeto CSR e o desempenho acadêmico obtido.

Já para Hanushek (1998), é improvável encontrar avaliações confiáveis sobre um projeto a custo de um bilhão de dólares anuais. Nesse sentido, Chingos (2010) expõe que o CSR gerou a redução na qualificação docente, apesar de produzir efeitos positivos com relação ao tamanho da turma.

Jones (2015) afirma que, nos dois primeiros anos do Califórnia CSR, os alunos das classes menores perceberam poucas mudanças em ganhos acadêmicos. De acordo com Hanushek (1998), o projeto dificilmente terá efeito benéfico no desempenho geral dos alunos.

Por fim, conforme Stecher e Bohrnstedt (2002), projetos paralelos mostraram ter influência considerável sobre os resultados. Sobre isso, Jones (2015) afirma que os formuladores de políticas públicas devem implementar esse tipo de pesquisa em escala razoável e dividida em vários anos, caso queiram encontrar resultados significativos.

2.6.1.5 Ontario Canadá – Redução da classe no ano 2007 e 2008

Segundo Bascia (2010), nos anos de 2007 e 2008, o Ministério da Educação da cidade de Ontário no Canadá, por intermédio da Associação Canadense de Educação, realizou um projeto para estudar a política de redução de tamanho de classe nas escolas de Ontário.

Mascall e Leung (2012) explicam que, naquela cidade, o financiamento da educação é baseado na alocação do custo por aluno e o financiamento de cada escola varia diretamente em relação à matrícula. Bascia (2010) informa que o projeto foi realizado por meio da Associação Canadense de Educação, em conjunto com um grupo de pesquisadores contratados pela Universidade de Toronto. Segundo ela, o estudo revisou a literatura, analisou dados estatísticos e conduziu pesquisas de campo em oito distritos escolares de Ontário.

A autora ainda afirma que a equipe de pesquisadores revisou a base de pesquisa e examinou os dados estatísticos coletados pelo Ministério da Educação entre os anos de 2003 e 2004 / 2007 e 2008. O estudo envolveu oito distritos escolares, 24 escolas e 84 salas de aula, e as observações realizaram-se em nível primário, jardim de infância e até a 3ª série.

No que tange à duração do projeto, segundo Bascia (2010), ele foi desenvolvido em quatro anos. A pesquisadora ressalta que o investimento foi suficiente para contratar 1.200 novos professores por ano entre 2004 e 2005 e entre 2007 e 2008. Ademais, as classes foram diminuídas nos primeiros anos para 23 alunos ou menos e, nos demais, para 20 alunos ou menos.

Em busca dos resultados, os pesquisadores levaram em conta a opinião dos professores e dos pais, com a justificativa de que a mudança para esses agentes proporcionaria uma visão mais abrangente. De acordo com a autora, para os professores, as principais mudanças percebidas foram:

1. A oportunidade de melhorar a interação entre professor e aluno;
2. A oportunidade para implantar inovações pedagógicas e práticas profissionais;
3. A melhora na qualidade do relacionamento entre professor e aluno;
4. O maior envolvimento dos estudantes na aprendizagem;
5. O aumento no tempo para ajudar os alunos a experimentar as dificuldades de aprendizagem;
6. possibilidade de realização de atividades intensivas, focadas e orientadas pelo professor;

7. A possibilidade de o professor encorajar a participação individual de cada aluno;
8. O aumento na capacidade de monitoramento dos alunos;
9. A facilidade de deslocamento do professor, em virtude das salas com mais espaço;
10. O ambiente menos estressante.

Ainda segundo Bascia (2010), ao avaliar o ambiente escolar, percebeu-se que, com turmas menores, a segurança infantil se mostrou mais fácil de ser aplicada, gerando uma resposta rápida em caso de emergência. Ela afirma que essa medida facilita a maneira de explicar os procedimentos emergenciais às crianças.

No que se refere aos estudantes com necessidades educacionais especiais, a pesquisadora argumenta que a diminuição da quantidade de alunos por turma nas classes primárias regulares facilitou o aprendizado principalmente em leitura e escrita. Além disso, ratifica que, a partir da realização de uma pesquisa on-line, foi possível verificar a opinião dos pais acerca das mudanças no comportamento de seus filhos. As percepções mais importantes foram:

1. Os filhos pareciam estar aprendendo mais;
2. Os filhos pareciam gostar mais da escola e se sentir mais à vontade na sala de aula;
3. Os pais perceberam que os filhos estavam com menos dificuldade de aprendizagem;
4. Os pais perceberam menos desafios comportamentais;
5. Os pais perceberam interações sociais mais positivas e maior envolvimento na aprendizagem.

A iniciativa de redução de tamanho de turma realizada em Ontário cumpriu grande parte de sua promessa. Bascia (2010) lembra que, por meio da oferta de um ambiente em que os professores pudessem interagir com os alunos individualmente com mais frequência, bem como usar uma variedade de estratégias instrucionais, o resultado se mostrou promissor. A pesquisadora também enfatizou que o resultado do estudo mostrou que os estudantes que aprendem mais são mais engajados e menos indisciplinados.

Flessa (2012) argumenta que a redução do tamanho da classe primária de Ontário está entre duas tendências políticas. Para esse autor, a pesquisa é uma política centralizada e provincial, e as diretrizes e resultados estão fortemente

ligados a incentivos e sanções. Nesse sentido, a tendência de centralização do projeto de Ontário acontece em razão do reconhecimento de que qualquer mudança na escola é mais desafiadora que qualquer nova política.

Mascall e Leung (2012) acrescentam que os distritos com maiores recursos foram capazes de coordenar a implantação local do projeto de redução de classes. Em contrapartida, os desprovidos de recursos foram incapazes de gerenciar a implementação. Dessa forma, a redução de classe não foi tão positiva sobre os professores e alunos como poderia ter sido.

2.6.2 Europa

Os estudos apresentados a seguir, foram os encontrados no continente europeu.

2.6.2.1 Estudo Turquia

Segundo Yigit *et al.* (2017), o objetivo do estudo realizado na Turquia foi examinar as percepções dos alunos do Ensino Médio sobre o ambiente de aprendizagem em sala de aula no curso de Ciências, levando em consideração as variáveis de localização da escola e o tamanho da turma.

Ece e Neslihan (2014) salientam que, na Turquia, a pré-escola é dada principalmente em instituições formais, exclusivas para isso. Erdal (2010) argumenta que o objetivo do ensino primário turco é garantir que toda criança adquira conhecimentos, habilidades, comportamentos e hábitos que as tornem bons cidadãos.

Sobre isso, Yigit *et al.* (2017) informam que o estudo optou por utilizar o questionário de Avaliação do Ambiente de Aprendizagem Construtivista - AAAC. Os autores afirmam que essa ferramenta foi utilizada para mapear as percepções dos estudantes sobre o ambiente de aprendizagem em sala de aula. A amostra de dados incluiu 1.882 alunos das grades 6 a 8, selecionados aleatoriamente dentre 20 escolas em uma província do norte de Turquia.

Os pesquisadores expõem que, para a avaliação dos dados, utilizou-se a análise multivariada, cujo intuito era identificar as diferenças de localização das classes e as percepções dos alunos sobre os ambientes de aprendizagem em uma sala de aula construtivista. Os dados para essa pesquisa foram coletados em maio de 2011, e todos os instrumentos foram administrados no horário de aula regular, sob a supervisão do professor de Ciências num período de 30 minutos.

Na visão de Yigit *et al.* (2017), a análise preliminar dos dados revelou que houve diferenças significativas entre as percepções dos alunos de escolas rurais e urbanas, tanto de turmas pequenas quanto de turmas grandes. Mais especificamente, os estudantes de turmas pequenas e de escolas rurais relataram melhoria na aprendizagem, em comparação àqueles que estavam em turmas grandes e escolas urbanas.

Em contrapartida, Ece e Neslihan (2014) ressaltam que, na Turquia, no Ensino Fundamental, as salas são superlotadas, aspecto muito criticado pelos educadores do país. Analisando esse contexto, os pesquisadores afirmam que as instituições formais de ensino não dispõem de estruturas organizacionais para que as crianças sejam colocadas em salas menores.

2.6.2.2 Long-Term Effects of Class Size -Suécia

De acordo com Fredriksson, Ockert e Oosterbeek (2011), o estudo feito na Suécia avaliou os efeitos em longo prazo do tamanho das turmas nas escolas primárias entre os anos de 2007 a 2009. Para tanto, utilizaram-se dados administrativos suecos, com o objetivo de explorar a variação no tamanho das classes e criar uma regra de tamanho máximo para elas.

Os autores expõem que o sistema escolar obrigatório na Suécia tem várias camadas organizacionais, sendo a unidade primária do sistema, a escola. Assim, há pelo menos um distrito escolar por município, e cada um deles administra formalmente as escolas obrigatórias. Além disso, os municípios não têm autonomia de decisão sobre os recursos repassados pelo governo central, tampouco sobre os regulamentos educacionais.

A decisão sobre os recursos e regulamentos, conforme Fredriksson, Ockert e Oosterbeek (2011), cabe aos conselhos escolares, e foram eles os responsáveis pela implantação do projeto. Para a realização do estudo, os conselhos decidiram que o tamanho máximo das turmas seria de 25 alunos para turmas do primário e de 30 alunos para o secundário.

As bases de dados utilizadas neste estudo originam-se do projeto chamado UGU, executado pelo departamento de educação da Universidade de Gotemburgo. Os dados dessa pesquisa contêm notas de testes cognitivos feitos para crianças com 13 anos de idade e uma amostra de aproximadamente 10% de crianças nascidas entre 1967, 1972 e 1982.

Para a utilização destes dados, os pesquisadores combinaram as informações de registros mantidas pelo Departamento de Estatística da Suécia sobre o tamanho da turma, com as informações dos pais e os resultados de médio e de longo prazo.

A partir da avaliação dos resultados, Fredriksson, Ockert e Oosterbeek (2011) concluem que, em classes menores, houve benefícios para os resultados cognitivos. Eles afirmam, ainda, que turmas menores proporcionaram o aumento dos salários dos alunos no pós-escola. Ademais, analisou-se o custo-benefício, e descobriu-se que uma redução no tamanho das turmas de 25 alunos para 20 alunos resultou numa taxa interna de retorno de 20%.

Em primeira linha, os autores consideram que os resultados aferidos sugerem que uma redução no tamanho das classes aumenta os anos de escolaridade. A estimativa é apenas estatisticamente significativa para as mulheres. Além disso, identificou-se que as variações do tamanho da classe não afetaram a probabilidade de esses alunos encontrarem emprego no pós-escola.

No entanto, Biesta e Säfström (2011) salientam que, na Suécia, as escolas são um dos negócios mais lucrativos do país, com menor risco e maior retorno. Nesse sentido, Åberg-Bengtsson (2009) ressalta que, naquele país, os gastos altos por aluno e a diminuição da turma em povoados muito pequenos aumentam o risco de fechamento de instituições de ensino..

2.6.2.3 Primary and Secondary Schools in England and Wales

Blatchford, Basset e Brown (2011) afirmam que a atual política do governo da Inglaterra e do País de Gales é manter um total de no máximo 30 alunos por sala de aula, com idades entre quatro e sete anos. Na Escócia, o objetivo é ter ainda menos alunos por turma. De acordo com os autores, o principal intuito desse trabalho foi pesquisar, de forma sistemática, as relações entre o tamanho da turma e os processos em sala de aula.

Com intuito de proporcionar a melhoria da educação na Inglaterra e no País de Gales, o estudo comparou os efeitos sobre o engajamento em sala de aula e a interação entre professor e aluno. Os pesquisadores ressaltam que a intenção era verificar se os resultados variavam de acordo com o nível de realização dos alunos na comparação entre escolas primárias e secundárias.

Blatchford, Basset e Brown (2011) informam que foram realizadas observações sistemáticas com 686 alunos em 49 escolas, utilizando métodos multiníveis para examinar a relação entre o tamanho da classe e as medidas aplicadas, controlando variáveis intervenientes que potencialmente pudessem interferir na realização do aluno. Ademais, esclarecem que os resultados encontrados nos níveis primários e secundários mostraram que, nas turmas menores, os alunos receberam mais atenção individual por parte dos professores e, além disso, houve um nível maior de interação. Os pesquisadores salientam que o engajamento em sala de aula diminuiu nas classes maiores, mas, ao contrário das expectativas dos cientistas, isso foi particularmente recorrente para alunos com menor desempenho no nível secundário.

A conclusão foi a de que, em turmas menores de nível secundário, os estudantes com baixa escolaridade puderam se beneficiar em termos de atenção individual e de facilitação da aprendizagem. Para Beckmann, Cooper e Hill (2009), tanto na Inglaterra quanto no País de Gales, a diminuição do número de alunos por turma pode contribuir para que os professores produzam melhores resultados na aprendizagem discente.

2.6.3 América do Sul

A seguir, apresentam-se apresentados os trabalhos realizados na América do Sul.

2.6.3.1 O efeito do tamanho da turma sobre o desempenho escolar: Uma Avaliação do impacto da “enturmação” no Ensino Fundamental do Rio Grande do Sul

Segundo Camargo (2014), o objetivo do estudo foi verificar o impacto do tamanho das turmas no desempenho acadêmico dos alunos das escolas públicas do Rio Grande do Sul. Para isso, o método utilizado foi o de variáveis instrumentais e de regressão com descontinuidade, e a amostra utilizada foi a referente ao ano de 2007.

A pesquisadora salienta que o governo do estado do Rio Grande do Sul, em 2007, promoveu algumas medidas cujo intuito era a melhoria da qualidade da educação estadual. Dentre as medidas, destacam-se:

1. Reorganizar a rede de ensino recompondo as coordenadorias regionais de educação;
2. Remanejar professores que estavam em apoio pedagógico, para a atividade de docência;
3. Reorganizar as turmas.

De acordo com a pesquisadora, pouco tempo após essa medida, as coordenadorias regionais de educação implementaram um estudo para analisar o número de alunos por classe nas escolas públicas estaduais. Os dados foram obtidos a partir de coletas de escola por escola, e de município por município.

Por meio da coleta de dados, ainda segundo Camargo (2014), observou-se que, das quase 54 mil turmas do estado, pelo menos 5 mil continham um número baixo de alunos. Das 2.856 escolas analisadas na época, 745 delas reorganizaram suas turmas, algo em torno de 26% do total. As turmas foram reorganizadas seguindo o parecer do Conselho Estadual de Educação do Rio Grande do Sul e divididas com base nos seguintes parâmetros: Máximo de 50 alunos por turma;

1. Máximo de 40 alunos para o Ensino Fundamental por turma
2. Máximo de 45 alunos para o Ensino Médio por turma.

Os resultados obtidos, tal como informa Camargo (2014), utilizaram o modelo causal de Rubin para avaliar seus impactos. Com relação às bases de dados, utilizaram-se os microdados da Prova Brasil de coordenação do INEP, e os resultados foram divulgados por escola, no Sistema de Avaliação do Rendimento Escolar do estado do Rio Grande do Sul. A autora ressalta, ainda, que os dados utilizados para a realização do projeto se referem aos alunos da quarta e da oitava séries do Ensino Fundamental das escolas públicas da rede estadual do estado em questão.

Com base nos achados da pesquisa, Camargo (2014) informa que a maior parte das turmas tinha entre 20 e 30 alunos, tanto para a quarta série, quanto para a oitava série. Ademais, a autora constatou que o tamanho médio das turmas encontradas foi de 25 alunos por classe para ambas as séries. Assim, apesar da maioria dos coeficientes estimados apresentarem um sinal negativo, indicando uma associação negativa entre o tamanho da turma e o desempenho dos alunos, os coeficientes não são estatisticamente significantes. Diante disso, afirma-se que o estudo não encontrou evidências estatisticamente significantes de que a reorganização das turmas, ocorrida na metade do ano de 2007 tivesse qualquer impacto no desempenho acadêmico dos alunos.

Em contrapartida, Hanushek (1998) afirma que é provável haver situações com professores específicos, grupos específicos de alunos e assuntos específicos, nas quais pequenas turmas podem ser muito benéficas para o desempenho discente.

Para Oliveira, Menezes-Filho e Komatsu (2018), com o aumento dos investimentos em educação no Brasil nos últimos anos, bem como com a constante baixa na qualidade da educação, a questão principal é saber quais fatores influenciam o desempenho escolar que poderiam ser aplicados.

Segundo Camargo (2014), um estudo futuro poderia ser realizado para investigar se, em um intervalo de tempo maior, essas reorganizações de turmas no Rio Grande do Sul teriam um resultado mais significativo no desempenho acadêmico dos estudantes pesquisados.

2.6.3.2 Custos e Efetividade de Políticas de Redução do Tamanho da Classe e Ampliação da Jornada Escolar: Uma aplicação de Estimadores de *Matching*

Trata-se de um projeto apresentado por meio de uma dissertação de mestrado em Economia, que obteve o segundo lugar no 31º Prêmio BNDES de Economia. A pesquisa é da autoria de Jaqueline Maria de Oliveira e foi orientada pelo professor Naércio Aquino de Menezes-Filho, no Rio de Janeiro, em 2008.

Oliveira (2008) afirma que concorreram ao Prêmio BNDES de Economia 44 trabalhos, inscritos por 14 universidades brasileiras de pós-graduação em Economia. O objetivo do estudo foi verificar se a redução do tamanho da classe, bem como a ampliação da jornada escolar teriam influências sobre o desempenho na disciplina de matemática dos estudantes da 4ª série do Ensino Fundamental nas escolas brasileiras. As escolas avaliadas no trabalho estavam localizadas apenas nas áreas urbanas.

De acordo com a autora, o trabalho se valeu da metodologia de *Propensity Score Matching* (PSM) para o tamanho da classe e o estimador *Nearest Neighbor Matching* (NNS) para a jornada escolar. A base de dados foi o Sistema de Avaliação da Educação Básica (SAEB), realizado pelo INEP, referente ao ano de 2005.

Segundo Garcia *et al.* (2016), a educação básica no Brasil é monitorada principalmente pelo Índice de Desenvolvimento da Educação Básica (IDEB). Esse índice é importante, pois combina o desempenho dos estudantes no Sistema de Avaliação da Educação Básica (SAEB) com a Prova Brasil.

Oliveira (2008) ressalta que os resultados da análise dos dados obtidos pelo projeto mostraram um efeito positivo do aumento da jornada no desempenho escolar, sendo que o aumento de quatro horas de jornada para cinco horas de jornada gerou um aumento de 8,36 pontos na proficiência em Matemática ao se avaliar as notas obtidas pelos alunos.

Para a pesquisadora, a diminuição de alunos por turma, de 38 alunos para 30, gerou um efeito positivo, estimado em 10,67 pontos na distribuição de proficiência. Ademais, os achados da pesquisa revelam que ampliação da jornada escolar alcançou um resultado positivo na relação custo- benefício, ao passo que diminuição de alunos se mostrou mais custosa.

Nesse sentido, a autora sinaliza que as políticas de aumento da jornada escolar dos estudantes são mais efetivas do que políticas de redução do tamanho da classe, quando as turmas têm até 34 alunos. Do contrário, a pesquisadora afirma que a redução do tamanho da classe se mostra mais efetiva.

A pesquisadora destaca que os dados obtidos favorecem a conclusão de que os recursos escolares têm influência no desempenho escolar. No entanto, apesar de ambas as políticas terem impacto sobre o desempenho escolar, é preciso considerar os custos de implantação da política mais adequada.

2.6.3.3 Tamanho das Classes na Rede Estadual Paulista: A Gestão de Rede Pública à Margem das Desigualdades Educacionais

Travitzki e Cássio (2017) assinalam que o objetivo de sua pesquisa foi verificar se os tamanhos das classes nos anos finais do Ensino Fundamental da rede estadual paulista influenciavam no desempenho dos estudantes na Prova Brasil. Isso posto, o trabalho investigou a influência da variável “tamanho da classe” para a disciplina de Matemática, tendo, como referência, o ano de 2013.

Os autores afirmam que, a partir do resultado encontrado, foi apresentada uma série histórica sobre o tema. Além disso, discutiram-se tendências para política de tamanho de classes para o futuro. Na pesquisa, utilizou-se o método de pareamento em escore de propensão generalizado. As análises dos dados foram realizadas por meio do *software* R, de livre acesso e código aberto.

Segundo os pesquisadores, os dados utilizados continham 337.524 registros, representando 418.589 alunos, distribuídos em 12.582 classes, em um total de 3.473 escolas. A análise valeu-se da nota de Matemática como variável-resposta, ao passo que a variável de tratamento foi o tamanho da classe. Para a construção da série histórica, usou-se o Censo Escolar para o nível das classes e o Cadastro de Escolas da SEE-SP para o nível das escolas via lei de acesso à informação, referente aos meses de maio de 2014 a maio de 2016. O mês de maio foi escolhido devido ao fato de ser o mês de referência para consolidação dos dados do Censo Escolar no estado.

Para os autores, a análise dos dados proporcionou verificar a existência de uma relação significativa entre o tamanho da classe e o desempenho escolar. Eles afirmam que tal relação é não linear, ou seja, sugere que há um ponto ótimo para o

tamanho das classes em relação ao aumento do desempenho escolar. Isso prova a necessidade de as classes não serem nem muito grandes, nem muito pequenas.

Ademais, Travitzki e Cássio (2017) informam que o ponto ótimo varia em diferentes contextos socioeconômicos. De forma geral, os resultados obtidos com a amostra total de alunos se assemelham mais aos resultados obtidos para o terço superior de nível socioeconômico do que para os dois terços inferiores.

De acordo com Barros *et al.* (2001), no Brasil, o nível educacional é fortemente ligado ao nível socioeconômico e, se nada for feito para melhorar a qualidade da educação, os nascidos pobres hoje serão os pobres de amanhã.

Conforme os dados apurados, Travitzki e Cássio (2017) ressaltam que as políticas de tamanho de turma da rede paulista, ainda que fossem executadas com perfeição, atingiriam apenas um terço dos estudantes, precisando haver um melhor controle de variáveis intervenientes. Com base nos dados coletados, observou-se uma relação não linear entre o tamanho da classe e o desempenho escolar, em que é possível visualizar o ponto ótimo. Assim, o ponto considerado ótimo para os anos finais do Ensino Fundamental é de 34 alunos, que corresponde exatamente à média do tamanho da classe. Vistas dessa forma, as políticas educacionais paulistas se mostraram consistentemente fundamentadas em evidências, pelo menos quando se fala em média.

