



UNIVERSIDADE FEDERAL
DO ESPÍRITO SANTO

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE EDUCAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO

TALITA EMÍDIO ANDRADE SOARES

**O SAEB EM DEBATE: AVANÇOS E DESAFIOS DA POLÍTICA NACIONAL
DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA**

**Vitória
2022**



Centro de Educação

Programa de Pós-Graduação em Educação

TALITA EMIDIO ANDRADE SOARES

**O SAEB EM DEBATE: AVANÇOS E DESAFIOS DA POLÍTICA NACIONAL
DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA**

Dissertação apresentada ao Programa de Pós-Graduação em Educação da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do título de mestre em Educação, na área de concentração Educação, Formação Humana e Políticas Públicas.

Orientador: Prof. Dr. Wagner dos Santos

Vitória, 2022

Ficha catalográfica disponibilizada pelo Sistema Integrado de Bibliotecas - SIBI/UFES e elaborada pelo autor

S676s Soares, Talita Emidio Andrade, 1996-
O Saeb em debate : avanços e desafios da política nacional de avaliação da educação básica / Talita Emidio Andrade Soares. - 2022.
145 f. : il.

Orientador: Wagner dos Santos.
Dissertação (Mestrado em Educação) - Universidade Federal do Espírito Santo, Centro de Educação.

1. Sistema de Avaliação da Educação Básica. 2. Avaliação Educacional. 3. Política de Avaliação. I. dos Santos, Wagner. II. Universidade Federal do Espírito Santo. Centro de Educação. III. Título.

CDU: 37

TALITA EMIDIO ANDRADE SOARES

**O SAEB EM DEBATE: AVANÇOS E DESAFIOS DA POLÍTICA NACIONAL
DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA**

Dissertação apresentada ao Programa de Pós-Graduação em Educação da Universidade Federal do Espírito Santo, como requisito parcial para avaliação, na linha de pesquisa Educação, Formação Humana e Políticas Públicas.

COMISSÃO EXAMINADORA

Prof. Dr. Wagner dos Santos
Universidade Federal do Espírito Santo
Orientador

Prof. Dra. Sandra Soares Della Fonte
Universidade Federal do Espírito Santo

Prof. Dr. Gabriel Dias de Carvalho Júnior
Instituto Federal de Minas Gerais

Prof. Dr. Daniel Eduardo Ríos Muñoz
Universidad de Santiago de Chile

À minha vovó Maria.

AGRADECIMENTOS

À Deus por estar sempre à frente do meu caminho.

Aos meus pais, Francisco e Sônia, por terem me dado a liberdade e a confiança para escolher onde gostaria de estar e pelo apoio nas decisões mais difíceis.

Aos meus sogros, Denilson e Eliane, pelo carinho comigo e por sempre me incluírem em suas orações.

Aos amigos, que souberam compreender minhas ausências no período de realização desta dissertação. Em especial, a Thaís Coelho Freitas, minha irmã de coração.

Ao meu marido Denilson, por ser compreensivo e um exemplo de profissional da área, motivo pelo qual me orgulho tanto.

Ao professor Dr. Wagner dos Santos, pela orientação e direcionamento desta pesquisa, que sem dúvida, contribui para o meu crescimento acadêmico, profissional e pessoal.

Aos professores que compuseram a banca, Dra. Sandra Soares Della Fonte, Dr. Gabriel Dias de Carvalho Júnior e Dr. Daniel Eduardo Ríoz Muñoz, pelas relevantes contribuições e pela disponibilidade em colaborar com esta pesquisa.

Aos membros do Proreitoria, Daniela, Dilson, Gabriel, Heitor, Jean, Lucas, Matheus, Rodrigo, Sayonara e Wagner Zeferino. Em especial, a Zilka, Ronildo e Pablo pelas conversas e palavras de conforto.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), pela concessão de bolsa durante o período do mestrado.

A todos que contribuíram, direta ou indiretamente, obrigada.

A persistência é o caminho do êxito

Charles Chaplin

RESUMO

Esta pesquisa foi desenvolvida com o objetivo de analisar a trajetória histórica do Saeb ao longo dos mais de trinta anos de implementação, identificando os princípios e as concepções que fundamentam o sistema. Para responder a esse objetivo, assume, como questão norteadora: de que maneira o Saeb se insere historicamente como uma política de avaliação da educação básica brasileira? Utiliza uma metodologia qualitativa e exploratória, a pesquisa bibliográfica, a análise crítico-documental e o paradigma indiciário como abordagens teórico-metodológicas. Assume como fontes: publicações acadêmicas relacionadas com o Saeb, entrevistas jornalísticas com os sujeitos envolvidos no processo e documentos governamentais produzidos com o intuito de regulamentar e direcionar as ações dos governos brasileiro em prol das políticas de avaliação da educação básica brasileira. Como resultado, evidencia o fortalecimento do uso regulativo e punitivo do sistema em detrimento de uma perspectiva avaliativa do próprio papel desempenhado pelo Estado como condutor das políticas educacionais, sobretudo a partir do surgimento da Prova Brasil que, ao implementar a avaliação censitária no cenário nacional, contribuiu para as primeiras iniciativas de responsabilização, embora brandas, no país. Contudo, as recentes atribuições definidas para o Saeb estão associadas a mecanismos de responsabilização forte, com a previsão do seu uso no repasse de recursos financeiros e como possibilidade de acesso ao ensino superior, sinalizando para uma nova mudança de perspectiva do sistema.

Palavras-chave: Sistema de Avaliação da Educação Básica. Avaliação Educacional. Política de Avaliação.

ABSTRACT

This research was developed with the objective of analyzing the historical trajectory of Saeb to the over thirty years of implementation, identifying the principles and concepts that underlie the system. For this, it assumes as guiding question: in which way does the Saeb insert itself historically as a policy of evaluation of Brazilian basic education? Thus, this work uses a qualitative and exploratory methodology, bibliographic research, critical-documentary analysis and the evidence as theoretical-methodological approaches. Academic publications studies related to the Saeb, journalistic interviews with the subjects involved in the process and government documents produced in order to regulate and direct the actions of Brazilian governments in favor of education evaluation policies Brazilian basic were used as data. As a result, it highlights the strengthening of regulatory and punitive use of the system to the detriment of an evaluative perspective of the role played by the State as a driver of educational policies, especially since the emergence of Prova Brasil, which, by implementing the census assessment on the national scene, contributed for the first, albeit mild, accountability initiatives in the country. However, the recent attributions defined for the Saeb are associated with mechanisms of strong accountability, with the forecast of its use in the transfer of financial resources and as a possibility of access to higher education, signaling a new change in system perspective.

Keywords: Basic Education Evaluation System. Educational Assessment. Evaluation Policy.

LISTA DE FIGURAS

FIGURA 1 – Fluxograma da seleção dos artigos de acordo com o método Prisma.....	33
FIGURA 2 – Classificação dos temas.....	34
FIGURA 3 – Curva Característica do item.....	84
FIGURA 4 – Distribuição normal e escala de proficiência da Prova Brasil.....	86

LISTA DE QUADROS

QUADRO 1 – Documentos Normativos.....	24
QUADRO 2 – Caracterização do <i>corpus</i> textual dos artigos seleccionados.....	34
QUADRO 3 – Distribuição dos descritores da Matriz de Referência de Língua Portuguesa do Saeb 1997.....	91
QUADRO 4 – Distribuição dos descritores da Matriz de Referência de Língua Portuguesa do Saeb 2001.....	92
QUADRO 5 – Distribuição dos descritores da Matriz de Referência de Matemática do Saeb 1997.....	94
QUADRO 6 – Distribuição dos descritores da Matriz de Referência de Matemática do Saeb 2001.....	95

LISTA DE SIGLAS

- AGIM – Agência de Garantia de Investimentos Multilaterais
- AID – Associação Internacional de Desenvolvimento
- ANA – Avaliação Nacional de Alfabetização
- ANEB – Avaliação Nacional da Educação Básica
- ANRESC – Avaliação Nacional do Rendimento Escolar
- BID – Banco Interamericano de Desenvolvimento
- BIRD – Banco Internacional de Reconstrução e Desenvolvimento
- BNCC – Base Nacional Comum Curricular
- CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
- CCI – Curva Característica do Item
- CEPAL – Comissão Econômica para América Latina e Caribe
- CESGRANRIO – Centro de Seleção de Candidatos do Grande Rio
- CFI – Cooperação Financeira Internacional
- CIADI – Centro Internacional para Arbitragem de Disputas sobre Investimentos
- CONSED – Conselho Nacional de Secretários de Educação
- DCNs – Diretrizes Curriculares Nacionais
- EAP – *Expected A Posteriori*
- EDURURAL – Projeto Nordeste de Educação Básica I
- ENADE – Exame Nacional de Desempenho dos Estudantes
- ENCCEJA – Exame Nacional para Certificação de Competências de Jovens e Adultos
- ENEM – Exame Nacional do Ensino Médio
- EUA – Estados Unidos da América
- FCC – Fundação Carlos Chagas
- FGV – Fundação Getúlio Vargas

FHC – Fernando Henrique Cardoso

FNDE – Fundo Nacional de Desenvolvimento da Educação

IAFD – Indicador de Adequação da Formação Docente

IDEB – Índice de Desenvolvimento da Educação Básica

IICA – Instituto Interamericano de Cooperação Agrícola

INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

INSPER – Instituto de Ensino e Pesquisa

LDB – Lei de Diretrizes e Bases da Educação Básica

MEC – Ministério da Educação

ML3P – Modelo logístico de três parâmetros

NBE – *National Bureau of Education*

NSE – Nível socioeconômico

OCDE – Organização para a Cooperação e Desenvolvimento Econômico

ONU – Organização das Nações Unidas

OREALC – *Oficina Regional de Educacion para a America Latina y el Caribe*

PAEBES – Programa de Avaliação da Educação Básica do Espírito Santo

PAEBES TRI – Avaliação Interna Trimestral Diagnóstica da Aprendizagem

PISA – *Programme for International Student Assessment*

PCN – Parâmetros Curriculares Nacionais

PDE – Plano de Desenvolvimento da Educação

PNAIC – Pacto Nacional pela Alfabetização na Idade Certa

PNE – Plano Nacional de Educação

PNUD – Programa das Nações Unidas para o Desenvolvimento

PRISMA – *Preferred Reporting Items for Systematic Reviews and Meta-Analyses*

PROJETO NORDESTE – Projeto Nordeste de Educação Básica II

PUC/RJ – Pontifícia Universidade Católica do Rio de Janeiro

SAEB – Sistema de Avaliação da Educação Básica

SAEP – Sistema Nacional de Avaliação do Ensino Público de 1º Grau

SCIELO – *Scientific Electronic Library Online*

SEAMA – Sistema Estadual de Avaliação do Maranhão

SEAPE – Sistema Estadual de Avaliação da Aprendizagem Escolar

SENEB – Secretaria Nacional de Educação Básica

SIMAVE – Sistema Mineiro de Avaliação e Equidade da Educação Pública

SINAEB – Sistema Nacional de Avaliação da Educação Básica

SINAES – Sistema Nacional de avaliação da Educação Superior

SPAECE – Sistema Permanente de Avaliação da Educação Básica do Ceará

SRE – *School Effectiveness Research*

SUDENE – Secretaria Executiva de Superintendência do Desenvolvimento do Nordeste

TCT – Teoria Clássica dos Testes

TIMSS – *Trends in International Mathematics and Science Study*

TRI – Teoria de Resposta ao Item

UEMG – Universidade Estadual de Minas Gerais

UFCE – Universidade Federal do Ceará

UFU – Universidade Federal de Uberlândia

UNDIME – União Nacional dos Dirigentes Municipais de Educação

UNESCO – Organização das Nações Unidas para a Educação, a Ciência e a Cultura

USP – Universidade de São Paulo

WOS – *Web of Science*

SUMÁRIO

INTRODUÇÃO	15
TEORIA E MÉTODO.....	19
ORGANIZAÇÃO DO ESTUDO	26

CAPÍTULO I

1 SISTEMA DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA: REVISÃO SISTEMÁTICA DA LITERATURA	28
1.1 INTRODUÇÃO.....	28
1.2 METODOLOGIA.....	31
1.3 RESULTADOS	32
1.3.1 A origem do Saeb mediante as reformas educacionais da década de 1990	37
1.3.2 O Saeb como política de responsabilização (<i>accountability</i>).....	41
1.3.3 Os fatores associados aos resultados do Saeb	44
1.4 CONSIDERAÇÕES PARCIAIS	47

CAPÍTULO 2

2 EXPERIÊNCIAS INICIAIS DE AVALIAÇÃO EDUCACIONAL QUE SERVIRAM COMO BASE PARA O DESENHO DO SAEB	51
2.1 INTRODUÇÃO.....	51
2.2 A INFLUÊNCIA DO BANCO MUNDIAL.....	52
2.3 O PROJETO NORDESTE DE EDUCAÇÃO BÁSICA I (EDURURAL)	56
2.4 PROJETO NORDESTE DE EDUCAÇÃO BÁSICA II (PROJETO NORDESTE)..	60
2.5 A AVALIAÇÃO DO RENDIMENTO ESCOLAR DE ALUNOS DE ESCOLAS DE 1º GRAU DA REDE PÚBLICA	64
2.6 SISTEMA NACIONAL DE AVALIAÇÃO DO ENSINO PÚBLICO DE 1º GRAU (SAEP)	68
2.7 CONSIDERAÇÕES PARCIAIS	72

CAPÍTULO 3

3	O PROCESSO DE DESENVOLVIMENTO E INSTITUCIONALIZAÇÃO DO SAEB	75
3.1	INTRODUÇÃO.....	75
3.2	O PROCESSO DE DESENVOLVIMENTO E DE INSTITUCIONALIZAÇÃO DO SAEB (1990-1995).....	76
3.3	A ADOÇÃO DA TEORIA DE RESPOSTA AO ITEM (TRI).....	80
3.3.1	Teoria de Resposta ao Item: nova metodologia de construção do teste e análise de resultados	81
3.3.2	A Escala de Proficiência do Saeb	85
3.4	A ADOÇÃO DAS MATRIZES DE REFERÊNCIA	87
3.4.1	Língua Portuguesa.....	90
3.4.2	Matemática.....	93
3.5	CONSIDERAÇÕES PARCIAIS	96

CAPÍTULO 4

4	O SAEB CENSITÁRIO E A EFETIVAÇÃO DE UMA CULTURA POLÍTICA DE AVALIAÇÃO PAUTADA NA RESPONSABILIZAÇÃO DA EDUCAÇÃO BÁSICA BRASILEIRA	99
4.1	INTRODUÇÃO.....	99
4.2	OS INSTRUMENTOS DO SAEB	101
4.2.1	A Avaliação Nacional da Educação Básica (Aneb) e a Avaliação Nacional do Rendimento Escolar (Anresc/Prova Brasil).....	102
4.2.2	A Avaliação Nacional da Alfabetização (ANA)	105
4.2.3	Síntese das avaliações do Saeb.....	107
4.3	A INCORPORAÇÃO DO MODELO DE PRESTAÇÃO DE CONTAS NA EDUCAÇÃO BÁSICA BRASILEIRA	108

4.4	O SAEB E A EFETIVAÇÃO DE UMA CULTURA POLÍTICA DE AVALIAÇÃO PAUTADA NA RESPONSABILIZAÇÃO ESCOLAR.....	113
4.5	CONSIDERAÇÕES PARCIAIS	120
	CONSIDERAÇÕES FINAIS.....	123
	REFERÊNCIAS.....	129

INTRODUÇÃO

No início do século XX, segundo Vianna (1998), os Estados Unidos da América (EUA) e a Inglaterra começaram a desenvolver pesquisas que procuravam associar o processo socioeconômico a valores e conhecimentos transmitidos por intermédio da educação. Por esse motivo, fez-se necessário um “sistema de controle” das atividades que estavam relacionadas com o processo educacional.

Nesse âmbito, destacam-se os estudos pioneiros do psicólogo Edward Lee Thorndike (1874-1949). Seu trabalho, conforme Sousa (1995), enfatizava a importância da mensuração do comportamento humano, o que resultou no desenvolvimento de testes padronizados para medir habilidades e aptidões dos alunos.

Vale a pena destacar que, até a década de 1920, os conceitos de avaliação e medidas educacionais se apresentavam confundidos, pois não havia uma definição formal, adotada pela literatura especializada, que os diferenciasses (VIANNA, 1998). Desse modo, o uso da avaliação era restrito como medida, cuja concepção era baseada na Psicologia Comportamental, desenvolvida a partir da contribuição de renomados pesquisadores, como Burrhus Frederic Skinner (1904-1990) e John Broadus Watson (1878-1958).

Centrada, acima de tudo, no estabelecimento das diferenças individuais entre pessoas, utilizava como técnica predominante, e quase excludente, a aplicação de testes, tanto no âmbito individual quanto no coletivo (bateria de testes). Desse modo, a avaliação tinha pouco a ver com os programas que eram desenvolvidos nas escolas (CASTILLO ARREDONDO; DIAGO, 2009, p. 31).

Alguns anos mais tarde, tem-se a influência do educador Ralph W. Tyler (1902-1994). A ele se atribui a origem do termo avaliação educacional, em 1934, concomitante com o surgimento do movimento “educação por objetivos”. Resumidamente, o princípio era formular metas e verificar se eram atingidas. Para ele, a avaliação,

[...] ofereceria elementos para uma crítica fundamentada da instituição, baseada em dados empíricos, e permitiria, ao mesmo tempo, uma discussão sobre a eficiência da sua atuação. A avaliação, ainda de acordo com o pensamento de Tyler, possibilitaria o aprimoramento dos programas, com a eliminação do inoperante e o desenvolvimento daqueles aspectos que se tivessem revelado positivos (VIANNA, 1998, p. 77).

Considerado o “pai” da avaliação educacional (CASTILLO ARREDONDO; DIAGO, 2009), Tyler, após duas décadas, não só teve grande atuação em atividades associadas aos testes,

medidas e avaliação, como também influenciou toda a educação norte-americana. Em meados dos anos 1960, projetou o *National Assessment of Educational Progress* (Naep), avaliação periódica dos EUA que, de acordo com Vianna (1998), fornecia elementos para as correções necessárias requeridas pelo sistema educacional estadunidense e auxiliava na identificação de pontos de excelência ou críticos a serem por ele enfrentados.

Seu trabalho também repercutiu em outras regiões do mundo, inclusive, no próprio Brasil, pois, consoante com Waiselfisz e Horta Neto (2016), o Naep foi utilizado como referência para o desenho do Sistema Nacional de Avaliação no país. Em comum, havia a aplicação de testes padronizados para os alunos e questionários para professores e gestores, desenvolvidos na tentativa de abranger o maior número possível de fatores contextuais que explicassem as dificuldades dos alunos.

Nesse período, durante as décadas de 1930 e 1940, graças a Tyler, “[...] a educação passou a ser concebida como um processo sistemático, destinado a produzir mudanças na conduta dos alunos por meio da instrução” (CASTILLO ARREDONDO; DIAGO, 2009, p. 31). Para ele, “[...] educar consistiria em gerar e/ou mudar padrões de comportamento, devendo, em consequência, o currículo ser construído com base na especificação de habilidades desejáveis expressas em objetivos a serem alcançados” (VIANNA, 1998, p. 76). Desse modo, para Castillo Arredondo e Diago (2009), a avaliação passou a ser considerada como mecanismo que permite comprovar o grau de consecução dos objetivos propostos.

Posteriormente, nas décadas de 1960 e 1970, ainda segundo os autores, a avaliação era vista na totalidade do âmbito educacional. Tal concepção também foi desenvolvida nos EUA, consequência de um movimento denominado “responsabilidade escolar” (CASTILLO ARREDONDO; DIAGO, 2009), que surgiu devido ao descontentamento do país com a escola pública, pois, em teoria, volumosos recursos eram destinados a ela.

A nova concepção que se tem da avaliação neste tempo faz com que não somente o rendimento dos alunos seja afetado, mas todos aqueles fatores que convergem em um programa educacional, isto é, professor, recursos, conteúdos, atividades, organização, métodos, programas etc. (CASTILLO ARREDONDO; DIAGO, 2009, p. 32).

Nessa época, é importante destacar a contribuição de outros dois autores: o psicólogo Lee Cronbach (1916-2001) e o filósofo Michael Scriven (1928-), que foi influenciado pelo primeiro. Conforme Castillo Arredondo e Diago (2009), Cronbach enfatizava a avaliação do processo e, ao mesmo tempo, a avaliação por meio de objetivos previamente estabelecidos.

Por outro lado, para Scriven, avaliar os resultados reais independia das metas ou critérios preestabelecidos. O que interessava, na verdade, eram as atitudes geradas pelos programas nas pessoas implicadas.

Nos anos 1970, também houve a intensificação de novos enfoques e tendências na área de avaliação educacional. Como discutido, a responsabilidade da escola estava em alta e por isso a avaliação era orientada de maneira interna (avaliação escolar) e externa (avaliação de programa ou método). Esta última era organizada por agentes externos à escola para servir como subsídios ao processo de tomada de decisões.

Nas décadas seguintes, 1980 e 1990, tem-se a proliferação de modelos avaliativos que podem ser classificados conforme os métodos de abordagem e metodologia: quantitativo e qualitativo. De acordo com Castillo Arredondo e Diago (2009), apesar de diferentes, ambos coexistiram em muitos modelos, demonstrando que não existe uma forma única de avaliar. Os autores destacam ainda a incorporação de novos elementos de outras disciplinas, o que contribuiu para o aperfeiçoamento do instrumento e, conseqüentemente, para sua capacidade de diagnóstico.

Desse modo, é plausível dizer que o conceito de avaliação educacional evoluiu com o passar dos anos conforme as necessidades e os interesses da época. Com isso surgiram novas modalidades, como: a avaliação de aprendizagens, de currículos e programas, de sistemas educativos, de profissionais e de políticas educacionais (STADLER, 2017). Em todas, a avaliação tem o papel crucial de averiguar a qualidade da educação por meio do acompanhamento do processo educativo ou da análise de seus resultados.

No Brasil, percebemos uma maior preocupação com o desenvolvimento de testes cognitivos a partir dos anos 1960, concomitante com a criação das fundações de pesquisa em avaliação, como a Fundação Carlos Chagas (FCC), Fundação Getúlio Vargas (FGV) e Centro de Seleção de Candidatos do Grande Rio (Cesgranrio). O interesse por esse tipo de teste contribuiu para o surgimento dos exames vestibulares, o que lhes deu maior visibilidade.

Contudo, a avaliação realizada por agentes externos à escola, dita avaliação externa, ganhou notoriedade no país quase três décadas depois, influenciada pela onda das reformas neoliberais na América Latina. Um dos fatores contribuintes foi a recomendação de organizações multilaterais internacionais, como o Banco Mundial. Dentre as suas exigências,

centrava-se, sobretudo, o uso da avaliação externa como política educacional (GATTI; VIANNA; DAVIS, 1991).

Ao final da década de 1980, propôs-se no Brasil que dez capitais fossem tomadas para um estudo piloto cujo objetivo seria verificar a viabilidade de um projeto de avaliação em nível nacional. A partir daí, com a disseminação em seminários e demais eventos sobre a temática, expandiu-se o quantitativo de pesquisas que propiciaram a criação de um Sistema de Avaliação da Educação Básica (Saeb), em 1990, que se constitui como objeto de análise desta dissertação.

Durante esses mais de trinta anos de existência, o Saeb passou por diversos aprimoramentos, adaptações e alterações metodológicas. Nessa vertente, partimos da **hipótese** de que as várias mudanças no desenvolvimento do sistema conduziram ao aperfeiçoamento avaliativo técnico-procedimental, como abrangência, regularidade, metodologia etc. Por outro lado, também contribuíram para o uso do sistema como política de responsabilização (do inglês, *accountability*), no qual a responsabilidade pelos maus resultados da avaliação recaem sobre a escola, professores e alunos.

Dessa forma, o **objetivo geral** desta pesquisa é analisar a trajetória histórica do Saeb ao longo dos mais de trinta anos de implementação, identificando os princípios e as concepções que fundamentam o sistema. Para responder a esse objetivo, assumimos como **questão norteadora**: de que maneira o Saeb se insere historicamente como uma política de avaliação da educação básica brasileira?