Segundo Barros *et al.* (2001), no Brasil, os indicadores educacionais estão abaixo dos padrões mundiais, e o que mais preocupa os autores é que, entre os estudantes de baixa renda, o efeito é ainda pior. Por isso, estudos sobre as diminuições do número de alunos por classe, como meio de garantir o aumento do desempenho escolar, são muito importantes.

2.6.3.4 Efeitos de Tamanho da Sala no Desempenho dos Alunos: Evidências para São Paulo

Matavelli (2018) afirma que o objetivo do estudo foi avaliar o impacto do número de alunos por classe no desempenho acadêmico. Para isso, utilizaram-se as notas de Português e Matemática da Prova Brasil do ano de 2018. Segundo a autora, as séries avaliadas foram o quinto ano e o nono ano, sendo o quinto ano municipal da cidade de São Paulo e o nono ano estadual de São Paulo.

Barros *et al.* (2001) destacam que é surpreendente como muitos tipos de determinantes, além da qualidade dos professores, têm influência no desempenho acadêmico. Nesse sentido, levantam a dúvida acerca da importância da diminuição do número de estudantes por turma nesse debate.

Segundo Matavelli (2018), para a obtenção do resultado, utilizou-se a metodologia de regressão descontínua “*Fuzzy*”. Buddelmeyer e Skoufias (2004) explicam que a regressão descontínua “*Fuzzy*” resolve temas causais a partir de dados não experimentais. O modelo considera a probabilidade de um indivíduo ser parte do grupo tratado.

Matavelli (2018) ressalta que os resultados encontrados mostram não existirem evidências estatisticamente significantes de que o tamanho da sala tenha impactado nas notas dos alunos. A fim de aumentar a confiabilidade do resultado, a pesquisadora ainda utilizou dados do quinto e nono ano estaduais de Minas Gerais e de Santa Catarina, bem como o quinto ano estadual de São Paulo, chegando exatamente à mesma conclusão.

2.6.4 África

A seguir, apresentam-se os trabalhos encontrados no continente Africano.

2.6.4.1 Effects of Assignment and Class Size on Secondary School Students Achievement in Mathematics. (Ogbadigbo – Nigéria)

Segundo Sule (2016), o projeto verificou os efeitos da atribuição e do tamanho da turma no desempenho acadêmico dos alunos de Matemática em escolas secundárias previamente selecionadas da área governamental de Ogbadigbo, pertencente ao estado de Benue, na Nigéria.

Sooter (2013) afirma que, a partir do ano de 2013, o governo federal nigeriano, com base em sua política nacional de educação, reconheceu oficialmente a necessidade de investir em educação, e isso deu origem a uma expansão sem precedentes na educação do país. Sule (2016) lembra que o estudo adotou, como metodologia, o delineamento experimental e a análise de dados e, como instrumento de coleta, foram utilizados um questionário e um teste de desempenho em Matemática.

O autor esclarece que a amostra do estudo foi composta por 150 alunos do Ensino Médio, selecionados de forma aleatória em escolas secundárias. Os resultados encontrados mostraram que os estudantes de turmas pequenas tiveram desempenho significativamente melhor, em comparação com os de turmas maiores. A pesquisa também revelou que os alunos expostos a tarefas regulares tiveram um desempenho melhor do que aqueles que não foram expostos.

Para Sooter (2013), a diminuição do número de alunos por turma pode ser algo positivo para a melhoria da educação nigeriana. O pesquisador argumenta que as escolas da Nigéria atualmente têm professores de nível acadêmico baixo.

Nesse sentido, Sule (2016) ressalta que a recomendação, a partir dos resultados encontrados, foi a de que mais salas devessem ser construídas, com número máximo de 30 estudantes, facilitando o ensino e a aprendizagem. Para o autor, a pesquisa chegou à conclusão de que é necessário fazer com que o número de alunos a serem ensinados na disciplina de Matemática seja moderado, para que o professor seja capaz de gerenciar a classe de forma eficaz, dando atenção individual a cada aluno e principalmente àquele com dificuldade de aprendizagem.

2.6.4.2 Impact of Class Size on Learning, Behavioral and General Attitudes of Students in Secondary Schools in Abeokuta, Ogun State Nigeria

Yusuf, Onifade e Bello (2016) expõem que o projeto foi realizado com intuito de investigar o impacto do tamanho das classes nas atitudes dos alunos. Segundo eles, foram utilizadas quatro escolas secundárias selecionadas na cidade de Abeokuta no estado de Ogun na Nigéria.

Os autores afirmam que o objetivo do estudo foi verificar o possível efeito do tamanho da classe sobre as atitudes de aprendizagem e o comportamento em geral dos estudantes. Segundo os pesquisadores, esse efeito também foi correlacionado com outras variáveis intervenientes. Eles explicam que foram administrados questionários para o levantamento dos dados, aplicados a um total de 360 alunos, selecionados de maneira uniforme. Sendo assim, das quatro escolas selecionadas, duas eram privadas, e as outras duas, públicas.

O estudo verificou 10 hipóteses e chegaram a esses principais resultados:

1. O tamanho da turma é altamente significativo sobre o impacto das atitudes dos alunos do ensino secundário na aprendizagem;
2. Os alunos foram afetados positivamente na pontualidade, motivação e na participação;
3. O resultado revelou que o ruído excessivo é uma característica comumente associada a classes grandes;
4. Turmas pequenas geram maior motivação dos alunos para frequentar a escola.

Yusuf, Onifade e Bello (2016) afirmam que, na Nigéria, em geral, o tamanho das classes nas escolas secundárias está muito acima dos 30 ou 40 recomendados. A maioria das escolas excede o número de 50 alunos por turma, sendo que algumas têm 80 alunos. Ademais, é possível também encontrar classes com 100 alunos ou mais.

Com o número de alunos por sala de aula num patamar tão alto, Sooter (2013) considera que os objetivos das políticas de educação nigerianas são voltados basicamente ao ensino de normas sociais, ao incentivo à criatividade e ao senso de cooperação. Para o autor, diante desse cenário, a implantação de políticas públicas específicas fica comprometida.

2.6.4.3 Influência do tamanho da Turma na Sala de Aula dos Alunos Com Relação à Disciplina, Engajamento e Comunicação: Um Estudo de Caso de Escolas Secundárias no Estado de Ekiti, Nigéria.

Segundo O. Ajayi, Audu e E. Ajayi (2017), a ideia do estudo foi verificar se há a influência do tamanho da turma no engajamento, na disciplina e na comunicação dos estudantes em sala de aula nas escolas secundárias do estado de Ekiti, na Nigéria. Os autores informam que a amostra pesquisada contava com 128 professores secundários seniores, selecionados dentre uma população de professores de 4.529 pessoas. Sendo assim, a pesquisa adotou um delineamento descritivo de pesquisa *Survey* e usou, como instrumento de captação de informação, um questionário de envolvimento, disciplina e comunicação da sala de aula. Segundo os pesquisadores, esse instrumento tem um grau de confiabilidade de 78%.

O. Ajayi, Audu e E. Ajayi (2017) também destacam que foram formuladas quatro questões de pesquisa e três hipóteses nulas para a orientação do estudo. As questões de pesquisa foram respondidas utilizando a pontuação média e o desvio padrão, enquanto as hipóteses nulas foram testadas usando o chi-quadrado.

Os autores afirmam que o estudo revelou que o tamanho da turma tem influência significativa na disciplina, no envolvimento e na comunicação da turma do ensino secundário superior. Ademais, ressaltam que, a partir dos resultados encontrados, recomendou-se que as escolas adotassem um número máximo de um professor para até 40 estudantes, o que é considerado uma turma pequena. A ideia era, de acordo com os pesquisadores, incentivar o maior envolvimento e a melhor comunicação em sala de aula.

O. Ajayi, Audu e E. Ajayi (2017) relatam que, na pesquisa, foram categorizadas como tamanho de classe pequena as classes de 1 a 40 alunos por turma para um professor; acima de 41 alunos, as turmas foram definidas como grandes. Para a obtenção dos resultados, o estudo examinou a influência do tamanho da turma em sala de aula do ensino secundário, avaliando a disciplina, o engajamento e a comunicação dos estudantes.

Além disso, o trabalho adotou um delineamento descritivo, em pesquisa *Survey* de escala Likert com cinco respostas, e procurou obter informações que divulgassem fenômenos comuns, perguntando ao indivíduo sobre suas percepções, atitudes, comportamentos e crenças.

Sobre os resultados, O. Ajayi, Audu e E. Ajayi (2017) informa que houve evidências de que o tamanho da turma influencia significativamente a disciplina, o engajamento e a comunicação. Por implicação, os autores afirmam que, quando se coloca em pauta a turma grande, é recorrente uma classe ineficaz em disciplina, engajamento e comunicação. Contudo, ressaltam que, por meio do estudo, foi possível perceber que o ensino insuficiente, a infraestrutura inadequada e as restrições dos recursos orçamentários prejudicam mais que o tamanho da classe em si.

Em contrapartida, Sooter (2013) pondera que a diminuição de alunos por turma no contexto nigeriano encontra entraves, pois a efetivação dessa política necessitaria do emprego de mais professores. Desse modo, a proporção de alunos para professor na educação infantil da Nigéria depende dos recursos disponíveis.

O. Ajayi, Audu e E. Ajayi (2017) informam que, para a implantação dos resultados do projeto da Nigéria, os administradores de escolas públicas e privadas deveriam ser encorajados a adotar um máximo de 40 estudantes para cada professor, em cada disciplina, além de envolvimento e comunicação eficazes em sala de aula.

Nesse sentido, Sooter (2013) ressalta o fato de o setor educacional nigeriano ser dominado por empresários com o objetivo de lucrar, demonstrando que a diminuição de alunos por turma pode encontrar diversas barreiras. Além disso, o autor afirma que a proporção de 25 alunos para um professor tem dificuldades de implantação devido à falta de supervisão e monitoramento.

O. Ajayi, Audu e E. Ajayi (2017) sinalizam a necessidade de o governo aumentar o financiamento orçamentário do setor educacional, com o intuito de melhorar as infraestruturas escolares e, assim, contratar docentes mais qualificados, construindo mais salas de aulas, a fim de controlar a incidência de salas grandes.

2.6.5 Oriente Médio

O estudo a seguir foi o único encontrado no Oriente Médio, sobre o tema aqui proposto.

2.6.5.1 É menos Realmente Mais? A Relação Entre o Tamanho da Turma e a Realização Educacional em Israel

Shafrir, Shavit e Blank (2016) revelam que, no ano de 2014, uma organização de pais de alunos das escolas primárias israelenses, organizou uma campanha para reduzir o número de alunos por turma. Os autores explicam que o esforço foi apelidado de “Protesto da Sardinha”, pois os fatos da época mostravam que, segundo os pais, as salas superlotadas prejudicavam os alunos.

Segundo os pesquisadores, em Israel, até o ano de 2008, o tamanho da classe era de 40 alunos em todos os níveis. Eles afirmam que, a partir desse ano, uma resolução do governo israelense pediu a redução gradual do tamanho das turmas para 27 alunos no ensino primário e para 32 alunos no ensino médio.

O objetivo da pesquisa, de acordo com Shafrir, Shavit e Blank (2016), era o de responder à questão do impacto do tamanho da turma, a partir de uma análise

hierárquica das pontuações dos alunos nos exames israelenses de Meitzav, para o período de 2006 e 2009.

Para isso, a base de dados utilizada teve como fundamento a análise dos exames de Meitzav. Esses exames são administrados pela Autoridade Nacional de Mediação e Avaliação em Educação, com o intuito de determinar o grau em que os alunos satisfazem os requisitos curriculares israelenses em quatro disciplinas centrais, quais sejam: Inglês, Artes Linguísticas (Hebraico ou Árabe), Matemática, Ciência / Tecnologia.

Além dos resultados do exame, Shafrir, Shavit e Blank (2016) examinaram uma base de arquivos do Ministério da Educação, que continha dados sobre os alunos (nível de escolaridade dos pais, origem étnica entre outros). A fim de encontrar os resultados dos modelos hierárquicos, eles utilizaram a metodologia de *Statistical Package for the social Sciences - SPSS Mixed Procedure*. Segundo os pesquisadores, a análise foi realizada usando três modelos alinhados:

Modelo I: incluiu apenas o tamanho da turma;

Modelo II: controlou as variáveis de fundo do aluno e a composição socioeconômica da classe, conforme indicado pela educação dos pais. O controle dessas variáveis tornou possível estimar qual parte da relação entre o tamanho e a realização é explicada pela alocação seletiva dos alunos com certos atributos de classes de tamanhos diferentes;

Modelo III: analisou se existiam diferenças na relação entre tamanho de classe e a realização entre diferentes dados, e entre alunos com maior desempenho e alunos com menor desempenho.

Os autores concluem que, embora a relação entre o tamanho da turma e a realização seja realmente positiva, quando avaliado que os alunos de classes maiores têm um desempenho melhor, o resultado fica um tanto quanto subjetivo, pois acaba sendo afetado por outros fatores ou por variáveis intervenientes. Nesse sentido, Ben-David e Kimhi (2017) afirmam que, para medir o desempenho por aluno em Israel, há a necessidade de levar em consideração a variável interveniente de gasto por aluno, pois o país é um dos que mais gasta nesse quesito. Por isso, o controle do número de alunos por turma pode ser uma política pública, em termos de custos.

Por fim, Shafrir, Shavit e Blank (2016) afirmam que, quando se controlou o resultado por conquistas anteriores ao estudo e para os níveis educacionais de Israel, não se encontrou nenhuma relação significativa entre o tamanho da turma e o desempenho.

3 MÉTODO

Nesta seção, apresentam-se as características técnicas da pesquisa empírica, com destaque para o tipo de pesquisa realizada, o período de referência, a abrangência dos dados, a abordagem das variáveis e a técnica a ser utilizada na apuração dos dados.

3.1 Tipo de pesquisa

As pesquisas em educação, segundo Muijis (2011), são normalmente divididas em duas principais vertentes, os métodos quantitativos e os métodos qualitativos.

Punch e Oancea (2014) explicam que o método quantitativo é aquele em que, na pesquisa empírica, os dados são apresentados na forma de números ou na forma de medições. Já no método qualitativo, os dados não são apresentados em números, e sim em palavras, mas podem aparecer também em outras formas, por exemplo, imagens.

Segundo Willis (2008), é imprescindível que algumas características sejam observadas no trabalho científico. Dentre elas, destacam-se: :

1. **Objetivo:** é o modo de abordar a pesquisa. Ele deve ser claro, a fim de evitar a contaminação dos dados com qualquer forma de subjetividade que possa levar a determinadas objeções.
2. **Empírico:** é a forma de tratamento dos dados. Eles devem ser analisados com métodos estatísticos, de forma que possam ser replicados e testados por outros pesquisadores.
3. **Linear, Planejado e Estruturado:** é o cronograma de execução do estudo. Deve ser projetado com antecedência e executado de acordo com o plano.
4. **Preferência por Métodos Experimentais Controlados:** para o autor, é sempre melhor conduzir experimentos em condições controladas, caso não seja possível, ele aconselha usar métodos correlacionais que tentem se aproximar o máximo possível do real.

5. Busca pelas Leis e Regras de Comportamento Universais: Assim, segundo o autor, é possível prever e controlar como um organismo tende a se comportar.

Diante do exposto, este trabalho optou pelo uso do método quantitativo e apresenta as características que Willis (2008) determinou como imprescindíveis.

Para o desenvolvimento da presente pesquisa, serão utilizados dados secundários de primeira mão, disponibilizados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP, compostos pelos microdados dos Questionários do Diretor, do Professor, da Escola e do Aluno (tanto para os Anos Iniciais quanto para os Anos Finais do Ensino Fundamental) da Prova Brasil, para o ano de 2017.

Sob essa perspectiva, Boeije e Hox (2005) afirmam que os dados primários são coletados com foco no problema de pesquisa em questão, a partir do uso de procedimentos que melhor se adaptam ao problema de pesquisa.

De acordo com Franco (2021), esses dados coletados pelos próprios pesquisadores por meio de atividades experimentais ou observacionais. Para o autor, nem sempre os dados de primeira mão são os mais apropriados para a investigação de determinada situação, sendo os dados secundários também uma fonte muito importante.

Nesse sentido, Johnston (2013) demonstra que a análise de dados secundários coletados por outra pessoa, para algum outro propósito, oferece uma opção viável e econômica para os pesquisadores realizarem suas análises empíricas.

Complementando Johnston (2013), Hair Jr, Paige e Brunsveld (2020) afirmam que, apesar de os dados secundários serem originalmente coletados para algum outro propósito, podem servir para um projeto de pesquisa subsequente. Eles podem aparecer em diversos formatos, como tabelas, gráficos, imagens ou mesmo na forma bruta. Contudo, isso não impede que ofereçam uma visão significativa e vantajosa para estudos posteriores.

3.2 Abrangência de dados

A pesquisa aqui proposta tem como abrangência o universo de escolas públicas do Ensino Fundamental brasileiro, tanto para os Anos Iniciais quanto para os Anos Finais.

Segundo dados do Ministério da Educação, o SAEB 2017 é composto de mais de 70 mil escolas e conta com a participação de aproximadamente 5,4 milhões de estudantes, avaliando sua proficiência nas disciplinas de Língua Portuguesa e Matemática. Além disso, fornece questionários capazes de medir toda a estrutura educacional e o nível socioeconômico destes estudantes.

3.3 Período de Referência

A aplicação da Prova Brasil do ano de 2017 ocorreu entre outubro e novembro de 2017 em todas as unidades da Federação. Os dados colhidos se referem ao ano de 2017.

3.4 Método de Preparação e Análise dos dados

Para a busca do objetivo proposto, optou-se pelo uso de um modelo estatístico chamado de *Propensity Score Matching*. Segundo Thoemmes (2012), esse modelo é uma ferramenta para inferência causal em estudos de característica não randomizada, que permite a observância de grandes conjuntos de variáveis intervenientes.

Para chegar aos resultados deste trabalho, é necessário controlar diversas variáveis intervenientes para que não influenciem o resultado a ser encontrado. Assim, ao utilizar o *Propense Score Matching*, pode-se analisar um conjunto grande de variáveis intervenientes ao mesmo tempo.

Thoemmes (2012) explica que o modelo estatístico serve para ajustar o efeito de um tratamento em variáveis intervenientes que podem contaminar o resultado. O autor defende que o modelo é uma alternativa habitualmente utilizada para o ajuste de regressão.

A lógica desse modelo, segundo o autor, está justamente em alcançar o equilíbrio das variáveis intervenientes em um único escore de propensão. Para ele, a propensão é a probabilidade das variáveis intervenientes de receberem tratamento. Isso posto, a expressão matemática que define o modelo é: $e(x) = P(Z=1|X)$, sendo, $e(x)$: o *propensity score*, P : a probabilidade, $Z = 1$: a indicação de tratamento com valores 1 para tratamento e 0 para controle, “|”: a condicional e o X : o conjunto de variáveis intervenientes.

Ao encontrar o equilíbrio das variáveis intervenientes, evita-se a contaminação dos resultados, garantindo que apenas o número de alunos por turma tenha influência no desempenho escolar. De acordo com Thoemmes (2012), o *Propensity Score Matching*, ao alcançar o equilíbrio entre as variáveis intervenientes, faz com que elas não possam mais intervir no efeito que se pretenda alcançar. Ademais, o autor destaca que, para o uso desse modelo, o pesquisador em questão seleciona um conjunto de variáveis intervenientes de pré-teste.

A fim de definir quais variáveis seriam consideradas intervenientes, fez-se uma seleção de acordo com o referencial teórico, conforme ilustrado na tabela 4:

Tabela 4: Relação de autores sobre os determinantes de desempenho educacional

Determinantes do Desempenho Educacional
Hanushek (2019) - Capital humano, desempenho escolar e a qualidade, educação da família, experiência no mercado de trabalho são determinantes para o sucesso educacional. Além do capital do conhecimento que é o mais importante.
Andrade e Laros (2007) - As famílias, as escolas, os bairros e até os grupos de amigos.
Menezes-Filho (2007) - Características da família, a educação da mãe, o número de livros, a existência de um computador em casa, entre outros.
Nascimento (2007) - Características socioeconômicas, as habilidades anteriormente adquiridas pelos estudantes e os recursos destinados à educação.
Sousa (2018) - Existe uma relação intrínseca entre os aspectos socioeconômicos e a qualidade educacional, e o fator de maior influência no desempenho escolar é o nível de formação acadêmica dos pais.
Menezes-Filho (2014) - Atributos tanto dos alunos como das famílias, o nível de escolaridade da mãe, a cor da pele, o trabalho fora de casa, o atraso escolar e a reprovação prévia, além do número de livros e a presença de computadores em casa.
Oliveira (2015) - Adesão dos Gestores tem resultados melhores em avaliação de larga escala

Soares (2005) - (I) - Aluno: - Raça, Sexo, Trajetória Escolar, Motivação, Talentos, Saúde, Atitudes em relação à escola, Proficiência; (II) - Família: Renda Familiar, Nível de Escolaridade dos pais, Envolvimento dos pais ao processo educacional, Estrutura familiar, Aspectos Culturais; (III) - Estrutura Escolar: Nível de escolaridades dos professores, Experiência dos professores, Engajamento dos professores, Gestão Escolar, Recursos didáticos, Infraestrutura escolar, Projeto Pedagógico, Recursos Financeiros / Investimentos, Gastos por aluno.

França e Gonçalves (2008) - os pais são a principal referência para os filhos em termos de conduta na escola, e aqueles pais com capital humano elevado têm filhos com nível educacional semelhante ou superior.

Fonte: Elaboração própria.