Como objetivos específicos, temos: (a) realizar uma revisão sistemática da literatura sobre o Saeb, de modo a compreender o que tem sido estudado e apontar para as lacunas existentes, indicando possibilidades de pesquisa no campo da avaliação educacional; (b) compreender as experiências iniciais de avaliação estandardizada que decorreram no país e que serviram como base para a implementação do sistema; (c) analisar a teoria adotada, o quê e como se avalia, e como esses aspectos contribuíram para o processo de desenvolvimento e institucionalização do Saeb; e (d) realizar um levantamento sobre as primeiras iniciativas de uma política de avaliação nacional pautada nos processos de prestação de contas e de responsabilização, a partir dos resultados do Saeb.

O desenvolvimento desta dissertação também se justifica pelo interesse pessoal em estudar o objeto. O primeiro contato que tivemos com o Saeb foi na educação básica, por meio da Prova

Brasil, como estudante. Desde então, temos nos deparado com as avaliações externas aplicadas em larga escala. Realizamos o Exame Nacional do Ensino Médio (Enem) para ter acesso ao ensino superior e o Exame Nacional de Desempenho dos Estudantes (Enade), durante o curso de Licenciatura em Matemática da Universidade Federal de Viçosa (UFV), quando nos graduamos. Mais tarde, como professora, tivemos a oportunidade de participar da aplicação do Saeb na escola em que atuávamos, o que intensificou nosso interesse em conhecer os processos que permeiam o sistema e que dão subsídios para sua operacionalização.

Do ponto de vista político, esta pesquisa se insere no conjunto de ações desenvolvidas pelo Instituto de Pesquisa em Educação e Educação Física (Proteoria),¹ que tem se dedicado a produzir investigações acerca das avaliações externas aplicadas em larga escala no Brasil e no exterior (MARQUES; STIEG; SANTOS, 2020; MARQUES *et al.*, 2020; MARQUES, 2020; SANTOS *et al.*, 2021; SOARES, D.; SOARES, T.; SANTOS, 2020, 2021; SOARES, T.; SOARES, D.; SANTOS, 2021a, 2021b).

Ademais, localiza-se na linha de pesquisa *Educação, Formação Humana e Políticas Públicas* do Programa de Pós-Graduação em Educação da Universidade Federal do Espírito Santo (Ufes), articulando-se com os estudos que estão sendo desenvolvidos sobre o uso das avaliações padronizadas como políticas educacionais (SOARES, 2023;² TEIXEIRA, 2024³).

TEORIA E MÉTODO

Apesar dos diferentes conceitos de avaliação existentes na literatura especializada, há uma estrutura característica que, conforme expresso por Castillo Arredondo e Diago (2009), com sua ausência, não seria possível conceber uma avaliação de forma autêntica. Essa estrutura servirá como suporte conceitual para as discussões propostas ao longo desta dissertação.

¹ O Proteoria foi criado em 1999 com a finalidade de realizar pesquisas sobre problemas relevantes nas interseções entre os campos da História, da Educação e da Educação Física. Maiores informações sobre o Instituto podem ser encontradas no site em <http://proteoria.org/>.

² Tese em desenvolvimento.

³ Tese em desenvolvimento.

Inicialmente, assumimos a avaliação como “[...] um processo dinâmico, aberto e contextualizado que se desenvolve ao longo de um período de tempo; não é uma ação pontual ou isolada” (CASTILLO ARREDONDO; DIAGO, 2009, p. 38). Essa definição também vale para a avaliação educacional, pois os assuntos referentes à educação estão entrelaçados no cenário sociopolítico. Sendo assim, os autores chamam a atenção para as três dimensões consideradas fundamentais e irrenunciáveis para qualquer avaliação: a) a obtenção de informações; b) a formulação de juízos de valor; e c) a tomada de decisões.

Na primeira, a obtenção de informações acontece por meio da “[...] aplicação de procedimentos válidos e confiáveis para conseguir dados e informação sistemática, rigorosa, relevante e apropriada que fundamente a consistência e a segurança dos resultados da avaliação” (CASTILLO ARREDONDO; DIAGO, 2009, p. 39). Trata-se, portanto, dos dados brutos gerados ao final do processo avaliativo.

Na segunda, “[...] os dados obtidos devem permitir fundamentar a análise e a avaliação dos fatos que se pretende avaliar, para que se possa formular um juízo de valor o mais acertado possível” (CASTILLO ARREDONDO; DIAGO, 2009, p. 39). Ou seja, é uma espécie de julgamento (ou juízo) realizado a partir das informações disponíveis.

Nesse sentido, Cotta (2001) alerta para o objetivo geral do Saeb, que consiste em fornecer um diagnóstico do desempenho escolar dos estudantes e não, propriamente, explicar os resultados obtidos. A autora destaca que,

De maneira geral, o impacto das informações de avaliação é cumulativo, ou seja, tais informações nem sempre provocam mudanças imediatas em políticas e programas, mas elas servem, via de regra, para aumentar o conhecimento acumulado sobre uma dada situação ou forma de intervenção na realidade social (COTTA, 2001, p. 108).

Desse modo, compete aos principais interessados, como professores, diretores e gestores educacionais, o compromisso de se inteirar das informações oferecidas pela avaliação e, a partir disso, formular os juízos de valor cabíveis a cada caso. Por outro lado, é preciso levar em consideração que ela não pode ser a única base de decisão (ABRAMOWICZ, 1994).

Como discutido, existem outras informações e valores que fazem parte do cenário educacional e que não são, necessariamente, mensurados pela avaliação externa. Por isso também é preciso estar atento às questões inerentes à própria escola, as quais envolvem os processos, práticas e subjetividades, que só são perceptíveis por aqueles que pertencem ao seu cotidiano.

Ainda assim, instruir-se da realidade educacional e formular um juízo de valor, por si só, não é suficiente para promover mudanças significativas no contexto educativo. Do mesmo modo, de acordo com Castillo Arredondo e Diago (2009), é necessário estabelecer previamente os critérios a serem seguidos, a fim de formular juízos de valor educacionais coerentes e assertivos, capazes de auxiliar o processo de tomada de decisões.

Na medida em que a educação desencadeia ações que podem transformar a realidade social dos sujeitos, segundo Abramowicz (1994), cabe à avaliação firmar esse “compromisso” com a realidade. Embora seja notória sua importância, ela não dá conta de todos os desafios no âmbito do sistema educacional. Por esse motivo, “[...] deverá vir acompanhada de uma série de outras medidas que criem condições possibilitadoras de se encaminhar uma melhoria da qualidade de ensino” (ABRAMOWICZ, 1994, p. 82).

Seguindo esse raciocínio, a terceira dimensão se refere à tomada de decisões, em que providências pertinentes poderão ser tomadas de acordo com os juízos de valor que foram emitidos sobre a informação relevante disponível. A avaliação deve, portanto, ser conduzida para subsidiar esse processo, especialmente, “[...] no sentido de garantir que políticas públicas validem os recursos despendidos e produzam os efeitos esperados” (ABRAMOWICZ, 1994, p. 90).

Posto isso a avaliação educacional, em seu sentido mais amplo, pode ser definida como um “[...] processo que visa à coleta e ao uso de informações que permitam tomar decisões sobre um programa educacional” (VIANNA, 1982, p. 66). Para Cronbach (2000), esse programa pode ser um conjunto de materiais instrucionais distribuídos em nível nacional, atividades de uma única escola ou experiências educacionais de um único aluno.

Nesse sentido, o autor sinaliza três tipos de tomadas de decisões: o primeiro é a melhoria do curso com o objetivo de decidir quais matérias e métodos de instrução são satisfatórios e em que parte a mudança é necessária; o segundo são as decisões sobre os indivíduos, cujo intuito é identificar as necessidades do aluno para planejar sua instrução, avaliar seu mérito para fins de agrupamento e familiarizá-lo com seus próprios avanços e dificuldades; o terceiro, refere-se ao regulamento administrativo, com a finalidade de analisar o sistema escolar, bem como os professores, alunos, diretores etc.

Em resumo, a avaliação permitiria decisões que levariam ao aperfeiçoamento do currículo dos alunos submetidos a determinado programa e à discussão de assuntos de natureza

administrativa (VIANNA, 1982). Vale a pena destacar que existe um senso comum na literatura de que essas sejam as principais funções da avaliação (ABRAMOWICZ, 1994; CRONBACH, 2000; CASTILLO ARREDONDO; DIAGO, 2009; TAVARES, 2012), sobretudo com relação aos exames estandardizados que, em teoria, objetivam a melhoria da qualidade do ensino (BRASIL, 2020a).

Conforme discutido por Cotta (2001), o impacto da utilização dos resultados da avaliação na tomada de decisões dependerá tanto da capacidade de avaliar quanto da disposição em utilizar os resultados. Em geral, os países que possuem uma cultura avaliativa consolidada tendem a fazer uso desse processo de maneira mais espontânea em relação aos que ainda estão em desenvolvimento. Por exemplo, o pioneirismo dos Estados Unidos da América (EUA) e da Inglaterra, que foram os países que iniciaram tal movimento e contribuíram para sua expansão em escala global (SOUSA, 2014).

De todo o modo, o sistema avaliativo precisa ser provocador de ações de melhoria, pois, caso contrário, será reduzido a um simples instrumento de coleta de dados. Em concordância com Tavares (2012, p. 63) se, porventura, “[...] os resultados de desempenho dos alunos e contextuais das escolas não forem utilizados pelo sistema na tomada de decisão e no planejamento de ações, não faz sentido a implantação da avaliação de larga escala”.

Subsidiada nos conceitos supracitados e para alcançar os objetivos propostos, esta pesquisa foi desenvolvida, essencialmente, sob as possibilidades teóricas e metodológicas do paradigma indiciário proposto por Carlo Ginzburg (1989). Trata-se de uma proposta de um método heurístico centrado nos dados pormenorizados, ou seja, nos detalhes assumidos como pistas, indícios ou sinais que podem apontar para um caminho pelo qual a pesquisa pode ser conduzida.

Na coletânea *Mitos, emblemas, sinais: morfologia e história*, especificamente no texto *Sinais: raízes de um paradigma indiciário*, o autor italiano critica a História tradicional, que ocultou uma série de detalhes marginalizados. Embora considerados de pequena importância, no contexto geral, poderiam ser reveladores para a compreensão e problematização dos fatos históricos. Ginzburg (1989) defende, então, a construção do raciocínio por sinais/indícios que, quando articulados, permitiriam a obtenção de conclusões mais amplas, remontando às causas por meio dos efeitos percebidos.

Ademais, o autor defende que as percepções e crenças dos sujeitos em frente aos acontecimentos históricos sejam consideradas e, assim, o historiador poderia trabalhar com pistas, para além dos fatos explícitos. Isso permitiria a captação de aspectos da realidade que seriam inatingíveis a partir dos meios clássicos de investigação, pois se “[...] a realidade é opaca, existem zonas privilegiadas – sinais, indícios – que permitem decifrá-la” (GINZBURG, 1989, p. 152). Para ele, “[...] o que caracteriza esse saber é a capacidade, a partir de dados aparentemente negligenciáveis, de remontar a realidade complexa não experimentável diretamente” (GINZBURG, 1989, p. 152).

Ainda segundo Ginzburg (1989), a investigação baseada em indícios já era realizada como técnica pelos caçadores do período Neolítico.⁴ Para sobreviver, os indivíduos aprenderam a decifrar pistas, como pegadas na lama, odores, tufo de pelos, entre outros eventos. Assim, o paradigma indiciário resgata a função do caçador, que fareja, registra, interpreta e classifica as pistas deixadas no caminho (pegadas, fezes, pelos) e vai desenvolvendo uma narrativa coerente mediante tais vestígios.

Nessa vertente, utilizamos o processo de caça como objeto de estudo, na busca por compreender e problematizar as influências exercidas pelos sujeitos e instituições multilaterais no desenho e implementação do Saeb. Por meio dos vestígios deixados pelos agentes que participaram dessa ação, analisaremos os seus desdobramentos para as políticas educacionais brasileiras contemporâneas, que foram subsidiadas pelo desenvolvimento desse sistema de avaliação. Nosso intuito, então, é localizar as ações humanas no tempo responsáveis por promover alterações que contribuíssem para o desenvolvimento técnico-procedimental do Saeb, bem como o fortalecimento do uso regulativo do sistema.

Para tanto, este estudo, de natureza qualitativa, assume as pesquisas documental e bibliográfica em seus aspectos teóricos e metodológicos. A primeira caracteriza-se pela busca de informações em fontes que não receberam nenhum tipo de tratamento científico. A segunda remete às contribuições realizadas por diferentes sujeitos, cujas fontes investigadas já possuem um reconhecimento de domínio científico (OLIVEIRA, 2007).

A pesquisa documental é muito próxima da pesquisa bibliográfica. O elemento diferenciador está na natureza das fontes: a pesquisa bibliográfica remete para as contribuições de diferentes autores sobre o tema, atentando para as fontes secundárias, enquanto a pesquisa documental recorre a materiais que ainda não

⁴ Período entre os anos de 7 mil a.C. e 2500 a.C, compreendendo uma das fases da Pré-História, caracterizado, sobretudo, pela sedentarização humana e produção agrícola (GOSDEN, 2012).

receberam tratamento analítico, ou seja, as fontes primárias (SÁ-SILVA; ALMEIDA; GUINDANI, 2009, p. 6).

Desse modo, utilizamos, como fontes primárias, as publicações acadêmicas relacionadas com o Saeb e entrevistas jornalísticas com os sujeitos envolvidos no processo, disponibilizadas, na íntegra, na internet. Além disso, assumimos como fontes os documentos governamentais produzidos com o intuito de regulamentar e direcionar as ações dos governos brasileiro em prol das políticas de avaliação da educação básica brasileira. Esses documentos estão descritos a seguir no Quadro 1.

QUADRO 1 – Documentos Normativos

Ano	Documento	Resumo
1959	Lei n°. 3.692, de 15 de dezembro de 1959	Institui a Superintendência do Desenvolvimento do Nordeste e dá outras providências
1980	Decreto n°. 85.287, de 23 de outubro de 1980	Cria o Programa de Expansão e Melhoria da Educação no Meio Rural do Nordeste (Edurural – NE)
1982	Ministério da Educação	Programa de Expansão e Melhoria da Educação no Meio Rural do Nordeste
1986	Decreto n°. 92.435, 3 de março de 1986	Dispõe sobre a Estrutura Básica da Secretaria Executiva da Superintendência do Desenvolvimento do Nordeste (Sudene)
1986	Lei n°. 7.499, de 25 de junho de 1986	Aprova as diretrizes do Primeiro Plano de Desenvolvimento do Nordeste da Nova República
1987	Ministério da Educação	Alcance e funcionamento operacional do Projeto BRA-86-002
1988	Ministério da Educação	Constituição da República Federativa do Brasil de 1988
1993	Ministério da Educação	Plano Decenal de Educação para Todos 1993-2003
1994	Ministério da Educação	Sistema Nacional de Avaliação da Educação Básica: objetivos, diretrizes, produtos e resultados
1994	Portaria n°. 1.795, de 27 de dezembro de 1994	Cria o Sistema de Avaliação da Educação Básica (Saeb)
1996	Lei n°. 9.934, de 20 de dezembro de 1996	Lei de Diretrizes e Bases da Educação Nacional (LDB)
1998	Ministério da Educação	Matrizes Curriculares de Referência para o Saeb/97
2001	Lei n°. 10.172, de 9 de janeiro de 2001	Aprova o Plano Nacional de Educação (PNE 2001)
2005	Portaria n°. 931, de 21 de março de 2005	O Saeb passa a ser composto por dois processos de avaliação: a Avaliação Nacional da Educação Básica (Aneb) e a Avaliação Nacional do Rendimento Escolar (Anresc)
2005	Portaria n°. 69, de 4 de maio de 2005	Estabelece a sistemática para a realização da Avaliação Nacional do Rendimento Escolar (Anresc)
2005	Portaria n°. 89, de 25 de maio de 2005	Estabelece a sistemática para a realização da Avaliação Nacional da Educação Básica (Aneb)

2007	Decreto nº. 6.094, de 24 de abril de 2007	Implementação do Plano de Metas Compromisso Todos pela Educação
2013	Portaria nº. 482, de 7 de junho de 2013	Dispõe sobre o Saeb
2014	Lei nº. 13.005, de 25 de junho de 2014.	Aprova o Plano Nacional de Educação (PNE 2014)
2017	Portaria nº. 564, de 19 de abril de 2017	Altera a Portaria MEC nº 482 que dispõe sobre o Saeb
2018	Ministério da Educação	Base Nacional Comum Curricular (BNCC)
2019	Portaria nº. 366, de 29 de abril de 2019	Estabelece as diretrizes de realização do Saeb em 2019
2020	Portaria nº. 458, de 5 maio de 2020	Institui normas complementares necessárias ao cumprimento da Política Nacional de Avaliação da Educação Básica
2020	Ministério da Educação/Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira	Matrizes de Referência de Língua Portuguesa e Matemática do Saeb: documento de referência do ano de 2001
2021	Portaria nº. 10, de 8 de janeiro de 2021	Estabelece parâmetros e fixa diretrizes gerais para implementação do Saeb, no âmbito da Política Nacional de Avaliação da Educação Básica
2021	Portaria nº. 250, de 5 de julho de 2021	Estabelece as diretrizes de realização do Saeb em 2021

Fonte: Elaboração dos autores.

Como fontes secundárias, para a compreensão do objeto, utilizamos as produções teóricas e acadêmicas das organizações multilaterais internacionais e dos sujeitos que contribuíram para a criação e o desenvolvimento do Saeb. Para a seleção desse material, realizamos uma busca exploratória de relatórios técnicos e artigos nos *sites* oficiais dessas organizações e em bases de dados, utilizando descritores específicos.⁵

Em síntese, subsidiados pelo paradigma indiciário (GINZBURG, 1989), procuramos descobrir, a partir da análise dessas fontes, informações para além do que elas pretendiam revelar. Esses dados podem ter sido tratados apenas como um fato posto pelos sujeitos que produziram esses materiais, mas, quando submetidos a uma análise semiótica, podem revelar muito mais do que o testemunho assumido simplesmente como um dado.

Ademais, referenciamos-nos na análise crítico-documental, proposta por Bloch (2001), para investigar o que está ou não presente nessas fontes, considerando que, mesmo que elas estejam aparentemente claras e complacentes, nada dirão se não soubermos interrogá-las. Para

⁵ O Capítulo 1 realiza uma revisão sistemática da literatura sobre o Saeb. Para seleção das fontes, utiliza as diretrizes de elegibilidade e análise do modelo Prisma (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) e a Análise Temática (BRAUN; CLARKE, 2006).

ele é preciso problematizar a noção de História, considerando-a como as “[...] ações dos homens no tempo” (BLOCH, 2001, p. 55), e não mais como o estudo de fatos isolados.

Bloch (2001) afirma, ainda, ao escrever sobre a observação histórica, que o conhecimento de todos os fatos humanos se deu por meio dos vestígios e marcas deixados pelos povos que viveram no passado: “[...] tudo o que o homem diz ou escreve, tudo o que fabrica, tudo o que toca, pode e deve informar sobre ele” (p. 79). Assim, interrogamos as fontes para compreender o que está presente ou ausente nos documentos analisados, assumindo-as como artefatos culturalmente estabelecidos e dotados de intencionalidades pelos grupos que a originaram.

Nesse processo de análise das fontes, apropriamo-nos dos conceitos de *centro e periferia* (GINZBURG, 2002) e *cultura política* (BERSTEIN, 2009). Para Ginzburg (2002), o *centro* se constitui como um lugar de poder, em que se dita o que se faz e o que se produz. Já a *periferia* é um lugar sem direito de decisão, o que não significa a inexistência de práticas de resistência ao *centro*. Para Berstein (2009), uma cultura política nasce quando uma sociedade anseia por mudanças em frente aos problemas e crises enfrentados ao longo de sua história, fornecendo inspiração para projetos políticos direcionados ao futuro.

Desse modo, assumimos, como *centro*, as instituições com significativa influência na criação do Saeb e, como *periferia*, o sistema educacional brasileiro. O constante conflito das relações de força existente entre esses lugares auxiliou para a construção de uma *cultura política* nacional focada em exames estandardizados, o que será discutido, minuciosamente, no decorrer desta dissertação.

ORGANIZAÇÃO DO ESTUDO

Esta dissertação está organizada em quatro capítulos, todos de natureza qualitativa e exploratória e se estrutura no designado “modelo escandinavo”, ou seja, no formato de artigos, desenvolvidos de modo a se articularem com o objetivo geral do trabalho.

No primeiro capítulo, que se caracteriza como uma revisão sistemática da literatura, analisamos a produção científica, considerando os principais estudos que já foram realizados

no Brasil sobre o Saeb. Dessa forma, investigamos a fundamentação teórica da produção analisada e o modo como propõe as discussões sobre a temática.

No segundo capítulo, efetuamos um levantamento das experiências iniciais de avaliação que serviram como base para o desenho do Saeb, apresentando os princípios que nortearam a sua criação. Para isso, consideramos o período compreendido entre os anos de 1960 e 1990, momento em que ocorreram os primeiros processos avaliativos aplicados em âmbito nacional.

No terceiro capítulo, analisamos as primeiras modificações que contribuíram para o processo de desenvolvimento e institucionalização do sistema, como alterações metodológicas, o que e como ele avalia. Dessa forma, estabelecemos como recorte temporal o período compreendido entre 1990 e 2001, quando sucedeu a adoção da Teoria de Resposta ao Item (TRI) e das matrizes de referência.

No quarto capítulo, discutimos sobre a configuração dos processos avaliativos do Saeb e a incorporação do modelo de prestação de contas na educação básica brasileira. Dessa maneira, consideramos o período de implementação da Prova Brasil, a partir de 2005, em que surgiram iniciativas de responsabilização no país.

Por fim, apresentamos nossas considerações finais, nas quais retornamos ao objetivo geral para realizar uma síntese do que foi discutido e apresentamos novas possibilidades de estudos na área.

CAPÍTULO I

SISTEMA DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA: REVISÃO SISTEMÁTICA DA LITERATURA

1.1 INTRODUÇÃO

Os sistemas de avaliação da educação, organizados por agentes externos à escola (ditas avaliações externas),⁶ ganharam maior visibilidade na América Latina a partir da década de 1980, fruto da ascensão das políticas internacionais no continente que defendiam o gerenciamento estatal da educação, por meio de sua regulação e monitoramento (UCZAK, 2014; MENDONCA, 2014).

Nesse período, propuseram-se no Brasil, as primeiras iniciativas externas de avaliação aplicada em larga escala, com destaque para o Projeto Nordeste de Educação Básica I (Edurural), o Projeto Nordeste de Educação Básica II (Projeto Nordeste) e a Avaliação do Rendimento Escolar de Alunos de Escolas de 1º Grau da Rede Pública. Esta última, realizada por iniciativa do Ministério da Educação (MEC) com a colaboração da Fundação Carlos Chagas (FCC), com o objetivo de verificar a viabilidade de um projeto de avaliação em nível nacional, servindo de base para o desenho inicial do Sistema Nacional de Avaliação do Ensino Público de 1º Grau (Saep).

A primeira aplicação do Saep ocorreu no final de 1988. Com ela, foi possível verificar a pertinência dos instrumentos e procedimentos de medida até então utilizados nas avaliações externas nacionais, com vistas para a ampliação da proposta, que culminaria na criação do Sistema de Avaliação da Educação Básica (Saeb). Dessa forma, alguns estudos referem-se ao Saep como um projeto piloto do Saeb (BONAMINO; FRANCO, 1999; PESTANA, 2016).

As primeiras aplicações do Saeb ocorreram em 1990 e 1993 e assumiram os currículos de sistemas estaduais como subsídio para a elaboração das provas, que avaliaram a

⁶ Essas avaliações geralmente são aplicadas em larga escala, ou seja, a um grande número de estudantes. Dessa forma, é comum referir-se a elas como avaliações externas em larga escala ou avaliações estandardizadas (MARQUES; STIEG; SANTOS, 2020).

aprendizagem de uma amostra de estudantes, matriculados nas 1^a, 3^a, 5^a e 7^a séries do ensino fundamental, em língua portuguesa, matemática, ciências naturais e redação.

Entretanto, o Saeb só foi regulamentado quatro anos após o primeiro ciclo, a partir da publicação da Portaria n.º. 1.795, de 27 de dezembro de 1994 (BRASIL, 1994). Conforme o Art. 1º, constituíam os objetivos gerais do referido sistema de avaliação:

1. contribuir para o desenvolvimento, em todos os níveis do sistema educativo, de uma cultura avaliativa que estimule a melhoria dos padrões de qualidade e equidade da educação básica e adequados controles sociais de seus resultados; 2. programar e desenvolver, em articulação com as Secretarias de Educação, processos permanentes de avaliação, apoiados em ciclos regulares de aferições das aprendizagens e competências obtidas pelos alunos e do desempenho dos estabelecimentos que compõem os sistemas de educação básica; 3. mobilizar os recursos humanos, técnicos e institucionais do Ministério, das Secretarias e de universidades e centros de estudos e pesquisas sociais e educacionais, para gerar e difundir os conhecimentos, técnicas e instrumentos requeridos pelas práticas de aferição e avaliação educacional; 4. proporcionar, aos responsáveis pela gestão educacional e pelo ensino, às famílias e aos cidadãos em geral, informações seguras e oportunas a respeito do desempenho e dos resultados dos sistemas educativos. (BRASIL, 1994, p. 1).

Nesses 30 anos de existência, o Saeb passou por alguns aprimoramentos, adaptações e alterações metodológicas, que auxiliaram na consolidação e na expansão do sistema (PESTANA, 2016). Entre essas modificações, destacamos o levantamento de dados contextuais, por meio de questionários, a partir do ciclo de 1995; a adoção de Matrizes de Referência para a elaboração dos itens, a partir do ciclo de 1997; e sua reestruturação por meio da publicação da Portaria Ministerial n.º. 931, de 21 de março de 2005 (BRASIL, 2005a), fazendo com que o sistema passasse a ser composto pela Avaliação Nacional da Educação Básica (Aneb) e pela Avaliação Nacional do Rendimento Escolar (Anresc), mais conhecida como Prova Brasil.

A Aneb manteve o delineamento amostral, até então adotado pelo Saeb, condicionado ao mínimo de dez estudantes por turma avaliada. A Prova Brasil, por sua vez, inaugurou a avaliação censitária no contexto nacional, passando a avaliar todas as unidades que atendessem aos critérios de no mínimo 30 matrículas em cada etapa dos anos iniciais (4^a série/5º ano) ou dos anos finais (8^a série/9º ano) do ensino fundamental de escolas públicas, permitindo (re)conhecer as particularidades de cada escola e município (BRASIL, 2005a). Para a 3^a série do ensino médio a avaliação tornou-se censitária apenas a partir de 2017, com a publicação da Portaria n.º. 564, de 19 de abril de 2017 (BRASIL, 2017a).

Além disso, a partir de 2007, as médias de desempenho dos estudantes, calculadas no Saeb, passaram a ser combinadas com as taxas de aprovação, reprovação e abandono, apuradas no Censo Escolar, mediante criação do Índice de Desenvolvimento da Educação Básica (Ideb), instituído pela publicação do Decreto n°. 6.094, de 24 de abril de 2007 (BRASIL, 2007a). O Ideb passou a ser utilizado como “[...] referência para o repasse e a distribuição dos recursos do Fundo Nacional de Desenvolvimento da Educação (FNDE)” (PESTANA, 2016, p. 80), ampliando o impacto do Saeb nas políticas educacionais brasileiras.

Em 2013, como uma das ações previstas na portaria que institui o Pacto Nacional pela Alfabetização na Idade Certa (Pnaic), a Avaliação Nacional de Alfabetização (ANA) passou a compor o Saeb, visando a realizar um diagnóstico amplo do processo de alfabetização nas escolas públicas brasileiras. Desde então, o Saeb passou a ser composto por três avaliações: ANA, Aneb e Prova Brasil.

Em 2020, o Saeb passou por outra reestruturação, mediante a publicação da Portaria n°. 458, de 5 de maio de 2020 (BRASIL, 2020a). A avaliação, até então aplicada bienalmente, passou a ser anual, abrangendo todas as séries a partir do 2° ano do ensino fundamental. Além disso, os resultados das avaliações realizadas no âmbito do ensino médio passaram a ser utilizados como uma nova possibilidade de acesso ao ensino superior (BRASIL, 2020a).

Considerando esse percurso histórico e as transfigurações do sistema nas três últimas décadas, faz-se necessário compreender a dinâmica dos estudos que consideram o Saeb como objeto de análise, no sentido de contribuir para o estado da arte inerente ao tema. Nessa perspectiva, o presente capítulo se insere no conjunto de estudos que se destinam a elaborar uma análise da produção acadêmica no campo da educação, sobretudo aquelas que se dedicam às avaliações de sistemas (GATTI; VIANNA; DAVIS, 1991; GATTI, 2009; BAUER, 2010) ou os exames standardizados (BAUER; ALAVARSE; OLIVEIRA, 2015; MARQUES; STIEG; SANTOS, 2020).

Posto isso, este capítulo objetiva realizar uma revisão sistemática da literatura sobre o Saeb, de modo a compreender o que tem sido estudado e apontar para as lacunas existentes, indicando possibilidades de pesquisa no campo da avaliação educacional.

1.2 METODOLOGIA

Neste capítulo, de natureza qualitativa do tipo exploratória, realizou-se uma revisão sistemática da literatura sobre o Saeb. Para isso, foram utilizadas as diretrizes de elegibilidade e análise do modelo Prisma (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) e a Análise Temática proposta por Braun e Clarke (2006).

O modelo Prisma, elaborado como ponto de partida para o desenvolvimento de uma revisão de literatura, auxilia, por meio de seus métodos sistemáticos, na identificação, seleção e avaliação crítica a respeito das pesquisas mais relevantes relacionadas ao tema em estudo (MOHER *et al.*, 2009). Trata-se de um *checklist* que direciona o processo disposto em quatro fases: identificação, triagem, elegibilidade e inclusão.

A análise temática, por sua vez, segundo Souza (2019), contribui para estreitar o campo de visão analítico do pesquisador. É um instrumento estruturado em seis fases para que certas decisões, tradicionalmente implícitas, se tornem evidentes na realização da análise, tornando-se eficaz para interpretar e relatar padrões (temas) a partir dos dados qualitativos.

Para compor esta revisão, optou-se por uma consulta expandida ao Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), considerando SciELO (*Scientific Electronic Library Online*), Scopus e Web of Science (WoS) como bases de dados eletrônicas. Essas bases foram escolhidas por serem as principais indexadoras para a obtenção de melhores classificações nas avaliações do Qualis Periódicos da Área da Educação⁷ (COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR – CAPES, 2019).

O prazo final adotado para a indexação dos artigos nas bases consideradas foi o ano de 2019, não havendo delimitação inicial. Além disso, utilizou-se como descritor o “Sistema de Avaliação da Educação Básica”. Como critérios de inclusão, optou-se por periódicos revisados por pares, disponíveis em qualquer idioma.

⁷ Para mais informações, acessar o relatório disponível no *site*: <http://www.gov.br/capes/pt-br/centrais-de-conteudo/relatorio-qualis-educacao-pdf>. Acesso em: 16 abr. 2020.

Excluíram-se os artigos duplicados e os vinculados a periódicos cujo foco não fosse a educação. Em seguida, realizou-se uma leitura dos resumos dos artigos previamente selecionados, excluindo aqueles cuja temática principal não era o objeto desta pesquisa.

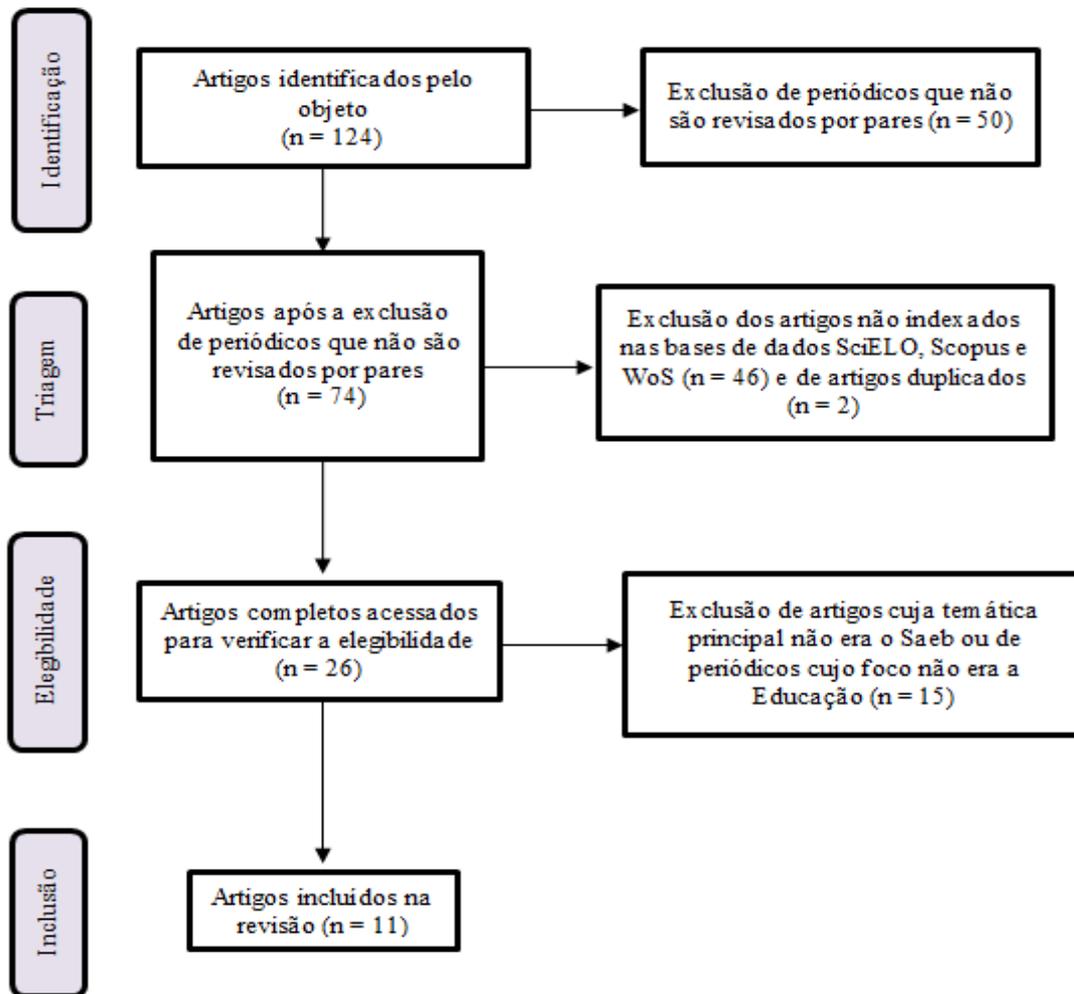
Por fim, foram selecionados os artigos para compor essa revisão, que foram lidos na íntegra por todos os autores e agrupados de acordo com a análise temática, conforme suas similaridades, que são apresentadas neste capítulo. Também é realizada uma discussão sobre seus autores, que abarca o vínculo institucional, os grupos de pesquisa aos quais participam, os parceiros e os referenciais bibliográficos adotados nas publicações. Para isso, utilizou-se como fonte a plataforma Lattes, cujas informações foram obtidas em junho de 2020.

1.3 RESULTADOS

Mediante as estratégias de pesquisa, foram identificados, inicialmente, 124 artigos, dos quais 74 pertenciam a periódicos revisados por pares. Destes, 28 estavam indexados nas bases de dados definidas (SciELO, Scopus e WoS). Com a leitura dos resumos, optou-se pela exclusão daqueles cuja temática principal não era o objeto desta pesquisa ou que foram publicados em periódicos cujo foco não era a educação. Entre os artigos excluídos, destacam-se a presença de discussões sociais, curriculares, estruturais e de estudos que consideram instituições de ensino específicas, restando 11 artigos para as análises propostas neste estudo.

A Figura 1 apresenta um fluxograma da seleção dos artigos analisados, que foi realizada de acordo com as diretrizes do método Prisma.

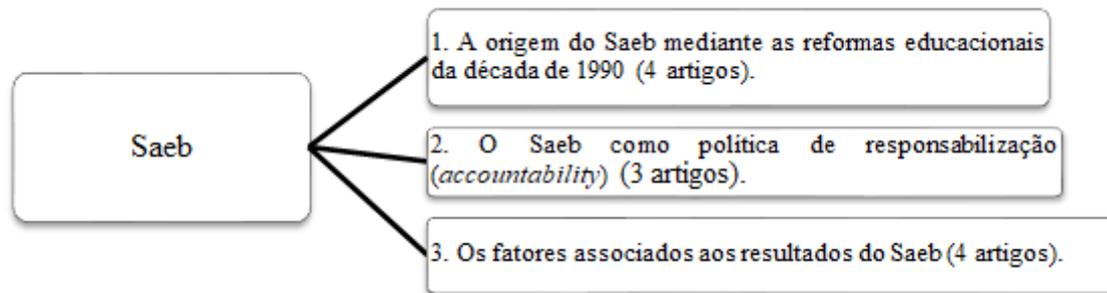
FIGURA 1 – Fluxograma da seleção dos artigos de acordo com o método Prisma



Fonte: Elaboração dos autores.

Após a leitura integral dos artigos selecionados, os temas foram classificados em três categorias (1, 2 e 3). A Figura 2 apresenta como se deu esta classificação, que foi utilizada na construção do Quadro 2, que também apresenta o ano de publicação, o título e o(s) autor(es) dos artigos selecionados pela metodologia do modelo Prisma.

FIGURA 2 – Classificação dos temas



Fonte: Elaboração dos autores.

QUADRO 2 – Caracterização do *corpus* textual dos artigos selecionados

ANO	Título	Autor (es)	Categoria
1999	Avaliação e política educacional: o processo de institucionalização do Saeb.	Alicia Bonamino, Creso Franco.	1
2008	Vinte anos de avaliação da educação básica no Brasil: aprendizagens e desafios.	Maria Inês de Matos Coelho.	1
2008	Políticas educacionais e desempenho escolar nas capitais brasileiras.	Fátima Alves.	3
2010	Sistemas estaduais de avaliação: uso dos resultados, implicações e tendências.	Sandra Zákia Sousa, Romualdo Portela de Oliveira.	1
2012	Uma análise de <i>rankings</i> de escolas brasileiras com dados do Saeb.	Ana Maria Paiva Franco, Naércio Menezes Filho.	2
2015	A educação brasileira está melhorando? Evidências do Pisa e do Saeb.	Martin Camoy, Tatiana Khavenson, Izabel Fonseca, Leandro Costa, Luana Marotta.	3
2016	O surgimento da avaliação da qualidade na educação básica brasileira.	Jaakko Kauko, Vera Gorodski Centeno, Helena Candido, Eneida Shiroma, Anni Klutas.	1
2016	A Avaliação Nacional da Alfabetização no contexto do sistema de avaliação da educação básica e do Pacto Nacional pela Alfabetização na Idade Certa: responsabilização e controle.	Adriana Dickel.	2
2016	Jovens em contextos sociais desfavoráveis e sucesso escolar no ensino médio.	Vanessa Gomes de Castro, Fernando Tavares Júnior.	3
2017	A avaliação no ensino de ciências naturais nos documentos oficiais e na literatura acadêmica: uma temática com muitas questões em aberto.	Claudio Rejane da Silva Dantas, Neusa Teresinha Massoni, Flávia Maria Teixeira dos Santos.	2
2018	Estudo longitudinal sobre eficácia educacional no Brasil: comparação entre resultados contextualizados e valor acrescentado.	Maria Eugénia Ferrão, Gabriela Thamara de Freitas Barros, Alvana Maria Bof, Adolfo Samuel de Oliveira.	3

Fonte: Elaboração dos autores.

Ressalta-se que entre os artigos analisados, nenhum autor apresenta mais de uma contribuição para a realização deste estudo. Ademais, a autoria dos trabalhos que compõem essa revisão se dá de forma individual ($n = 3$), em dupla ($n = 4$) ou em grupo ($n = 4$).

Os trabalhos individuais são realizados por professoras universitárias (ALVES, 2008; COELHO, 2008; DICKEL, 2016). No que diz respeito à autora Adriana Dickel, professora titular da Universidade de Passo Fundo (UPF), destaca-se que uma de suas áreas de pesquisa é o estudo de políticas e práticas de alfabetização no primeiro ciclo do ensino fundamental. Vale ressaltar que não foram encontradas ligações de Dickel com os demais autores, nem outras produções relacionadas ao Saeb.

Quanto à autora Maria Inês de Matos Coelho, professora titular da Universidade Estadual de Minas Gerais (UEMG), ressaltamos sua atuação como consultora, desde 2009, do Programa das Nações Unidas para o Desenvolvimento (Pnud). Apesar de Coelho (2008) ser citada por Kauko *et al.* (2016) e por Dantas, Massoni e Santos (2017) e da avaliação educacional ser uma de suas linhas de pesquisa, verificou-se a ausência de outras produções da autora diretamente relacionadas ao Saeb.

A autora Fátima Alves desenvolveu seu artigo quando era professora associada da Pontifícia Universidade Católica do Rio de Janeiro (PUC/RJ). Esse período se deu após a conclusão de seu doutorado em educação sob a orientação do professor Creso Franco, autor de outra produção analisada no presente capítulo. É com Franco que Alves tem o maior volume de publicações, seguido por Alícia Bonamino, também autora de um dos artigos. Além de apresentar outras produções relacionadas ao Saeb, Alves (2008) é citada por Kauko *et al.* (2016) e já prestou consultoria para instituições como a Organização das Nações Unidas para a Educação, a Ciência e a Cultura (Unesco) e o Banco Mundial.

As produções realizadas em dupla são oriundas de professores universitários ($n = 3$) (BONAMINO; FRANCO, 1999; SOUSA; OLIVEIRA, 2010; FRANCO; MENEZES-FILHO, 2012) e alunos de uma mesma instituição ($n = 1$) (CASTRO; TAVARES JÚNIOR, 2016). Dois dos artigos publicados por professores são de profissionais que compartilham o mesmo local de trabalho (BONAMINO; FRANCO, 1999; SOUSA; OLIVEIRA, 2010).

O artigo dos professores da PUC/RJ (BONAMINO; FRANCO, 1999) foi produzido durante o doutorado da professora associada Alícia Maria Catalano de Bonamino, também sob

orientação do professor Creso Franco, que desde o final da década de 1990 coordenou projetos que analisam a situação da educação brasileira, baseando-se nos dados do Saeb.

Entre os projetos de pesquisa coordenados pelo professor Creso Franco, ressalta-se a atuação das autoras Alícia Bonamino e Fátima Alves na busca por fatores da eficácia escolar, promoção de equidade na escola e no desenvolvimento e validação de instrumentos adequados para surveys educacionais nos anos iniciais do ensino fundamental. Por comporem os mesmos projetos de pesquisa, as autoras Bonamino e Alves têm trabalhos em conjunto e também com Creso Franco.

Ressalta-se que além de Bonamino e Franco (1999) serem citados por Sousa e Oliveira (2010), os autores também têm produções individuais referenciadas em alguns dos artigos analisados (COELHO, 2008; ALVES, 2008; SOUSA; OLIVEIRA, 2010; KAUKO *et al.*, 2016; DICKEL, 2016). Dessarte, Bonamino e Franco se mostram referências nas discussões acerca do Saeb.

Em relação aos professores da Universidade de São Paulo (USP) (SOUSA; OLIVEIRA, 2010) destaca-se que a professora colaboradora Sandra Zákia Sousa trabalhou no MEC, no período de 1982 e 1991, atuando também na gerência do Projeto de Capacitação de Recursos Humanos do Programa Edurural. O professor titular aposentado Romualdo Luiz P. Oliveira, por sua vez, atuou como colaborador da Unesco. Apesar de não ser o caso do presente trabalho, ambos são citados por outros dois artigos (KAUKO *et al.*, 2016; DICKEL, 2016). Verificou-se que, além dos autores apresentarem um considerável volume de publicações em conjunto, há também a participação de Adriana Bauer, referência em avaliação educacional e integrante de um dos projetos de pesquisa dos quais os autores fizeram parte.

Com respeito ao artigo dos professores de instituições distintas (FRANCO; MENEZES-FILHO, 2012), podemos ressaltar que Naércio Aquino Menezes-Filho, do Instituto de Ensino e Pesquisa (Insper), orientou a professora adjunta Ana Maria de Paiva Franco, da Universidade Federal de Uberlândia (UFU), durante seu doutorado em educação, sendo com ela suas únicas produções relacionadas ao Saeb. Menezes-Filho é citado em duas produções analisadas (ALVES, 2008; CASTRO; TAVARES JÚNIOR, 2016). Ana Maria de Paiva Franco, por sua vez, atuou como analista e consultora de instituições como Itaú (Unibanco), Banco Mundial e Comissão Econômica para América Latina e Caribe (Cepal). Além disso, a autora tem publicações com o economista Renato Baumann, seu orientador de mestrado e

diretor do Escritório Cepal no Brasil (1995-2010), período no qual a autora prestou serviços técnicos.

Por último, encontram-se os artigos publicados por grupos no âmbito nacional ($n = 1$) e internacional ($n = 3$). Ressalta-se que enquanto a produção nacional se restringe a um tema específico, que diz respeito à avaliação no ensino de ciências naturais (DANTAS; MASSONI; SANTOS, 2017), a produção internacional (CARNOY *et al.*, 2015; KAUKO *et al.*, 2016; FERRÃO *et al.*, 2018) abrange temas mais amplos como a economia (CARNOY *et al.*, 2015), formulação de políticas públicas (KAUKO *et al.*, 2016) e resultados de pesquisas realizadas pelo Inep (FERRÃO *et al.*, 2018).

Entre os autores das produções realizadas em grupo, é importante acentuar a relevância de Martin Carnoy, autor com mais citações entre os artigos analisados (BONAMINO; FRANCO, 1999; ALVES, 2008; CARNOY *et al.*, 2015; KAUKO *et al.*, 2016; FERRÃO *et al.*, 2018). Economista e professor da Universidade de Stanford, Carnoy também trabalhou como consultor do Banco Mundial e da Unesco. Além disto, o autor é conhecido por seus debates acerca da escola charter, por suas pesquisas em relação ao impacto das políticas de responsabilização na aprendizagem, eficácia e eficiência nas escolas e também pelos efeitos da globalização nos sistemas educacionais.

A seguir, apresenta-se uma exposição para cada categoria, identificada com base na análise temática deste estudo.

1.3.1 A origem do Saeb mediante as reformas educacionais da década de 1990

Nesta categoria se encontram quatro artigos que discutem o processo de criação do Saeb mediante as reformas educacionais que ocorreram durante a década de 1990, e que subsidiaram sua consolidação enquanto sistema de avaliação da educação básica brasileira.

À vista disso, apresentam-se três vertentes pertinentes relacionadas ao Saeb. A princípio, os artigos de Bonamino e Franco (1999) e Kauko *et al.* (2016) discutem quais foram as implicações para a pesquisa educacional brasileira, abordando a cooperação internacional e as primeiras experiências em avaliações, que influenciaram o processo de institucionalização do Saeb.

Em seguida, os artigos de Bonamino e Franco (1999) e Sousa e Oliveira (2010) analisam os objetivos inicialmente estabelecidos para o Saeb e sua influência na formulação de políticas públicas, por exemplo, a implantação de um currículo comum.

Por fim, o artigo de Coelho (2008) analisa a construção dos fatores de qualidade, eficiência, equidade e produtividade. Mediante uma análise temporal em maior nível, o artigo também aborda os enfoques e aperfeiçoamentos do Saeb ao longo do projeto histórico.

A primeira experiência em avaliação bem documentada, de acordo com Kauko *et al.* (2016) e Bonamino e Franco (1999), se refere ao Edurural, que envolveu três estados da região Nordeste,⁸ inicialmente apresentado em 1977. Os objetivos do Edurural eram voltados à ampliação do acesso ao ensino primário, redução das taxas de evasão e retenção, e melhoria da qualidade da educação.