A partir do levantamento teórico, optou-se pela escolha das seguintes variáveis intervenientes dos questionários do SAEB referente ao ano de 2017, de acordo com as tabelas 5, 6 e 7:

Tabela 5: Alunos – Variáveis escolhidas

Tabela Alunos - Anos iniciais e Anos Finais	
Variável	Descrição
Qual é o seu sexo?	Sexo do aluno
Qual é a sua cor ou raça?	Raça do aluno
Qual a sua idade?	Idade do aluno
Na sua casa tem computador?	Existência de computador na casa do aluno
Você mora com sua mãe?	Se o aluno mora com a mãe
Até que série sua mãe, ou a mulher responsável por você, estudou?	Escolaridade da mãe do aluno
Sua mãe, ou a mulher responsável por você, sabe ler e escrever?	Mãe do aluno sabe ler ou escrever
Você mora com seu pai?	Aluno mora com o pai ou responsável
Até que série seu pai, ou o homem responsável por você, estudou?	Escolaridade do pai do aluno
Seu pai, ou homem responsável por você, sabe ler e escrever?	Pai do aluno sabe ler ou escrever
Com qual frequência seus pais, ou responsáveis por você, vão à reunião de pais?	Frequência dos pais do aluno na reunião de pais

Seus pais ou responsáveis incentivam você a estudar?	Incentivo dos pais com relação ao estudo dos filhos
Seus pais ou responsáveis incentivam você a fazer o dever de casa e/ou os trabalhos da escola?	Incentivo dos pais para fazer a lição de casa
Seus pais ou responsáveis incentivam você a ler?	Incentivo dos pais para que os filhos leiam
Seus pais ou responsáveis incentivam você a ir a escola e/ou não faltar às aulas?	Incentivo dos pais para ida dos filhos à escola
Seus pais ou responsáveis conversam com você sobre o que acontece na escola?	Conversa dos pais com os alunos sobre a escola
Você utiliza a biblioteca ou sala de leitura da sua escola?	Frequência de utilização da biblioteca pelos alunos
Classe Econômica - Critério Brasil	Classe econômica dos alunos por meio do Critério Brasil

Fonte: Elaboração própria.

Na tabela Alunos, optou-se pela escolha de diversas variáveis que vão desde as características físicas do aluno, o seu nível socioeconômico, suas atitudes com relação à escola, até as suas interações com seus pais. A seguir, observam-se quais foram as variáveis escolhidas na tabela Professor:

Tabela 6: Professor – Variáveis escolhidas

Tabela Professor	
Variável	Descrição
Qual é o seu sexo?	Sexo do professor
Você poderia nos dizer a sua faixa etária?	Idade do professor
Qual é o mais alto nível de escolaridade que você concluiu (até a graduação)?	Mais alta graduação que o professor tem
Indique o curso de pós-graduação de mais alta titulação que você possui.	Mais alta titulação que o professor tem
Além da atividade como docente nesta escola, você exerce outra atividade que contribui para sua renda pessoal?	O professor exerce alguma atividade além da docência
Há quantos anos você trabalha como professor?	Período de tempo que trabalha como professor
Há quantos anos você trabalha como professor nesta escola?	Período de tempo que trabalha como professor nesta escola

Em quantas escolas você trabalha?	Quantidade de escolas em que trabalha como professor
Durante os últimos dois anos, você participou de qualquer um dos seguintes tipos de atividades de desenvolvimento profissional, e qual foi o impacto dessa atividade em sua prática como professor? Cursos/oficinas sobre metodologias de ensino na sua área de atuação.	Recodificado: Participação do professor em Cursos e Oficinas na área de atuação (SIM ou NÃO)
Durante os últimos dois anos, você participou de qualquer um dos seguintes tipos de atividades de desenvolvimento profissional, e qual foi o impacto dessa atividade em sua prática como professor? Cursos/oficinas sobre outros tópicos em educação.	Recodificado: Participação do professor em Cursos e Oficinas sobre outros tópicos (SIM ou NÃO)
Durante os últimos dois anos, você participou de qualquer um dos seguintes tipos de atividades de desenvolvimento profissional, e qual foi o impacto dessa atividade em sua prática como professor? Curso de especialização (mínimo de 360 horas) ou aperfeiçoamento (mínimo de 180 horas) sobre metodologias de ensino na sua área de atuação.	Recodificado: Participação do professor em Curso de Especialização ou Aperfeiçoamento na área de atuação (SIM ou NÃO)
Durante os últimos dois anos, você participou de qualquer um dos seguintes tipos de atividades de desenvolvimento profissional, e qual foi o impacto dessa atividade em sua prática como professor? Curso de especialização (mínimo de 360 horas) ou aperfeiçoamento (mínimo 180 horas) sobre outros tópicos em educação.	Recodificado: Participação do professor em Curso de Especialização ou Aperfeiçoamento sobre outros tópicos (SIM ou NÃO)

Fonte: Elaboração própria.

Optou-se, na tabela Professor, pelas características físicas do docente, sua escolaridade, seu tempo de atuação e seus aperfeiçoamentos em nível profissional. Já na Tabela Diretor, optou-se pelo uso das seguintes variáveis:

Tabela 7: Diretor – Variáveis escolhidas

Tabela Diretor	
Variável	Descrição
Qual é o seu sexo?	Sexo do Diretor
Você poderia nos dizer a sua faixa etária?	Faixa etária do Diretor
Qual é o mais alto nível de escolaridade que você concluiu?	Mais alta graduação que o diretor tem
Indique o curso de pós-graduação de mais alta titulação que você tem	Mais alta titulação que o diretor tem
Indique a área temática do curso de pós-graduação de mais alta titulação que você tem	Temática da mais alta graduação que o diretor tem
Além da direção desta escola, você exerce outra atividade que contribui para sua renda pessoal?	O diretor exerce alguma atividade além de diretor
Há quantos anos você exerce funções de direção?	Período de tempo que trabalha como diretor
Durante os últimos dois anos, você participou de alguma atividade de desenvolvimento profissional?	Nos últimos dois anos, houve a participação do diretor em atividades de desenvolvimento profissional.
O Conselho Escolar é um colegiado geralmente constituído por representantes da escola e da comunidade que tem como objetivo acompanhar as atividades escolares. Neste ano, quantas vezes se reuniu o Conselho Escolar?	Recodificado para: Existência de Conselho Escolar
Neste ano, qual foi o principal critério utilizado para a formação das turmas nesta escola?	Critério de formação das turmas pelo diretor

Fonte: Elaboração própria.

Assim, nesta tabela, optou-se pelas características do diretor, sua faixa etária, sua escolaridade, seu aperfeiçoamento e, de forma mais importante, qual foi o critério utilizado pelo diretor para formação de turmas no ano de 2017. Por fim, na tabela Escola, as variáveis escolhidas estão ilustradas na tabela 8:

Tabela 8: Escola – Variáveis escolhidas

Tabela Escola	
Variável	Descrição
Indique se nesta escola existem ou não os recursos apontados e quais são suas	Recodificado para: Existência de computadores para uso dos alunos (SIM ou NÃO)

condições de uso. Computadores para uso dos alunos.	
Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Acesso à internet para uso dos alunos.	Recodificado para: Existência de acesso à internet para uso dos alunos (SIM ou NÃO)
Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Computadores para uso dos professores.	Recodificado para: Existência de computadores para uso dos professores (SIM ou NÃO)
Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Acesso à internet para uso dos professores.	Recodificado para: Existência de acesso à internet para uso dos professores (SIM ou NÃO)
Indique se nesta escola existem ou não os recursos apontados e quais são suas condições de uso. Computadores exclusivamente para o uso administrativo.	Recodificado para: Existência de computadores exclusivamente para uso dos professores (SIM ou NÃO)
Indique a existência e as condições de uso dos seguintes espaços da escola: Biblioteca.	Recodificado para: Existência de biblioteca (SIM ou NÃO)
Indique a existência e as condições de uso dos seguintes espaços da escola: Quadra de esportes.	Recodificado para: Existência de quadra de esportes (SIM ou NÃO)
Indique a existência e as condições de uso dos seguintes espaços da escola: Laboratório de informática.	Recodificado para: Existência de laboratório de informática (SIM ou NÃO)
Indique a existência e as condições de uso dos seguintes espaços da escola: Laboratório de Ciências.	Recodificado para: Existência de laboratório de Ciências (SIM ou NÃO)

Fonte: Elaboração própria.

Assim, observa-se que, na tabela Escola, as variáveis intervenientes escolhidas têm relação direta com a infraestrutura disponibilizada para o aluno em questão.

Thoemmes (2012) alerta que esta etapa é considerada crítica, pois a credibilidade do *Propensity Score Matching* depende da escolha adequada das variáveis intervenientes.

Para Shadish, Clark e Steiner (2008), somente o uso de variáveis intervenientes de conveniência como sexo e idade, por exemplo, não é suficiente para encontrar o objetivo esperado. Eles afirmam que o pesquisador deve se esforçar para encontrar todas as variáveis que podem influenciar no tratamento.

Da mesma forma que sexo e idade podem influenciar os resultados deste trabalho, variáveis intervenientes como o nível socioeconômico do estudante, a escolaridade de seus pais, existência de uma biblioteca na escola, entre outros, precisam ser controladas, pois sua incidência contamina e altera os resultados. Ao comparar turmas de tamanhos diferentes sem controlar essas variáveis, a relação desempenho e tamanho da turma não tem nexos de causalidade, pois outras variáveis podem ser responsáveis por qualquer diferença de desempenho.

Dessa forma, Cannas e Arpino (2015) ressaltam que uma estimativa de propensão incorreta pode contaminar o balanceamento. Para que isso não ocorra, os autores determinam, como solução, o agrupamento das unidades individuais em *clusters* ou grupos por grau de semelhança.

De acordo com Mishra e Guha (2016), o *clustering* é uma ferramenta muito útil e basicamente onipresente na análise de dados estatísticos. Eles afirmam que o *clustering* nada mais é do que um agrupamento de dados sob a definição de similaridade, em que itens semelhantes ficam em um mesmo grupo, e itens diferentes ficam em grupos diferentes.

Ao agrupar as turmas em *clusters* por grau de semelhança, é possível compará-las entre si. Assim, comparam-se turmas com aproximadamente a mesma incidência de variáveis intervenientes, evitando-se que a contaminação das variáveis intervenientes prejudique o resultado.

Berkin (2006) diz que o *clustering* é a divisão de dados em grupos de objetos semelhantes, em que alguns detalhes são desconsiderados em prol da simplificação dos dados. O autor explica que o armazenamento de dados em *clusters* é uma técnica de modelagem de dados que fornece resumos concisos e agrupados.

Já para Mishra e Guha (2016), o aspecto mais importante, ao utilizar o *cluster*, é a definição do que constitui esse *cluster*, a fim de verificar por que um é melhor do que o outro. Os autores alertam que a qualidade do *cluster* depende da profunda análise de todos os pontos a ele atribuídos. Ao definir o limite de constituição dos

clusters, é possível definir quais turmas compõem cada um deles, de acordo com o seu escore de propensão definido no Propensity Score Matching.

Ainda nessa perspectiva, Prearo (2013) destaca a importância da análise de *cluster* para a efetiva detecção de *outliers* multivariados. O autor salienta que os chamados *outliers* multivariados são casos detectados em grupos formados por apenas um elemento, partindo do princípio de que nenhum outro caso foi considerado igual ou similar a ele.

Ao realizar o agrupamento das turmas em *clusters* percebe-se que algumas não se encaixam em nenhum *cluster*, ou seja, não são aproximadamente semelhantes a nenhuma outra dentro do universo analisado. Estes são os chamados *outliers* multivariados e, como eles não servem para nenhuma comparação, são excluídos.

Segundo Prearo (2008), é por meio da distância euclidiana que os *clusters* são mais bem agrupados. Para que isso ocorra, o autor demonstra ser imprescindível que o pesquisador especifique com antecedência o número desejado de agrupamentos.

Conforme explicam Liberti *et al.* (2014), a distância euclidiana é o estudo da geometria com base no conceito de distância, útil para diversas aplicações. Os autores consideram que, na distância euclidiana, os dados de entrada consistem em um agrupamento incompleto de distâncias, e os dados de saída, em um agrupamento de pontos no espaço.

Madhulatha (2012) descreve a distância euclidiana como uma medida comumente utilizada, calculada com base no encontro do quadrado da distância entre cada variável, somando os quadrados encontrados e, ao final, calculando raiz quadrada dessa soma.

Sob esta ótica, Huang (2008) afirma que a distância euclidiana é uma métrica-padrão em problemas geométricos. Por ser a distância normal entre dois pontos, ela pode ser facilmente medida com uma régua em duas ou três dimensões. Dessa forma, métrica da distância euclidiana é amplamente usada em problemas relacionados ao agrupamento por *clusters*.

Neste momento, agrupam-se as turmas por grau de semelhança, a partir do princípio de que todas as variáveis intervenientes foram controladas. Assim, cada turma apresenta uma quantidade específica de incidência dessas variáveis

intervenientes e, por essa razão, é possível agrupá-las em *clusters* por grau de semelhança, definida pelo uso da distância euclidiana e do Nearest Neighbor Matching.

Segundo Thoemmes (2012), após a estimativa do *Propense Score Matching*, começa a correspondência entre os escores obtidos. Essa correspondência pode ser realizada de muitas maneiras. A mais usada – e que será utilizada neste trabalho – é a técnica de correspondência *Nearest Neighbor Matching* ou vizinho mais próximo. O autor ainda explica que essa técnica é utilizada para selecionar participantes aproximadamente semelhantes dentro do banco de dados aferido e, para garantir boas correspondências, é necessário definir o máximo de diferença permitida.

Nesse sentido, Dong, Xu e Zeng (2020) ressaltam que a técnica *Nearest Neighbor Matching* é adequada para ser usada quando há a necessidade de obter um grande número de pontos-chave correspondentes. Tal como Prearo (2013), os autores ressaltam a necessidade de observar com atenção os *outliers* multivariados.

Assim, neste estudo optou-se pela utilização do *Propense Score Matching* para mapear a influência das variáveis intervenientes nas turmas investigadas, utilizando-se também a técnica de *clustering* e da distância euclidiana. Para o pareamento dos *clusters*, optou-se pelo *Nearest Neighbor Matching*, e, ainda, excluíram-se os *outliers* multivariados.

3.5 Tamanho final da amostra

O tamanho final da amostra é de 17.533 turmas de quinto ano e de 13.608 turmas de nono ano. O total de escolas analisadas foi de 16.607, com um diretor por escola. Nessas escolas, o total de professores foi de 101.379, com o total 437.413 alunos de quinto ano e 370.918 alunos de nono ano.

3.6 Preparação dos Dados

Os dados do SAEB 2017, disponibilizados pelo INEP, contêm informações sobre os questionários aplicados nas provas e sobre os resultados dos alunos nas disciplinas de Língua Portuguesa e Matemática. Além disso, é possível encontrar o dicionário dos dados, as escalas de proficiência, os *inputs*, os documentos técnicos e as matrizes de referência.

Para o desenvolvimento deste trabalho, utilizaram-se o questionário do Diretor, do Professor, da Escola e do Aluno (tanto para Anos Iniciais quanto para Anos Finais do Ensino Fundamental). O dicionário dos dados foi utilizado para mapear as variáveis a serem consideradas.

O *software* escolhido foi o SPSS. Segundo Ong e Puteh (2017), ele foi desenvolvido pela *IBM Corporation* e é um dos mais utilizados pelos pesquisadores no mundo todo. Os autores ressaltam que esse *software* é capaz de realizar diversas análises estatísticas, entre elas, análises univariadas, bivariadas e multivariadas, tanto para técnicas paramétricas como para não paramétricas.

Para iniciar a análise dos dados do estudo, foi necessário converter os dados dos questionários do formato *Comma-separated Values* – CSV – para o formato do SPSS. Após a conversão, o primeiro passo foi utilizar o questionário do diretor, pois nele consta uma questão sobre o critério de formação de turmas utilizado pelo diretor naquele ano, conforme ilustrado na tabela 9:

Tabela 9: Diretor – Questão 39 – Critério de formação de turmas

Questionário do Diretor - SAEB 2017	
Questão 39 - Neste ano, qual foi o principal critério utilizado para a formação das turmas nesta escola?	A - Homogeneidade quanto à idade (alunos com a mesma idade)
	B - Homogeneidade quanto ao rendimento escolar (alunos com rendimento similar)
	C - Heterogeneidade quanto à idade (alunos com idades diferentes)
	D - Heterogeneidade quanto ao rendimento escolar (alunos com rendimentos diferentes)
	E - Outro critério
	F - Não houve critério

Fonte: Elaboração própria.

A partir do filtro pela questão 39, letra F "Não houve critério", no questionário do diretor, chegou-se ao número de 16.607 escolas. Essa quantidade representa a amostra de escolas que não aplicaram nenhuma metodologia para divisão das turmas e, por isso, não sofreram com o viés de seleção, estando aptas a serem examinadas.

De acordo com Nunan, Bankhead e Aronson (2018), o viés de seleção se caracteriza pela comparação de grupos participantes diferentes, não apenas contaminados pelas variáveis intervenientes presentes na pesquisa em si, mas também por pré-seleção anterior à análise dos dados.

Assim, os bancos de dados dos alunos Anos Iniciais, dos alunos Anos Finais, dos professores e da escola também foram limpos, mantendo-se, em seu conteúdo, apenas essas 16.607 escolas.

Ao observar os bancos de dados e comparar com os determinantes encontrados no referencial teórico, foi possível identificar as variáveis intervenientes correspondentes ao indicado pelos autores.

Das variáveis encontradas, apenas a classe econômica precisou ser criada a partir de outras variáveis já constantes do banco de dados. Para isso, foi utilizado o Critério Brasil como mecanismo de medição dos aspectos socioeconômicos dos alunos.

Segundo Alves, Soares e Xavier (2014), o Critério de Classificação Econômica Brasil, mais conhecido como Critério Brasil, foi adotado pela Associação

Brasileira de Empresas de Pesquisa – ABEP, com o intuito de estimar o poder de compras das famílias brasileiras. Os autores explicam que, por intermédio de simples perguntas sobre bens e serviços disponíveis no domicílio e ainda com questionamentos sobre a escolaridade do chefe da família, é possível situar os indivíduos em classes econômicas. Isso se verifica nas tabelas 10 e 11, a seguir:

Tabela 10: Pontuação atribuída a cada variável, pelo Critério Brasil

Variável	Quantidade de Pontos Equivalentes				
	0	1	2	3	4
Televisão em Cores	0	1	2	3	4
Rádio	0	1	2	3	4
Banheiro	0	4	5	6	7
Automóvel	0	4	7	9	9
Empregada Mensalista	0	3	4	4	4
Máquina de Lavar	0	2	2	2	2
Videocassete e/ou DVD	0	2	2	2	2
Geladeira	0	4	4	4	4
Freezer (aparelho independente ou parte da geladeira duplex)	0	2	2	2	2

Fonte: Elaboração própria.

Pode-se observar que cada variável representa uma quantidade diferente de pontos em relação à quantidade de itens presentes em casa.

Tabela 11: Pontuação atribuída a cada variável, pelo Critério Brasil

Grau de Instrução do Responsável Financeiro da Família	Pontuação
Analfabeto / Fundamental I - Incompleto	0
Fundamental I - Completo / Fundamental II - Incompleto	1
Fundamental II - Completo / Médio – Incompleto	2
Médio - Completo / Superior – Incompleto	4
Superior completo	8

Fonte: Elaboração própria.

Nesta tabela, percebe-se que o nível socioeconômico do responsável pela família é medido pelo Critério Brasil por intermédio de sua escolaridade. Já na tabela 12, podem-se verificar quais intervalos de pontos correspondem a cada classe econômica.

Tabela 12: Pontuação atribuída a cada classe econômica, pelo Critério Brasil

Classe Econômica	Intervalo de Pontos
A	de 35 a 46 pontos
B	de 23 a 34 pontos
C	de 14 a 22 pontos
D	de 8 a 13 pontos
E	de 0 a 7 pontos

Fonte: Elaboração própria.

Desse modo, atingidos os pontos, é possível adequar os alunos em classes econômicas, conforme a quantidade de pontos obtida.

Neste trabalho, optou-se pelo uso do Critério Brasil referente ao ano de 2014, pois este foi o último ano em que a ABEP utilizou a mesma matriz de dados que o SAEB, o que possibilitou o cruzamento perfeito das informações.

Para juntar os bancos de dados do SAEB em dois únicos bancos – um referente aos anos iniciais (5º ano do Ensino Fundamental) e outro referente aos anos finais (9º ano do Ensino Fundamental) –, foi necessário transformar todas as variáveis disponibilizadas em questões, de letra para número e de número para porcentagem.

Ao fazer isso, definimos a variável Código da Turma como a *Break Variable* ou variável de agrupamento. Com isso, foi possível obter a quantidade de 17.533 turmas de 5º ano ou de Anos Iniciais e de 13.608 turmas de 9º ano ou de Anos Finais, no universo de 16.607 escolas.

A variável Critério de Divisão de Turmas realizado pelo diretor foi excluída do modelo estatístico. Ela já havia sido aplicada no início, ao limpar todos os bancos de dados. Tal como essa variável, a Temática de Título do Diretor também foi excluída, já que, segundo o artigo 64 da LDB, todos os diretores devem ter em sua formação a base comum nacional em educação. De acordo com o Art. 64 da Lei nº 9.394, de 20 de dezembro de 1996,

A formação de profissionais de educação para administração, planejamento, inspeção, supervisão e orientação educacional para a educação básica, será feita em cursos de graduação em pedagogia ou em nível de pós-graduação, a critério da instituição de ensino, garantida, nesta formação, a base comum nacional.” (BRASIL, 1996).

Diante disso, a variável Temática de Título do Diretor permanece constante em todas as turmas e pode ser excluída do modelo.

Por intermédio do SPSS, como conceito de *Propensity Score Matching*, foram utilizados a distância euclidiana e o *Nearest Neighbor*, a fim de agrupar as turmas duas a duas. A princípio, havia o interesse de apenas calcular a distância euclidiana de todas as turmas, em pares, e selecionar, para cada uma delas, seu par com menor distância. Como não foi possível a obtenção de um recurso computacional para esse cálculo demasiadamente grande – por exemplo, com 10.000 turmas, calcular-se-ia uma matriz de distância de 100.000.000 de células –, optou-se pela utilização do método *k-means* de clusterização.

Por oportuno, registre-se ainda que, nesse arcabouço de escolhas, foi necessário arbitrar um número de *cluster* inicial. A decisão foi de $N/2$, na hipótese de haver $N/2$ *clusters* com dois casos cada, ou seja, pares perfeitos nas variáveis intervenientes. Assim, para as 17.533 turmas de 5º ano foi arbitrado um desenho de 8.766 *clusters* e, para as 13.608 turmas de 9º ano, foi arbitrada a construção de 6.804 *clusters*.

4 RESULTADOS

A análise dos resultados teve seu início a partir da exclusão das turmas que não encontraram nenhum agrupamento com outra turma, ou seja, formaram um *cluster* com apenas 1 caso.

Em contrapartida, ainda que a maioria dos *clusters* tenha apresentado dois casos, como se desejava, cerca de 15% dos agrupamentos revelaram mais de 2 casos – entre 3 e 4, em sua maioria. Diante disso, criaram-se tantos *clusters* 2 a 2 quanto possíveis no cruzamento das turmas – por exemplo, um *cluster* de 4 turmas formou 6 pares de turmas, resultado da combinação de 4 casos, a 2 a 2 – $C_{4,2}$.