O Edurural, segundo Bonamino e Franco (1999), foi financiado pelo Banco Mundial e teve sua atuação graças ao apoio da FCC e da Fundação Cearense de Pesquisa e Cultura, ambientada na Universidade Federal do Ceará (UFCE). Os resultados do projeto, que apontaram a precariedade do ensino nas escolas participantes, foram amplamente divulgados no cenário internacional, por meio dos relatórios do Banco Mundial. No Brasil, restringiu-se apenas às discussões no ambiente acadêmico (KAUKO *et al.*, 2016).

Após a conclusão do Edurural, percebeu-se a necessidade de ampliação do projeto para os demais estados da região Nordeste. Assim, em parceria com o MEC, o Banco Mundial passou a financiar o Projeto Nordeste de Educação Básica II, em 1984 (BONAMINO; FRANCO, 1999; KAUKO *et al.*, 2016).

O Projeto Nordeste, como ficou conhecido, envolveu os nove estados⁹ da região. A necessidade de desenvolvimento de um sistema de avaliação, para avaliar seu impacto no segmento educacional, foi o que “[...] levou as primeiras iniciativas que redundaram na criação do Sistema Nacional de Avaliação do Ensino Público de 1º Grau (Saep)” (BONAMINO; FRANCO, 1999, p. 110).

O Saep, em 1988, não apenas foi o primeiro esquema de avaliação sistemática, mas também preparou o terreno para a avaliação da educação básica brasileira, que ganha destaque na

⁸ Os estados participantes do Edurural eram Ceará, Pernambuco e Piauí (BONAMINO; FRANCO, 1999).

⁹ Os estados participantes do Projeto Nordeste eram: Alagoas, Bahia, Ceará, Maranhão, Paraíba, Pernambuco, Piauí, Rio Grande do Norte e Sergipe (PEREIRA, 2007).

década 1990 (KAUKO *et al.*, 2016). Nesse período, houve ampla discussão acerca das teorias de reprodução cultural e social, o que demonstrou as funções seletivas e excludentes da educação brasileira (COELHO, 2008).

Dessa forma, por meio desse cenário favorável e com a crescente utilização dos exames estandardizados, a partir de 1990, o Saeb passou a substituir o Saep, tendo sido criado com o objetivo de contribuir para o monitoramento da educação básica brasileira e para oferecer subsídios aos formuladores de políticas públicas (BONAMINO; FRANCO, 1999; SOUSA; OLIVEIRA, 2010).

Percebeu-se, pelas produções analisadas (BONAMINO; FRANCO, 1999; SOUSA; OLIVEIRA, 2010; COELHO, 2008), que além de passar por várias modificações, impostas por organismos e/ou estudos internacionais, o Saeb também influenciou outros sistemas de avaliação no âmbito nacional.

De acordo com Sousa e Oliveira (2010), destaca-se a influência do Saeb nos sistemas estaduais¹⁰ de avaliação, tornando-se uma referência para os gestores estaduais. Nota-se sua influência tanto como modelo de desenho a ser seguido quanto como parâmetro para a proposição de alternativas (SOUSA; OLIVEIRA, 2010). A citar, por exemplo, a Matriz de Referência do Saeb,¹¹ instituída em 1997, que também serviu como subsídio para a formulação dos currículos estaduais (BONAMINO; FRANCO, 1999; SOUSA; OLIVEIRA, 2010).

Durante seus primeiros vinte anos, por meio de um estudo realizado por Coelho (2008), apresentam-se as principais modificações sofridas pelo Saeb neste período. Dentre elas, destaca-se o levantamento de dados sobre o nível socioeconômico (NSE). Bonamino e Franco (1999), por sua vez, apontam os resultados do Relatório Coleman (1986) como essenciais para essa composição do Saeb. Esse relatório apontou o NSE dos alunos como o fator primordial para o desempenho escolar, o que levou à valorização de pesquisas de caráter social junto aos exames estandardizados (BONAMINO; FRANCO, 1999).

O Saeb, desde 1995, levanta dados referentes ao contexto social dos estudantes por meio da aplicação dos questionários contextuais, respondidos por alunos, professores e diretores,

¹⁰No artigo de Sousa e Oliveira (2010) analisam-se os seguintes estados: Bahia, Ceará, Minas Gerais, Paraná e São Paulo.

¹¹Essa pode ter sido a primeira experiência de implantação de um currículo comum no contexto nacional.

como observa Coelho (2008). Além disso, em concordância com a autora, a adoção de técnicas mais modernas para medição do desempenho do aluno, a redefinição das séries avaliadas e o uso das matrizes de referência também estão entre as principais alterações do Saeb.

Com base nessas alterações, conforme apresenta Coelho (2008), o Saeb também se consolida por seus estudos em relação à efetividade da escola (*School Effectiveness Research – SRE*) ou da “escola eficaz”. Essas pesquisas, por sua vez, visavam à elevação dos padrões educacionais, o que já era comum em países como Estados Unidos e Reino Unido (COELHO, 2008).

De acordo com Coelho (2008), constatou-se que várias produções internacionais na linha de estudo da “efetividade escolar” são referenciadas nos documentos do Saeb, bem como identificadas nos estudos realizados com base nos seus dados. Diante disso, o Saeb justifica-se como um forte contribuinte para que as mediações dos resultados de aprendizagem venham a ser realizadas por recursos quantitativos, bem como quanto ao uso da avaliação como instrumento de regulação e da administração gerencial e competitiva (COELHO, 2008).

Apesar de sua relevância, e de haver “[...] condições para correlacionar a medida de resultado com as características socioeconômicas dos alunos e com as características da escola” (COELHO, 2008, p. 241), persiste a dificuldade em esclarecer quais os efeitos dos fatores escolares (qualidade, eficiência, equidade e produtividade). A título de exemplo, sublinha-se a implicação das práticas docentes de sala de aula no desempenho dos alunos, conforme apresenta Coelho (2008), apuradas pelos questionários contextuais.

Encerram-se as discussões demonstrando que a avaliação passou a ser um dos eixos centrais das políticas públicas, em geral, e também das políticas educacionais (COELHO, 2008), no qual o Saeb determina-se a monitorar tais políticas e a qualidade da educação (BONAMINO; FRANCO, 1999), apesar de que para Sousa e Oliveira (2010), essa ainda não ser uma realidade.

Além disto, atenta-se ao fato de que a comunidade educacional, suas ações e debates foram estabelecidos por meio da adoção e/ou referência a ideias, práticas e políticas estrangeiras (KAUKO *et al.*, 2016), priorizadas de acordo com as condições impostas pelos acordos.

Diante disso, alerta-se para a necessidade de melhor conhecimento a respeito dos fatores escolares brasileiros (COELHO, 2008; KAUKO *et al.*, 2016) e de pesquisas que analisem

afundo a problematização local das regiões e/ou das escolas (SOUSA; OLIVEIRA, 2010; KAUKO *et al.*, 2016). É preciso considerar a extensão da diversidade da população brasileira, uma vez que se priorizou a ideia de a construção de um roteiro global apoiado no fato do esquema do exame standardizado brasileiro ser orientado para o *Programme for International Student Assessment* (Pisa) (KAUKO *et al.*, 2016).

Dessarte, os artigos dessa categoria creem na avaliação como além da execução de um projeto, podendo atender a várias demandas, dos gestores educacionais aos diversos setores da sociedade, como colocam Bonamino e Franco (1999). Por isso, análises mais complexas são necessárias, para dar continuidade às investigações de melhoria na qualidade da educação, como concluem Kauko *et al.* (2016).

Entre as aproximações dos artigos analisados, nota-se que apesar de apresentarem os documentos oficiais da legislação brasileira, Bonamino e Franco (1999) e Kauko *et al.* (2016) não consideram, em grande parte, os estudos dos pesquisadores em avaliação educacional do período em que se deu a implementação do Saeb. Mostrou-se um considerável volume trabalhos produzidas e/ou publicadas pelas fundações participantes, em especial pela FCC.

1.3.2 O Saeb como política de responsabilização (*accountability*)

Nessa categoria são agrupados três artigos que discutem o Saeb como política de responsabilização (*accountability*), servindo para monitorar e regular os sistemas de ensino por meio do desempenho escolar obtido.

Isso posto, apresentam-se duas concepções acerca desta categoria. Primeiramente, os artigos de Franco e Menezes-Filho (2012) e Dantas, Massoni e Santos (2017) discutem o impacto dos exames standardizados na constituição de políticas educacionais e como elas direcionam a política de avaliação.

Depois, o artigo de Dickel (2016) avalia esse impacto por meio de um recorte, no qual analisa a relação entre o Pnaic e a ANA, um dos componentes do Saeb.

A avaliação, de acordo com Dantas, Massoni e Santos (2017), tem se pautado na avaliação formativa, contínua e acumulativa, bem como na valorização dos aspectos qualitativos, conforme disposto nos documentos oficiais. No entanto, segundo os autores, o que se perpetua na prática é a cultura de uma avaliação meramente quantitativa. Por esse motivo,

comprova-se um grau de distanciamento entre o que a legislação estabelece e o que de fato venha ser praticado em sala de aula (DANTAS; MASSONI; SANTOS, 2017).

Além disso, de acordo com os autores, a avaliação na literatura segue duas posições contraditórias. Uma defende a avaliação para orientar políticas públicas e apoiar as práticas de ensino e de aprendizagem escolar, buscando a qualidade da educação para amparar o desenvolvimento curricular. A outra, por sua vez, denuncia a avaliação por monitorar e regular os sistemas de ensino por meio da política de responsabilização, o que consequentemente interfere nas práticas escolares (DANTAS; MASSONI; SANTOS, 2017).

Nesse caso, Dickel (2016) examina as possibilidades da avaliação para alcançar a melhoria da qualidade do ensino. Segundo a autora, a ANA apresenta-se como uma ferramenta à disposição do professor, com o intuito de monitorar a aprendizagem das crianças e auxiliá-lo em seu trabalho.

Em contrapartida, sua institucionalização também contribui para maior regulação pedagógica (DICKEL, 2016). A autora, portanto, critica algumas ações das avaliações do Saeb. Entre elas, pelo próprio Pnaic, priorizam-se as aprendizagens acumuladas num determinado período, em que são acompanhadas inclusive de um projeto curricular explícito, ditado pela avaliação. Reforça-se, portanto, a responsabilização do professor e da escola pelo resultado obtido, uma vez que estes são os encarregados do ensino dos conteúdos avaliados.

Diante disso, assim como Dantas, Massoni e Santos (2017), Franco e Menezes-Filho (2012) discutem o Saeb como política de responsabilização, no qual se referem à preocupação do Brasil de melhorar a qualidade do ensino por meio da atribuição de metas educacionais. Estas metas são propostas em indicadores baseados no desempenho dos alunos nos exames estandardizados, como as provas do Saeb e os resultados do Pisa (FRANCO; MENEZES-FILHO, 2012).

Desse modo, Franco e Menezes-Filho (2012) também apresentam o Ideb como um propulsor para as políticas de responsabilização, uma vez que, desde 2007, pôde servir para que governos conseguissem monitorar as redes de ensino e, a partir disso, exigir ações em busca de um avanço esperado no indicador (FRANCO; MENEZES-FILHO, 2012).

Destaca-se que por meio do desempenho médio dos alunos nos testes, divulgados pelo Ideb, as escolas estão suscetíveis a uma dinâmica de ranking, e, por isso, ocasiona-se o detrimento

de umas em relação às outras, similarmente ao que acontece com empresas na concorrência dos mercados (FRANCO; MENEZES-FILHO, 2012).

Nesse caso, Franco e Menezes-Filho (2012) discordam de tal dinâmica e apresentam diversos fatores que podem interferir no desempenho escolar. Além disso, os autores defendem que a elaboração de políticas educacionais não deve ser baseada exclusivamente nos resultados, sendo necessário realizar uma análise qualitativa acerca dos fatores que envolvam a comunidade escolar.

Posto isto, ainda de acordo com esses autores, o ideal seria uma avaliação pensada para acompanhar o progresso do aluno ao longo de vários anos.¹² No entanto, verificou-se que um dos desafios de tal situação é que grande parte dos países em desenvolvimento ainda não tem um sistema de avaliação da qualidade da educação que consiga tal êxito.

Em conclusão, constata-se que, entre os artigos desta categoria, existe fortemente a crítica quanto ao uso que se tem feito dos resultados das avaliações do Saeb, sobretudo no que diz respeito às suas ações como política de responsabilização (*accountability*) (FRANCO; MENEZES-FILHO, 2012; DANTAS; MASSONI; SANTOS, 2017). No entanto, Franco e Menezes-Filho (2012) não apresentam as primeiras iniciativas dessa responsabilização na educação brasileira, nem tampouco seus respectivos contextos.

Evidencia-se, portanto, a necessidade de pesquisas que considerem as resoluções do Saeb e/ou da legislação brasileira que ocasionaram em tal responsabilização, por exemplo, o Plano de Desenvolvimento da Educação (PDE), o Plano Nacional de Educação (PNE) e a própria ANA, já mencionada.

Além disso, os artigos de Dickel (2016) e Dantas, Massoni e Santos (2017) criticam algumas das ações do Saeb, como o ranqueamento grosseiro das escolas, mas não apresentam possíveis alternativas de mudanças, nem tampouco para aperfeiçoamento do sistema.

Entre as poucas soluções apresentadas pelos artigos, justifica-se o uso de análises qualitativas acerca dos fatores da escola em conjunto com a quantitativa, porém não se apresenta o modo como seriam feitas e nem quais seriam os responsáveis por sua realização (FRANCO; MENEZES-FILHO, 2012; DANTAS; MASSONI; SANTOS, 2017).

¹²De acordo com Brasil (2020), o Novo Saeb pretende avaliar todos os anos escolares a partir do 2º ano do ensino fundamental das escolas públicas e privadas, permitindo-se acompanhar o progresso do aluno no decorrer dos alunos da vida escolar.

Outra discussão apresentada pelos artigos refere-se à imposição de um currículo por meio da avaliação (DICKEL, 2016; DANTAS; MASSONI; SANTOS, 2017). Apesar da relevância desse fato, os artigos não consideram o movimento da adoção das matrizes de referência do Saeb nem a influência da Base Nacional Comum Curricular (BNCC) que, de acordo com Brasil (2020a), será utilizada pelo Saeb para avaliar todos os anos escolares a partir do 2º ano do ensino fundamental.

Em síntese, exterioriza-se a necessidade de pesquisas que busquem abarcar algumas das questões expostas anteriormente, de maneira a considerar tanto o contexto das mudanças como o de seus agentes influenciadores. Desse modo, é necessário realizar uma investigação profunda de como o Saeb se torna um instrumento de monitoramento da qualidade da educação básica.

1.3.3 Os fatores associados aos resultados do Saeb

Nesta categoria encontram-se quatro artigos em que se analisam, por meio dos microdados do Saeb, disponibilizados na plataforma do Inep, os fatores associados aos resultados do Saeb. Dessa forma, os artigos apresentam duas discussões. A princípio, o artigo de Alves (2008) discorre sobre a organização dos sistemas de ensino e como os resultados são associados a eles. Posteriormente, os demais artigos correlacionam os resultados do Saeb com indicadores de contexto, no qual se destaca o NSE (CARNOY *et al.* 2015; CASTRO; TAVARES JÚNIOR, 2016; FERRÃO *et al.*, 2018).

Primeiramente, Alves (2008) discute acerca da organização do ensino em redes: municipal, estadual, federal e privada. De acordo com a autora, grande parte das diferenças de desempenho entre essas redes está relacionada com suas próprias especificidades.

Para as redes municipal e estadual, enfatizadas por Alves (2008), são apontados os serviços educacionais e as condições de acesso como os principais impulsionadores dessas diferenças, que são confirmadas em um estudo de caso que considerou o estado de São Paulo.

Verificou-se que enquanto a rede municipal investiu em aspectos gerais da política social, como uniforme e merenda escolar, a rede estadual investiu em uma agenda estritamente educacional, como o tempo de permanência dos alunos nas escolas (ALVES, 2008). Com isso, a autora percebeu um rendimento menor nas notas do Saeb do município em relação ao estado.

Além disso, a autora defende que o fato de o estudante ter cursado a pré-escola tem influência significativa em seu desempenho escolar. Entretanto, critica a situação do Brasil, em que se garante o acesso à educação infantil sem se preocupar com a qualidade do ensino ofertado.

Ademais, ainda segundo Alves (2008), outro aspecto determinante para um bom desempenho escolar no Saeb está relacionado à formação dos professores. Este, por sua vez, é analisado como uma particularidade dos países em desenvolvimento, como o Brasil, em que nem sempre os docentes têm formação adequada às disciplinas que lecionam.

Por fim, conclui-se que apesar da modesta contribuição dessas políticas educacionais, nas redes municipal e estadual, para o aumento do desempenho dos alunos, ao que tudo indica, essas medidas ainda são insuficientes para oferecer a promoção da magnitude que as escolas brasileiras necessitam (COELHO, 2008).

Outra variável determinante para o desempenho escolar, de acordo com Carnoy *et al.* (2015) e Ferrão *et al.* (2018), é o NSE dos alunos. Além disso, a relevância que a família dá ao estudo e seu grau de escolaridade também influenciam na permanência dos alunos na escola e nos resultados que eles obtêm.

Não obstante, conforme os resultados obtidos por Castro e Tavares Júnior (2016), os ganhos no Pisa e no Saeb têm sido, em geral, maiores para os estudantes com menos recursos. Similarmente, os resultados de Carnoy *et al.* (2015) sugerem que esses estudantes não estão necessariamente atrás dos que possuem maiores recursos. Isto significa que, apesar da influência do background¹³ familiar, ele não é exclusivo para determinar a realização escolar do aluno (CASTRO; TAVARES JÚNIOR, 2016).

Por isso, os problemas nos sistemas de ensino deveriam ser resolvidos desde a pré-escola (CASTRO; TAVARES JÚNIOR, 2016; FERRÃO *et al.*, 2018), de modo a auxiliar os alunos ao longo de sua trajetória escolar, evitando o comprometimento de sua escolarização futura (FERRÃO *et al.*, 2018).

Conclui-se que, como a educação é uma possibilidade de ascensão social, em decorrência do valor atribuído à educação escolarizada, o entendimento do background familiar é crucial para

¹³O *background* familiar, de acordo com Castro e Tavares Júnior (2016) é o conjunto de características socioeconômicas mensuráveis (o nível de instrução materno, o nível de instrução dos progenitores, artefatos ou livros no domicílio, entre outros) que proporcionam melhor aproveitamento dos recursos disponíveis.

a implementação de projetos e promoção de políticas e/ou ações educacionais destinadas à melhoria da qualidade da educação e à equidade no sistema de ensino (FERRÃO *et al.*, 2018).

Constata-se, portanto, mais uma vez, a necessidade de um sistema de avaliação que monitore os alunos ao longo dos anos, em especial aqueles com menos recursos (FERRÃO *et al.*, 2018). Dessa forma, o sistema forneceria subsídios às decisões políticas para um planejamento com medidas de caráter focalizado, pois conhecer melhor o perfil desses estudantes também ajudaria as escolas a adaptarem melhor esses alunos (CASTRO; TAVARES JÚNIOR, 2016; FERRÃO *et al.*, 2018).

Ademais, o sistema de avaliação permitiria “[...] a avaliação do impacto de políticas, programas e medidas na atenuação das desigualdades sociais e educacionais” (FERRÃO *et al.*, 2018, p. 290). Contudo, é preciso analisar os dados com cautela e atentar-se à generalização exacerbada dos resultados, pois o próprio Pisa apresentou uma superestimava dos ganhos (CARNOY *et al.*, 2015).

Percebe-se, também, que os artigos dessa categoria se preocupam em validar, mediante pesquisas empíricas, concepções relevantes para o desempenho escolar dos estudantes, como a influência do NSE e do background familiar (CARNOY *et al.*, 2015; CASTRO; TAVARES JÚNIOR, 2016; FERRÃO *et al.*, 2018). Em contrapartida, não se atentam para a descoberta de novos fatores que possam interferir no rendimento dos estudantes no Saeb, como a formação e a valorização docente, a complexidade da gestão e a infraestrutura das unidades escolares. Acredita-se que ponderar esses fatores contextuais é essencial para compreender as múltiplas faces do sistema educacional brasileiro e suas implicações para o processo de ensino e aprendizagem escolar.

Além disso, os artigos dessa categoria utilizam dados de edições específicas, sem considerar as modificações sofridas pelo Saeb, inclusive de uma edição para outra (ALVES, 2008; CARNOY *et al.*, 2015; CASTRO; TAVARES JÚNIOR, 2016; FERRÃO *et al.*, 2018). Vale ressaltar que a adoção pela análise de um período específico, sem associá-la à implantação de políticas públicas educacionais, pode se tornar vaga.

Como exemplo, pode-se citar o artigo de Alves (2008), que considera os dados do Saeb das edições de 1999, 2001 e 2003, mas não trata da adoção do programa de progressão continuada, apontada como um dos fatores que corroboraram não só para a ampliação de

matrículas, mas que também produziu um efeito significativo nas taxas de abandono e de aprovação escolar (BRASIL, 1999; MENEZES-FILHO *et al.*, 2009).

Em suma, os artigos desta categoria indicam que apesar de o Saeb ser um sistema de avaliação importante, pouco se produz em relação aos efeitos diretos que ele oferece à sua população (ALVES, 2008; CARNOY *et al.*, 2015; CASTRO; TAVARES JÚNIOR, 2016; FERRÃO *et al.*, 2018). Evidencia-se, portanto, a necessidade de maiores investimentos em pesquisas que também abarquem o processo de tomada de decisões dos governos a partir dos resultados do Saeb, e não somente o juízo de valores, como tem se estabelecido.

1.4 CONSIDERAÇÕES PARCIAIS

Neste capítulo, apresentou-se uma revisão sistemática da literatura sobre o Saeb. Observou-se que apesar de sua criação datar do ano de 1990, a publicação de artigos com temáticas diretamente relacionadas ao sistema, no âmbito da educação, acontece quase uma década depois (BONAMINO; FRANCO, 1999), ganhando notoriedade a partir da criação da Prova Brasil, que inaugurou o caráter censitário nas avaliações estandardizadas no cenário nacional.

Foram selecionados 11 artigos, classificados em três categorias de acordo com a análise temática e que abordam: a origem do Saeb mediante as reformas educacionais da década de 1990; o Saeb enquanto política de responsabilização (*accountability*); e os fatores associados aos resultados do Saeb.

Percebeu-se que o seu surgimento é abordado de forma generalizada pelos artigos e que vários porquês não são explicados ao longo dos textos, por exemplo, quais são os sujeitos que contribuíram para a elaboração/desenho desse sistema de avaliação? Como se deram as primeiras experiências de avaliação no âmbito nacional e quais foram suas justificativas? O que elas têm em comum com o Saeb? Qual o cenário de sua efetivação? E quais metodologias foram/são utilizadas para a avaliação e o levantamento de dados?

Apesar de destacarem as primeiras experiências de avaliação que culminaram no surgimento do Saeb, como o Edurural e o Projeto Nordeste, os artigos não abordam o que de fato se retirou de proveito dessas avaliações. Assim, aponta-se para a necessidade de estudos que possam abarcar a similaridade do Saeb com estes projetos de avaliação, de maneira a

compreender suas prioridades, como a escolha dos anos e das disciplinas avaliadas, a preferência pelas metodologias utilizadas, a seleção das competências e habilidades e as instituições responsáveis por sua aplicação.

Observou-se também que essas discussões poderiam abarcar os interesses e as inclinações políticas do Brasil, durante a implantação do sistema, de maneira mais centralizada. As ações da FCC, pioneira na gestão das avaliações no contexto nacional, por exemplo, são pouco exploradas nos estudos que se referem ao surgimento do Saeb, embora haja um amplo acervo de documentos e relatórios do período, produzidos por pesquisadores da instituição.

Verifica-se, ainda, a presença de perguntas em aberto relacionadas à atuação internacional no contexto de criação do sistema de avaliação brasileiro: quem financiou sua implantação? Quais os possíveis interesses por trás disso? Quais foram as condicionantes e os desdobramentos dessa ação? Embora alguns artigos debatam a influência do Banco Mundial na consolidação do Saeb, não foi encontrado um estudo capaz de responder a todos esses questionamentos, que, por sua vez, constituem um importante subsídio para a compreensão da origem do sistema.

Ademais, ressalta-se a necessidade de estudos que considerem o processo sócio-histórico e epistemológico que culminou, para além da criação, na consolidação do Saeb, abarcando as principais modificações sofridas durante os seus primeiros trinta anos de existência.

Nessa vertente, deve-se considerar as diversas mudanças na estruturação do sistema, como a criação das matrizes de referência, que proporcionou uma discussão inicial acerca da implementação de um currículo comum no âmbito nacional e a adoção da Teoria de Resposta ao Item (TRI) para a construção dos testes, que proporcionou análises longitudinais a partir da comparabilidade dos resultados no decorrer dos anos. Essas análises, que consideram o desenvolvimento da educação ao longo dos anos, também precisam ser mais bem exploradas nos estudos, por oferecerem subsídios para o acompanhamento da efetividade das políticas públicas educacionais adotadas.

Quanto aos fatores externos que oferecem impacto no desempenho escolar dos estudantes, algumas das produções analisadas destacaram o NSE e o background familiar (CARNOY *et al.*, 2015; FERRÃO *et al.*, 2018) e a formação docente (ALVES, 2008). No entanto, percebe-se a ausência de estudos que tratem de outros fatores contextuais de ampla discussão na literatura especializada, como a infraestrutura das escolas (SOARES NETO *et al.*, 2013;

MATOS; RODRIGUES, 2016; SOARES; SOARES; SANTOS, 2020) e a complexidade da gestão, no que se refere ao porte da escola, ao número de estudantes nela matriculados e a quantidade de etapas ofertadas (OLIVEIRA; CARVALHO, 2018; GOBBI *et al.*, 2020). Acredita-se que esses fatores também possam influenciar a aprendizagem dos estudantes.

Como referência para estudos futuros, algumas das produções analisadas (ALVES, 2008; COELHO, 2008; SOUSA; OLIVEIRA, 2010; CARNOY *et al.*, 2015; KAUKO *et al.*, 2016; DANTAS; MASSONI; SANTOS, 2017; FERRÃO *et al.*, 2018) sinalizam para autores que deverão ser consultados, considerando a relevância de seus trabalhos relacionados ao Saeb. Dentre eles, destacam-se: Heraldo M. Vianna; Bernardete A. Gatti; Maria I. Pestana; Ruben Klein; e Julio J. Waiselfisz.

Além disso, revela-se a importância da compreensão do papel das fundações que os autores supracitados fizeram parte, sobretudo no que diz respeito a suas pesquisas em avaliação educacional. Evidencia-se, dessa forma, a necessidade de pesquisas que utilizem como artifícios não só as produções acadêmicas, que se mostraram escassas, mas também documentos, relatórios oficiais e entrevistas com os agentes diretamente relacionados ao processo de consolidação do sistema.

É digno de nota que os trabalhos que se propõem a estudar esta temática, em geral, situam-se entre duas dimensões fundamentais no campo da avaliação: o juízo de valor e a tomada de decisão. Desse modo, percebeu-se que tanto o cenário acadêmico como o da política educacional concentram suas discussões em torno dos juízos de valor que se estabelecem a partir dos resultados do sistema.

Em compensação, como política educacional de avaliação, não se pode abrir mão da tomada de decisão, ação pela qual o Saeb se destina a fazer. Demonstra-se, portanto, a importância de maiores investimentos nessa dimensão, analisando quais políticas foram ou estão sendo estabelecidas, a partir de seus resultados, com o objetivo de qualificar as redes e os sistemas de ensino no Brasil.

Por último, ressalta-se que a nova configuração do Saeb, mediante a publicação da Portaria nº. 458, de 5 maio de 2020 (BRASIL, 2020a), que trata da expansão do sistema para todos os anos escolares, a partir do 2º ano do ensino fundamental, trará novas possibilidades de estudos e permitirá o acompanhamento do aluno em toda sua trajetória escolar. Também a possibilidade do uso dos resultados do sistema para o acesso ao ensino superior apresenta-se

como uma ação que necessita ser investigada, considerando os impactos que pode trazer ao sistema educacional brasileiro.

Espera-se, portanto, que por meio dos apontamentos apresentados neste trabalho, possa haver um maior interesse, por parte de pesquisadores educacionais, na investigação de temáticas que permeiam o Saeb e, a partir disso, auxiliar na construção de uma nova agenda de estudos relacionados à avaliação educacional, que poderá trazer contribuições significativas para a melhoria da qualidade da educação brasileira.

CAPÍTULO 2

EXPERIÊNCIAS INICIAIS DE AVALIAÇÃO EDUCACIONAL QUE SERVIRAM COMO BASE PARA O DESENHO DO SAEB

2.1 INTRODUÇÃO

Desde o início do século XX, têm-se desenvolvido estudos sobre a avaliação da aprendizagem em diversos países do mundo, principalmente nos Estados Unidos e na Inglaterra, berços de Edward Lee Thorndike (1874-1949) e Karl Pearson (1857-1936), respectivamente, pioneiros nas pesquisas da área. As contribuições desses autores, segundo Dalben (2008, p. 9), “[...] acarretaram um grande desenvolvimento de testes padronizados para medir as habilidades e aptidões dos alunos”, resultando no incremento da relevância de testes e medidas no campo educacional.

Outro educador pioneiro na temática foi Ralph W. Tyler (1902 - 1994). A ele se atribui a origem do termo “avaliação educacional”, em 1934, concomitante com o surgimento do movimento “educação por objetivos”, cujo princípio era formular metas e verificar se elas eram atingidas. Para ele, a avaliação,

[...] ofereceria elementos para uma crítica fundamentada da instituição, baseada em dados empíricos, e permitiria, ao mesmo tempo, uma discussão sobre a eficiência da sua atuação. A avaliação, ainda de acordo com o pensamento de Tyler, possibilitaria o aprimoramento dos programas, com a eliminação do inoperante e o desenvolvimento daqueles aspectos que se tivessem revelados positivos (VIANNA, 1998, p. 77).

Esses estudos precursores e seus resultados foram introduzidos no Brasil no âmbito da Escola Nova¹⁴ que, centrada no desenvolvimento de uma educação dita “científica”, fomentou a padronização de medidas para mensurar o grau de conhecimento dos estudantes. Com isso a

¹⁴Movimento pedagógico que surgiu na Inglaterra no século XIX que considerava o aluno em toda a sua singularidade e a escola como instrumento para a edificação da sociedade (ABRÃO, 2006). Suas ideias inspiraram intelectuais promovendo reformas educacionais em todo o mundo. No Brasil, Anísio Teixeira e outros educadores retomaram esse movimento nos anos de 1950 e 1960 (SILVA; FILGUEIRAS; VIEIRA, 2020).

qualidade da educação passou a ser monitorada por meio da análise dessas medidas (DALBEN, 2008).

Dessa forma, desde a década de 1960, têm-se percebido, no Brasil, iniciativas que demonstram a preocupação com a avaliação de desempenho escolar, bem como o desenvolvimento de avaliações estandardizadas que resultaram, em 1990, na criação do Sistema de Avaliação da Educação Básica (Saeb).

Considerando a importância atribuída ao Saeb, que atualmente se configura como o principal instrumento de avaliação diagnóstica da qualidade da educação brasileira, este capítulo foi desenvolvido com o objetivo de compreender as experiências iniciais de avaliação estandardizada que decorreram no país e que serviram como base para a implementação do sistema. Para tanto, considerou-se o período compreendido entre 1960 e 1990, momento em que se deram esses primeiros processos avaliativos. O interesse, assim, é responder à seguinte questão: de que modo essas experiências foram utilizadas para subsidiar a criação de um sistema de avaliação em nível nacional?

Para a exposição, além desta introdução, este capítulo se estrutura em outras seis seções. Na primeira, aborda a influência do Banco Mundial na educação básica brasileira por meio dos seus empréstimos. Na segunda e na terceira, discute acerca do Projeto Nordeste de Educação Básica I (Edurural) e do Projeto Nordeste de Educação Básica II (Projeto Nordeste), respectivamente, que, embora assumissem um caráter regional, ofereceram subsídios para o aperfeiçoamento das pesquisas posteriores. Na quarta e na quinta seções, abarcam-se dois ensaios em âmbito nacional: a Avaliação do Rendimento Escolar de Alunos de Escolas de 1º Grau da Rede Pública e o Sistema Nacional de Avaliação do Ensino Público de 1º Grau (Saep). Este último é assumido como estudo piloto, constituindo-se como a primeira designação do Saeb (HORTA NETO, 2007; FREITAS, 2013; PESTANA, 2016). Na quinta e última seção, apresentam-se as considerações finais obtidas.

2.2 A INFLUÊNCIA DO BANCO MUNDIAL

Preocupados com o estabelecimento de uma nova ordem internacional no período pós-guerra, os países centrais se reuniram em Bretton Woods para uma conferência. Como consequência, “[...] buscaram criar instituições capazes de conferir maior estabilidade à economia mundial

de forma a impulsionar o crescimento e evitar a emergência de novas crises internacionais” (SOARES, 1996, p.18). Assim, em 1944, criou-se o Banco Mundial.¹⁵

Desde a sua criação, os Estados Unidos da América (EUA) esteve à frente da gestão do referido banco. Afinal, segundo Soares (1996), a influência nas decisões e votações é proporcional à participação no aporte capital. Como esse país sempre foi o maior acionista, tornou-se o membro mais fluente. Por esse motivo, conforme Pereira (2012), a nova ordem monetária era baseada no dólar, uma proposta essencialmente norte-americana.

Inicialmente, seu objetivo era extrair lucros provendo garantias e empréstimos a longo prazo para a reconstrução e o desenvolvimento dos países-membros afetados pela guerra. Porém, apenas três anos após a Conferência de Bretton Woods, esse intuito foi posto de lado “[...] frente à urgência de blindar a Europa contra o contágio comunista” (PEREIRA, 2012, p. 400).

De acordo com Fonseca (1996), até os anos 1960, o projeto de desenvolvimento do Banco Mundial era pautado nas metas de crescimento econômico. Por outro lado, percebeu-se que a pobreza e as desigualdades socioeconômicas geravam tensões políticas e sociais no mundo, principalmente, nos países subdesenvolvidos.

Diante disso, em consoante com a autora, o banco definiu como prioridade atuar no que poderia representar uma ameaça à economia dos países desenvolvidos. Ou seja, no crescimento descontrolado da pobreza nos países periféricos. Ao final da década, sua definição de desenvolvimento foi alterada.

Em teoria, para Pereira (2012), o desenvolvimento era visto como um fator de “amortecimento” e “estabilização” das contradições entre grupos e classes sociais. Mas na realidade, era uma maneira do banco de conter o comunismo em escala internacional. Afinal, ele nunca apoiou ou financiou qualquer iniciativa voltada para a redistribuição de riqueza.

A partir daí, ainda segundo Pereira (2012), o Banco Mundial assume retórica reformista, que se efetivou somente nas décadas de 1960 e 1970, pois havia certo receio em efetuar empréstimos para países não rentáveis. Até 1957, o autor destaca que 52,7% do seu financiamento total eram direcionados aos países capitalistas mais industrializados.

¹⁵Constituem o Grupo Banco Mundial: o Banco Internacional para Reconstrução e Desenvolvimento (BIRD), a Associação Internacional de Desenvolvimento (IDA ou AID), Cooperação Financeira Internacional (IFC ou CFI), Agência de Garantia de Investimentos Multilaterais (MIGA ou AGIM) e o Centro Internacional para Arbitragem de Disputas sobre Investimentos (ICSID ou CIADI). Fonte: <http://web.worldbank.org/>.

Dessarte, o banco passou a efetuar empréstimos a longo prazo e a prestar assistência técnica e administrativa aos países em desenvolvimento. Em geral, sua assessoria era realizada por meio de projetos de cunho educacional, econômico ou social. A partir dos anos 1970, esses projetos se voltam para o incentivo da produtividade, em especial, no setor agrícola, considerado como um dos fatores para conter a pobreza (FONSECA, 1996).

Apesar da política de crédito do Banco Mundial se autodenominar como “cooperação”, ela é um empréstimo do tipo convencional com encargos pesados (ALTMANN, 2002). Ou seja, qualquer atraso na execução de projetos resulta no aumento significativo desse encargo, além dos juros e dos ajustes cambiais (FONSECA, 1996). Com isso, a crise do endividamento desses países deu margem para uma ampla transformação do papel desempenhado pelo banco.

De um banco de desenvolvimento, indutor de investimentos, o Banco Mundial tornou-se o guardião dos interesses dos grandes credores internacionais, responsável por assegurar o pagamento da dívida externa e por empreender a reestruturação e abertura dessas economias, adequando-as aos novos requisitos do capital globalizado (SOARES, 1996, p. 21).

Em razão dessa situação de crise e vulnerabilidade, conforme discutido por Soares (1996), o Banco Mundial começou a impor uma série de condicionantes para a concessão de novos empréstimos. Sem o seu aval, não haveria fontes de crédito internacional, o que tornou difícil a resistência de governos, eventualmente, insatisfeitos com essa ordem. Dessa forma, ele passou a exercer amplo controle sobre as políticas, desempenhando um papel estratégico no processo de reestruturação neoliberal dos países em desenvolvimento.

O objetivo dos programas de ajuste é assegurar o pagamento da dívida e transformar a estrutura econômica dos países de forma a fazer desaparecer características julgadas indesejáveis e inconvenientes ao novo padrão de desenvolvimento (neoliberal): protecionismo, excesso de regulação, intervencionismo, elevado grau de introversão, entre outras (SOARES, 1996, p. 23).

A ideia central, ainda de acordo com Soares (1996), era de que a maior parte das dificuldades desses países se encontrava neles próprios, sobretudo na rigidez de suas economias. A autora destaca que, para o Banco Mundial, as reformas em suas instituições e políticas passaram a ser mais importantes do que o alívio da dívida dos países.

No que tangencia a América Latina, as propostas estiveram aliadas à execução de programas sociais focalizados na população mais vulnerável economicamente (SOARES, 1996). Desse

modo, a ênfase recai sobre a educação, apontada pelo banco como instrumento de redução da pobreza e crescimento econômico (BANCO MUNDIAL, 1993, 1996).

A partir da década de 1970, surgiram os primeiros financiamentos para programas e projetos educacionais no continente. É importante dizer que, esses convênios foram estabelecidos em um momento crítico para tais países, marcado pela transição entre a ditadura militar e a construção do regime democrático (UCZAK, 2014). A educação era, portanto, um fator essencial para a formação de “capital humano” adequado aos requisitos do novo padrão de acumulação (SOARES, 1996).

No Brasil, o primeiro empréstimo¹⁶ do Banco Mundial ao setor educacional ocorreu no ano de 1971, em apoio ao projeto de “[...] expansão e/ou melhoria de 16 instituições federais para o treinamento de técnicos industriais e agrícolas nos níveis secundário e pós-secundário” (BANCO MUNDIAL, 1974, p. 1). Como resultado, evidenciou-se as desigualdades educacionais do país, sobretudo nas Regiões Norte e Nordeste, que careciam de educação básica. Segundo Fonseca (1996), essa etapa era considerada como pré-requisito para a preparação de mão de obra de nível médio, o que contribuiu para o setor produtivo do país. Isso explica a ênfase dada ao ensino profissionalizante no interior dos projetos.

Diante disso, avaliou-se a implementação de um plano de reforma educacional em oito estados¹⁷ brasileiros. Esse foi o segundo empréstimo do banco para a educação brasileira. De acordo com Banco Mundial (1974), o projeto previa a criação de cursos de nível médio que oferecessem educação geral, bem como uma oportunidade para os estudantes desenvolverem habilidades práticas em outras áreas, como artes industriais, agricultura e economia doméstica.

Nesse período, final dos anos 1970, o Brasil já era conhecido pelos acentuados contrastes socioeconômicos de sua população. Inclusive, conforme o relatório *Educational Performance of the Poor* do Banco Mundial, os Estados da Região Nordeste se encontravam entre os mais pobres do mundo. Os dados brasileiros¹⁸ evidenciaram que a região era educacionalmente desfavorecida se comparada com o restante da nação. Dois terços de sua população rural ou

¹⁶De acordo com o Banco Mundial (1974), este empréstimo bancário é intitulado como 755-BR, sendo equivalente a US\$8,4 milhões.

¹⁷Segundo o Banco Mundial (1974), os estados participantes são Amazonas, Pará, Maranhão, Piauí, Rio Grande do Norte, Paraíba, Alagoas e Sergipe.

¹⁸De acordo com Harbison e Hanusek (1992), utilizaram-se como indicadores: frequentou escola nos últimos 12 meses, saber ler e escrever, nunca frequentou a escola e a média de anos de escolaridade.

eram analfabetos ou possuíam, em média, um ano de escolaridade. Além disso, quando realizadas as comparações de renda, verificou-se que estavam entre os mais vulneráveis economicamente do país.

Como consequência, introduziu-se um programa de intervenção no setor educacional da Região Nordeste, custeado pelo Governo Brasileiro e pelo Banco Mundial, que ficou conhecido como Projeto Nordeste de Educação Básica I.

2.3 O PROJETO NORDESTE DE EDUCAÇÃO BÁSICA I (EDURURAL)

O Programa de Expansão e Melhoria da Educação no Meio Rural do Nordeste (Edurural) ou Projeto Nordeste de Educação Básica I, como é intitulado nos documentos do Banco Mundial, foi planejado entre os anos de 1978 e 1979. Esse foi o primeiro empréstimo do Banco Mundial para os anos iniciais do ensino fundamental brasileiro. De acordo com Gatti, Vianna e Davis (1991), o projeto partiu do compromisso dos técnicos das Secretarias de Educação com o Ministério da Educação (MEC), cuja preocupação era a forma como as equipes locais equacionavam a problemática do ensino rural naquela região.

O Edurural foi instituído pelo Decreto n.º 85.287, de 23 de outubro de 1980 (BRASIL, 1980), para executar o acordo¹⁹ firmado entre o Governo Brasileiro e o Banco Internacional de Reconstrução e Desenvolvimento (Bird), uma das agências que integram o Banco Mundial. Conforme o art. 2º, o Edurural objetivava a “[...] expansão das oportunidades educacionais e a melhoria das condições da educação no meio rural do Nordeste, bem como o fortalecimento do processo de planejamento e administrações educacionais” (BRASIL, 1980, p. 21.250).

Especificamente, os objetivos do Edurural se concentravam em expandir o acesso das crianças da zona rural ao ensino fundamental, para reduzir o “desperdício” de recursos educacionais inerentes à repetição e ao abandono escolar e concentrar recursos suficientes nas áreas desfavorecidas para que conseguisse fazer diferença (HARBISON; HANUSHEK, 1992). Para isso, o programa privilegiava a melhoria do ensino rural por meio de várias ações que deveriam conduzir não só à capacitação pessoal e assistência técnica, como também à produção de material didático (GATTI; VIANNA; DAVIS, 1991).

¹⁹ Refere-se ao contrato MEC/Bird n.º 1.867/BR/1980 celebrado em 18 de julho de 1980.

Dentre os investimentos do Banco Mundial, em concordância com Harbison e Hanushek (1992), destacam-se: a construção e reforma de escolas; o treinamento de professores; o desenvolvimento de currículos especialmente adaptados ao ambiente rural; o fornecimento de livros didáticos e outros materiais de aprendizagem; e o fortalecimento do aparato administrativo da escola municipal por meio de uma nova estrutura institucional.

Além disso, os autores também destacam a injeção de insumos provenientes de outros projetos na região, como: merenda escolar; bonificação salarial e/ou técnicas de aperfeiçoamento ao professor; e o fornecimento de móveis e reabilitação do espaço físico e de livros. É digno de nota que esses insumos variavam de um projeto para outro.

Conforme explicam Gatti, Vianna e Davis (1991), o Banco Mundial arcou com cerca de um terço²⁰ dos custos do projeto, o que, conseqüentemente, resultou em reformulações e exigências no que dizia respeito à sua execução. Ressalta-se, portanto, que, dentre essas, exigiu-se uma avaliação externa do projeto que fosse capaz de “[...] fornecer subsídios para a definição dos programas de educação rural” (GATTI; VIANNA; DAVIS, 1991, p. 9).

Para sua execução, em conformidade com o art. 8º do referido decreto, o MEC foi autorizado a estabelecer as normas complementares necessárias à implantação e desenvolvimento do programa. Em relação aos estudos e projetos de avaliação do Edurural,²¹ enquanto a Fundação Cearense de Pesquisa e Cultura da Universidade Federal do Ceará (UFCE) se encarregou dos aspectos institucionais, a FCC se concentrou naqueles referentes ao rendimento escolar. Dessa forma, segundo Gatti (2013), a avaliação acompanhou toda a implementação e desenvolvimento do programa sob vários aspectos.

A avaliação do Projeto Edurural, considerada um marco na história da avaliação de políticas e programas educacionais (GATTI, 2009), foi um estudo desenvolvido por meio da coleta de dados em 1981, 1983 e 1985. Nesses mesmos anos, de acordo com Gatti (2013), também foram apuradas informações sobre a população avaliada. Por exemplo, características das escolas, a forma de gerenciamento, o corpo docente, as famílias etc.

Ceará, Pernambuco e Piauí compuseram uma amostra aleatória de escolas rurais em que a avaliação foi desenvolvida para que fosse mais tarde aplicada nos demais Estados (GATTI,

²⁰De acordo com Harbison e Hanushek (1992), os custos totais foram de US\$92 milhões e US\$32 milhões foram concedidos graças ao empréstimo do Banco Mundial.

²¹Nessa época, o Edurural passa a ser reconhecido como um dos programas de mais complexa execução realizados até então na América Latina (BRASIL, 1982).

1994; VIANNA, 1995; GATTI, 2013). A escolha desses Estados, no caso de Piauí e Pernambuco, deu-se por representarem melhor as características da região; o Ceará, por ser o local da sede da equipe central da pesquisa (GATTI; VIANNA; DAVIS, 1991).

Dessarte, os resultados do ano de 1981 serviram como base para verificar se haveria ou não avanços na aprendizagem escolar dos estudantes em razão dos investimentos na região que ocorreriam nos anos posteriores (GATTI; VIANNA; DAVIS, 1991; HARBISON; HANUSHEK, 1992). Assim, como uma das dimensões da qualidade do ensino, tomou-se o desempenho dos alunos, aferido por testes de Língua Portuguesa e Matemática nas 2ª e 4ª séries do ensino fundamental (GATTI; VIANNA; DAVIS, 1991; HARBISON; HANUSHEK, 1992; GATTI, 1994; VIANNA, 1995).

Paralelamente a tal sistemática, procurou-se também compreender a forma de atuação da escola em seu contexto, respeitando os limites das relações socioeconômico-culturais locais (GATTI; VIANNA; DAVIS, 1991). Dessa forma, optou-se pela elaboração de seis estudos de caso, a fim de verificar as variáveis conjunturais atuantes no processo de ensino-aprendizagem da população avaliada.

O delineamento da amostra para os estudos de caso, de acordo com Gatti, Vianna e Davis (1991), foi efetuado em duas etapas. Na primeira, determinou-se o tamanho desejável da amostra total, feita por Estado e por modalidade de programas,²² chegando ao número de 603 escolas, que correspondia a 6% do universo total. Na segunda, foram selecionadas as escolas dentro de cada Estado. Algumas seriam atendidas especificamente pelo Edurural e outras por outros programas de maneira aleatória.

Como o processo avaliativo visava a oferecer, prioritariamente, subsídios às Secretarias Estaduais, utilizaram-se as seguintes fontes para seleção dos conteúdos:

[...] i) os planos construídos pelas Secretarias e o currículo orientador dos programas desenvolvidos nas escolas das zonas rurais dos Estados da amostra, quando estes existiam; ii) técnicos e pessoal envolvido com os órgãos municipais de ensino; iii) professoras trabalhando na zona rural; e iv) diários de classe, livros adotados nas áreas de Português e Matemática, planos de aula e cadernos dos alunos (GATTI; VIANNA; DAVIS, 1991, p. 11).

²²De acordo com Gatti, Vianna e Davis (1991) e Harbison e Hanushek (1992), seriam considerados os beneficiados pelos recursos advindos do Edurural e também aqueles que utilizavam insumos provenientes de outros programas.