Assim, após essa construção, resultaram 20.362 *clusters* – ou pares de turmas – no 5º ano e 13.038 *clusters* – ou pares de turmas – no 9º ano.

De posse dos pares de turmas, pareadas pelos critérios expostos na seção 3, a fim de comparar o desempenho dos estudantes entre as turmas, foi necessária a seleção de um teste de hipótese comparativo para amostras pareadas.

Para a seleção desse teste, foi necessário verificar a condição de aderência à distribuição normal dos dados para, então, definir se eles ensejam uma análise paramétrica ou uma análise não paramétrica. Neste trabalho, optou-se pelo uso do teste Kolmogorov Smirnov para verificação de tal aderência.

Van den Berg (2019) explica que existem dois testes Kolmogorov Smirnov, e o primeiro deles serve para testar se uma variável segue uma dada distribuição em uma população. O autor afirma que essa distribuição é geralmente – mas nem sempre – uma distribuição normal. Portanto, esse primeiro teste serve para identificar se a distribuição é ou não é normal. O segundo teste serve para testar se uma variável tem distribuição idêntica entre duas amostras diferentes. Ambos os testes, segundo o autor, podem ser utilizados no SPSS. Neste trabalho, usou-se o primeiro.

Dessa forma, ao utilizar o referido teste, percebeu-se a rejeição da hipótese nula de normalidade. Verificou-se ser necessária a utilização de uma análise não paramétrica, conforme ilustrado nas tabelas 13 e 14:

Tabela 13: Teste Kolmogorov Smirnov referente aos dados do 5º ano

Teste de normalidade Kolmogorov Smirnov - Quinto Ano		
	Estatística	p-valor
Notas de Língua Portuguesa	0,010	< 0,01
Notas de Matemática	0,018	< 0,01

Fonte: Elaboração própria.

Percebe-se que as notas de Língua Portuguesa e de Matemática, na coluna estatística, apresentam um valor maior que 0,01, indicando a rejeição da hipótese nula de aderência à curva normal e ensejando uma análise não paramétrica. O resultado se repete tanto na tabela 13 como na tabela 14.

Tabela 14: Teste Kolmogorov Smirnov referente aos dados do 9º ano

Teste de normalidade Kolmogorov Smirnov - Nono Ano		
	Estatística	p-valor
Notas de Língua Portuguesa	0,031	< 0,01
Notas de Matemática	0,015	< 0,01

Fonte: Elaboração própria.

Diante do exposto, o teste não paramétrico escolhido foi o de Wilcoxon para duas amostras pareadas.

Singh, Roy e Tripathi (2013) explicam que um teste estatístico pode ser dividido em paramétrico e não paramétrico. O teste paramétrico é aquele que faz suposições sobre os parâmetros ou propriedades de distribuição dos dados analisados. Já o teste não paramétrico não faz tais suposições, sendo no sentido estrito, essencialmente nulo.

Segundo Rosner, Glynn e Lee (2005), o teste de Wilcoxon é um teste de característica não paramétrica frequentemente utilizado para dados emparelhados, como medições de pré e pós-tratamento, baseando-se em análises de unidades independentes. Cyprian, Oyeka e Ebu (2012) complementam afirmando que o teste de Wilcoxon é normalmente utilizado para testar a hipótese nula de que a mediana de uma distribuição é igual a algum valor pré-definido.

Nesse sentido, Backes *et al.* (2016) consideram que o teste Wilcoxon está entre os testes de hipóteses mais populares da estatística. Os autores afirmam que esses testes são aplicados com o intuito de avaliar se os resultados têm significância estatística ou se podem ser explicados pelo acaso. No geral, o teste Wilcoxon exige um alto esforço combinatório para grandes cortes de amostra, principalmente as que contêm um número significativo de empates. Por conseguinte, o valor P é frequentemente próximo a uma distribuição normal.

Bridge e Sawilowsky (1999) sinalizam a importância de os pesquisadores compreenderem o impacto dos testes estatísticos nos resultados de suas pesquisas. Os autores mencionam que a aplicação de modelos estatísticos ineficientes não apenas aumenta o uso de recursos, como também a possibilidade de se cometer algum erro. Ressaltam, ainda, que o teste Wilcoxon é um teste imparcial e uniformemente poderoso, principalmente em distribuições relativamente simétricas com caudas mais pesadas ou inclinações extremas.

4.1 Apresentação dos resultados

De forma geral, é possível perceber pouca influência do número de alunos por turma no desempenho acadêmico. Além disso, observa-se que, quando há essa influência, as turmas maiores tendem a se desempenhar melhor.

Para facilitar a demonstração dos resultados obtidos, as tabelas são apresentadas da seguinte forma: em primeiro lugar, os Anos Finais, na disciplina de Língua Portuguesa e na disciplina de Matemática; e em segundo lugar, os Anos Iniciais, também nas disciplinas de Língua Portuguesa e na disciplina de Matemática.

A fim de definir a quantidade de alunos por turmas comparadas entre si por meio dos *clusters*, utilizaram-se, como base, as legislações sobre o tema. No âmbito federal, pode-se observar que o Projeto de Lei 4628 de 2016 definiu o número de 20 alunos para os Anos Iniciais do Ensino Fundamental e o número de 25 alunos para os Anos Finais. Já o Projeto de Lei 4731 de 2012 define que o número de alunos dos Anos Iniciais deve ser de 25 alunos e, dos Anos Finais, de 35 alunos. Em contrapartida, no âmbito estadual, observa-se que o Projeto de Lei 666 de 2015, do Rio de Janeiro, define que o número de alunos por turma deve ser de até 25 para os Anos Iniciais e de até 35 para os Anos Finais.

Nota-se que os projetos de lei, em geral, diferem entre si, em múltiplos de cinco. Da mesma forma, optou-se por comparações que variam de cinco em cinco alunos por turma.

4.1.1 Anos Finais ou 9º ano

Como já dito, para analisar os Anos Finais, inicialmente focaliza-se a disciplina de Língua Portuguesa; posteriormente, segue a disciplina de Matemática.

4.1.1.1 Língua Portuguesa 9º Ano

Optou-se por começar a análise com um pequeno número de comparações, a fim de verificar se o modelo aplicado se mostrava viável para este estudo. Desse modo, chegou-se aos dados ilustrados na tabela 15:

Tabela 15: Pequena Comparação de Turmas – Língua Portuguesa – 9º Ano

Língua Portuguesa 9º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
Até 15 alunos	400	246,35	-3,09	0,020 (p<0,05)
Acima de 15 alunos		249,44		
Até 20 alunos	1.199	252,54	0,07	0,052 (p>0,05)
Acima de 20 alunos		252,47		
Até 25 alunos	2.322	255,69	1,38	0,850 (p>0,05)
Acima de 25 alunos		254,31		
Até 30 alunos	2.927	254,67	-2,49	0,000 (p<0,05)
Acima de 30 alunos		257,16		
Até 35 alunos	2.207	255,92	-1,24	0,949 (p>0,05)
Acima de 35 alunos		257,16		

Fonte: Elaboração própria.

Já nessa amostra, é possível observar que, que tanto na comparação entre turmas de 15 alunos e turmas de mais de 15 alunos, como na comparação entre turmas de até 30 alunos e turmas de mais de 30 alunos, o teste de Wilcoxon demonstrou haver significância estatística no modelo, indicando que turmas maiores obtêm melhores desempenhos.

Com o apontamento positivo desta primeira análise, optou-se por examinar as comparações de *clusters* por série, ou seja, utilizou-se um espaço maior entre o até e o acima, conforme ilustrado na tabela 16:

Tabela 16: Comparações por Série – Seção 1 – Língua Portuguesa - 9º Ano

Língua Portuguesa 9º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
Até 15 alunos	3	277,08	6,49	0,109 (p>0,05)
Acima 45 alunos		270,59		
Até 15 alunos	16	233,2958	-29,7564	0,039 (p<0,05)
Acima 40 alunos		263,0522		
Até 15 alunos	71	241,4758	-16,0406	0,001 (p<0,05)
Acima 35 alunos		257,5164		
Até 15 alunos	180	247,1768	-8,5933	0,002 (p<0,05)
Acima 30 alunos		255,7701		
Até 15 alunos	284	248,3826	-4,0115	0,003 (p<0,05)
Acima 25 alunos		252,3941		
Até 15 alunos	364	246,3518	-4,615	0,004 (p<0,05)
Acima 20 alunos		250,9675		
Até 20 alunos	13	247,22	-4,48	0,917 (p>0,05)
Acima 45 alunos		251,70		
Até 20 alunos	61	245,30	-5,33	0,407 (p>0,05)
Acima 40 alunos		250,63		
Até 20 alunos	299	254,03	-0,16	0,444 (p>0,05)
Acima 35 alunos		254,19		
Até 20 alunos	662	254,03	-1,33	0,077 (p>0,05)
Acima 30 alunos		255,36		
Até 20 alunos	970	253,15	0,08	0,186 (p>0,05)
Acima 25 alunos		253,07		

Fonte: Elaboração própria.

Já na seção 1, observa-se que turmas maiores se desempenham melhor em todas as séries estatisticamente significantes, segundo o teste de Wilcoxon. É possível verificar que, na série de até 15 alunos, em comparação com a série acima de 40, essa diferença de desempenho se apresenta de forma mais acentuada. Assim, faz-se necessário observar atentamente a seção 2, ilustrada na tabela 17:

Tabela 17: Comparações por Série – Seção 2 – Língua Portuguesa - 9º Ano

Língua Portuguesa 9º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 25 alunos	36	247,80	-5,21	0,293 (p>0,05)
acima 45 alunos		253,01		
até 25 alunos	193	248,37	-2,45	0,294 (p>0,05)
acima 40 alunos		250,82		
até 25 alunos	833	254,21	-0,27	0,905 (p>0,05)
acima 35 alunos		254,48		
até 25 alunos	1.728	255,69	-0,58	0,238 (p>0,05)
acima 30 alunos		256,27		
até 30 alunos	64	252,2074	-7,25	0,010 (p<0,05)
acima 45 alunos		259,4615		
até 30 alunos	406	250,6738	-0,51	0,031 (p<0,05)
acima 40 alunos		251,1812		
até 30 alunos	1.600	254,21	-1,88	0,067 (p>0,05)
acima 35 alunos		256,09		
até 35 alunos	89	252,9939	-6,89	0,005 (p<0,05)
acima 45 alunos		259,8842		
até 35 alunos	626	252,5033	-1,64	0,030 (p<0,05)
acima 40 alunos		254,1467		
até 40 alunos	111	252,8724	-8,98	0,001 (p<0,05)
acima 45 alunos		261,8486		

Fonte: Elaboração própria.

Com menos intensidade, as séries de até 30 alunos, em comparação com as acima de 40 e acima de 45 alunos, juntamente com as séries de até 35 alunos em comparação com acima de 40 e acima de 45 alunos e por fim na comparação de até 40 alunos com acima de 45 alunos, demonstram haver significância estatística na comparação dos *clusters*. Destaque-se que sempre as turmas maiores apresentam melhor desempenho.

Com o intuito de determinar se o número de alunos por turma influencia ou não no desempenho escolar e ainda verificar como esse desempenho aumenta ou diminui, a depender da quantidade de alunos por turma, foi necessário detalhar ainda mais as comparações, conforme ilustrado na tabela a seguir

Tabela 18: Comparação detalhada – Seção 1 – Língua Portuguesa - 9º Ano

Língua Portuguesa 9º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 12 alunos	97	244,11	-10,14	0,970 (p>0,05)
acima 12 alunos		254,25		
até 13 alunos	194	240,01	-7,53	0,063 (p>0,05)
acima 13 alunos		247,54		
até 14 alunos	286	245,02	-2,88	0,281 (p>0,05)
acima 14 alunos		247,90		
até 15 alunos	400	246,3518	-3,0929	0,020 (p<0,05)
acima 15 alunos		249,4447		
até 16 alunos	475	247,97	-2,66	0,074 (p>0,05)
acima 16 alunos		250,63		
até 17 alunos	674	252,67	1,09	0,957 (p>0,05)
acima 17 alunos		251,58		
até 18 alunos	840	250,04	-1,53	0,379 (p>0,05)
acima 18 alunos		251,57		
até 19 alunos	1.048	253,14	0,79	0,235 (p>0,05)
acima 19 alunos		252,35		
até 20 alunos	1.199	252,54	0,07	0,052 (p>0,05)
acima 20 alunos		252,47		
até 21 alunos	1.398	253,48	1,73	0,718 (p>0,05)
acima 21 alunos		251,75		

Fonte: Elaboração própria.

Na seção 1, é possível observar que houve significância estatística em turmas de até 15 alunos, em comparação com turmas com mais de 15 alunos. Nota-se, mais uma vez, que as turmas maiores se desempenham melhor.

Na seção 2, conforme ilustrado na tabela 19, observa-se que, quando comparadas as turmas de até 28, 29, 30 e 31 alunos com turmas maiores, as classes maiores apresentam melhor desempenho.

Tabela 19: Comparação detalhada – Seção 2 – Língua Portuguesa - 9º Ano

Língua Portuguesa 9º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 22 alunos	1.629	255,15	2,64	0,822 (p>0,05)
acima 22 alunos		252,51		
até 23 alunos	1.873	253,48	-0,02	0,534 (p>0,05)
acima 23 alunos		253,50		
até 24 alunos	2.142	255,69	1,51	0,490 (p>0,05)
acima 24 alunos		254,18		
até 25 alunos	2.322	255,69	1,38	0,850 (p>0,05)
acima 25 alunos		254,31		
até 26 alunos	2.500	255,69	0,33	0,706 (p>0,05)
acima 26 alunos		255,36		
até 27 alunos	2.661	255,69	-0,01	0,351 (p>0,05)
acima 27 alunos		255,70		
até 28 alunos	2.816	255,4747	-0,96	0,017 (p<0,05)
acima 28 alunos		256,4335		
até 29 alunos	2.948	254,6743	-2,3237	0,000 (p<0,05)
acima 29 alunos		256,9980		
até 30 alunos	2.927	254,6743	-2,49	0,000 (p<0,05)
acima 30 alunos		257,1623		
até 31 alunos	2.884	255,6780	-1,84	0,002 (p<0,05)
acima 31 alunos		257,5164		
até 32 alunos	2.870	256,29	-1,22	0,114 (p>0,05)
acima 32 alunos		257,51		

Fonte: Elaboração própria.

Já na seção 3, verifica-se que o desempenho de turmas maiores se mostrou de forma mais acentuada nas comparações com classes de até 39, 40 e 41 alunos, conforme se vê na tabela 20:

Tabela 20: Comparação detalhada – Seção 3 – Língua Portuguesa - 9º Ano

Língua Portuguesa 9º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 33 alunos	2.732	256,68	-0,94	0,396 (p>0,05)
acima 33 alunos		257,62		
até 34 alunos	2.483	256,31	-1,21	0,359 (p>0,05)
acima 34 alunos		257,52		
até 35 alunos	2.207	255,92	-1,24	0,949 (p>0,05)
acima 35 alunos		257,16		
até 36 alunos	1.808	255,88	-1,30	0,797 (p>0,05)
acima 36 alunos		257,18		
até 37 alunos	1.356	255,19	1,05	0,130 (p>0,05)
acima 37 alunos		254,14		
até 38 alunos	1.045	254,08	-0,09	0,540 (p>0,05)
acima 38 alunos		254,17		
até 39 alunos	780	252,0462	-3,6615	0,007 (p<0,05)
acima 39 alunos		255,7077		
até 40 alunos	604	252,2976	-4,8647	0,010 (p<0,05)
acima 40 alunos		257,1623		
até 41 alunos	327	251,2653	-6,3632	0,047 (p<0,05)
acima 41 alunos		257,6285		
até 42 alunos	208	251,70	-1,30	0,281 (p>0,05)
acima 42 alunos		253,00		

Fonte: Elaboração própria.

Na seção 4, observa-se que, nas comparações entre turmas de até 44 e 45 alunos e turmas maiores, novamente as classes maiores apresentam melhor desempenho, tal como se registra na tabela 21:

Tabela 21: Comparação detalhada – Seção 4 – Língua Portuguesa - 9º Ano

Língua Portuguesa 9º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 43 alunos	169	254,21	-5,25	0,525 (p>0,05)
acima 43 alunos		259,46		
até 44 alunos	116	252,7799	-9,0687	0,000 (p<0,05)
acima 44 alunos		261,8486		
até 45 alunos	72	251,3027	-8,1588	0,003 (p<0,05)
acima 45 alunos		259,4615		
até 46 alunos	35	252,68	-6,78	0,272 (p>0,05)
acima 46 alunos		259,46		
até 47 alunos	29	252,87	-6,59	0,294 (p>0,05)
acima 47 alunos		259,46		
até 48 alunos	26	252,11	-7,35	0,454 (p>0,05)
acima 48 alunos		259,46		
até 49 alunos	19	252,87	-6,59	1,000 (p>0,05)
acima 49 alunos		259,46		

Fonte: Elaboração própria.

Na comparação detalhada, é possível fazer diversas constatações interessantes. Logo no início, pode-se verificar que há diferença significativa de desempenho nas turmas com até 15 alunos, quando comparadas com turmas acima de 15 alunos. Nesse caso, tem-se uma diferença entre medianas de 3,0929, representando que as salas maiores têm desempenho melhor.

Constata-se também que, quando a turma tem mais de 15 alunos, o desempenho permanece constante, sem diferenças estatisticamente significantes até a quantidade de 27 alunos. Ou seja, entre turmas de 16 alunos e turmas de até 27 alunos, é possível manter o mesmo desempenho. Esse resultado facilita a decisão do gestor de políticas públicas em educação, ao formar uma turma de Língua Portuguesa dos Anos Finais, que tenha entre 16 e 27 alunos. Em outras palavras, é possível manter o desempenho dos estudantes com uma margem segurança de 16 até 27 alunos por turma.

Observa-se também que, nas comparações entre turmas de até 28, 29, 30 e 31 alunos e turmas maiores, novamente há a indicação de que as maiores apresentam melhor desempenho. Assim, a partir de 32 alunos por turma até 38 alunos, é possível manter o desempenho sem mudanças estatisticamente significantes.

Além disso, verifica-se que, na comparação entre turmas de até 39, 40 e 41 e turmas maiores, estas últimas igualmente apresentam melhor desempenho. Por sua vez, ele se mantém estável em turmas com até 42 e 43 alunos.

Em turmas com até 44 e 45 alunos, observa-se, mais uma vez, que salas maiores apresentam melhores desempenhos. Nas classes de 46 a 49 alunos, o desempenho permanece estatisticamente estável.

A maior mediana encontrada está em turmas maiores que 44 alunos, com 261,8486 pontos. Desse modo, por meio do resultado encontrado, pode-se oferecer, como parâmetro ao gestor de políticas públicas em educação, a hipótese de que turmas de Língua Portuguesa dos Anos Finais do Ensino Fundamental, a partir de 45 alunos, se desempenham melhor que turmas com menos estudantes.

Em decorrência de possíveis restrições orçamentárias por parte do ente da Federação, o estudo serve de parâmetro para a hipótese de que turmas com 45 alunos ou mais – consideradas “grandes” pelo referencial teórico – são uma opção viável para aumentar o desempenho escolar na disciplina e no ano em questão. Ademais, trata-se de uma opção mais barata e melhor em desempenho do que turmas menores.

4.1.1.2 Matemática 9º Ano

Diferentemente do ocorrido na disciplina de Língua Portuguesa do 9º Ano, o pequeno número de comparações da disciplina de Matemática do 9º ano não apresentou nenhum resultado estatisticamente significativo. Isso pode ser observado na tabela 22:

Tabela 22: Pequeno número de comparações – Matemática - 9º Ano

Matemática 9º Ano				
Lógica	n	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
Até 15 alunos	400	243,85	-2,41	0,060 (p>0,05)
Acima de 15 alunos		246,26		
Até 20 alunos	1.199	249,11	-0,86	0,772 (p>0,05)
Acima de 20 alunos		249,97		
Até 25 alunos	2.322	252,68	1,04	0,172 (p>0,05)
Acima de 25 alunos		251,64		
Até 30 alunos	2.927	252,76	-0,78	0,862 (p>0,05)
Acima de 30 alunos		253,54		

Até 35 alunos	2.207	253,36	0,16	0,107 ($p>0,05$)
Acima de 35 alunos		253,20		

Fonte: Elaboração própria.

Para confirmar essa prerrogativa, optou-se novamente pelo uso das comparações em séries, tal como se mostra na seguinte tabela:

Tabela 23: Comparações por Séries – Seção 1 – Matemática - 9º Ano

Matemática 9º Ano				
Lógica	n	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 15 alunos	3	236,51	-35,14	0,109 ($p>0,05$)
acima 45 alunos		271,65		
até 15 alunos	16	237,02	-15,23	0,121 ($p>0,05$)
acima 40 alunos		252,25		
até 15 alunos	71	245,9906	-6,1155	0,024 ($p<0,05$)
acima 35 alunos		252,1061		
até 15 alunos	180	247,75	0,89	0,157 ($p>0,05$)
acima 30 alunos		246,86		
até 15 alunos	284	247,57	1,42	0,223 ($p>0,05$)
acima 25 alunos		246,15		
até 15 alunos	364	243,52	-3,10	0,071 ($p>0,05$)
acima 20 alunos		246,62		
até 20 alunos	13	237,30	-2,24	0,972 ($p>0,05$)
acima 45 alunos		239,54		
até 20 alunos	61	240,96	0,91	0,685 ($p>0,05$)
acima 40 alunos		240,05		
até 20 alunos	299	249,11	2,00	0,602 ($p>0,05$)
acima 35 alunos		247,11		

até 20 alunos	662	250,66	1,48	0,689 (p>0,05)
acima 30 alunos		249,18		
até 20 alunos	970	249,32	0,15	0,592 (p>0,05)
acima 25 alunos		249,17		

Fonte: Elaboração própria.