Tentava-se, portanto, construir um conjunto de provas adequadas àquela realidade, buscando, assim, maior validade para os dados de rendimento escolar (GATTI, 2013). Por esse motivo, conforme Gatti (1994), os testes foram baseados em um levantamento das atividades curriculares desenvolvidas pelas escolas da região analisada, cujos itens das provas tencionavam avaliar o quão distante as crianças estavam dos conhecimentos tomados como básicos nas duas séries avaliadas. Esperava-se, portanto, que os alunos alcançassem 100% dos acertos ou valores próximos a isso (GATTI; VIANNA; DAVIS, 1991; GATTI, 1994), porém verificou-se que os resultados foram aquém do mínimo esperado.

Os resultados dos seis estudos de caso sobre a atuação das escolas participantes, desenvolvidos pela FCC, revelaram que os alunos apresentavam uma aprendizagem básica bastante prejudicada e segmentada (GATTI, 1994; VIANNA, 1995). O ensino oferecido aos menos favorecidos socioeconomicamente, aparentemente, era mais precário do que se imaginava ou do que as pesquisas podiam demonstrar (GATTI; VIANNA; DAVIS, 1991).

De acordo com Harbison e Hanushek (1992), um fator preocupante era que as escolas, quando disponíveis, não possuíam o necessário para atender às necessidades básicas humanas, como disponibilidade de água e saneamento básico. Ou seja, era comum não apresentarem os atributos mínimos para alcançar um contexto mais próspero.

Durante os anos de atuação do Projeto Edurural, verificou-se que não houve melhorias no desempenho dos alunos, pelo menos não detectável estatisticamente, nem mesmo em sua última avaliação, em 1985, “[...] apesar dos vultosos investimentos em construções, reformas, treinamentos, produção de material didático, etc.” (GATTI, 1994, p. 68).

Em contrapartida, o Edurural possibilitou identificar algumas razões que explicavam o baixo rendimento dos alunos. Dentre elas, podem-se destacar: rotatividade de professores; instabilidade na distribuição de livros, material e merenda escolar; baixos salários; condições precárias das escolas; e as frequências irregulares dos alunos (GATTI; VIANNA; DAVIS, 1991; GATTI, 1994; VIANNA, 1995).

Em conclusão, evidenciou-se a precariedade da qualidade do ensino nas escolas rurais do Nordeste por meio da avaliação estandardizada, considerada uma das características mais importantes do projeto (HARBISON; HANUSHEK, 1992; BANCO MUNDIAL, 1993). Por esse motivo, Harbison e Hanushek (1992) discutem sobre a importância de que a avaliação

adequada seja incorporada desde o início aos programas, para verificar se realmente aconteceram mudanças.

Por fim, o Edurural evidenciou a relevância de estudos sobre as informações do nível socioeconômico (NSE) dos estudantes, coletadas durante a atuação do projeto, que revelaram que os antecedentes familiares parecem ter efeitos diretos no desempenho dos alunos, por exemplo, o nível de escolaridade da mãe. Conseqüentemente, o projeto “[...] representou um exemplo do que se poderia fazer com estudos dessa natureza na direção de propugnar por uma escola mais condizente com as necessidades das populações menos favorecidas socialmente” (GATTI, 2009, p. 10).

No final, discutiu-se, portanto, a necessidade de implantação de um projeto que continuasse abarcando a Região Nordeste, o que influenciou as discussões do Projeto Nordeste de Educação Básica II.

2.4 PROJETO NORDESTE DE EDUCAÇÃO BÁSICA II (PROJETO NORDESTE)

O Projeto Nordeste de Educação Básica II ou Projeto Nordeste, como ficou conhecido, segundo Pereira (2007), foi pensado durante a década de 1980, período em que o regime militar chegava ao fim. O projeto é formulado em meio à crise política, social e econômica enfrentada por todo o país, em especial na Região Nordeste, como revelou o Projeto Edurural (1980-1987).

Como discutido, em razão da precariedade do ensino ofertado pelas escolas rurais da Região Nordeste, passou-se a discutir a necessidade de um plano para o desenvolvimento da região, estabelecendo políticas que atendessem às suas particularidades prioritárias (PEREIRA, 2007). Nesse período, são aprovadas as diretrizes do Primeiro Plano de Desenvolvimento do Nordeste na Nova República, por meio da Lei n.º. 7.499, de 25 de junho de 1986 (BRASIL, 1986b). Um dos objetivos gerais do plano consistia no atendimento das necessidades básicas da população regional por meio de estratégias, como a ampliação e a melhoria dos serviços sociais e a democratização e descentralização do aparelho político-institucional (BRASIL, 1986b).

Tangenciando a educação, destaca-se que a adoção de uma política de educação destinada à ampliação dos direitos sociais “[...] tem por finalidade elevar as aspirações da população por melhores condições de vida, fortalecer a base democrática do país e preparar o indivíduo para o exercício da cidadania” (BRASIL, 1986b, p. 22). Assim, as políticas e linhas de estratégias estabelecidas foram: a universalização da educação básica, propiciando acesso ao ensino de 1º grau; e a expansão e melhoria da qualidade do ensino nos 1º e 2º graus.

Inicialmente, as negociações relacionadas com a educação eram coordenadas pela Secretaria Executiva de Superintendência do Desenvolvimento do Nordeste (Sudene),²³ em razão da sua importância para a região. Desse modo, o Decreto n.º. 92.435, de 3 de março de 1986 (BRASIL, 1986a), dispôs sobre sua estrutura básica, tendo em vista a execução do Plano do Desenvolvimento do Nordeste.

Desde 1984, o MEC, por sua vez, adentrou-se nas discussões educacionais do projeto, que objetivava a criação de um calendário para a formulação das diretrizes que iriam nortear as propostas para cada Estado, o que, conseqüentemente, criou certo embate com a Sudene (PEREIRA, 2007). Insatisfeito por ocupar uma posição secundária na definição das políticas educacionais, o MEC, aos poucos, vai conquistando espaços na coordenação dos trabalhos, principalmente por ser o responsável pela promoção de discussões sobre o setor educacional (CABRAL NETO, 1997).

Por outro lado, conforme explica Cabral Neto (1997), os secretários estaduais de Educação se apresentaram contrários às diretrizes apresentadas pelo MEC e pelo Banco Mundial, visto que estas os excluía das operações. Desse modo, evidenciou-se que o foco, na verdade, não passava de uma disputa política, uma competição pelo controle das diretrizes educacionais, e não uma mudança significativa na qualidade da educação na região, como o Projeto Nordeste se propunha fazer (PEREIRA, 2007). Diante de tal embate, o Banco Mundial recuou e só retomou a discussão do projeto na década de 1990, sob a responsabilidade do MEC (BANCO MUNDIAL, 1993).

Baseando-se na experiência do Edurural, de acordo com o Banco Mundial (1993), o projeto também preparava uma educação básica do Nordeste, que não só seguia o mesmo desenho, mas também poderia ser desenvolvida com maior amplitude em todos os Estados da região. A

²³A Sudene foi criada pela Lei n.º. 3.692, de 15 de dezembro de 1959 (BRASIL, 1959), para fomentar o desenvolvimento na região, sobretudo devido às secas enfrentadas.

estratégia de assistência do Banco Mundial ao Brasil, basicamente, consistia em “[...] apoiar as políticas e os investimentos que promovam o crescimento econômico e o desenvolvimento social e maior estabilidade macroeconômica” (BANCO MUNDIAL, 1993, p. 2). Assim, o projeto foi desmembrado em duas partes: nacional, perante a responsabilidade do MEC; e estadual, com a participação dos nove Estados nordestinos.

Enfatizando o aumento da eficiência dos gastos públicos, o foco das estratégias se concentrou no investimento da educação primária, já que, na época, o Brasil se encontrava distante da universalização desse ensino e porque esta possui, de acordo com as diretrizes do Banco Mundial, melhor “retorno social” e maior potencial para aumentar a equidade de oportunidades (BANCO MUNDIAL, 1993, 1996).

Dentre os objetivos assinalados, em nível estadual, no qual o Banco Mundial financiava 97% dos custos do projeto, destacam-se:

[...] i) simplificar as burocracias estaduais de educação e ajudá-los a alcançar uma alocação mais eficiente dos gastos; ii) aumentar o acesso de estudantes primários estaduais e municipais a livros didáticos e outros materiais instrucionais; iii) melhorar as habilidades dos professores estaduais e municipais, a eficácia da sala de aula e as habilidades de gerenciamento dos diretores das escolas; iv) reabilitar e ampliar escolas primárias estaduais e municipais; e v) estimular novas promessas que aproximam-se do ensino primário nos níveis estadual e municipal por meio de um Fundo de Inovações (BANCO MUNDIAL, 1993, p. 2).

Quanto ao nível federal, no qual o Banco Mundial custeava 3% dos custos, sublinham-se o fortalecimento e a capacidade do MEC para empenhar as principais funções de supervisão e suporte, como: testes padronizados para estudantes; planejamento, política e estabelecimento de normas; e transferências financeiras para igualar os gastos com educação nos Estados (BANCO MUNDIAL, 1993). Logo, fez-se encargo do MEC conduzir uma avaliação do progresso da educação de cada Estado.

Com a visita da equipe do Banco Mundial ao Brasil, em 1988, para discussão do projeto, o MEC precisou apresentar uma proposta de avaliação. Para isso, reuniu-se uma equipe da Secretaria Nacional de Educação Básica (Seneb) com o apoio do Instituto Interamericano de Cooperação Agrícola (Iica). O Iica é uma organização internacional que oferece suporte aos países-membros para alcançar o desenvolvimento agrícola e o bem-estar rural. Além da sua aproximação com a região de abrangência do projeto, para o Banco Mundial, o setor agrícola era considerado um dos fatores para conter o crescimento da pobreza (FONSECA, 1996). Isso explica a participação do Iica no projeto.

Nessa época, o professor Júlio Jacobo Waiselfisz ocupava um cargo no Iica que tinha convênio com o MEC. Por esse motivo, o Projeto Nordeste contou com a cooperação e consultoria do professor, que desenvolveu a metodologia de avaliação (PESTANA, 1992; WAISELFISZ; HORTA NETO, 2016).

A proposta de avaliação, de acordo com a entrevista do professor Waiselfisz, concedida ao renomado pesquisador João Luiz Horta Neto, diferentemente do projeto Edurural, não utilizava o desenho experimental, nem quase experimental e nem pré-teste ou pós-teste (WAISELFISZ; HORTA NETO, 2016). Considerava o planejamento de maneira integral, envolvendo diagnóstico, execução e avaliação, ações que, segundo Waiselfisz, não poderiam existir separadas.

Nessa entrevista, Waiselfisz afirma que não acreditava na eficiência da montagem de projetos, método utilizado pelo Banco Mundial, cujas análises consideravam as escolas beneficiadas e o grupo de escolas que não recebia tais benefícios. Para ele, a intervenção era necessária no sistema educacional como um todo, com um propósito maior de avaliar a educação (WAISELFISZ; HORTA NETO, 2016). Desse modo, a proposta apresentada era tão abrangente que, segundo Pestana (1992), que na época era gerente de avaliação do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) – autarquia federal do MEC – a Seneb decidiu aplicá-la em todo o Brasil. Esse também era o desejo de Waiselfisz.

Segundo o entrevistado, foram Júlio Correia²⁴ e Madalena Rodrigues dos Santos²⁵ que colocaram em prática a ideia de discutir a avaliação em nível nacional. Os dois já haviam participado da experiência de planejamento participativo e autonomia das escolas que Waiselfisz tinha implantado em Pernambuco. Ambos não só acreditavam nessas ideias, como também eram contrários aos métodos do Banco Mundial de avaliar somente o projeto (WAISELFISZ; HORTA NETO, 2016).

Em conclusão, a proposta de avaliação ocasionou a montagem do Sistema Nacional de Avaliação do Ensino Público de 1º Grau (Saep) que, por sua vez, se baseou em uma pesquisa intitulada como Avaliação do Rendimento Escolar de Alunos de Escolas de 1º Grau da Rede Pública.

²⁴ Secretário adjunto da Secretaria de Ensino de Primeiro e Segundo Grau (SepS) do MEC que mais tarde se tornou reitor da Universidade Federal de Pernambuco (Ufpe).

²⁵ Chefe de Gabinete da SepS e que anos depois passou a trabalhar no Bird.

2.5 A AVALIAÇÃO DO RENDIMENTO ESCOLAR DE ALUNOS DE ESCOLAS DE 1º GRAU DA REDE PÚBLICA

O final da década de 1980, como discutido, foi marcado por várias discussões a respeito dos problemas enfrentados pelos sistemas educacionais que, segundo Gatti (2009), atingiu seu ápice devido ao debate público sobre as informações reveladas pelos indicadores, que evidenciaram as altas taxas de repetência e de evasão escolar. Dentre as discussões da época, conforme a autora, destaca-se o questionamento sobre a inexistência de dados sobre o rendimento escolar dos alunos em nível de sistema.

Nesse contexto, para Gatti (2013), em uma viagem aos EUA, o ministro interino da Educação, Aluísio Guimarães Sotero, foi apresentado a *National Bureau of Education* (NBE) – a Secretaria Nacional de Educação do Governo Estadunidense. Essa unidade era responsável pela coleta de informações educacionais do país. Por isso, era seu dever gerenciar as avaliações educacionais e divulgar seus resultados para a sociedade. O NBE, conforme a autora, apresentava uma estrutura descentralizada, na qual realizava diversas formas de acompanhamento e monitoramento. Para isso, utilizavam-se de instituições especializadas em avaliação ou de especialistas de universidades.

Partindo dessa realidade americana, o ministro passou a avaliar se seria viável ao MEC atuar de forma semelhante com relação à educação básica no Brasil. Dessa forma, no início do ano de 1987, convocou os pesquisadores da FCC para discutir tal possibilidade (GATTI, 2013).

Dentre os pesquisadores participantes dessa reunião, ainda segundo Gatti (2013), encontravam-se Heraldo Marelim Vianna e Bernadete Gatti, que discutiram a possibilidade da realização de estudos que viabilizassem uma avaliação de desempenho escolar das redes educacionais de ensino brasileiras. Dentre as pautas da reunião, a autora destaca as finalidades e o papel da avaliação, bem como a experiência já acumulada pelo país e a *expertise* quanto à logística de implantação de um sistema de tal natureza.

Algumas sugestões foram levantadas e, apesar da troca de ministro em que Hugo Napoleão do Rego Neto assumiu o MEC, mantiveram-se as preocupações discutidas. Diante disso, “[...] optou-se pelo desenvolvimento de estudos exploratórios para analisar-se a viabilidade prática de um sistema de avaliação de redes escolares, formas de construção das provas e quais referenciais curriculares seriam adotados” (GATTI, 2013, p. 51), uma vez que não existia um

currículo único nacional. A ideia, portanto, consistia em fundamentar uma avaliação diagnóstica.

No final do ano de 1987, por meio do Inep, o MEC deu início a um amplo programa de avaliação do rendimento dos alunos de escolas de 1º grau da rede pública em toda a extensão do território brasileiro (VIANNA, 1989b; GATTI; VIANNA; DAVIS, 1991; GATTI, 1994; VIANNA, 1995; GATTI, 2009). A autora Gatti (2009) salienta que, com a elaboração da Constituição Federal de 1988 (BRASIL, 1988), houve a possibilidade da instituição de uma nova estrutura política, indicando, inclusive, a necessidade de mudança na atuação do MEC em relação ao ensino fundamental e médio. Pretendia-se, portanto, que esse viesse “[...] a ter um papel orientador e um papel avaliador como referência para políticas e avanços na área da educação básica” (GATTI, 2009, p. 11).

O projeto de pesquisa, conforme Vianna (1989b), surgiu também devido à sugestão do então ministro Hugo Napoleão, que almejava a criação de um programa externo de avaliação do 1º grau que subsidiasse as Secretarias Estaduais de Educação com informações a respeito dos problemas de aprendizagem escolar. Por outro lado, indiretamente, pretendia-se também um maior envolvimento das Secretarias com o processo de avaliação em todo o sistema (VIANNA, 1989b; GATTI; VIANNA; DAVIS, 1991; GATTI, 1994; VIANNA, 1995). Assim, com o apoio das Secretarias Estaduais, estabeleceu-se um convênio entre elas e o MEC para a realização da pesquisa.

Para o desenho da avaliação e implementação no território nacional, de acordo com Gatti, Vianna e Davis (1991), o MEC solicitou a colaboração da FCC, que, devido à complexidade, exigiu um desdobramento em três fases distintas: um estudo de caráter experimental envolvendo 10 cidades (capitais de Estados do país); um estudo de 20 cidades; e um estudo de 39 cidades. Ao final, no conjunto, foram avaliados 27.455 alunos de 238 escolas em 69 cidades, localizadas nos vários Estados existentes durante a época de investigação.

Os objetivos da pesquisa, segundo Vianna (1989b), foram, inicialmente, discutidos com administradores escolares e professores do 1º grau, com o intuito de particularizar a área de investigação e estabelecer os procedimentos para o tratamento e análise dos dados. Procurou-se, portanto, uma definição de objetivos que não só atendessem às Secretarias, mas também fornecessem aos especialistas em currículo elementos para reflexão; e aos professores, dados que complementassem aqueles obtidos em suas próprias avaliações (VIANNA, 1989b).

Diante disso, a pesquisa Avaliação do Rendimento Escolar de Alunos de Escolas de 1º Grau da Rede Pública possuía os seguintes objetivos:

[...] a) identificar, na diversidade do quadro educacional brasileiro, os pontos curriculares mais críticos, que deveriam merecer consideração de professores e administradores, a fim de solucionar problemas de aprendizagem que criam dificuldades ao longo do processo instrucional; b) verificar o desempenho dos alunos de 1ª, 3ª, 5ª e 7ª séries em determinados aspectos cognitivos que seriam, na opinião de professores militantes, fundamentais à aprendizagem, no currículo da escola de 1º grau; e c) fornecer subsídios para os professores atuarem no seu dia-a-dia, recuperando os alunos, ao longo do processo, em suas possíveis deficiências (VIANNA, 1989b, p. 6).

À vista disso, por meio do duplo aspecto da avaliação (interno e externo), de acordo com Vianna e Gatti (1988), pretendia-se ir à busca de elementos qualitativos e quantitativos que possibilitassem o processo de tomada de decisões e que, conseqüentemente, levariam ao aperfeiçoamento do processo institucional.

Desse modo, enquanto a avaliação interna estava relacionada com a eficiência do microsistema (escola), a avaliação externa buscava o conhecimento do sistema educacional como um todo, contribuindo, inclusive, para a definição de novas políticas educacionais. Além disso, os autores destacam que:

[...] as diferentes formas de avaliação - interna e externa - não se revestem de um caráter de julgamento, não visam à condenação do aluno, do professor, da escola e nem a culpabilização do sistema, mas sim a identificação de pontos críticos que precisam ser repensados e redefinidos, possibilitando, assim, que novas orientações sejam impressas ao processo de ensino/aprendizagem (VIANNA; GATTI, 1988, p. 6).

Posto isso, conforme explicam Vianna e Gatti (1988), a avaliação deveria ser tratada não só como uma possibilidade de acompanhar a aprendizagem do aluno, mas também permitir um diálogo entre alunos, professores, familiares e administradores escolares, conduzindo à melhoria da qualidade do ensino.

Ademais, a premissa básica da pesquisa, segundo os autores, consistia no entendimento de que a escola era responsável pela transmissão de um conhecimento mínimo cultural comum, envolvendo saberes e habilidades que todos os indivíduos tinham o direito de possuir. Por esse motivo, o conhecimento mínimo deveria ser assegurado, pois “[...] é indispensável ao exercício dos direitos do homem e do cidadão” (VIANNA; GATTI, 1988, p. 6).

A primeira fase da Avaliação do Rendimento Escolar de Alunos de Escolas de 1º Grau da Rede Pública, um estudo envolvendo dez cidades,²⁶ foi uma aplicação experimental para “[...] verificar a viabilidade do processo, como as administrações e as escolas receberiam esse tipo de avaliação, se as provas seriam adequadas, etc.” (GATTI, 2009, p.11). De acordo com Vianna e Gatti (1988), também foram coletadas informações sobre as características das escolas e de seus estudantes. Esses dados passaram a ser úteis na compreensão dos resultados alcançados.

Com a primeira fase aplicada, expandiu-se para a segunda em um estudo envolvendo 20 cidades²⁷ que, segundo Vianna (1989a), objetivava a coleta de novos dados para uma possível intervenção eficiente de professores e administradores na realidade educacional.

Nessa fase, é importante destacar que houve um trabalho de revisão de todos os itens utilizados. Além do mais, passou-se a escolher os locais de aplicação da pesquisa, considerando: a existência de uma rede estruturada com todas as séries envolvidas; a representatividade educacional das cidades no conjunto do Estado; o funcionamento das várias séries em diversos turnos; a posição geoeconômica das cidades com um quadro bastante diversificado; o total de alunos participantes da rede de ensino; e a facilidade de acesso para os aplicadores dos instrumentos (VIANNA, 1989a).

Na terceira fase, em um estudo de 39 cidades,²⁸ que seguiu o formato anterior, várias reflexões se mostraram relevantes sobre esse processo de avaliação. A princípio, desde a fase experimental, já se pôde constatar que a escola estava “falhando” em sua atividade específica que, de acordo com Vianna e Gatti (1988), consistia na transmissão de conhecimentos aos alunos. Aliás, um fator curioso apresentado pelos autores é que normalmente se espera um

²⁶Para Vianna e Gatti (1988), as cidades foram: Belém, São Luís, Natal, Recife, Aracaju, Belo Horizonte; Niterói, Florianópolis, Cuiabá e Brasília.

²⁷De acordo com Vianna (1989a), as cidades foram: Santarém e Castanhal (Pará), Imperatriz e São José do Ribamar (Maranhão), Mossoró e Currais Novos (Rio Grande do Norte), Caruaru e Timbaúba (Pernambuco), Estância e Itabaiana (Sergipe), Volta Redonda e Campos (Rio de Janeiro), Montes Claros e Barbacena (Minas Gerais), Blumenau e Criciúma (Santa Catarina), Rondonópolis e Sinop (Mato Grosso) e Plano Piloto e Taguatinga (Brasília).

²⁸Para Vianna (1989b), as cidades participantes foram: Manaus, Tefé e Parintins (Amazonas), Macapá (Amapá), Rio Branco (Acre), Porto Velho (Rondônia), Teresina, Floriano e Parnaíba (Piauí); Fortaleza, Sobral e Crato (Ceará); João Pessoa, Campina Grande e Patos (Paraíba), Maceió, Palmeira dos Índios e Penedo (Alagoas), Salvador, Jacobina e Vitória da Conquista (Bahia), Vitória, Cachoeiro do Itapemirim e Colatina (Espírito Santo), São Paulo, Ribeirão Preto e Marília (São Paulo), Curitiba, Maringá e Cascavel (Paraná), Porto Alegre, Santa Maria e Pelotas (Rio Grande do Sul), Goiânia, Jataí e Ceres (Goiás) e Campo Grande, Ponta Porã e Corumbá (Mato Grosso do Sul).

melhor desempenho dos alunos que conseguem superar a barreira das avaliações escolares ao avançar as séries. Entretanto, isso não ocorreu, o que revelou ser uma particularidade que merecia investigação.

Segundo Vianna (1989a), a rede pública de ensino de 1º grau mostrou ser frequentada, predominantemente, pelos segmentos menos favorecidos da sociedade, circunstância que também afetou o desempenho dos alunos. Além disso, como afirma o autor, eram evidentes as dificuldades enfrentadas pelos alunos para conseguirem acesso à escola e, quando conseguiam, eles se deparavam com a precariedade das condições materiais dos estabelecimentos.

Ao final das avaliações de rendimento escolar, realizadas no período compreendido entre 1988 e 1991, mais uma vez se revelou a precariedade da educação básica brasileira. Assim, ficou evidente a importância da criação de um sistema de avaliação que “[...] informe aos responsáveis pelo ensino/educação os problemas da realidade pedagógica que ocorrem na escola e que se refletem nos diferentes níveis de capacitação cognitiva” (GATTI; VIANNA; DAVIS, 1991, p. 11).