Observa-se que, na comparação entre turmas de até 15 alunos e turmas com mais de 35 alunos, as classes maiores obtêm melhor desempenho do que as menores. Já na tabela 24, a seção 2 mostra que, na comparação entre salas de até 30 alunos e salas com mais de 45 alunos, entre salas de até 35 alunos e salas com mais de 45 alunos, e entre salas de até 40 alunos e salas com mais de 45 alunos, novamente as classes com maior quantidade de estudantes apresentam desempenho melhor:

Tabela 24: Comparações por Séries – Seção 2 – Matemática - 9º Ano

Matemática 9º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 25 alunos	36	240,89	-4,13	0,396 (p>0,05)
acima 45 alunos		245,02		
até 25 alunos	193	247,21	3,69	0,645 (p>0,05)
acima 40 alunos		243,52		
até 25 alunos	833	252,60	3,41	0,089 (p>0,05)
acima 35 alunos		249,19		
até 25 alunos	1.728	252,76	0,67	0,577 (p>0,05)
acima 30 alunos		252,09		
até 30 alunos	64	257,2109	-7,9249	0,004 (p<0,05)
acima 45 alunos		265,1358		
até 30 alunos	406	247,76	1,34	0,303 (p>0,05)
acima 40 alunos		246,42		
até 30 alunos	1.600	252,64	0,81	0,410 (p>0,05)
acima 35 alunos		251,83		
até 35 alunos	89	247,5086	-22,4040	0,001 (p<0,05)
acima 45 alunos		269,9126		
até 35 alunos	626	249,41	2,07	0,885 (p>0,05)
acima 40 alunos		247,34		
até 40 alunos	111	247,2052	-22,71	0,000 (p<0,05)
acima 45 alunos		269,9126		

Fonte: Elaboração própria.

A comparação por séries, diferentemente do pequeno número de comparações, demonstrou resultados estatisticamente significantes em turmas de até 15 alunos, quando comparadas com turmas de mais de 35 alunos, revelando que salas maiores têm melhor desempenho. A mesma significância estatística ocorreu na comparação de turmas com até 30 alunos e turmas com mais de 45 alunos: novamente as classes maiores apresentam o melhor desempenho.

Também houve significância estatística na série referente às turmas de até 35 alunos, em comparação com turmas com mais de 45 alunos, assim como nas turmas de até 40 alunos, em comparação com turmas com mais de 45 alunos. Nesse caso, é possível observar que as turmas maiores obtêm melhor desempenho, de forma muito mais acentuada que nas outras comparações.

Com o intuito de verificar se o número de alunos por turma na disciplina de Matemática do 9º ano influenciava o desempenho escolar, foi necessário, à luz dos resultados já encontrados, detalhar as comparações de *clusters*, conforme apresentado na tabela a seguir:

Tabela 25: Comparação detalhada – Seção 1 – Matemática - 9º Ano

Matemática 9º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 12 alunos	97	250,82	1,04	0,265 (p>0,05)
acima 12 alunos		249,78		
até 13 alunos	194	243,23	-2,45	0,126 (p>0,05)
acima 13 alunos		245,68		
até 14 alunos	286	243,51	-2,10	0,510 (p>0,05)
acima 14 alunos		245,61		
até 15 alunos	400	243,85	-2,41	0,060 (p>0,05)
acima 15 alunos		246,26		
até 16 alunos	475	247,75	-0,10	0,216 (p>0,05)
acima 16 alunos		247,85		
até 17 alunos	674	249,17	-0,06	0,505 (p>0,05)
acima 17 alunos		249,23		
até 18 alunos	840	247,01	-2,32	0,858 (p>0,05)
acima 18 alunos		249,33		
até 19 alunos	1.048	249,10	-0,71	0,498 (p>0,05)
acima 19 alunos		249,81		
até 20 alunos	1.199	249,11	-0,86	0,772 (p>0,05)
acima 20 alunos		249,97		

até 21 alunos	1.398	249,78	0,89	0,187 (p>0,05)
acima 21 alunos		248,89		

Fonte: Elaboração própria.

Na seção 1, observa-se que, em nenhuma das comparações, foi possível identificar relevância estatística. Já na seção 2, é possível observar que, na comparação entre turmas com até 32 alunos e turmas com mais de 32 alunos, as turmas menores, mesmo que ligeiramente, obtêm melhor desempenho. Isso se verifica na tabela 26:

Tabela 26: Comparação detalhada – Seção 2 – Matemática - 9º Ano

Matemática 9º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 22 alunos	1.629	250,51	1,06	0,685 (p>0,05)
acima 22 alunos		249,45		
até 23 alunos	1.873	251,07	1,24	0,537 (p>0,05)
acima 23 alunos		249,83		
até 24 alunos	2.142	252,66	1,46	0,223 (p>0,05)
acima 24 alunos		251,20		
até 25 alunos	2.322	252,68	1,04	0,172 (p>0,05)
acima 25 alunos		251,64		
até 26 alunos	2.500	252,76	0,69	0,527 (p>0,05)
acima 26 alunos		252,07		
até 27 alunos	2.661	252,76	0,62	0,706 (p>0,05)
acima 27 alunos		252,14		
até 28 alunos	2.816	252,76	-0,43	0,753 (p>0,05)
acima 28 alunos		253,19		
até 29 alunos	2.948	252,60	-0,94	0,248 (p>0,05)
acima 29 alunos		253,54		
até 30 alunos	2.927	252,76	-0,78	0,862 (p>0,05)
acima 30 alunos		253,54		
até 31 alunos	2.884	253,27	-0,79	0,715 (p>0,05)
acima 31 alunos		254,06		
até 32 alunos	2.870	253,8103	0,04	0,013 (p<0,05)
acima 32 alunos		253,7727		

Fonte: Elaboração própria.

Na tabela 27, nota-se que houve significância estatística nas comparações de turmas até 33, 34, 36, 37 e 38 alunos, em relação a salas maiores. Nessa seção, ficou claro que salas menores obtêm melhor desempenho, ainda que ligeiramente.

Há significância estatística também na comparação entre turmas de até 41 alunos e turmas maiores, no entanto verifica-se que turmas maiores apresentam melhor desempenho.

Tabela 27: Comparação detalhada – Seção 3 – Matemática - 9º Ano

Matemática 9º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 33 alunos	2.732	254,5607	0,7880	0,002 (p<0,05)
acima 33 alunos		253,7727		
até 34 alunos	2.483	253,8576	0,6578	0,012 (p<0,05)
acima 34 alunos		253,1998		
até 35 alunos	2.207	253,36	0,16	0,107 (p>0,05)
acima 35 alunos		253,20		
até 36 alunos	1.808	253,4177	0,3099	0,038 (p<0,05)
acima 36 alunos		253,1078		
até 37 alunos	1.356	252,2840	4,1928	0,001 (p<0,05)
acima 37 alunos		248,0912		
até 38 alunos	1.045	250,5095	2,2620	0,017 (p<0,05)
acima 38 alunos		248,2475		
até 39 alunos	780	249,57	-2,68	0,472 (p>0,05)
acima 39 alunos		252,25		
até 40 alunos	604	248,71	-5,40	0,141 (p>0,05)
acima 40 alunos		254,11		
até 41 alunos	327	247,2052	-7,2269	0,034 (p<0,05)
acima 41 alunos		254,4321		
até 42 alunos	208	247,71	-0,52	0,315 (p>0,05)
acima 42 alunos		248,23		

Fonte: Elaboração própria.

É possível observar também que, quando se focaliza o ilustrado na tabela 28, as comparações entre turmas de até 44, 45 e 47 alunos e turmas maiores, revelam que as classes maiores obtêm melhor desempenho, de forma acentuada.

Tabela 28: Comparação detalhada – Seção 4 – Matemática – 9º Ano

Matemática 9º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)

até 43 alunos	169	249,89	-4,54	0,475 ($p>0,05$)
acima 43 alunos		254,43		
até 44 alunos	116	247,1770	-20,35	0,000 ($p<0,05$)
acima 44 alunos		267,5242		
até 45 alunos	72	243,5252	-21,62	0,003 ($p<0,05$)
acima 45 alunos		265,1358		
até 46 alunos	35	242,97	-4,21	0,55 ($p>0,05$)
acima 46 alunos		247,18		
até 47 alunos	29	244,0724	-29,1185	0,028 ($p<0,05$)
acima 47 alunos		273,1909		
até 48 alunos	26	242,60	-21,48	0,58 ($p>0,05$)
acima 48 alunos		264,08		
até 49 alunos	19	244,07	-29,12	0,147 ($p>0,05$)
acima 49 alunos		273,19		

Fonte: Elaboração própria.

A comparação detalhada de *clusters* trouxe um resultado peculiar, assim como o ocorrido na disciplina Língua Portuguesa dos Anos Finais: as turmas maiores apresentaram melhor desempenho. No entanto, no caso da disciplina de Matemática do 9º ano, turmas menores, com até 32 alunos, mantiveram desempenho estatisticamente estável, experimentando perdas de desempenho quando a quantidade muda para de 33 a 40 alunos por turma. Isso mostra que classes com até 32 alunos têm desempenho melhor do que as de 33 a 40 alunos.

A partir de 41 alunos, é possível observar que turmas maiores apresentam melhor desempenho. Contudo, na comparação entre turmas de até 47 alunos e turmas de mais 47 alunos, verifica-se a maior mediana das comparações, no valor de 273,19 pontos. Logo, é possível oferecer como parâmetro para o gestor de políticas públicas em educação a hipótese de que salas de 41 alunos ou mais se desempenham melhor que salas menores.

Sendo assim, o número ideal de alunos por turma para a disciplina de Matemática dos Anos Finais deve ser maior que 47 alunos, número que apresenta o valor máximo das medianas.

O resultado encontrado mostra, ainda, que é aceitável – caso o gestor de políticas públicas em educação não disponha da infraestrutura necessária para montar uma sala que comporte 47 alunos ou mais – optar por formar uma sala de até 32 alunos. Mais uma vez, este trabalho demonstrou que, nos Anos Finais, turmas maiores se desempenham melhor, favorecendo a economia de recursos.

4.1.2 Anos Iniciais ou 5º ano

Também nos Anos Iniciais, inicialmente, analisa-se a disciplina de Língua Portuguesa; posteriormente, focaliza-se a a disciplina de Matemática.

4.1.2.1 Língua Portuguesa 5º Ano

Para analisar a disciplina de Língua Portuguesa referente ao 5ºano ou Anos Iniciais do Ensino Fundamental , optou-se por seguir a mesma sequência de comparações utilizadas para o 9º ano.

Assim, o primeiro passo foi realizar uma quantidade pequena de comparações, a fim de verificar como o modelo estatístico se comportava nessa disciplina, conforme ilustrado na tabela a seguir:

Tabela 29: Pequeno número de comparações – Língua Portuguesa – 5º Ano

Língua Portuguesa 5º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
Até 15 alunos	588	210,56	-1,71	0,403 (p>0,05)
Acima de 15 alunos		212,27		
Até 20 alunos	2.318	210,5574	-0,67	0,023 (p<0,05)
Acima de 20 alunos		211,2237		
Até 25 alunos	4.435	212,05	-0,65	0,560 (p>0,05)
Acima de 25 alunos		212,70		
Até 30 alunos	4.068	214,66	0,22	0,352 (p>0,05)
Acima de 30 alunos		214,44		
Até 35 alunos	1.221	212,93	4,10	0,002 (p<0,05)
Acima de 35 alunos		208,83		

Fonte: Elaboração própria.

Logo na pequena comparação, foi possível verificar que, tanto em salas de até 20 alunos, comparadas com salas de mais de 20 alunos, quanto em salas de até 35 alunos, comparadas com salas de mais de 35 alunos, houve indícios de significância estatística. A amostra apontou, inicialmente, para salas maiores com

melhores desempenhos e, em seguida, para salas menores com melhores desempenhos.

Por isso, foi necessário abrir as comparações em séries para verificar, de maneira mais direcionada, em que partes as comparações de *clusters* mostrariam ganhos ou perdas de desempenho, conforme se pode observar nesta outra tabela:

Tabela 30: Comparações por Séries – Seção 1 – Língua Portuguesa - 5º Ano

Língua Portuguesa 5º Ano				
Lógica	n	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 15 alunos	6	215,93	10,73	0,463 (p>0,05)
acima 40 alunos		205,20		
até 15 alunos	49	205,10	-1,71	0,644 (p>0,05)
acima 35 alunos		206,81		
até 15 alunos	155	211,81	1,95	0,962 (p>0,05)
acima 30 alunos		209,86		
até 15 alunos	348	209,97	-0,74	0,556 (p>0,05)
acima 25 alunos		210,71		
até 15 alunos	510	210,56	-1,81	0,394 (p>0,05)
acima 20 alunos		212,37		
até 20 alunos	39	203,45	1,69	0,586 (p>0,05)
acima 40 alunos		201,76		
até 20 alunos	253	208,67	1,23	0,929 (p>0,05)
acima 35 alunos		207,44		
até 20 alunos	851	210,56	0,44	0,088 (p>0,05)
acima 30 alunos		210,12		
até 20 alunos	1.672	210,1994	-0,50	0,004 (p<0,05)
acima 25 alunos		210,6946		

Fonte: Elaboração própria.

Na seção 1 da comparação por séries, é possível verificar que apenas a comparação entre turmas de até 20 alunos e turmas de mais de 25 alunos mostrou significância estatística, embora de forma ligeira. Ademais, verificou-se que turmas maiores se desempenham melhor.

Assim, observa-se, na tabela 31, que apenas na comparação entre turmas de até 30 alunos e turmas de mais de 35 alunos houve significância estatística, revelando que turmas menores apresentam melhor desempenho.

Tabela 31: Comparações Séries – Seção 2 – Língua Portuguesa - 5º Ano

Língua Portuguesa 5º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 25 alunos	89	203,45	-1,07	0,773 (p>0,05)
acima 40 alunos		204,52		
até 25 alunos	690	211,38	2,75	0,371 (p>0,05)
acima 35 alunos		208,63		
até 25 alunos	2.584	213,95	2,04	0,798 (p>0,05)
acima 30 alunos		211,91		
até 30 alunos	149	207,79	2,87	0,889 (p>0,05)
acima 40 alunos		204,92		
até 30 alunos	1.309	212,9235	2,26	0,006 (p<0,05)
acima 35 alunos		210,6638		
até 35 alunos	195	206,62	1,69	0,575 (p>0,05)
acima 40 alunos		204,93		

Fonte: Elaboração própria.

Na análise por séries, identificaram-se apenas duas comparações com significância estatística, a saber: a comparação entre turmas com até 20 alunos e turmas de mais de 25 alunos; e a comparação entre turmas com até 30 alunos e turmas de mais de 35 alunos.

Apenas com o resultado da comparação teste e da comparação por séries não foi possível identificar, de forma clara, se o número de alunos por turma teria ou não influência no desempenho escolar dos estudantes de 5ºano na disciplina de Língua Portuguesa. Por isso, foi necessário utilizar a comparação detalhada para o alcance do resultado, conforme apresentado a seguir:

Tabela 32: Comparação detalhada – Seção 1 – Língua Portuguesa - 5º Ano

Língua Portuguesa 5º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 08 alunos	5	224,77	0,01	0,686 (p>0,05)
acima 08 alunos		224,76		
até 09 alunos	12	235,65	9,67	0,209 (p>0,05)
acima 09 alunos		225,98		
até 10 alunos	13	235,65	10,89	0,133 (p>0,05)
acima 10 alunos		224,76		
até 11 alunos	22	224,77	14,19	0,758 (p>0,05)

acima 11 alunos		210,58		
até 12 alunos	102	205,09	-4,99	0,932 (p>0,05)
acima 12 alunos		210,08		
até 13 alunos	211	211,08	0,99	0,730 (p>0,05)
acima 13 alunos		210,09		
até 14 alunos	369	211,41	1,55	0,071 (p>0,05)
acima 14 alunos		209,86		
até 15 alunos	588	210,56	-1,71	0,403 (p>0,05)
acima 15 alunos		212,27		
até 16 alunos	881	210,5574	-3,32	0,012 (p<0,05)
acima 16 alunos		213,8777		
até 17 alunos	1.210	210,5574	-2,43	0,017 (p<0,05)
acima 17 alunos		212,9862		

Fonte: Elaboração própria.

Pode-se observar, na tabela 32, que, na comparação de turmas de até 16 e 17 alunos com turmas maiores, estas últimas apresentam melhor desempenho.

Na tabela 33, observa-se que turmas de até 20 alunos, quando comparadas com turmas maiores, mostram que as maiores se desempenham melhor. Em contrapartida, a comparação entre turmas de até 21, 22 e 23 alunos e turmas maiores revelou que as classes menores se desempenham melhor.

Tabela 33: Comparação detalhada – Seção 2 – Língua Portuguesa - 5º Ano

Língua Portuguesa 5º Ano				
Lógica	n	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 18 alunos	1.540	210,55	-0,85	0,275 (p>0,05)
acima 18 alunos		211,40		
até 19 alunos	1.923	210,35	-0,33	0,071 (p>0,05)
acima 19 alunos		210,68		
até 20 alunos	2.318	210,5574	-0,67	0,023 (p<0,05)
acima 20 alunos		211,2237		
até 21 alunos	2.798	211,4544	0,25	0,009 (p<0,05)
acima 21 alunos		211,2091		
até 22 alunos	3.288	212,0640	0,50	0,040 (p<0,05)
acima 22 alunos		211,5619		
até 23 alunos	3.736	212,4137	0,49	0,036 (p<0,05)
acima 23 alunos		211,9233		
até 24 alunos	4.182	213,69	1,13	0,416 (p>0,05)
acima 24 alunos		212,56		
até 25 alunos	4.435	212,05	-0,65	0,560 (p>0,05)

acima 25 alunos		212,70		
até 26 alunos	4.769	214,04	0,72	0,223 (p>0,05)
acima 26 alunos		213,32		
até 27 alunos	4.807	214,94	0,86	0,665 (p>0,05)
acima 27 alunos		214,08		
até 28 alunos	4.658	215,68	0,96	0,879 (p>0,05)
acima 28 alunos		214,72		

Fonte: Elaboração própria.

Ao comparar as turmas de até 34 e 35 alunos com turmas maiores, verifica-se que as classes menores apresentam melhor desempenho:

Tabela 34: Comparação detalhada – Seção 3 – Língua Portuguesa - 5º Ano

Língua Portuguesa 5º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 29 alunos	4.482	215,38	1,29	0,270 (p>0,05)
acima 29 alunos		214,09		
até 30 alunos	4.068	214,66	0,22	0,352 (p>0,05)
acima 30 alunos		214,44		
até 31 alunos	3.562	213,70	0,41	0,294 (p>0,05)
acima 31 alunos		213,29		
até 32 alunos	2.828	213,73	-0,16	0,189 (p>0,05)
acima 32 alunos		213,89		
até 33 alunos	2.328	212,83	0,23	0,146 (p>0,05)
acima 33 alunos		212,60		
até 34 alunos	1.740	213,0719	2,37	0,002 (p<0,05)
acima 34 alunos		210,6946		
até 35 alunos	1.221	212,9300	4,10	0,002 (p<0,05)
acima 35 alunos		208,8341		
até 36 alunos	760	211,56	2,99	0,150 (p>0,05)
acima 36 alunos		208,57		
até 37 alunos	441	208,04	3,12	0,985 (p>0,05)
acima 37 alunos		204,92		
até 38 alunos	293	208,02	-0,09	0,223 (p>0,05)
acima 38 alunos		208,11		
até 39 alunos	204	205,99	-1,60	0,365 (p>0,05)
acima 39 alunos		207,59		

Fonte: Elaboração própria.

Observe-se que nenhuma das comparações desta seção apresentou significância estatística. Isso se percebe melhor na próxima tabela:

Tabela 35: Comparação detalhada – Seção 4 – Língua Portuguesa - 5º Ano

Língua Portuguesa 5º Ano				
Lógica	n	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 40 alunos	127	208,93	4,01	0,364 (p>0,05)
acima 40 alunos		204,92		
até 41 alunos	69	205,77	8,25	0,749 (p>0,05)
acima 41 alunos		197,52		
até 42 alunos	57	205,77	13,27	0,682 (p>0,05)
acima 42 alunos		192,50		
até 43 alunos	29	209,06	-6,97	0,456 (p>0,05)
acima 43 alunos		216,03		
até 44 alunos	23	207,34	-8,69	0,927 (p>0,05)
acima 44 alunos		216,03		
até 45 alunos	18	204,92	-17,67	0,085 (p>0,05)
acima 45 alunos		222,59		

Fonte: Elaboração própria.

Com a comparação detalhada, verificou-se que, na comparação entre turmas de até 16 alunos e turmas com mais de 16 alunos, as turmas maiores apresentaram melhor desempenho. O resultado se repete na comparação entre turmas de até 17 e mais de 17 alunos.

As turmas maiores novamente demonstram ter melhor desempenho na comparação entre turmas com até 20 alunos e turmas com mais de 20 alunos. Em contrapartida, na comparação entre turmas de até 21 alunos e turmas com mais de 21 alunos, as classes menores se desempenham melhor, dando mais robustez a esse resultado. Isso também acontece nas comparações entre turmas de até 22 alunos e turmas com mais de 22 alunos por turma, bem como nas turmas de até 23 alunos e nas turmas com mais de 23 alunos, revelando que classes menores se desempenham melhor.

Novamente, na comparação entre turmas de até 34 e até 35 alunos e turmas de mais de 34 e mais de 35 alunos, detectou-se que classes menores se desempenham melhor que as maiores.

Portanto, ao observar a comparação de turmas com 20 alunos e a comparação de turmas com 21 alunos, percebe-se que, no primeiro caso, as salas maiores ou com mais de 20 alunos apresentam melhor desempenho. Ademais, a comparação de 21 alunos mostra que as classes menores ou com até 21 alunos apresentam melhor desempenho.

Dessa forma, pode-se chegar à conclusão de que o número ideal de alunos por turma na disciplina de Língua Portuguesa dos Anos Iniciais é de 21 alunos. Esse resultado garante que futuras políticas públicas em educação definam, com base estatística, o número ideal de estudantes na disciplina e no ano em questão.

4.1.2.2 Matemática – 5º Ano

Para analisar a disciplina de Matemática referente aos anos iniciais, optou-se por utilizar a mesma sequência de testes de comparações, que se mostraram promissores nas outras análises. Isso posto, o primeiro passo é analisar a pequena quantidade de comparações, conforme a tabela:

Tabela 36: Pequeno número de comparações – Matemática - 5º Ano

Matemática 5º Ano				
Lógica	n	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
Até 15 alunos	588	217,56	-2,73	0,056 (p>0,05)
Acima de 15 alunos		220,29		
Até 20 alunos	2.318	219,58	0,06	0,278 (p>0,05)
Acima de 20 alunos		219,52		
Até 25 alunos	4.435	222,45	-0,06	0,062 (p>0,05)
Acima de 25 alunos		222,51		
Até 30 alunos	4.068	223,00	1,09	0,031 (p<0,05)
Acima de 30 alunos		221,91		
Até 35 alunos	1.221	220,88	7,19	0,001 (p<0,05)
Acima de 35 alunos		213,68		

Fonte: Elaboração própria.

Pode-se observar que houve significância estatística na comparação entre turmas com até 35 alunos e turmas com mais de 35 alunos, indicando preliminarmente que turmas menores apresentam melhores desempenhos.

Com o intuito de verificar se o resultado da pequena comparação se mantém ou se altera, optou-se por analisar a comparação por séries, como se verifica a seguir:

Tabela 37: Comparações por Séries – Seção 1 – Matemática - 5º Ano

Matemática 5º Ano				
Lógica	n	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 15 alunos	6	227,09	6,32	0,753 (p>0,05)
acima 40 alunos		220,77		
até 15 alunos	49	215,43	3,31	0,574 (p>0,05)
acima 35 alunos		212,12		
até 15 alunos	155	217,58	-0,09	0,966 (p>0,05)
acima 30 alunos		217,67		
até 15 alunos	348	217,57	-2,49	0,398 (p>0,05)
acima 25 alunos		220,06		
até 15 alunos	510	217,57	-3,18	0,062 (p>0,05)
acima 20 alunos		220,75		
até 20 alunos	39	216,44	13,65	0,387 (p>0,05)
acima 40 alunos		202,79		
até 20 alunos	253	216,89	3,40	0,113 (p>0,05)
acima 35 alunos		213,49		
até 20 alunos	851	219,24	0,19	0,955 (p>0,05)
acima 30 alunos		219,05		
até 20 alunos	1.672	219,23	0,06	0,206 (p>0,05)
acima 25 alunos		219,17		

Fonte: Elaboração própria.

Na tabela 37, observa-se que nenhuma comparação demonstrou haver significância estatística, diferentemente do que ocorre na seção 2, representada na tabela 38. Nela se observa que, na comparação entre turmas de até 25 alunos e turmas de mais de 35 alunos, entre turmas de até 25 alunos e turmas de mais de 30 alunos, e na comparação entre turmas de até 30 alunos e turmas de mais de 35 alunos, as turmas menores apresentam melhor desempenho que as maiores.

Tabela 38: Comparações por Séries – Seção 2 – Matemática - 5º Ano

Matemática 5º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 25 alunos	89	215,73	12,94	0,365 (p>0,05)
acima 40 alunos		202,79		
até 25 alunos	690	219,9094	5,4955	0,008 (p<0,05)

acima 35 alunos		214,4139		
até 25 alunos	2.584	222,8159	2,7908	0,024 (p<0,05)
acima 30 alunos		220,0251		
até 30 alunos	149	216,44	10,81	0,075 (p>0,05)
acima 40 alunos		205,63		
até 30 alunos	1.309	221,5522	6,8332	0,000 (p<0,05)
acima 35 alunos		214,7190		
até 35 alunos	195	215,72	9,04	0,135 (p>0,05)
acima 40 alunos		206,68		

Fonte: Elaboração própria.

Com intuito de verificar se o número de alunos por turma influenciava no desempenho escolar, optou-se por detalhar ainda mais as comparações de *clusters*, conforme a tabela 39:

Tabela 39: Comparação detalhada – Seção 1 – Matemática - 5º Ano

Matemática 5º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 08 alunos	5	223,80	6,13	0,686 (p>0,05)
acima 08 alunos		217,67		
até 09 alunos	12	225,62	-11,40	0,158 (p>0,05)
acima 09 alunos		237,02		
até 10 alunos	13	225,62	-9,54	0,249 (p>0,05)
acima 10 alunos		235,16		
até 11 alunos	22	223,80	5,46	0,661 (p>0,05)
acima 11 alunos		218,34		
até 12 alunos	102	210,27	-7,10	0,379 (p>0,05)
acima 12 alunos		217,37		
até 13 alunos	211	214,64	-5,24	0,158 (p>0,05)
acima 13 alunos		219,88		
até 14 alunos	369	219,90	1,59	0,410 (p>0,05)
acima 14 alunos		218,31		
até 15 alunos	588	217,56	-2,73	0,056 (p>0,05)
acima 15 alunos		220,29		
até 16 alunos	881	220,1121	-2,3398	0,021 (p<0,05)
acima 16 alunos		222,4519		
até 17 alunos	1.210	218,6385	-3,3244	0,033 (p<0,05)
acima 17 alunos		221,9629		

Fonte: Elaboração própria.

Aqui se verifica que turmas de até 16 alunos e até 17 alunos, quando comparadas com turmas maiores, revelam que estas últimas obtêm melhores resultados que as menores.

Já na tabela 40, observa-se que nenhuma das comparações feitas demonstrou significância estatística:

Tabela 40: Comparação detalhada – Seção 2 – Matemática - 5º Ano

Matemática 5º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 18 alunos	1.540	219,14	-0,76	0,693 (p>0,05)
acima 18 alunos		219,90		
até 19 alunos	1.923	218,98	-0,13	0,438 (p>0,05)
acima 19 alunos		219,11		
até 20 alunos	2.318	219,58	0,06	0,278 (p>0,05)
acima 20 alunos		219,52		
até 21 alunos	2.798	221,25	1,73	0,117 (p>0,05)
acima 21 alunos		219,52		
até 22 alunos	3.288	222,19	2,15	0,720 (p>0,05)
acima 22 alunos		220,04		
até 23 alunos	3.736	222,19	1,68	0,894 (p>0,05)
acima 23 alunos		220,51		
até 24 alunos	4.182	222,89	2,04	0,151 (p>0,05)
acima 24 alunos		220,85		
até 25 alunos	4.435	222,45	-0,06	0,062 (p>0,05)
acima 25 alunos		222,51		
até 26 alunos	4.769	223,41	1,83	0,405 (p>0,05)
acima 26 alunos		221,58		
até 27 alunos	4.807	223,62	1,63	0,174 (p>0,05)
acima 27 alunos		221,99		
até 28 alunos	4.658	224,07	1,66	0,179 (p>0,05)
acima 28 alunos		222,41		

Fonte: Elaboração própria.

Nota-se que houve significância estatística na tabela 41, diferentemente do que ocorreu na tabela 40, nas comparações entre turmas de até 29 alunos, até 30 alunos, até 31 alunos, até 32 alunos, até 33 alunos, até 34 alunos, até 35 alunos e até 36 alunos e turmas maiores. Nesse intervalo, foi possível verificar que as turmas menores obtiveram melhor desempenho.

Tabela 41: Comparação detalhada – Seção 3 – Matemática - 5º Ano

Matemática 5º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 29 alunos	4.482	223,6133	1,7013	0,017 (p<0,05)
acima 29 alunos		221,9120		
até 30 alunos	4.068	222,9976	1,0856	0,031 (p<0,05)
acima 30 alunos		221,9120		
até 31 alunos	3.562	221,8295	1,1659	0,002 (p<0,05)
acima 31 alunos		220,6636		
até 32 alunos	2.828	221,5416	0,8910	0,000 (p<0,05)
acima 32 alunos		220,6506		
até 33 alunos	2.328	220,7266	1,3429	0,000 (p<0,05)
acima 33 alunos		219,3837		
até 34 alunos	1.740	221,1040	5,6629	0,000 (p<0,05)
acima 34 alunos		215,4411		
até 35 alunos	1.221	220,8753	7,1905	0,000 (p<0,05)
acima 35 alunos		213,6848		
até 36 alunos	760	218,3807	6,9254	0,000 (p<0,05)
acima 36 alunos		211,4553		
até 37 alunos	441	215,73	8,29	0,052 (p>0,05)
acima 37 alunos		207,44		
até 38 alunos	293	215,55	8,67	0,372 (p>0,05)
acima 38 alunos		206,88		
até 39 alunos	204	214,63	7,75	0,148 (p>0,05)
acima 39 alunos		206,88		

Fonte: Elaboração própria.

A tabela 42, diferentemente das outras, demonstrou significância estatística em apenas uma comparação: entre turmas com até 43 alunos e turmas maiores, indicando, novamente, que turmas menores apresentam melhor desempenho, conforme ilustrado a seguir:

Tabela 42: Comparação detalhada – Seção 4 – Matemática - 5º Ano

Matemática 5º Ano				
Lógica	N	Mediana	Diferença mediana (até – acima)	p-valor (teste de Wilcoxon)
até 40 alunos	127	216,55	9,67	0,373 (p>0,05)
acima 40 alunos		206,88		
até 41 alunos	69	214,41	10,21	0,098 (p>0,05)
acima 41 alunos		204,20		
até 42 alunos	57	214,41	16,64	0,055 (p>0,05)
acima 42 alunos		197,77		
até 43 alunos	29	215,7309	8,8435	0,039 (p<0,05)
acima 43 alunos		206,8874		
até 44 alunos	23	214,41	7,52	0,191 (p>0,05)
acima 44 alunos		206,89		
até 45 alunos	18	213,42	6,54	0,647 (p>0,05)
acima 45 alunos		206,88		

Fonte: Tabela produzida a partir dos dados da análise estatística deste trabalho.

Ao observar a comparação detalhada para a disciplina de Matemática nos Anos Iniciais, verifica-se que, na comparação entre turmas de até 16 alunos e turmas com mais de 16 alunos, bem como na comparação entre turmas com até 17 alunos e mais de 17 alunos, as turmas maiores apresentam melhor desempenho em relação às menores.

Em contrapartida, na comparação entre turmas de até 29, 30, 31, 32, 33, 34, 35, 36 e 43 alunos e turmas maiores, observa-se que as menores obtêm melhor desempenho.

Diante disso, é possível oferecer, como parâmetro, a hipótese de que o número de alunos por turma, para a disciplina da Matemática nos Anos Iniciais, deve ser maior que 17 alunos e menor ou igual a 29. Acrescente-se que a melhor mediana encontrada no modelo se refere a turmas com até 29 alunos, no valor de 223,6133. Esse resultado é capaz oferecer ao gestor de políticas públicas uma margem de segurança ao montar uma turma para o ano e para a disciplina em questão.

5 PRODUTO FINAL: PLANO DE AÇÃO COM FOCO NA IMPLEMENTAÇÃO DE POLÍTICAS PÚBLICAS

O desenvolvimento e implementação deste plano de ação ocorrerão com base nos resultados obtidos nesta pesquisa. Sendo assim, o uso dos resultados observará um ciclo de políticas públicas em educação.

Assim, a partir da necessidade de se criar uma política pública em educação, por exemplo, a manipulação do número de alunos por sala de aula, o presente plano de ação será responsável por direcionar o modo como se dará a criação dessa política pública por etapas.

Por meio de uma ferramenta dinâmica para computador, serão formuladas etapas para que o gestor de políticas públicas siga e desenvolva a política pública que desejar.

Além disso, esta ferramenta estabelecerá um meio de monitorar e avaliar o trabalho desenvolvido com vistas a atingir metas e a obter sucesso nas políticas públicas em educação.

6 CONCLUSÃO

Este estudo teve, por objetivo geral, verificar a influência do número de alunos por sala de aula no seu desempenho escolar, tanto nos Anos Iniciais quanto nos Anos Finais do Ensino Fundamental. .

Para isso, utilizou-se da literatura nacional e internacional, a fim de mapear grande parte dos estudos acerca da manipulação dos alunos por turma, bem como do nexo de causalidade entre tal manipulação e o desempenho escolar discente.

Em sua maioria, os estudos não encontraram relação direta entre o tamanho da turma e o desempenho escolar. Mesmo assim, optou-se por testar se esta hipótese se aplicaria para o Brasil, pois, segundo Hanushek (1998), é provável que, em situações com professores específicos, grupos de estudantes específicos e assuntos específicos, turmas menores possam contribuir para a melhoria do desempenho escolar.

Nesse sentido, para evitar que outras variáveis – além do número de alunos por turma – pudessem intervir nessa hipótese, mapearam-se todas as variáveis intervenientes que poderiam influenciar o desempenho escolar, tais como a escolaridade dos pais, o nível acadêmico dos professores, a infraestrutura escolar, entre outros. As legislações sobre o tema também foram mapeadas. O intuito foi verificar se alguma delas determinava o número de alunos por turma, evitando que a pré-seleção pudesse servir de variável interveniente.

Assim, utilizou-se um sofisticado modelo estatístico formado pela combinação do *Propense Score Matching*, da distância euclidiana e do *Nearest Neighbor Matching*.

Como resposta aos objetivos específicos “a” e “b”, com foco no Brasil, verificou-se que há uma disparidade alta entre a quantidade de estudantes por turma, se observadas a iniciativa privada e a pública.

Ao detalhar os dados em nível regional, verificou-se que, exceto a região Sul – onde essa disparidade é pequena – as demais regiões acompanham o resultado nacional. Desse modo, nota-se uma grande diferença entre a quantidade de alunos por turma da iniciativa privada, em relação à iniciativa pública.

Para responder ao objetivo geral e ao objetivo “c”, utilizou-se o modelo estatístico, e o resultado dessa análise mostrou que o número de alunos por turma tem pouca influência no desempenho escolar discente, tanto nos Anos Iniciais quanto nos Anos Finais. Além disso, quando essa influência existe, turmas maiores tendem a se desempenhar melhor que as menores.

Ao comparar os resultados encontrados nesta dissertação com os trabalhos citados no referencial teórico (WORD, 1990; MAIER *et al.*1997; GILMAN, 1998; STECHER; BOHRNSTEDTST, 2002; BASCIA, 2010; YIGIT *et al.*, 2017; FREDRIKSSON; OCKERT; OOSTERBEEK, 2011; BLATCHFORD; BASSET; BROWN, 2011; CAMARGO, 2014; OLIVEIRA, 2008; TRAVITZKI; CÁSSIO, 2017; MATAVELLI, 2018; SOOTER, 2013; YUSUF; ONIFADE; BELLO, 2016; O. AJAYI; AUDU; E. AJAYI 2017; SHAFRIR; SHAVIT; BLANK, 2016.) , observou-se que, nos Anos Finais do Ensino Fundamental, as salas maiores obtiveram melhor desempenho que as menores, quando houve significância estatística. Tanto na disciplina de Português como na disciplina de Matemática, percebeu-se não haver estudos que coincidem com esses resultados.

Tal como as pesquisas presentes no referencial teórico, este estudo contribuiu de forma única para a literatura sobre o tema. Nesse sentido, salienta-se que, tanto os estudos da América do Norte como o da Califórnia (STECHER; BOHRNSTEDTST, 2002), de Wisconsin (MAIER *et al.*1997) e do Canadá (BASCIA, 2010), quanto os estudos da Inglaterra, do País de Gales (BLATCHFORD; BASSET; BROWN, 2011) e da Nigéria (SOOTER, 2013; YUSUF; ONIFADE; BELLO, 2016; O. AJAYI; AUDU; E. AJAYI 2017) divergem do senso comum, ao demonstrar que a quantidade menor de alunos por turma é positiva para os professores, e não propriamente para o desempenho do estudante. Diante disso, conclui-se que cada estudo pôde contribuir de forma diferente para o tema desenvolvido por esta dissertação.

Quando o foco recaiu sobre os Anos Iniciais do Ensino Fundamental, a pesquisa mostrou que turmas menores, comparações estatisticamente significantes, tiveram melhor desempenho que turmas maiores.

No caso da disciplina de Língua Portuguesa, esse número é de até 21 alunos e, no caso de Matemática, de até 29 alunos. Assim, os trabalhos que corroboram esse resultado são o Tennessee Star Project (WORD, 1990), o Projeto Prime Time (GILMAN, 1988), e a dissertação “Custo-efetividade de políticas de redução do tamanho da classe e ampliação da jornada escolar: uma aplicação de estimadores de matching” (OLIVEIRA, 2008)

Hanushek (1998) e Hoxby (1998) afirmam que o Projeto Star apresentou efeito único e duradouro no desempenho dos alunos nos Anos Iniciais e que pequenas turmas têm uma melhora diferenciada na realização. Já o projeto Prime Time, segundo Gilman (1988), demonstra que os estudantes dos Anos Iniciais obtiveram melhor desempenho quando colocados em turmas menores, em comparação a quando pertenciam a turmas maiores. Sob essa ótica, Oliveira (2008) afirma que o seu trabalho encontrou um efeito positivo na diminuição do tamanho da turma dos Anos Iniciais, de 38 para 30 alunos, na disciplina de Matemática.

De forma mais abrangente, no que concerne ao resultado geral desta pesquisa – em que se verificou pouca influência do número de estudantes por turma no desempenho escolar – pode-se perceber semelhanças com os resultados encontrados nos estudos de São Paulo (MATAVELLI, 2018), Santa Catarina (MATAVELLI, 2018), Rio Grande do Sul (CAMARGO, 2014) e Israel (SHAFRIR; SHAVIT; BLANK, 2016) que não verificaram relação entre o número de alunos por turma e o desempenho escolar.

A partir do resultado encontrado nesta dissertação, é possível oferecer ao gestor de políticas públicas em educação, como produto final deste trabalho, dados que podem servir de parâmetro para a hipótese de manipulação de alunos por turma, como forma de melhorar o desempenho escolar.

Por meio dos dados aqui fornecidos, o gestor é capaz de verificar que o número de alunos por turma se comporta de formas diferentes, nos Anos Iniciais e nos Anos Finais. Além disso, é possível constatar que há diferenças de comportamento dos dados, quando se analisam as disciplinas de Português e Matemática.

Assim, esta pesquisa serve de contribuição para literatura educacional, no sentido de demonstrar que o número de alunos por turma não apresenta influência no desempenho escolar dos estudantes do Ensino Fundamental brasileiro, auxiliando tanto o gestor de políticas públicas em educação, quanto os diretores de escolas.

As limitações aqui encontradas relacionam-se à tentativa de utilizar tanto os dados do Censo Escolar, quanto os dados da Prova Brasil. No decorrer de sua execução, algo se mostrou inviável, uma vez que o código das turmas do Censo Escolar não era o mesmo utilizado pela Prova Brasil. Diante disso, foi necessária a escolha de apenas uma matriz de dados, que foi a Prova Brasil. A escolha dessa base deu-se em virtude da existência de uma pergunta presente no questionário do Diretor, concernente ao modo como ele tinha realizado, naquele ano, a divisão das turmas em sua escola. Um dos diferenciais deste trabalho está justamente na utilização dessa questão como filtro, pois é capaz de evitar a contaminação dos dados por uma prévia seleção feita pelo diretor da escola, antes mesmo da coleta dos dados.

Para futuros estudos, esta dissertação deixa como sugestão uma pesquisa focada na iniciativa privada, em escolas com metodologia diferenciada de ensino. Acredita-se que, dessa forma, seja possível verificar se o número de alunos por turma é capaz de influenciar, efetivamente, o desempenho escolar.

REFERÊNCIAS

ABERG-BENGTSSON, Lisbeth. The Smaller the Better? A Review of Research on Small Rural Schools in Sweden. **International Journal of Educational Research** Suécia, v.48. p. 100-108 fev. 2009.

AJAYI, Victor.; AUDU, Christina.; AJAYI, Evelyn Influence of Class Size on Student's Classroom Discipline, Engagement and Communication: A Case Study of Senior Secondary Schools in Ekiti State. **Sky Journal of Education** Nigéria, v.5. p. 34-41. nov 2017.

ALUFOHAI, Peace J. Influence of Teachers' Age, Marital Status and Gender on Students' Academic Achievement, **Asian Journal of Educational Research** Nigéria, v. 3 p. 60-66 n.4. dez. 2015

ALVES, Maria T. G.; SOARES, José F. Efeito-Escola e Estratificação Escolar: o Impacto da Composição de Turmas por Nível de Habilidade dos Alunos. **Educação em Revista**, Belo Horizonte, v. 45. p. 25-58. jun. 2007

ALVES, Teresa G.; SOARES, José F.; XAVIER, Flávia P. Índice Socioeconômico das Escolas de Educação Básica Brasileiras. **Avaliação em Políticas Públicas Educação**, Rio de Janeiro, v. 22 n.84, p. 671-704. 2014.

AMARAL, Nelson C. Financiamento da Educação Básica e o PNE Ainda e Sempre, Muitos Desafios. **Revista Retratos da Escola**, Brasília, v.8, n.15, p.293-311, jul/dez 2014.

AMARAL, Luiz F. L. E.; MENEZES-FILHO, Naércio. A Relação Entre Gastos Educacionais e Desempenho Escolar. **Research Papers in Economics**, São Paulo, nov. 2009.

AMORIM, Erica; PARANDEKAR, Suhas; WELSH, Andrea; "en breve" is produced by the Knowledge and Learning Team of the Operations Services Department of the Latin America and the Caribbean Region of the World Bank. **World Bank**, Washington, n.163, nov. 2008.

ANDRADE, Aldair O.; BELLO, Mess L. S.; SOUZA, Maria R.; REIS, Daiana S. LDB nº 9.394/96, da Formação Básica do Cidadão: Satisfação dos Estudantes do 9º ano na Disciplina de Língua Portuguesa em Uma Escola Estadual no Município de Humaitá/AM. **Série Educar**, Amazonas, v.38, c.4, p.29-36, jul. 2019.

ANDRADE, Edison P. O Sentido do Dever e o Direito da Criança e do Adolescente. **IV CONGRESSO INTERNACIONAL DE PEDAGOGIA SOCIAL** São Paulo, jul. 2012.

ANDRADE, Edson F. Regime de Colaboração no financiamento da Educação Básica no Brasil: A experiência dos Fundos Contábeis em Foco. **Revista Educação**, Santa Maria, v.38, n.2, p.389-402, mai/ago 2013.

ANDRADE, Josemberg M.; LAROS, Jacob A. Fatores Associados ao Desempenho Escolar: Estudo Multinível com Dados do SAEB/2001. **Psicologia: Teoria e Pesquisa**, Brasília, n.23 (1), nov.2007.

ANTUNES, Adriene B. D. O Debate da Avaliação em Larga Escala no Brasil: Características, Referências e Tendências. **Colóquio Luso-Brasileiro de Educação Portugal**, v.3, mai. 2018.

ANTUNES, Ana P.; CAROLINE, Larissa; LARA, Djenissis; SOFFA, Patricia. A Formação do Professor Frente a Educação Inclusiva. **Conhecimento Interativo**, Paraná, v.11, n.2, p. 91-104, jul. /dez. 2017.

ARAUJO, Edilson P. Uma Análise Sobre os Resultados do IDEB e a Visão dos Gestores das Escolas Públicas do Município de Salgueiro. **Revista OPARA**, Pernambuco, v. 8, p. 26-41, jul. 2018..