Desse modo, Gatti, Vianna e Davis (1991) apontaram para a necessidade de um esquema de avaliação que considerasse as múltiplas faces do sistema educacional, evitando, dessa maneira, a imposição de modelos que nem sempre são ajustáveis aos seus diversos aspectos, ainda mais considerando os múltiplos contextos em que a realidade educacional brasileira estava imersa. Diante de tal quadro, a avaliação educacional nacional avançou em direção a criação de um Sistema Nacional de Avaliação.

2.6 SISTEMA NACIONAL DE AVALIAÇÃO DO ENSINO PÚBLICO DE 1º GRAU (SAEP)

Um estudo realizado pela Organização das Nações Unidas para a Educação, a Ciência e a Cultura (Unesco), com a cooperação da Comissão Econômica para a América Latina e o Caribe (Cepal) para a Sexta Conferência Regional de Ministros da Educação da América Latina e Caribe, apontou, conforme Waiselfisz (1991), que o Brasil possuía a maior taxa de repetência escolar em relação ao ensino de 1º grau e uma elevada taxa de evasão escolar, como já revelavam as pesquisas anteriores.

Para o autor, evidenciou-se, portanto, que a melhoria e a qualidade da educação deveriam ser as prioridades centrais das políticas educacionais, como reconheceu, em 1984, o Comitê da Unesco. Dentre as estratégias propostas para a melhoria, enfatizou-se a busca por informações acerca do desempenho dos sistemas educacionais, para que essas fundamentassem o processo de tomada de decisões das políticas e de estratégias educacionais na tentativa de reverter aquele quadro, como reitera Waiselfisz (1991).

Nesse contexto, o MEC, por intermédio da Seneb, decide ampliar a abrangência proposta pelo Projeto Nordeste, por meio da criação de mecanismos de avaliação da educação que pudessem ser aplicados em todo o território nacional, o que deu origem ao Sistema Nacional de Avaliação do Ensino Público de 1º Grau (Saep)²⁹ – cuja finalidade era o estabelecimento de um sistema de controle dos conteúdos mínimos e dos padrões de qualidade no ensino básico, reforçando os arts. 210 e 214 da Constituição Federal de 1988 (WAISELFISZ, 1991; BRASIL, 1994; BONAMINO; FRANCO, 1999; HORTA NETO, 2007).

A proposta de avaliação do Saep, de acordo com Waiselfisz e Horta Neto (2016), embora mais complexa, referenciou-se no *National Assessment of Educational Progress* (Naep) dos EUA. Em comum, havia a aplicação de testes padronizados para os alunos, questionários para professores e gestores, desenvolvidos na tentativa de abranger o maior número possível de fatores contextuais que pudessem explicar as dificuldades dos alunos (WAISELFISZ; HORTA NETO, 2016).

O surgimento do Saep, em 1988, segundo Horta Neto (2007), foi influenciado por duas experiências pertinentes em avaliação: o Projeto Edurural e a Avaliação do Rendimento Escolar de Alunos de Escolas de 1º Grau da Rede Pública. Esta última serviu como base para o desenvolvimento do instrumento de avaliação utilizado. Ademais, os recursos que garantiram seu desenvolvimento foram obtidos mediante convênio firmado entre o Iica e o MEC, do qual Waiselfisz fez parte (WAISELFISZ, 1991; PESTANA, 1992; BRASIL, 1994a; HORTA NETO, 2007).

Desse modo, em novembro de 1988, houve uma aplicação piloto nos Estados do Paraná e Rio Grande do Norte, com o intuito de testar, em situação de campo, a pertinência dos instrumentos e procedimentos que visavam à aplicação nacional no ano seguinte (BRASIL,

²⁹Também conhecido como Sistema de Avaliação do Ensino Público de 1º Grau (HORTA NETO, 2007; FREITAS, 2013) e como Sistema de Avaliação da Educação Pública (PESTANA, 2016).

1994; BONAMINO; FRANCO, 1999; HORTA NETO, 2007; WASELFISZ; HORTA NETO, 2016). Essa aplicação, por sua vez, gerou um relatório em que constam: o nome de Waiselfisz como coordenador; as quatro pessoas que integravam a equipe técnica do MEC, entre elas, Maria Inês Pestana; e os cinco consultores responsáveis pelos estudos de custo-aluno, gestão e professor referentes ao teste de desempenho (WASELFISZ; HORTA NETO, 2016).

Waiselfisz e Horta Neto (2016) destacam, ainda, que, como a equipe do MEC era composta por poucos membros, era fundamental o trabalho cooperativo das Secretarias de Estados da Educação, como também reforça Vianna (1995). Essa construção do sistema, relativamente em parceria com as Secretarias Estaduais, foi um dos pontos de divergência entre o MEC e o Banco Mundial. Para o banco, o MEC deveria estar à frente das negociações do sistema e repassá-las às demais instâncias governamentais. Por isso, recusou o financiamento para a primeira avaliação nacional (ALTMANN, 2002). Segundo Waiselfisz, o Banco Mundial também não havia concordado em realizar a aplicação piloto. Foi de sua responsabilidade convencer o MEC a efetuar a avaliação por conta própria, já que havia recursos disponíveis (WASELFISZ; HORTA NETO, 2016).

No ano seguinte, em 1989, estava prevista a aplicação. Entretanto, devido às dificuldades orçamentárias e à ausência de financiamento, ela foi interrompida (WASELFISZ; HORTA NETO, 2016). Os trabalhos só foram retomados pela Seneb em agosto de 1990, via Projeto BRA/86/002³⁰ que, por meio dos fundos de empréstimo, foi financiado pelo Banco Mundial, permitindo a primeira avaliação em nível nacional (BRASIL, 1994a; COELHO, 2008).

Durante o planejamento das provas do Saep, mais uma vez a definição dos conteúdos mostrou ser um problema a ser enfrentado, já que, até então, segundo Horta Neto (2007), também não havia um currículo único nacional definido e adotado pelas instituições escolares. Para a elaboração de uma proposta comum curricular, semelhante a que foi realizada na Avaliação do Rendimento Escolar de Alunos de Escolas de 1º Grau da Rede Pública, conforme explica o autor, consultou-se uma amostra de professores com o objetivo de identificar os conteúdos que eram desenvolvidos em sala de aula e que auxiliassem na elaboração dos itens que integrariam as avaliações. Além disso, procurou-se identificar os aspectos comuns (foi

³⁰Segundo Brasil (1987), o Projeto MEC/PNUD BRA/86/002 tem como responsabilidade fundamental a melhoria do gerenciamento dos sistemas estaduais e territoriais da educação e a capacitação gerencial nos diferentes níveis da administração.

considerado comum o que aparecia em pelo menos 70% dos Estados) e, por meio de discussões com as equipes técnicas dos Estados, definiram-se os conteúdos que seriam avaliados e os demais instrumentos (questionário para diretores, professores e sobre a escola) (PESTANA, 1998).

Avaliou-se, assim, uma amostra de escolas públicas, tendo também como público-alvo os alunos das 1^a, 3^a, 5^a e 7^a séries do ensino fundamental nas disciplinas de Língua Portuguesa e Matemática para todas as séries, mais Ciências e Redação para as duas últimas. A escolha das séries, de acordo com Waiselfisz e Horta Neto (2016), fez-se considerando os estudos até então desenvolvidos (VIANNA; GATTI, 1988; VIANNA, 1989a; VIANNA, 1989b), e a escolha das disciplinas pode ser explicada pela ênfase dada à Língua Vernácula e Matemática pelo próprio Banco Mundial.

A implementação do Saep, segundo Waiselfisz (1991), objetivava o fornecimento de informações e de indicadores sobre a situação e os resultados do sistema público de ensino de 1º grau. Para isso, o sistema foi desenvolvido para responder às seguintes perguntas: até que ponto a gestão educacional está se tornando mais eficiente e democrática? Têm ocorrido mudanças nas condições de trabalho e na competência pedagógica do professor? Até que ponto as políticas adotadas, em nível nacional, estadual e regional, estão realmente facilitando o acesso à escolarização e promovendo a melhoria da qualidade de ensino? (BRASIL, 1994a). Diante disso, o sistema embarcou nos seguintes eixos: gestão do campo educacional; valorização do magistério; e universalização, eficiência e qualidade do ensino (WAISELFISZ, 1991; BRASIL, 1994a).

Por fim, no ano de 1991, os dados foram processados e seus resultados publicados em 1992, por meio de relatórios. Ainda nesse ano, o Saep passou a se chamar Sistema Nacional de Avaliação da Educação Básica (Saeb), tendo em vista que, futuramente, o sistema poderia se dedicar também à avaliação do segundo grau, em um trabalho conjunto com as Secretarias Estaduais de Educação (BRASIL, 1994). Dessa forma, processaram-se os primeiros dados nacionais, já pelo Saeb – um marco na educação que permitiu ao Governo Federal inteirar-se no tocante à qualidade da educação básica brasileira.

2.7 CONSIDERAÇÕES PARCIAIS

Este capítulo destinou-se a discutir o processo histórico que culminou na criação do Saeb, em 1990. Para isso, realizou-se um levantamento das experiências iniciais de avaliações externas, desenvolvidas a partir da década de 1960, que serviu como base para o desenho do sistema.

Quanto ao desenvolvimento dessas avaliações, ficou evidente que não seria possível sem a atuação de fundações de pesquisa em avaliação, com destaque para a FCC, que contribuiu significativamente na formação de pesquisadores educacionais para atuarem na implementação de um sistema nacional de avaliação.

Destaca-se, também, a influência do financiamento de instituições multilaterais internacionais para esse *boom* da avaliação no Brasil, como o Banco Mundial. Afinal, dentre as suas exigências, centrava-se, sobretudo, no uso da avaliação externa como política educacional (GATTI; VIANNA; DAVIS, 1991), sendo essa condicionante uma das dimensões mais importantes para o desenvolvimento dos projetos.

Além disso, a eficiência dos gastos públicos mostrou ser uma pauta relevante para o Banco Mundial que, por meio de sua “assistência” técnica e financeira, auxiliaria na alocação de recursos financeiros. No entanto, vários porquês não são explicados na execução dos projetos educacionais brasileiros, financiados pelo órgão. Por exemplo, quem coordenava as transações financeiras? Para onde os recursos eram alocados? O que exatamente era feito com esses recursos?

Quanto as primeiras experiências de avaliação que influenciaram o desenho e a montagem do Saeb, destacam-se: o Projeto Edurural; o Projeto Nordeste; a Avaliação do Rendimento Escolar de Alunos de Escolas de 1º Grau da Rede Pública; e o Saep. Essas, por sua vez, revelaram a inexistência de dados educacionais naquele período, principalmente em nível de sistema, além de se mostrarem influentes nos processos de avaliação posteriores.

O Projeto Nordeste, por exemplo, baseou-se no Projeto Edurural, que não só seguiu o mesmo desenho de avaliação, como também foi desenvolvido com maior amplitude nos Estados da região avaliada. Por meio desse projeto, percebe-se que, desde a década de 1980, havia interesse em avaliar o desempenho dos alunos que cursavam o final de cada ciclo escolar nas disciplinas Língua Portuguesa e Matemática, comuns as quatro experiências analisadas.

O Saep, por sua vez, fundamentou-se no Projeto Nordeste, cuja proposta de avaliação resultou na montagem inicial do sistema. Também o sistema teve como base o instrumento de avaliação utilizado pela Avaliação do Rendimento Escolar de Alunos de Escolas de 1º Grau da Rede Pública, além da escolha das séries e disciplinas avaliadas. Conforme evidenciado, essas quatro avaliações se mostraram de grande influência no processo de desenvolvimento do Saeb.

Ressalta-se, ainda, que alguns fatores se revelaram comuns a todas as experiências de avaliação analisadas neste capítulo. Primeiro, tem-se o foco dado à educação primária, visto que, no período estudado, o país se encontrava distante da universalização desse ensino. Em segundo, consta a composição das provas referentes às disciplinas avaliadas, que buscavam se adequar à realidade das escolas considerando o que era efetivamente adotado pelo sistema de ensino público. Por último, destaca-se o objetivo de subsidiar as Secretarias Estaduais e Municipais de Educação, atendendo aos professores, com informações sobre os problemas de aprendizagem escolar dos seus alunos com vistas à recuperação dos educandos.

Dessarte, entende-se que as experiências iniciais de avaliação se mostraram precursoras para o desenho do Saeb, principalmente para a escolha das séries e disciplinas avaliadas e para a construção do conteúdo. Desse modo, sinalizam-se para outros aspectos que também precisam ser investigados, como o porquê da adoção das disciplinas Língua Portuguesa e Matemática e qual o conteúdo avaliado. Essas questões são importantes para o entendimento do processo de elaboração das Matrizes de Referência do sistema.

Por outro lado, há indícios de uma disputa política para o controle das diretrizes educacionais que merece investigação, uma vez que ficaram evidentes alguns desentendimentos entre as Secretarias (estaduais e municipais) de Educação e o MEC, além da interferência de outras instâncias na configuração do sistema, como o Banco Mundial.

Em estudos futuros, parece-nos, ainda, interessante investigar a efetividade do Saeb ao longo dos seus mais de 30 anos de existência, sobretudo como política de avaliação. Além disso, aponta-se para a necessidade de se desenvolver estudos e pesquisas que possam discutir: como se deu o seu processo de desenvolvimento e institucionalização? Quais as contribuições do sistema para o desenvolvimento educacional brasileiro? Quais as possíveis implicações do uso de seus resultados para a educação?

Por fim, espera-se que este capítulo esclareça alguns dos fatores incorporados pelo sistema durante a sua criação, contribuindo como um instrumento de difusão de conhecimentos acerca do Saeb e para a promoção de discussões desse sistema no ambiente escolar e acadêmico.

CAPÍTULO 3

O PROCESSO DE DESENVOLVIMENTO E INSTITUCIONALIZAÇÃO DO SAEB

3.1 INTRODUÇÃO

Desde o final da década de 1990, o modelo de avaliação por meio de testagens em larga escala ganhou notoriedade no Brasil, adentrando-se na administração e gestão das redes de ensino (estaduais e municipais) do país (BAUER, 2019). Essa centralidade é evidenciada nos planos nacionais de educação, em que a avaliação é definida como uma forma de garantir a supervisão e inspeção da qualidade do ensino ofertado pelas escolas (BRASIL, 2001).

Nesse sentido, para o monitoramento e o cumprimento das metas estabelecidas pelos planos, utilizam-se os indicadores fornecidos pelo Sistema de Avaliação da Educação Básica (Saeb), cuja primeira aplicação aconteceu em 1990. O Saeb é um sistema de avaliação externa em larga escala, composto por um conjunto de instrumentos e realizado periodicamente pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) (BRASIL, 2021b), autarquia do Ministério da Educação (MEC).

As avaliações externas, conforme Vianna (2010, p. 18), “[...] são recomendáveis, na medida em que representam um trabalho não comprometido com a administração educacional e as políticas que o orientam”. Em outras palavras, são operacionalizadas por órgãos externos, uma vez que se pretende aferir uma visão de fora e supostamente isenta em relação aos sistemas educacionais (VIANNA, 2010).

Atualmente, o Saeb é aplicado à educação infantil e ao 2º ano do ensino fundamental de maneira amostral e às etapas finais de cada ciclo da educação básica (5º e 9º anos do ensino fundamental e 3º do ensino médio) com caráter censitário. O sistema avalia os estudantes por meio de testes nas disciplinas de Língua Portuguesa e Matemática, mas também avança para a aplicação em outras áreas do conhecimento, como Ciências da Natureza e Ciências Humanas.

Durante esses trinta anos, o Saeb passou por algumas modificações. A primeira foi em 1995. Nessa edição, o Inep encomendou uma avaliação externa do sistema que apontou uma série de questões a serem consideradas para o seu aperfeiçoamento. Dentre essas, destacam-se a

adoção de uma nova metodologia e a preocupação com os conteúdos mínimos comuns referenciados das provas.

Nessa perspectiva, este capítulo tem como objetivo analisar a teoria adotada, considerando o que e como se avalia, e também como esses aspectos contribuíram para o processo de desenvolvimento e institucionalização do Saeb. Procurou-se, ainda, esclarecer quais são os seus desdobramentos na constituição de uma política educacional. Para tal, estabeleceu-se, como recorte temporal, o período compreendido entre os anos de 1990 e 2001, marcado pelas principais modificações do sistema. O esforço, assim, é responder à seguinte questão: de que maneira essas alterações metodológicas contribuíram para a efetivação de uma política de avaliação e quais são suas implicações para a educação básica?

Desse modo, além desta introdução, este capítulo está estruturado em outras cinco seções: na primeira, apresenta-se a metodologia utilizada; na segunda, discute-se acerca do processo de desenvolvimento e institucionalização do Saeb; na terceira e na quarta seções, abarcam-se as duas principais modificações do período de estudo: a adoção da Teoria de Resposta ao Item (TRI) e a adoção das Matrizes de Referência; na quinta e última seção, apresentam-se os apontamentos finais específicos, encerrando assim o capítulo.

3.2 O PROCESSO DE DESENVOLVIMENTO E DE INSTITUCIONALIZAÇÃO DO SAEB (1990-1995)

Após a primeira aplicação nacional do Saeb, em 1990, o MEC definiu que novas avaliações aconteceriam a cada dois anos. Se, por um lado, a escolha desse intervalo se deu por esse ser um período suficiente para efetuar a coleta de dados e a divulgação dos resultados (WAISELFISZ; HORTA NETO, 2016), por outro, pretendia-se cumprir uma das condicionantes do Banco Mundial, que incentivou a adoção de sistemas de avaliação educacional com características amostrais e bianuais nos países latino-americanos em troca de financiamentos (MENDONÇA, 2014) que, por sua vez, eram visados pelo Governo Federal.

Entretanto, o segundo ciclo³¹ de avaliação do Saeb, que estava previsto para 1992, precisou ser reprogramado para o ano seguinte, pois não havia recursos financeiros suficientes para sua realização e nem financiamentos externos (HORTA NETO, 2007; WAISELFISZ; HORTA NETO, 2016).

Dessa forma, o segundo ciclo de avaliação ocorreu em 1993, financiado pelos próprios Estados da Federação. Entre as atribuições por eles assumidas, destaca-se o levantamento de dados que, segundo Pestana (2016, p. 75), foi o “[...] aspecto decisivo e o que mais contribuiu para a institucionalização do sistema e, principalmente, para introdução e disseminação de uma cultura de avaliação no setor educacional”. Isso gerou, nas Secretarias Estaduais de Educação, um maior interesse em dar continuidade à obtenção de informações, visando à melhoria da qualidade da educação (BRASIL, 1994a).

Diante disso, percebeu-se a necessidade de aprimorar os instrumentos de avaliação que constituíam o Saeb, tendo em vista a sua consolidação e aperfeiçoamento. Assim, o Saeb passou a receber apoio institucional do Plano Decenal de Educação para Todos 1993-2003 (BRASIL, 1993), que definiu o sistema como ação prioritária das diversas instâncias envolvidas, como Secretarias Estaduais e Municipais de Educação, universidades, conselhos e órgãos de classe (BRASIL, 1994a).

De acordo com o referido plano, o Saeb possuía a finalidade de “[...] aferir a aprendizagem dos alunos e o desempenho das escolas de primeiro grau e prover informações para avaliação e revisão de planos e programas de qualificação educacional” (BRASIL, 1993, p. 59). Além disso, seus objetivos,

[...] compreendem a pesquisa e desenvolvimento de métodos e instrumentos de avaliação escolar, a implementação, em universidades e centros de pesquisa, de núcleos ou centros regionais de estudos em avaliação, a formação e capacitação de recursos humanos para a área e, por fim, a produção de pesquisa aplicada para subsidiar políticas públicas de qualidade educacional (BRASIL, 1993, p. 59).

Esses objetivos aos poucos foram sendo alcançados. Os resultados obtidos pelo Saeb possibilitaram a consolidação de um banco de dados em nível nacional e por unidade federativa. Além disso, conforme Brasil (1994a), houve, também, a incorporação dos resultados na formulação de políticas públicas e no planejamento do ensino, bem como na

³¹Cada avaliação do Saeb, segundo Horta Neto (2007), é conhecida como ciclo de avaliação por apresentar um trabalho árduo: na elaboração e aplicação das provas; no processamento e análise dos dados; e na elaboração dos relatórios e informes para divulgação dos resultados.

elaboração de propostas curriculares desenvolvidas pelas Secretarias Estaduais e Municipais de Educação.

Conjuntamente, ainda segundo Brasil (1994a), deu-se a estruturação de núcleos regionais ligados a instituições universitárias e de pesquisa. A partir dos dados do Saeb e de outras pesquisas educacionais, esses núcleos contribuíram para o desenvolvimento teórico e metodológico da avaliação e para o treinamento de recursos humanos em nível de pós-graduação na área.

Verificou-se, ainda, uma progressiva valorização da avaliação na educação básica realizada por alguns órgãos, como: o Conselho Nacional de Secretários de Educação (Consed); a União Nacional dos Dirigentes Municipais de Educação (Undime); o Fórum dos Conselhos Estaduais de Educação; o Conselho Federal de Educação e outros (BRASIL, 1994a).

Desse modo, como as políticas educacionais deveriam oportunizar uma educação básica de qualidade com equidade e eficiência (BRASIL, 1993), para a sua efetividade, entendeu-se que era necessário o monitoramento permanente e a avaliação dos resultados dessas políticas. Para a realização desses processos (monitoramento e avaliação), o Saeb foi institucionalizado por meio da Portaria n°. 1.795, de 27 de dezembro de 1994 (BRASIL, 1994b), e passou a vigorar com os seguintes objetivos gerais:

[...] 1. contribuir para o desenvolvimento, em todos os níveis do sistema educativo, de uma cultura avaliativa que estimule a melhoria dos padrões de qualidade e equidade da educação básica e adequados controles sociais de seus resultados; 2. implementar e desenvolver, em articulação com as Secretarias de Educação, processos permanentes de avaliação, apoiados em ciclos regulares de aferições das aprendizagens e competências obtidas pelos alunos e do desempenho dos estabelecimentos que compõem os sistemas de educação básica; 3. mobilizar os recursos humanos, técnicos e institucionais do Ministério, das Secretarias e de universidades e centros de estudos e pesquisas sociais e educacionais, para gerar e difundir os conhecimentos, técnicas e instrumentos requeridos pelas práticas da aferição e avaliação educacional; 4. proporcionar, aos responsáveis pela gestão educacional e pelo ensino, às famílias e aos cidadãos em geral, informações seguras e oportunas a respeito do desempenho e dos resultados dos sistemas educativos (BRASIL, 1994b, p. 20.768).

Além disso, é preciso destacar que a referida portaria previa, também, a elaboração de um Plano Nacional de Avaliação da Educação Básica com diretrizes, normas básicas e objetivos específicos dos ciclos de avaliação, bem como seus padrões de desempenho e qualidade, cronogramas e orçamentos de execução (BRASIL, 1994b).

Para Mendonça (2014), o desenvolvimento do Saeb e a aprendizagem técnica oferecida por ele foram essenciais para a evolução das políticas educacionais durante a década de 1990. O

sistema contribuiu, inclusive, para o surgimento de iniciativas estaduais e municipais de avaliação externa aplicada em larga escala (GATTI, 2013).

Vale a pena destacar as iniciativas estaduais que, de acordo com Freitas (2013), utilizaram o desenho censitário (ou quase) com base no Saeb. Conforme a autora, além de servir como referência, o sistema também disponibilizava seus instrumentos, principalmente os itens das provas. Dessa maneira, reproduziam-se, portanto, a própria lógica utilizada pelo Saeb. Dentre elas, destacam-se as avaliações realizadas pelos Estados do Paraná, a partir de 1989; de Minas Gerais e São Paulo, a partir de 1992; e do Ceará, a partir de 1996 (GATTI, 2002, 2009, 2013).

Outra iniciativa estimulada pelo surgimento do Saeb foi a avaliação das escolas padrão, realizada pela Fundação Carlos Chagas (FCC), em 1992 e 1993. Encomendada pela Secretaria de Educação do Estado de São Paulo, avaliou-se o desempenho dos estudantes da 8ª série do ensino fundamental em Língua Portuguesa, Matemática, Conhecimentos Gerais e Redação. Segundo Gatti (2013), o objetivo era estudar o nível de habilidade envolvido, a dificuldade e discriminação dos itens e o seu comportamento, visando a futuras aplicações.