ARAUJO, Heleno.; GOUVEIA, Andrea B.; OLIVEIRA, João F. **Caderno de Avaliação das Metas do Plano Nacional de Educação: PNE 2014-2024**. Anpae, Brasília, 2018.

ARAUJO, Maria L. H. S. Avaliação Internacional: Concepções Inerentes ao PISA e Seus Resultados no Brasil. **Equidade no Ensino Superior**. Bahia, out. 2012.

ARAUJO, Monica R.; MELO, Sarita B.F. Ativamente: Proposta de Formação em Educação Empreendedora para os Instrutores do SENAT de Criciúma – SC. **Genesis**, Santa Catarina, 2017.

ARONSON, J.K.; BANKHEAD C.; NUNAN, D. Catalogue of Bias Collaboration. **BMJ Journal**,. Inglaterra v. 23, n.1, fev. 2018.

ARPINO, Bruno.; CANNAS, Massimo.; Comparing Different Approaches for Propensity Score Matching with Clustered Data: A Simulation Study. **RECSM – Research and Expertise Centre for Survey Methodology**, Itália, mar. 2015.

BARROS, Ricardo; DE MENDONÇA, Rosane P.; DOS SANTOS, Daniel D.; QUINTAES, Giovanni. Determinantes do Desempenho Educacional no Brasil, **IPEA**, Rio de Janeiro, out. 2001.

BASCIA, Nina. Reducing Class Size: What Do We Know? **Canadian Education Association**, Canadá, p. 22, 2010.

BECKER, Fernanda R. Avaliação Educacional em Larga Escala: A Experiência Brasileira. **Revista Iberoamericana de Educación/Revista Ibero-americana de Educação**, Espanha, v. 53, n.1 p. 1-10, jun. 2010.

BECKMANN, Andrea.; HILL, Dave.; COOPER, Charlie. Neoliberalization and Managerialization of Education in England and Wales - A Case for Reconstructing Education. **Journal for Critical Education Policy Studies**, Reino Unido, v.7, n.2, p. 310-345, nov. 2009.

BEM-DAVID, Dan.; KIMHI, Ayal. Economics of Education in Israel: Inputs, Outputs and Performance. **Shoresh Institution for Socioeconomic Research**, Israel, v.1, p. 65, dez. 2017.

BERKHIN, Pavel.; KOGAN, Jacob.; NICHOLAS, Charles.; TEBOULLE, Marc. A Survey of Clustering Data Mining Techniques, **Gruping Multidimensional Data**, Berlin, p. 25-71, 2006.

BERNADO, Elisangela S. Organização de Turmas: Uma Prática de Gestão Escolar em Busca de Uma Escola Eficaz. **Revista Educação e Cultura Contemporânea**, São Paulo, v.10, n. 21, 2013.

BERTHOLD, Michael R.; HOPNER, Frank. On Clustering Time Series Using Euclidean Distance and Pearson **Correlation**, **ArXiv – Computer Science Mathematics**, Alemanha, v.1, jan, 2016.

BIA, Michela.; MATTEI, Alessandra. Application of the Generalized Propensity Score: Evaluation of Public Contributions to Piedmont Enterprises, **Department of Public Policy and Public Choice – POLIS**, Itália, n. 89, p.92, abr. 2007.

BIESTA, Gert; SÄFSTRÖM, Carl A. A Manifesto For Education. **SAGE Journals**, Suécia, v.9, n.5, p.8, jan, 2011.

BLANK, Carmel; SHAFRIR, Reut; SHAVIT, Yossi. Is Less Really More? On the Relationship Between Class Size and Educational Achievement in Israel. **TAUB Center for Social Studies in Israel**, Israel, 2016.

BLATCHFORD, Peter; BASSET, Paul; BROWN, Penelope. Examining the Effect of Class Size on Classroom Engagement and Teacher-pupil Interaction: Differences in Relation to Pupil Prior Attainment and Primary Vs. Secondary Schools. **ELSEVIER**, Londres, v.21, n.6, dez 2010.

BOEIJE, Hennie R.; HOX, Joop J. Data Collection, Primary vs. Secondary. **Encyclopedia of social measurement**, Holanda, v.1, p.593 -599, 2005.

BOHRNSTEDT, George W.; STECHER, Brian M. Class Size Reduction in California: Findings From 1999–00 and 2000–01 **CSR Research Consortium**, Califórnia, v.1, p.160, fev. 2002.

BOLZAM, Angelina C.; KIM, Richard P. Direito à educação de qualidade e seus fundamentos jurídicos; **Cadernos de Direito**, Piracicaba, v.15, p.167-200, jul./dez. 2015.

BRASIL. **Constituição da República Federativa do Brasil**. Brasília, DF: Centro Gráfico, 1988.

BRASIL. **Decreto nº 6.425**, de 4 de abril de 2008. Dispõe sobre o censo anual da educação. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2008/Decreto/D6425.htm. Acesso em: 11 jan. 2021.

BRASIL. **Lei nº 13.005, de 25 de junho de 2014**. Aprova o Plano Nacional de Educação - PNE e dá outras providências. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l13005.htm. Acesso em: 11 jan. 2021

BRASIL. **Lei nº 8.069, de 13 de julho de 1990**. Dispõe sobre o Estatuto da Criança e do Adolescente e dá outras providências. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8069.htm. Acesso em: 11 jan. 2021

BRASIL. **Lei nº 9.394, de 20 de dezembro de 1996.** Estabelece as diretrizes e bases da educação nacional. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l9394.htm. Acesso em: 11 jan. 2021.

BRASIL. **Decreto-Lei n.º 75/2008, de 22 de abril.** Dispõe sobre a formação do Conselho Geral de Educação. Disponível em: http://bdjur.almedina.net/item.php?field=node_id&value=1313381. Acesso em: 11 jan. 2021.

BRASIL. **Portaria Nº 931, de 21 de março de 2005.** Institui o Sistema de Avaliação da Educação Básica - SAEB, que será composto por dois processos de avaliação: a Avaliação Nacional da Educação Básica - ANEB, e a Avaliação Nacional do Rendimento Escolar - ANRESC. Disponível em: <https://abmes.org.br/legislacoes/detalhe/489/portaria-mec-n-931>. Acesso em: 11 jan. 2021.

BRASIL, Presidência da República, Casa Civil. **Lei 9.394 – Estabelece as Diretrizes e Bases da Educação Nacional**, DF, 1996 Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l9394.htm. Acesso em 20 jul. 2021.

BRASIL. **Projeto de Lei 4.731 de 2012.** Disponível em: <https://www.camara.leg.br/noticias/421580-projeto-que-limita-numero-de-alunos-em-sala-de-aula-e-tema-de-debate-hoje/> Acesso em: 20 jul. 2021.

BRASIL. **Projeto de Lei nº 666 de 2015.** Disponível em: <http://alerjln1.alerj.rj.gov.br/scpro1519.nsf/012cfef1f272c0ec832566ec0018d831/4fc526c02cf4951683257e97006fbf62?OpenDocument> Acesso em: 20 jul. 2021.

BRASIL. **Projeto de Lei 280 de 2018.** Disponível em: <http://documentacao.saopaulo.sp.leg.br/iah/fulltext/projeto/PL0280-2018.pdf> Acesso em: 20 jul. 2021.

BRASIL. **Projeto de Lei 4.628 de 2018.** Disponível em: https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra;jsessionid=node01dthctcsonx3s1upp749kpsp8b6320142.node0?codteor=1440304&filename=PL+4628/2016 Acesso em: 20 jul. 2021.

BRIDGE, Patrick D.; SAWILOWSKY, Shlome S. Increasing Physicians' Awareness of the Impact of Statistics on Research Outcomes: Comparative Power of the t-test and Wilcoxon Rank-Sum Test in Small Samples Applied Research. **Journal of clinical epidemiology**, EUA, vol.52, mar. 1999.

BUDELMEYER, H.; SKOUFIAS, E. An Evaluation of The Performance of Regression Discontinuity Design on Progesa, **Policy Research Working Paper; World Bank** n.3386, set. 2004.

BUENO, Belmira O.; OLIVEIRA, Adolfo S. Reflexões sobre Dispositivos Escolares e Trabalho Docente a Partir dos Resultados da Prova Brasil: Progressão Continuada e Composição das Turmas. **Sociologia da Educação**, Brasília, 2014.

BUENO, José L. P.; COSTA, Regina A. C.; UTZIG, Gilmar. A Tecnologia da Informação Como Ferramenta de Acesso à Educação à Distância; **Revista Científica FAEST**, Mato Grosso, v.7, n. 2, p. 80 – 89, out. 2015.

BUENO, Maria C. D.; D'ANTONA, Álvaro O. A Geografia do Censo no Brasil: Potencialidades e Limitações na Execução de Análises Espaciais. **Geographia**, Rio de Janeiro, v. 19, n. 39, jan./abr. 2017.

CALDEIRA, Anita T. N. Enturmação e Desenvolvimento na Educação Infantil: um Estudo Exploratório em Duas Escolas da Rede Pública Municipal do Rio de Janeiro. **Pantheon UFRJ**, Rio de Janeiro, mar. 2018.

CAMBA, Mariangela.; TAVARES, Elisabeth S. As Avaliações de Larga Escala no Brasil: Desafios e Intencionalidades. **SEMINÁRIO INTERNACIONAL DE POLÍTICAS PÚBLICAS EM EDUCAÇÃO**, Santa Catarina, p.96, mai. 2017.

CANÁRIO, Rui. A Escola e As “Dificuldades de Aprendizagem” **Psicologia da Educação**, São Paulo, n.21, dez. 2005.

CAPRARA, Bernardo M. A influência do Capital Cultural no Desempenho Estudantil: Reflexões a partir do SAEB 2003. **LUME**, Porto Alegre, 2013.

CARIA, Neide P.; OLIVEIRA, Sandra M. S. S. Avaliação em Larga Escala e a Gestão da Qualidade da Educação. **Revista de Ciências Humanas**, Minas Gerais, v.16 n.26, p.22-40, jul. 2015.

CASTRO, Maria H. G. A Árdua Tarefa de Estabelecer Padrões de Desempenho. **Cadernos Cenpec**, São Paulo, v.2, n.3, 2007.

CHINGOS, Matthew. The Impact of a Universal Class-Size Reduction Policy: Evidence from Florida's Statewide Mandate. **Economics of Education Review**, EUA, v.31, n.5, p.543-562, out. 2012.

CLARK, M. H.; SHADISH, William R.; STEINER, Peter M. Can Nonrandomized Experiments Yield Accurate Answers? A Randomized Experiment Comparing Random and Nonrandom Assignments. **Journal of the American Statistical Association**. EUA, v.103, n.384, p.1334-1344, dez.2008.

CORREA, Vanessa C.; TROJAN, Rose M. Condições de Trabalho dos Professores e Desempenho Estudantil: uma análise crítica sobre as perspectivas da OCDE. **Revista de Políticas Públicas**, Brasil, v.19, n.2, jun. 2016.

CAMARGO, Juliana. **O efeito do tamanho da turma sobre o desempenho escolar: uma avaliação do impacto da “enturmação” no ensino fundamental do Rio Grande do Sul** / Juliana Camargo,- Porto Alegre: UFRGS/FCE/DERI, 2014. 31 p.: Il. -- (Texto para Discussão / Universidade Federal do Rio Grande do Sul, Faculdade de Ciências Econômicas; /2014)

COST, Scott.; SALZBERG, Steven. A Weighted Nearest Neighbor Algorithm for Learning with Symbolic Features, **Machine Learning**, Boston, v.10, p.57-78, jan. 1993.

CRATO, Nuno. Melhorar o Ensino de Matemática com Ferramentas do Século XXI. **Observatório de Educação Ensino Médio e Gestão**, Portugal, p.42, ago. 2010.

CURY, Carlos R. J. Educação Escolar e Educação no Lar: Espaços de Uma Polêmica. **Educação & Sociedade**, Campinas, v.27, n.96 p.667-688, out. 2006.

CURY, Carlos R. J. O Direito à Educação: Um Campo de Atuação do Gestor Educacional na Escola. **Escola de Gestores**, Brasília, p.24, 2005.

CURY, Carlos R. J. Por uma nova Educação de Jovens e Adultos. **Programa 2**, Campinas, p.4, 2006.

CURY, Carlos R. J. Qualidade em Educação. **Nuances – Estudos sobre Educação**, São Paulo, v. 17, n.18, p.15-31, jan./dez. 2010

DALLAZEM, Aline.; COELHO, Vera. O Desempenho Escolar na Voz dos Atores de Escolas Públicas Catarinenses, **Revista Teias**, Rio de Janeiro, v.20, n.56, jan./mar. 2019.

DINIZ, Ednar M. V. O Censo Escolar. **Revista Brasileira de Estudos Pedagógicos**, Brasília, v. 80, n.194, jan.1999.

DOURADO, Ednólia C. **Qualidade da educação em comunidades de aprendizagem: outro olhar sobre a prova Brasil**. 2018. 73 f. Dissertação (Mestrado Profissional em Educação, Linguagens, e Inovações Pedagógicas) – Faculdade de Educação, Universidade Federal da Bahia, Salvador, 2018.

ECE, Sahin B.; NESLIHAN, Dostoglu. Evaluation of Kindergarten Group Rooms in the Context of Size: Children and Teacher's Perspective in Turkey. **European Journal of Contemporary Education**, Turquia, v.10, n.4, p.253-264, 2014.

EHRENBERG, Ronald G., BREWER, Dominic J., GAMORAN, Adam., WILIAMS, Douglas. Class Size and Student Achievement. **Sage Journals**, Canadá, v.2, n.1, mai. 2001.

ENGLEHART, Joshua M. The Centrality of Context in Learning from Further Class Size. **Education Psychology Review**, v.19, n.4, p.455-467, dez. 2007.

ERDAL, Toprakci. The Reality of Primary Schools and Basic Education in Turkey. **International Journal of Education Research**, Turquia, n.1, v.1, ago.2010.

ESTEBAN, Maria T., Provinha Brasil: Desempenho Escolar e Discursos Normativos Sobre a Infância. **Revista Ciências da Educação**, Lisboa, n.9, p.47-56, 2019

FELICIO, Fabiana; FERNANDES, Reynaldo., O Efeito da Qualidade da Escola sobre o Desempenho Escolar: Uma Avaliação do Ensino Fundamental no Estado de São Paulo. **Econo Papers**, São Paulo, 2005.

FERREIRA, Livia A. A Avaliação no Plano Nacional de Educação (2014-2024). **Revista Meta Avaliação**, Rio de Janeiro, v.8, n.24, p. 410-439, dez.2016.

FERREIRA, Luiz Antonio Miguel. **O Estatuto da criança e do adolescente e o professor: reflexos na sua formação e atuação**. 2004. 223 f. Dissertação (Mestrado em Educação) – Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, São Paulo 2004.

FIGUEIRA, Cleonis V. Modelos de Regressão Logística. **Lume Repositório Digital**, Porto Alegre, 2006.

FINN, Jeremy D.; PANNOZZO, Gina M. The "Why's" of Class Size: Student Behavior in Small Classes. Estados Unidos, **Review of Education Research**, EUA, v.73, n.3 p. 321-368, set. 2003.

FLESSA, Joseph J. Principals as Middle Managers: School Leadership During the Implementation of Primary Class Size Reduction Policy in Ontario, **Leadership and Policy in Schools**, Canadá, v.11, n.3, ago.2012.

FONTANIVE, Nilma S.; KLEIN, Ruben. Avaliação em Larga Escala: Uma Proposta Inovadora. **Em aberto**, Brasília, v.15, n. 66, jan.1995.

FRANÇA, M. T. A.; GONÇALVES, F. Justiça Social no Ensino Fundamental Brasileiro: Transmissão Intergeracional de Desigualdade e Qualidade Educacional **Revista de Estudos Universitários**, São Paulo, v.34, n.1, jul. 2008.

FRANCO, Ana Maria de Paiva. **Os determinantes da qualidade da educação no Brasil. 2009**. Tese (Doutorado em Teoria Econômica) - Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2009.

FRANCO, Luiz Gustavo (org.). **Ensinando Biologia por Investigação: propostas para inovar a ciência na escola**. São Paulo: Na Raiz, 2021.

FREDRIKSON, P., OCKERT B., OOSTERBEEK H. Long-Term Effects of Class Size. **IZA Discussion Paper**, Suécia, n.5879, jul.2011.

GADDOTTI, Moacir. **Projeto Político-Pedagógico da Escola na perspectiva de uma educação para a cidadania**. Brasília, Cortez,1994.

GALLANT, A. R. **Nonlinear Statistical Models**.1.ed. Nova York: John Wiley & Sons e Inc., 1987.

GAMORAN, Adam; MILESI, Carolina., Effects of Class Size And Instruction on Kindergarten Achievement. **Sage Journals**, Wisconsin, v.28, n.4, dez 2006.

GARCIA, Paulo S.; PREARO, Leandro C.; ROMEIRO, MARIA C.; SECCO, Anderson.; BASSI, Marcos S. School Performance: An IDEB Analysis of the Seven Municipalities in the ABC Region. **REVEDUC**, São Paulo, v.10, n.2, 2016.

GATTI, Bernadete A. Políticas de Avaliação em Larga Escala e a Questão da Inovação Educacional. Série de Estudos – **Periódicos do programa de Pós-Graduação em Educação da UCDB**, Mato Grosso do Sul, n.33, jan/jul 2012.

GILMAN, David A. Prime Time in the First Grade at the North Gibson School Corporation: The First Four Years, **Resources in Education**, EUA, v.25, n.1, 1988.

GILRAINE, Michael; MACARTNEY, Hugh; MCMILLAN, Robert. Education Reform in General Equilibrium: Evidence from California's Class Size Reduction. **IDEAS REPEC**, Massachusetts, 2018.

GIMENES, Nelson; DA SILVA, Vandrê G.; PRINCÍPE, Lisandra M.; LOUZANO, Paula; MORICONI, Gabriela M. Além da Prova Brasil: Investimento em Sistemas Próprios de Avaliação Externa. **Estudos em Avaliação Educacional**, São Paulo, v.24, n.55, ago. 2013.

GLASS, Gene; SMITH, Mary. Meta-Analysis of Research on Class Size and Achievement. **Sage Journals**, EUA, v.1, n.1, jan.1979.

GOMES, Suzana S., PINTO, Thiago B. Egressos de Cursos de Graduação na Modalidade a Distância: O Que os Dados Revelam? **ESUD**, Rio Grande do Norte, nov. 2018.

GOMEZ, Lucas G. F., RAMOS, Lilian M. P. C. A Formação de Professores no Plano Nacional de Educação 2014-2024*. **Jornal de Políticas Educacionais**, Goiânia, v.13, n.18, mai. 2018.

GOODMAN, L. A., Kolmogorov-Smirnov tests for psychological research. **Psychol Bull**, USA, v.51, n.2:1, p.8-160, 1954.

GRACIANO, Mariângela (Coord.). **O Plano de Desenvolvimento da Educação (PDE)**. São Paulo: Ação Educativa, 2007.

GRAMANI, Maria C. Análise dos Determinantes de Eficiência Educacional do Estado do Ceará, **Avaliação de Políticas Públicas em Educação**, Rio de Janeiro, v.25, n.95, abr. 2017.

GRAUE, Elizabeth; HATCH, Kelly; RAO, Kalpana; OEN, Denise. The Wisdom of Class-Size Reduction. **Sage Journals**, Wisconsin, set. 2007.

GRAZZIOTINI, Luciane S. S.; MENDEZZI, Natalia P.; WAIMANNII, Moises. Trabalho e Formação Docente: Perfil dos(as) Educadores(as) no Cenário Pós-LDB (1996 – 2008). **CONGRESSO INTERNACIONAL DE FILOSOFIA E EDUCAÇÃO**, Caxias do Sul, mai. 2010.

GUARDA, Antônio., **Avaliação Metodológica para Determinar a Dinamicidade de Dados do Censo Demográfico**. 2010. 217f. Dissertação (Mestrado em Engenharia) - Programa de Pós-Graduação em Engenharia Civil, Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis, 2010.

GUHA, Sudipto.; MISHA, Nina.; Clustering Data Streams, **IEE Xplore- Proceedings 41st Annual Symposium on Foundations of Computer Science**, EUA, p.359-366, nov,2016.

GURGEL-GIANNETTI, Juliana., SIQUEIRA, Cláudia M. Mau Desempenho Escolar: Uma Visão Atual. **Revista Associação Médica Brasileira**, Minas Gerais, v.57, n.1, fev. 2014.

IBGE. **Conhecendo o Brasil**. Disponível em: <https://censo2021.ibge.gov.br/sobre/vonhecendo-o-brasil.html>. Acesso em: 11 jan. 2021.

IPEA. **ABEP – Associação Brasileira de Empresas de Pesquisa- Indicadores Chefe de Família – 2019 –** Disponível em: <https://www.ipea.gov.br/retrato/creditos.html>. Acesso em: 22/04/2021

HAIR JR, Joseph F., PAGE, Michael., BRUNSVELD, Niek. **Essentials of Business Research Methods**, Nova Iorque, 2020.

HANUSHEK, Eric. A. The Economic Value of Improved Schools. California, **NIET - National Institute for Excellence in Teaching**, EUA, p.59-72, abr. 2019.

HANUSHEK, Eric. A. The Evidence on Class Size. **Earning and learning: How schools matter**, EUA, p. 131-168, 1998.

HILDEBRANDT, Susan A.; EOM, Minhee. Teacher Professionalization: Motivational factors and the influence of age, **Teaching and Teacher Education**, EUA, v.27, n.2, p.416-423, fev. 2011.

HOXBY, Caroline M. The Effects of Class Size and Composition on Student Achievement: New Evidence from Natural Population Variation. **The Quarterly Journal of Economics**, Inglaterra, v.1, n.1, 1998.

HUANG, Anna., Similarity Measures for Text Document Clustering. **CiteSeerX**, Nova Zelândia, 2008.

HUNTER, Roger. Class Size: A Critical Implosion, *Critical Studies in Education*, **Melbourne Studies in Education**, Austrália, v.41, n.1, jan, 2000.

HYPOLITO, Álvaro M., JORGE, Tiago., OCDE, PISA e Avaliação em Larga Escala no Brasil: Algumas Implicações. **Sisyphus – Journal of Education**, Lisboa, v.8, n.1, p.10-27, fev. 2020.

JOHNSTON, Melissa P. Secondary Data Analysis: A Method of Which the Time Has Come. **Qualitative and Quantitative Methods in Libraries**, EUA, v.3, n.3, p.619-626, jan.2013

JONES, Colin A. The Right Size for Learning: The State of Class Sizes in Massachusetts. **Mass Budget**, EUA, p.11, 2015.

JORGE, Thiago A. S. Trabalho Docente e Desempenho Estudantil. **Trabalho & Educação**, Brasil, v.23, n.3, p.197-198, dez. 2014

KNIJNIK, Luiza F.; GIACOMONI, C.; STEIN, Lilian M. Teste de Desempenho Escolar: Um Estudo de Levantamento. **PSICO-USF**, Bragança Paulista, v.18, n.3, dez. 2013.

KOMATSU, Bruno K. Por que o Brasil vai Mal no PISA? Uma Análise dos Determinantes do Desempenho no Exame. **Policy Paper**, São Paulo, n.31, jun. 2018.

KOMATSU, Bruno; MENEZES-FILHO, Naércio A; OLIVIERA, Vitor. A Relação Entre a Qualidade da Gestão Municipal e o Desempenho Educacional no Brasil. **Observatório de Educação Ensino Médio e Gestão**, São Paulo, ago.2018.

LIBERTI, Leo; LAVOR, Carlile; MACULAN, Nelson. MUCHERINO, Antônio., Euclidean Distance Geometry and Applications, **ArXiv**, EUA, mai. 2014.

LIMA, Santos; SANTOS, Daniel D. Pai Importa? O Efeito da Presença Paterna Sobre o Desempenho Escolar. **EconoPapers ANPEC**, 2016.

LINDOSO, R. C. B.; SANTOS, A. L. F. Política Educacional e a Avaliação em Larga Escala como Elemento de Regulação da Educação. **Jornal de Políticas Educacionais**, Paraná, v.13, 2019.

LOPES, Raul H. C. The two-dimensional Kolmogorov-Smirnov test. **Proceedings of Science**, Holanda, abr. 2007.

MACHADO, Alessandra P. G. Prova Brasil: Uma Leitura da Avaliação. **ENCONTRO INTERNACIONAL DE FORMAÇÃO DE PROFESSORES FORUM PERMANENTE DE INOVAÇÃO EDUCACIONAL**, Sergipe, v.10, n.1, 2017.

MACHADO, Evelcy M. Pedagogia Social no Brasil: Políticas, Teorias e Práticas em Construção. **IX CONGRESSO NACIONAL DE EDUCAÇÃO – EDUCERE**, Paraná, p.11379-11392, out. 2014.

MACHADO, Franciele L. Censo Escolar e Sistec: As Mais Importantes Bases de Coleta de Dados para EPT. **Revista de Gestão e Avaliação Educacional**, Rio Grande do Sul, v.8, n.17, p.1-8, 2019.

MADHULATHA, Soni., An Overview on Clustering Methods, **Journal of Engineering**, EUA, v.2, n.2, p.719-725, mai.2012.

MAIA, Maria Z. B. Marco Regulatório da Educação à Distância no Brasil: Governo de Fernando Henrique Cardoso (1995 - 2002). **XII COLÓQUIO INTERNACIONAL SOBRE GESTÃO UNIVERSITÁRIA NAS AMÉRICAS** Santa Catarina, nov. 2014.

MAIER, Peter; MOLNAR Alex; PERCY Stephen; SMITH, Phillip; ZAHORIK, John. First Year Results of the Student Achievement Guarantee in Education Program. **SAGE Evaluation Team**, Wisconsin, dez. 1997.

MARASSI, Camila V. O Impacto da Composição de Turmas sobre a Desigualdade Social: (re) Produção ou Não? **Sala de Leitura**, Minas Gerais, 2010.

MARTINS, Claudia M. M; MARTINS, Josefa C. M. Os Programas Federais da Educação Como Ferramenta de Avaliação: Um Olhar sobre O Censo Escolar e a Prova Brasil. **VII FIPED**, Paraná, v.1, ed. 4, 2015.

MARTINS, Fernando J; PEREIRA, Noemi F. F. A Política de Educação Profissional no Plano Nacional de Educação 2014-2024. **Revista Eletrônica de Humanidades do Curso Ciências Sociais**, Macapá, v.10, n.1, p.35-55, jan./jun. 2017.

MARX, Alexander; BACKES, Christina; ECKART, Meese., LENHOF, Hans-Peter; KELLER, Andreas; EDISON-WMW: Exact Dynamic Programming Solution of the Wilcoxon–Mann–Whitney Test, **Genomics Proteomics Bioinformatics**, Alemanha, v.14, n.1, fev. 2016.

MASCAL, Blair; LEUNG, Joannie. District Resource Capacity and The Effectc of Educational Policy: The Case of Primary Class Size Reduction in Ontario, Leadership And Policy in Schools. **Leadership and Policy in Schools**, Canadá, v.11, n.3, ago. 2012.

MATAVELLI, Ieda R. Efeitos de Tamanho de Sala no Desempenho dos Alunos: Evidências para São Paulo. **Revista Brasileira de Economia**, São Paulo, v.74, n.3, jul./set. 2018.

MENDES, Lucas; LOZANO, Danielle; BAZON, Fernanda V. M. A Escolarização De Alunos Com Necessidades Educacionais Especiais No Campo Na Região De Campinas: Breve Análise Do Censo Escolar De 2011. **Educação e Pesquisa**, São Carlos, v.45, 2011.

MENEGÃO, Rita. C. S. G., Os impactos da avaliação em larga escala nos currículos escolares. **Dossiê: Políticas de Avaliação e Currículo**, Paraná, v.11, n.3, mar.2016.

MENEZES-FILHO, Naércio Aquino. **Os determinantes do desempenho escolar do Brasil**. [S.l: s.n.], 2007.

MENEZES FILHO, Naercio Aquino. Os determinantes do desempenho escolar do Brasil. In: **O Brasil e a ciência econômica em debate** [S.l: s.n.], v. 1., 2012.

MENEZES, Marcos Vinícius Marcondes de. **Ensino de Ciências e os sistemas de avaliação em larga escala na educação básica: processos formativos e aprendizagens profissionais da docência**. 2014. 174 f. Dissertação (Mestrado em Educação) - Faculdade de Ciências, Universidade Estadual Paulista Julio de Mesquita Filho, São Paulo 2014.

MENEZES-FILHO, Naércio A. Como Melhorar a Educação no Brasil? **Brasil Economia e Governo**, São Paulo, set.2015.

MINISTÉRIO DA EDUCAÇÃO. **Aplicação das provas do SAEB começa nesta segunda-feira para 7 milhões de estudantes.** Ministério da Educação. Brasília, 20 de novembro de 2019. Disponível em:

[http://portal.mec.gov.br/component/tags/tag/saeb#:~:text=Popula%C3%A7%C3%A3o%20Dalvo%20%E2%80%93%20O%20Saeb%202019,m%C3%A9dio%20\(tradicional%20e%20integrado\)](http://portal.mec.gov.br/component/tags/tag/saeb#:~:text=Popula%C3%A7%C3%A3o%20Dalvo%20%E2%80%93%20O%20Saeb%202019,m%C3%A9dio%20(tradicional%20e%20integrado).). Acesso em: 30 jul.2021

MINISTÉRIO DA EDUCAÇÃO. **INEP realizará audiência pública antes da aplicação da avaliação. Ministério da Educação.** Brasília, 20 de novembro de 2019. Disponível em:

<http://portal.mec.gov.br/component/tags/tag/saeb?start=20#:~:text=O%20Saeb%202019%20vai%20avaliar,outubro%20e%201%C2%BA%20de%20novembro.> Acesso em: 30 fev.2021

MINISTÉRIO DA EDUCAÇÃO. **Prova Brasil.** Disponível em: <http://portal.mec.gov.br/provabrasil>. Acesso em 11 jan.2021.

MONTEIRO, Joana. Gasto Público em Educação e Desempenho Escolar. **Revista Brasileira de Economia**, Rio de Janeiro, v.69, n.4, Rio de Janeiro, out./dez. 2015.

MORAES, Mariana., BECKER, Kalinca. O Efeito do Trabalho do Professor Sobre o Desempenho dos Alunos na Prova Brasil de 2011. **Gestão Pública Práticas e Desafios**, v.6, n.1, abr. 2015.

MOREIRA, Gisele A. M.; REZENDE, Maria G. N. A.; RANGEL, Tauã L. V. O Direito à Educação em Cenários de Crise. **Derecho y Cambio Social**, Rio de Janeiro, n.57, jul./set. 2019.

MUIJIS, Daniel. Doing Quantitative Research in Education with SPSS. **SAGE Publishing**, Inglaterra, 2011.

NASCIMENTO, Paulo A. M. M. Desempenho Escolar e Gastos Municipais por Aluno em Educação: Relação Observada em Municípios Baianos para o Ano 2000. **Avaliação Políticas Públicas em Educação**, Rio de Janeiro, v.15, n.56, set. 2007.

OANCEA, Alis; PUNCH, Keith F. Introduction to Research Methods in Education. **SAGE Publishing**, Inglaterra, 2014.

OECD. **PISA 2018 Results: Where All Students Can Succeed.** Vol. II. Paris: PISA, OECD Publishing, 2019. Disponível em: <https://doi.org/10.1787/b5fd1b8f-en>. Acesso em 09 mai. 2021

OECD. **What's PISA?** Disponível em: <https://www.oecd.org/pisa>. Acesso em: 15 dez. 2020.

OLIVEIRA, Ana Paula de Matos. **A Prova Brasil como política de regulação da rede pública do Distrito Federal**. 2011. 277 f. Dissertação (Mestrado em Educação) — Universidade de Brasília, Brasília, 2011.

OLIVEIRA, Daniela Vieira de. **Relação entre gastos educacionais e desempenho escolar nos municípios goianos**. 2016. 57 f. Dissertação (Mestrado em Administração) - Universidade Federal de Goiás, Goiânia, 2016.

OLIVEIRA, Jaqueline Maria de. **Custo-efetividade de políticas de redução do tamanho da classe e ampliação da jornada escolar: uma aplicação de estimadores de matching**. 2008. Dissertação (Mestrado em Teoria Econômica) - Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2008.

OLIVEIRA, M. P.; ROMANO, A. S. A Gestão Participativa e o Papel Da Liderança Do Diretor na Educação Profissional Técnica. **Revista Insiet**, São Paulo, v.2, n.2, 2015.

OLIVEIRA, Maria E. N. Indicadores de Desempenho dos Sistemas de Ensino e Determinantes da Qualidade de Ensino: Análise da Gestão Educacional de Municípios no Interior Paulista. **Revista Ambiente Educação**, São Paulo, v.1, p.66-75, 2015.

ONG, Mohd H. A.; PUTEH, Fadilah. Quantitative Data Analysis: Choosing Between SPSS, PLS and AMOS in Social Science Research. **International Interdisciplinary Journal of Scientific Research**, Malásia, v.3, n.1, p.13, jul. 2017.

ONU Organização das Nações Unidas. Principles and Recommendations for Population and Housing Censuses Revision 2. **United Nations Publication**, Nova Iorque, n.67, 2008.

OYEKA, Ikewelugo C. A., Ebuh, Godday U., Modified Wilcoxon Signed-Rank Test, **Open Journal of Statistics**, Nigéria, v.2, n.2, abr.2021.

PANASSOL, Paulo E. Gastos Educacionais e Desempenho Escolar em Municípios do Rio grande do Sul. **Informe Econômico**, Porto Alegre, v.14, n.2, jul./dez. 2018.

PASS, Ikaro T. **Evaluation of the quality determinants of basic education in 2013: a comparison of the South-Center and West mesoregions from Paraná.** 2016. 95 f. Dissertação (Mestrado em Desenvolvimento Econômico Regional e Economia do Agronegócio) - Universidade Estadual do Oeste do Paraná, Toledo, 2016.

PIRES, Luciana. Qualidade de vida de Adolescentes Modelos Profissionais, **Revista de Psicologia Teoria e Pesquisa**, Brasília, v. 28, n. 1, mar 2012.

POZZOBON, M.; FALCKE, D.; MARIN, A. Intervenção com Famílias de Alunos com Baixo Desempenho Escolar. **Ciências Psicológicas**, Uruguai, vol. 12, n.1, pp. 87-96, set. 2018.

PORTAL INEP. **Censo Escolar.** Disponível em: <http://portal.inep.gov.br/web/guest/censo-escolar>. Acesso em: 11 jan. 2021.

PORTAL INEP. **Microdados do INEP.** Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados>. Acesso em: 11 jan. 2021.

PORTAL INEP. **Resultados Finais das Escolas no SAEB 2017.** Disponível em: http://portal.inep.gov.br/artigo//asset_publisher/B4AQV9zFY7Bv/content/resultados-finais-das-escolas-no-saeb-2017-ja-estao-disponiveis-no-portal-do-inep/21206. Acesso em: 22 abr. 2021.

PORTAL INEP. **SAEB - Microdados do Sistema de Avaliação da Educação Básica** – Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/saeb> – Saeb. Acesso em: 11 jan. 2021.

PREARO, Leandro C., **O uso de técnicas estatísticas multivariadas em dissertações e teses sobre o comportamento do consumidor: um estudo exploratório** 2008. 116 f Dissertação (Mestrado em Economia) Universidade de São Paulo. – São Paulo, 2008.

PREARO, Leandro C., **Os Serviços Públicos e o Bem-estar Subjetivo da População: Uma Modelagem Multigrupos Baseada em Mínimos Quadrados Parciais.** 2013 274 f Tese (Doutorado em Ciências) Universidade de São Paulo São Paulo, 2013.

RAMAL, Andrea C. A LDB e o Processo de Renovação Pedagógica Jesuíta. **Revista de Educação CEAP**, São Paulo, n.17, p. 65 -78, jun.1997.

ROCHA, Paula D. P.; FERREIRA, Maria. Avaliação em Larga Escala e Indicativos de Qualidade na Educação: Como se Processa essa Relação? **Revista Educação em Questão**, Rio Grande do Norte, v.55, n.43, jan./mar. 2017

ROSENBAUM, Paul R.; RUBIN, Donald B. The Central Role of the Propensity Score in Observational Studies for Causal Effects. **Biometrics**, Chicago, v.70, n.1, p. 41-55, abr 1983.

ROSNER, Bernard.; GLYNN, Robert J.; LEE, Mei-Ling T. The Wilcoxon Signed Rank Test for Paired Comparisons of Clustered Data. **Biometrics**, EUA, v.62, n.1, p.185-192, mar.2005.

RUGGLES, Mary L. Is There a Relationship Between Class Size and Student Achievement? **Semantic Scholar**, Michigan, 2003.

RUIZ, Nathália F.; SOUZA, Vania F. M.; OLIVEIRA, Lucilia V.; YAEGASHI, Solange F. R. Plano Nacional de Educação (2014-2024): Reflexões Sobre a Alfabetização e o Letramento. **Revista Online de Gestão e Política Educacional**, Araraquara, v.24, n.1, jan./mar. 2020.

SANTOS, Almir P., CARBONERA, Vânia., Gestão Democrática da Educação e Avaliação em Larga Escala: Implicações para o Contexto Escolar. **Poliéses**, Santa Catarina, v.3, n.6, 2010.

SASSAKI, Alex H.; DI PIETRA, Giovanni A. C.; MENEZES-FILHO, Naércio A. Por que o Brasil Vai Mal no PISA? Uma análise os determinantes do desempenho no exame. **Inspere –Policy Papers**, São Paulo, n.31, p.27, jun. 2018.

SAVELI, Esméria L.; TENREIRO, Maria O. V. A Educação Enquanto Direito Social: Aspectos Históricos e Constitucionais. **Teoria Prática e Educação**, Paraná, v.15, n.2, mar. 2012.

SCHNEIDER, Marilda P. Políticas de Avaliação em Larga Escala e a Construção de um Currículo Nacional para a Educação Básica. **Eccos Revista Científica**, Brasil, n.30, p. 17-33, jan./abr. 2013.

SCHNEIDER, G.; GOUVEIA, A. B. Qualidade da Escola: uma proposta de índice para as condições materiais de escolas a partir de dados contextuais do Saeb. **Revista Brasileira de Pós-Graduação**, Brasília, v.8, n.1, dez. 2011.

SENA, P. O Financiamento da Educação de Qualidade. **Revista Educação e Política em Debate**, Minas Gerais, v.3, n.2, jul 2015.

SILVA, Antônio.; LOPES, Pedro. CASTRO, Amanda. Avaliação da Educação no Brasil: A Centralidade dos Testes em Larga Escala. **Holos**, Rio Grande do Norte, v.7, p.388-401, nov. 2016

SILVA, Vandré; GIMENES, Nelson.; LOUZANO, Paula. Por que as Secretarias de Educação Investem em Sistemas Próprios de Avaliação em Larga Escala? **Anpae**, Pernambuco, v.26, 2013.

SINGH, Nongmaithem., ROY, Aniruddha., TRIPATHI, Anil., Non-Parametric Test: Hands on SPSS, **ICAR RESEARCH**, set, 2013.

SOARES, José Efeito da Escola no Desempenho Cognitivo de Seus Alunos. **Revista Iberoamericana sobre Calidad, Eficacia y Cambio em Educacion**, Petrópolis, v.2, n.2, jul. 2005.

SOARES, Maria S. Avaliação da Educação em Larga Escala: (a) Prova Brasil? **Revista Tópicos Educacionais**, Recife, v.19, n.2, jul/dez 2013.

SOARES, Roberta R. B. S.; DE FARIAS, Otávia M. Reflexões Sobre o Discurso Docente Acerca da Indisciplina e do Mau Desempenho Escolar. **Entre Palavras**, Ceará, v.8, n.1, mar, 2018.

SOOTER, Tombowua. Early Childhood Education in Nigéria: Issues and Problems. Nigéria, **Journal of Education and Social Research**, Itália, v.3, n.5, ago. 2013.

SOUSA, Daniel G. de. **Determinantes da qualidade do ensino: um estudo na rede pública do estado de São Paulo. 2018.** Dissertação (Mestrado em Administração) - Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2018.

SOUZA, João. **Formação de professores para a Educação Básica: Dez anos da LDB.** 1.ed. Belo Horizonte: Autêntica, 2007.

SULE, Samuel S. Effects of Assignment and Class Size on Secondary School Student's Achievement in Mathematics, **ATBU Journal of Science, Technology and Education**, Nigéria, v.4, n.2, jun.2016.

SZATKOSKI, L. **A prova Brasil no cotidiano escolar**. 2014. 85 f. Dissertação (Mestrado em Educação) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2014.

THOEMMES, Felix. Propensity Score Matching in SPSS. **ArXiv**, Alemanha, p. 30, jan.2012.

TODOS PELA EDUCAÇÃO. **Cenários da Educação**. Disponível em: <https://www.todospelaeducacao.org.br/pag/cenarios-da-educacao>. Acesso em: 16 jul. 2020.

TRAVINCAS, Amanda C. T. Vinculação de Receita Mínima para a Satisfação do Direito à Educação no Brasil (cf, art. 212) Dirigismo Constitucional e Dever de Progressividade. **Justiça e Sociedade**, São Luís, v.1, n.1, 2016

TRAVITZKI, R.; CÁSSIO, F. L. Tamanho das Classes na Rede Estadual Paulista: A Gestão de Rede Pública à Margem das Desigualdades Educacionais. **ETD – Educação Temática Digital**, São Paulo, v.19, p.159-183, nov. 2017.

TROJAN, Rose M.; CORRÊA, Vanessa C. Condições de Trabalho dos Professores e Desempenho Estudantil: Uma análise crítica sobre as perspectivas da OCDE. **Revista de Políticas Públicas**, São Luís, v.19, n.2, 2015.

VAN DEN BERG, Ruben G. SPSS Kolmogorov-Smirnov Test for Normality - The Ultimate Guide, **SPSS Tutorials**, Holanda, 2019.

VANDENBERG, Kristy C. Class Size and Academic Achievement. **Electronic Theses and Dissertations**, Eua, 2012.

VASCONCELOS, Maria L. M. C. Progressão Continuada: Por Que a Revisão dos Ciclos? **Revista Lusófona de Educação**, Portugal, n.11, p. 77-84, 2008.

VIANNA, Heraldo M. Programas de Avaliação em larga escala: Algumas Considerações. **Estudos em Avaliação Educacional**, São Paulo, n.23, jun.2001.

VILLAR, Antônio. ZOIDO, Pablo. Challenges to quality and equity in educational performance for Latin America, a PISA 2012 perspective. **Revista Electronica de Investigacion y Evolucion Educativa**, Espanha, v.22, n.1, jun. 2016.

WILLIS, Jerry W. Qualitative Research Methods in Education and Educational Technology. **International Journal of Education**, EUA, v.3, n.2, ago.2008.

WILSON, Valerie. Reducing Class Size: Does the Evidence Support the Scottish Executive's Policy? **Scottish Educational Review**, v.39, n.2, jan 2007.

WOESSMANN, L.; WEST, M. Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in Timss. **European Economic Review**, Alemanha, v.50, n.3, p.675-736, abr.2002.

WORD, Elizabeth. **The State of Tennessee's Student/Teacher Achievement Ratio - STAR Project - Technical Report 1985-1990**. Tennessee, 1990.

YIGIT, Nevzat.; ALPASLAN, Muhammet M.; CINEMRE, Yasin.; BALCIN, Bilal. Examine Middle School Student's Constructivist Environment Perceptions in Turkey: School Location and Class Size. **TUSED**, Turquia, v.14, n.1, mar 2017.

YOUNG, Ian T. Proof Without Prejudice: Use of the Kolmogorov-Smirnov Test, **SAGE Journal**, USA, v.25, n.7, jul.1977.

YU, Kai; LIANG, Ji; ZHANG, Xuegong. Kernel Nearest-Neighbor Algorithm. **Neural Processing Letters**, Holanda, v.15, p.147-156, 2002.

YUSUF, Taofeek ; ONIFADE, Comfort.; BELLO, Oladipupo. Impact of Class Size on Learning, Behavioral and General Attitudes of Students in Secondary Schools in Abeokuta. **Journal of Research Initiatives**, Nigéria, v.2, n.1, p.19, 2016.

ZENG, Xianxian; XU, Wenchao; DONG, Ligu. Human Facial Feature Matching based on Motion Smoothness, **Computer Science**, China, v.526, n.1, p.6, 2020.

ZUANETTI, Patrícia A.; SCHNECK, Andréa P. C.; MANFREDI, Alessandra K. S. Consciência Fonológica e Desempenho Escolar, **Revista CEFAC**, São Paulo, v.10, n.2, jul. 2008.