Dando continuidade à trajetória do Saeb, para Pestana (1998), começar o sistema utilizando provas únicas em todo o território nacional não foi uma tarefa fácil. Apesar disso, a autora destaca que, a partir de 1995, todos os Estados e as suas respectivas redes de ensino passaram a participar voluntariamente do sistema.

Nesse mesmo ano, início da gestão do governo Fernando Henrique Cardoso (FHC), foram levantados questionamentos acerca da validade da prova realizada pelo sistema. Por esse motivo, o Inep/MEC encomendou a especialistas internacionais uma avaliação externa do próprio Saeb, que trouxe à tona algumas recomendações, sugerindo modificações nos instrumentos e nas técnicas utilizadas nos ciclos anteriores (PESTANA, 1998; HORTA NETO, 2007; CASTRO, 2016).

Com relação às recomendações do estudo, Mendonça (2014) ressalta que muitas foram enriquecidas por consultas a especialistas do Banco Mundial e a entidades brasileiras da área de avaliação educacional. Em especial, destacam-se a FCC e o Centro de Seleção de Candidatos do Grande Rio (Cesgranrio) que, além da experiência em exames standardizados, também já começavam a utilizar a TRI (CASTRO, 2016).

Dessa forma, graças às parcerias do Inep/MEC com organizações multilaterais internacionais³² e nacionais, como as Secretarias Estaduais e Municipais de Educação, o Consed e as UFs, passou-se a compartilhar perspectivas sobre o processo educacional dessas instituições (MENDONÇA, 2014). Como consequência, houve avanços significativos para o desenvolvimento do sistema avaliativo do país.

Como resultado, o ciclo de avaliação de 1995 passou por uma série de modificações, podendo ser considerado um marco significativo nas alterações do Saeb. A partir daí, o sistema: passou a avaliar as séries finais de cada etapa da educação básica da rede pública (4^a e 8^a séries do ensino fundamental e 3^o ano do ensino médio); incluiu as escolas da rede privada, cuja participação era restrita ao ensino fundamental; e procurou focar a avaliação do desempenho escolar dos estudantes nas disciplinas Língua Portuguesa e Matemática (BRASIL, 2020b).

Ademais, nesse ciclo foram incorporadas outras técnicas e instrumentos de avaliação, como o levantamento de dados sobre as características socioeconômicas e culturais dos estudantes por meio de questionários contextuais, preenchidos por eles e pelos seus professores. Além disso, o sistema também sofreu uma alteração metodológica na construção do teste e na análise dos seus resultados com a adoção da TRI, cuja principal contribuição consiste na possibilidade de comparar os resultados das avaliações ao longo do tempo (BRASIL, 2020b).

3.3 A ADOÇÃO DA TEORIA DE RESPOSTA AO ITEM (TRI)

Em busca de uma maior sofisticação técnica, que pudesse auxiliar no processo de ampliação do Saeb, a partir do ciclo de 1995, o sistema incorporou a TRI como metodologia para a construção das avaliações e análise dos seus resultados. Essa foi uma iniciativa inédita no cenário nacional.

Para tanto, considerando a ausência de profissionais especializados na área, a sofisticação técnica exigida para sua implementação e as dificuldades operacionais próprias de uma avaliação de porte nacional (HORTA NETO, 2007), o Inep optou por realizar parcerias com

³²As instituições, segundo Mendonça (2014), eram: a Organização das Nações Unidas para a Educação, a Ciência e a Cultura (Unesco), a *Oficina Regional de Educacion para a America Latina y el Caribe* (Orealc), a Organização para a Cooperação e Desenvolvimento Econômico (OCDE), o *Trends in International Mathematics and Science Study* (TIMSS), o Programa das Nações Unidas para Desenvolvimento/Organização das Nações Unidas (PNUD/ONU), o Bird e o Banco Interamericano de Desenvolvimento (BID).

fundações de pesquisas que já possuíam certa experiência na área (BONAMINO; FRANCO, 1999), como a FCC e a Cesgranrio. Cabe ressaltar que essas fundações contribuíram de forma significativa no desenvolvimento das primeiras experiências de avaliação educacional (em nível nacional) no país (GATTI, 2002, 2009; GATTI; VIANNA; DAVIS, 1991; VIANNA, 1995).

Desse modo, essas fundações ficaram responsáveis pela elaboração, aplicação e análise dos resultados das provas, bem como a definição de matrizes de conteúdo que, por sua vez, serviriam de base para a construção dos itens que passariam a compor as avaliações (BONAMINO; FRANCO, 1999; MENDONÇA, 2014; PESTANA, 2016).

De acordo com Ruben Klein (2013a), pesquisador da Cesgranrio e entusiasta do uso da TRI nos exames standardizados, a adoção dessa teoria no Saeb introduziu uma verdadeira revolução metodológica nas avaliações brasileiras. Essa revolução impulsionou outras iniciativas semelhantes, como: o Exame Nacional de Cursos (ENC), conhecido como Provão, a partir de 1996; o Exame Nacional do Ensino Médio (Enem), a partir de 1998; o Exame Nacional para Certificação de Competências de Jovens e Adultos (Encceja), a partir de 2001; o Exame de Desempenho dos Estudantes (Enade), que substituiu o Provão a partir de 2004; e o Sistema Nacional de Avaliação da Educação Superior (Sinaes), também a partir de 2004 (CASTRO, 2016; RABELO, 2013).

3.3.1 Teoria de Resposta ao Item: nova metodologia de construção do teste e análise de resultados

A adoção da TRI como metodologia para a construção de testes e análises de resultados implicou a transferência do objeto central das análises. Passou-se, então, a analisar cada item individualmente. Antes, porém, considerava-se o teste como um todo, cujas análises eram pautadas, exclusivamente, na Teoria Clássica dos Testes (TCT).

A TCT é uma vertente da Psicometria que, conforme Pasquali (2009, p. 993), procura explicar o escore total obtido em um teste, isto é, a soma das respostas dadas “[...] a uma série de tarefas, tipicamente chamadas de itens”. Essa teoria foi amplamente utilizada no início do século XX e ainda possui grande relevância nas avaliações internas, graças à simplicidade dos cálculos por ela exigidos e pela facilidade das interpretações dos parâmetros com ela

relacionados, que consideram, sobretudo, a dificuldade e a capacidade discriminatória dos itens.³³

O parâmetro de dificuldade é definido pela TCT como o percentual de acerto do item, de modo que, quanto maior for esse percentual, mais fácil ele será. Além disso, a nota (ou escore) de um respondente nada mais é do que o total de itens respondidos corretamente. Ou seja, o que é feito, majoritariamente, em todo o percurso da avaliação escolar.

Para avaliar o poder de discriminação de um item, utilizado para analisar a sua capacidade de identificar grupos com diferentes níveis de conhecimento,³⁴ a TCT apresenta algumas opções. Entre as mais utilizadas, destacam-se o coeficiente de correlação ponto-bisserial e o índice de discriminação.

O primeiro é derivado do coeficiente de correlação de Pearson e visa a mensurar o grau de associação entre a resposta dada a um determinado item e o escore do teste (PASQUALI, 2003; SOARES, 2018). Assim, quanto maior o valor mensurado, maior a discriminação do item. Segundo Tôres (2015), valores acima de 0,30 são considerados satisfatórios.³⁵

O índice de discriminação, por sua vez, é obtido com base na “regra dos 27” proposta por Kelley (1939). Para Pasquali (2003), trata-se da diferença entre o índice de dificuldade do item calculado para os 27% de respondentes com melhor (dito grupo superior) e pior (dito inferior) desempenho no teste. Em suma, quanto maior o valor dessa diferença, mais discriminativo será o item.

Embora ainda muito utilizadas em análise de itens, as estatísticas que envolvem a TCT apresentam limitações, principalmente no que se refere à comparabilidade de desempenho de indivíduos que se submetem a diferentes testes. Isso dificulta a realização de estudos longitudinais, que são úteis para avaliar o comportamento de uma política pública ao longo dos anos, por exemplo.

³³Existem outros parâmetros psicométricos que compõem a TCT que avaliam, por exemplo, a confiabilidade ou fidedignidade do teste como instrumento de medida. Maiores informações acerca desses parâmetros podem ser obtidas em Pasquali (2003).

³⁴Essa característica é muito útil em exames seletivos que precisam classificar os respondentes. Em suma, quanto maior a discriminação de um item, maior a sua capacidade em separar os estudantes com maior e menor habilidade.

³⁵De acordo com Andrade, Laros e Gouveia (2010, p. 426), no que se refere aos procedimentos metodológicos adotados no Saeb 2005: “[...] itens com coeficiente bisserial do gabarito menor ou igual a 0,15; itens com dois coeficientes bisseriais de distratores (alternativas erradas) maiores que 0,10 ou coeficiente bisserial de um distrator maior que a bisserial da alternativa correta, são encaminhados para a análise pedagógica”. Entretanto, não foram encontrados documentos oficiais mais recentes que tratam desses parâmetros.

Na TCT, os resultados das análises estatísticas, com fins de avaliar as características psicométricas dos itens do teste, como o índice de dificuldade e discriminação, são dependentes da amostra de sujeitos para os quais eles são calculados (subject-dependent) (PASQUALI, 2003, 2018). Dessa forma, um mesmo item pode ser considerado fácil para um determinado grupo com alto grau de conhecimento e difícil para um grupo com grau de conhecimento menor. Analogamente, um item pode apresentar problemas de discriminação se aplicado a uma amostra homogênea, composta por sujeitos com habilidades similares.

Seguindo essa vertente, a TRI foi desenvolvida no sentido de superar essas limitações. Entretanto, deve-se ressaltar que uma teoria não substitui a outra, ambas se complementam. No caso do Saeb, por exemplo, primeiramente os itens são analisados de forma exploratória pela TCT. Se apresentarem problemas, será realizada uma análise pedagógica por especialistas da área antes da estimação pela TRI (ANDRADE; LAROS; GOUVEIA, 2010).

Diferentemente da TCT, cujo interesse estava na produção de “testes de qualidade”, a TRI se inclina para a produção de “itens de qualidade”. De acordo com Andrade, Tavares e Valle (2000, p. 8), trata-se de “[...] um conjunto de modelos matemáticos que procuram representar a probabilidade de um indivíduo dar uma certa resposta a um item como função dos parâmetros do item e da habilidade (ou habilidades) do respondente”.

Contudo, esses modelos apresentam uma complexidade que torna o uso da TRI inacessível sem o auxílio de softwares computacionais especializados e de uma amostra de tamanho significativo. Isso justifica a sua utilização tardia (a partir de 1980), ao considerar que o seu surgimento, como teoria, data da década de 1950 (SOARES, 2018).

Além do mais, o uso da TRI pressupõe a presença de uma habilidade dominante que responde por todos os itens do teste (unidimensionalidade), bem como independência entre as respostas a diferentes itens do teste (independência local). Esses pressupostos são de difícil verificação sem a utilização de técnicas computacionais (ANDRADE; TAVARES; VALLE, 2000; PASQUALI, 2003; SOARES, 2018; PASQUALI, 2018).

Na prática, conforme Muñiz (1990), é comum o uso de modelos logísticos na TRI, justamente por permitirem um melhor tratamento matemático. Dentre esses modelos, são mais utilizados em exames standardizadas os de um (conhecido como modelo de Rasch), dois ou três parâmetros para itens dicotômicos, que consideram, respectivamente: a dificuldade; a dificuldade e a discriminação; a dificuldade, a discriminação e o acerto casual do item.

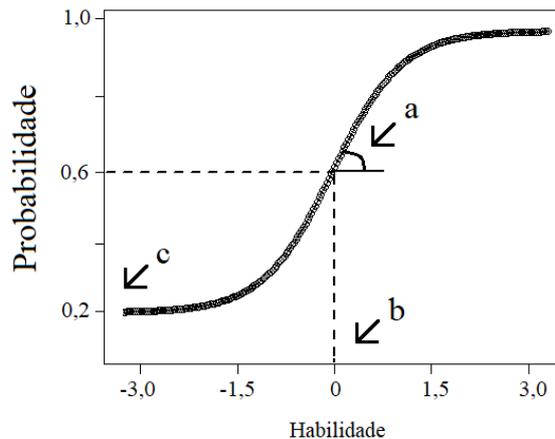
O modelo da TRI utilizado pelo Saeb é o logístico de três parâmetros (ML3P). Esse modelo, proposto por Birnbaum (1968), é expresso matematicamente pela equação:

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

Nesse caso, $P(\theta_j, a_i, b_i, c_i)$ representa a probabilidade de o indivíduo j com habilidade θ_j acertar o item i . Além disso, a_i , b_i e c_i são os parâmetros de discriminação, dificuldade e probabilidade de acerto ao acaso do item i , respectivamente.

Essa relação entre a probabilidade de acerto a um item, seus parâmetros psicométricos e a habilidade de o respondente pode ser representada por um gráfico denominado Curva Característica do Item (CCI), cujo comportamento é mostrado na Figura 3.

FIGURA 3 – Curva Característica do item



Fonte: Elaboração dos autores.

Desse modo, observa-se que os respondentes com maior (menor) habilidade possuem maior (menor) probabilidade de acerto do item, como esperado. Por meio de uma análise geométrica, é possível verificar que o parâmetro de discriminação do item (a_i) é proporcional à inclinação da CCI, no ponto em que a abscissa representa o parâmetro de dificuldade do item (b_i). Esse valor é obtido quando a probabilidade de resposta correta ao item é igual a $(1 + c_i)/2$, no qual c_i representa a probabilidade de acerto casual do item i .

É preciso destacar que os detalhes metodológicos que envolvem o cálculo dos parâmetros psicométricos dos itens do Saeb, bem como a estimação das notas de seus respondentes não

foram encontrados em documentos oficiais. Porém, sabe-se que para tal se utiliza o método Expected A Posteriori (EAP), estimador bayesiano, que atribui estimativas de proficiências a todos os padrões de respostas, inclusive, os certos ou errados (KLEIN, 2013a).

3.3.2 A Escala de Proficiência do Saeb

Uma das principais vantagens da TRI em relação à TCT está na comparabilidade de resultados de estudantes submetidos a provas distintas. Essa comparabilidade é possível mediante a construção de uma escala de proficiência, a partir do nível de proficiência de cada participante, estimado via TRI. A escala de proficiência do Saeb está baseada nos resultados do Saeb 1997, cuja média e desvio padrão para cada ciclo da educação básica estão representados na Tabela 1.

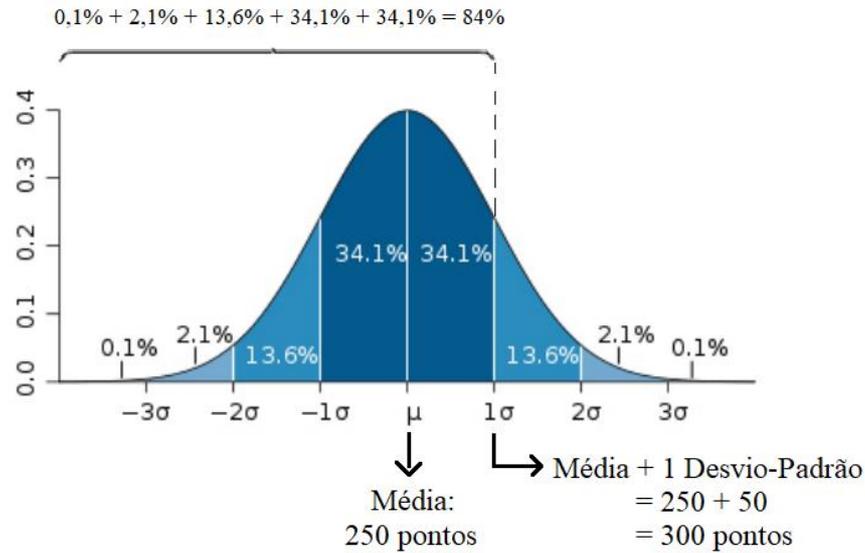
TABELA 1 – Limites inferiores e superiores e amplitudes das proficiências do Saeb 1997

Série	Matemática		Língua Portuguesa	
	Média	Desvio padrão	Média	Desvio padrão
4ª série/5º ano do EF	190,8	44	186,5	46
8ª série/9º ano do EF	250,0	50	250,0	50
3ª série do EM	288,7	59	283,9	56

Fonte: Nota Técnica Índice de Desenvolvimento da Educação Básica (BRASIL, 2009a, p. 3).

Em termos práticos, um estudante do 9º ano do ensino fundamental que obtiver nota 300 em Língua Portuguesa, por exemplo, estaria a um desvio padrão acima da média (SOARES, D.; SOARES, T.; SANTOS, 2020). Conforme ilustra a Figura 4, esse estudante estaria entre os 16% com melhor desempenho no exame.

FIGURA 4 – Distribuição normal e escala de proficiência da Prova Brasil



Fonte: Soares, Soares e Santos (2020, p. 5).

A partir da construção dessas escalas, são descritas as habilidades que os estudantes são capazes de demonstrar em cada um dos seus níveis. Assim, é possível compreender quais habilidades eles já construíram, estão desenvolvendo ou ainda precisam ser alcançadas. Segundo Rabelo (2013, p. 33), isso permite fazer “[...] comparações e acompanhar progressos de um sistema educacional ao longo dos anos”. Para isso, as provas de um determinado ano incluem alguns itens, ditos âncoras, aplicados no ano anterior.

A partir desses itens, é realizado um processo de equalização (equating),³⁶ uma espécie de transformação de medidas, de modo a alocar os resultados em uma mesma escala. Para Klein (2009), a metodologia de escolha e a identificação dos níveis âncoras no Saeb deram-se a partir de duas metodologias: a primeira para as edições de 1995 e 1997 e a segunda para as edições a partir de 1999.

Em 1995 e 1997, o Saeb utilizou a metodologia proposta por Beaton e Allen (1992). Nela os itens âncoras caracterizam os pontos ou níveis da escala em que: a grande maioria dos respondentes (65%) situados em torno do nível acerta o item; menos da metade dos estudantes no nível inferior também acerta o item; e a diferença entre os percentuais de acerto é maior do que 30%. Um dos problemas observados com a aplicação dessa metodologia, de acordo com

³⁶Segundo Klein (2013b), no Saeb tem sido usado o método EAP a priori, sendo a distribuição normal com média 0 e variância 1, $N \sim (0,1)$, para todos os grupos (séries/anos).

Klein (2009), era o não aparecimento de alguns descritores na descrição, pois, embora obtivessem bons parâmetros psicométricos, não configuravam como âncoras.

Assim, a partir do ciclo de 1999, uma nova metodologia foi aplicada, de modo a identificar itens âncoras e a verificar em quais níveis os estudantes estariam dominando o conteúdo do item. Essa identificação era realizada a partir do agrupamento em intervalos de 25 pontos de uma escala com média de 250 pontos e desvio padrão de 50 pontos.³⁷ Nessa nova abordagem, passou-se a considerar como âncora em um nível o item em que:

[...] i) O número de alunos no nível que respondeu ao item é maior que 50. ii) O percentual de acerto do item nos níveis anteriores é menor que 65%. iii) O percentual de acerto do item no nível considerado e nos níveis acima é maior que 65%. iv) O ajuste da curva é bom (KLEIN, 2009, p. 139).

Em síntese, conforme Klein (2013b), o conceito de interpretação da escala de proficiência está entre as contribuições mais importantes do Saeb, cujo intuito é descrever o que os estudantes sabem e são capazes de fazer em determinados níveis da escala. Para orientar o processo de elaboração dos itens da prova, permitindo a interpretação qualitativa dessas escalas, a partir do ciclo de avaliação de 1997, desenvolveram-se as Matrizes de Referência para o sistema, consideradas um grande marco para os elaboradores de instrumentos (FINI, 2009).

3.4 A ADOÇÃO DAS MATRIZES DE REFERÊNCIA

A ideia inicial para a construção de uma matriz de referência para o Saeb, a partir do ciclo de avaliação de 1997, centrava-se na necessidade de definir qual seria o conteúdo mínimo que o estudante deveria aprender em cada etapa escolar (WAISELFISZ; HORTA NETO, 2016). Em teoria, os resultados da avaliação externa seriam capazes de identificar o quão distante ou próximo os estudantes se encontravam do mínimo estabelecido.

Para a elaboração da primeira versão, o documento Matrizes Curriculares de Referência para o SAEB/97 (BRASIL, 1998) partiu de uma ampla consulta nacional sobre os conteúdos abordados nas escolas brasileiras. Além disso, consideraram: a Lei de Diretrizes e Bases da

³⁷Esses valores foram obtidos a partir dos resultados do Saeb 1997. A descrição dos itens da escala de proficiência do Saeb pode ser consultada no site: <http://portal.inep.gov.br/educacao-basica/saeb/matrizes-e-escalas>.

Educação Básica (LDB, Lei nº 9.394/96); a disseminação das Diretrizes Curriculares Nacionais (DCNs); e as reflexões dos professores e pesquisadores de cada área/disciplina avaliada, bem como as das equipes das Secretarias de Educação (PESTANA, 1998; BRASIL, 1998, 2002; ARAÚJO; LUZIO, 2005; CASTRO, 2016). Essa metodologia também foi adotada no ciclo de 1999.

Em 2001, mais uma vez o Saeb passou por modificações. Com a promulgação da LDB, marco na agenda de reformas do governo FHC, o Saeb referiu-se pela primeira vez à coleta de informações e à avaliação, cujo objetivo era atuar sobre a qualidade do ensino (HORTA NETO, 2007). Conforme consta no art. 9º, coube à União

[...] V - coletar, analisar e disseminar informações sobre a educação; VI - assegurar processo nacional de avaliação do rendimento escolar no ensino fundamental, médio e superior, em colaboração com os sistemas de ensino, objetivando a definição de prioridades e a melhoria da qualidade do ensino (BRASIL, 1996, p. 4).

Dessarte, tendo em vista a garantia do padrão de qualidade da educação básica (BRASIL, 1988, 1996), coube à avaliação o objetivo de “monitorar a qualidade” (LOCATELLI, 2002), sobretudo por meio do Saeb, que se tornou responsável por fornecer informações sobre o desempenho escolar e os fatores a ele associados. Por esse motivo, segundo Castro (2016), a influência da LDB foi fundamental, uma vez que se passou a exigir avaliações periódicas de todos os níveis do sistema educacional.

Para se adaptar às novas exigências, houve, portanto, uma atualização das Matrizes de Referência do Saeb, em 2001. Dessa forma, realizou-se uma nova consulta às UFs, para que suas equipes de ensino e professores regentes das séries/anos avaliados verificassem a compatibilidade entre as matrizes vigentes e o currículo proposto pelos sistemas estaduais de ensino (BRASIL, 2002). Partindo dessa realidade, o Inep e os grupos de especialistas das duas áreas (Língua Portuguesa e Matemática) se reuniram para a elaboração das novas matrizes.

É digno de nota que, em consoante com o documento SAEB 2001: novas perspectivas (BRASIL, 2002), as matrizes de referência não englobam a totalidade do currículo escolar. De acordo com Brasil (2002, p. 15), “[...] foi feito um recorte com base no que pode ser possível aferir por meio do tipo de instrumento de medida utilizado no Saeb e que, ao mesmo tempo, seja representativo do que está contemplado nos currículos vigentes no Brasil”. Por isso, essas matrizes não devem ser confundidas com os procedimentos, estratégias de ensino ou orientações metodológicas.

As Matrizes de Referência do Saeb, utilizadas para a elaboração dos itens das provas, reúnem o conteúdo a ser avaliado em cada disciplina e série/ano, informando as competências e habilidades esperadas dos estudantes (BRASIL, 2002). Pressupõe-se, portanto, que “[...] o aluno deve desenvolver determinadas competências cognitivas no processo de construção do conhecimento e apresentar habilidades a partir de tais competências” (ARAÚJO; LUZIO, 2005, p. 15).

Tal perspectiva, para Bonamino (2000), configura-se em um deslocamento da preocupação com o domínio dos conteúdos escolares –que demarcou os dois primeiros ciclos do Saeb, em direção à valorização do “saber” e do “saber fazer” do aluno, que se deu a partir do ciclo de 1997, com a adoção da matriz de referência. Desde então, conforme observado pela autora, as competências cognitivas e as habilidades instrumentais passaram a nortear a formulação dos itens dos testes de desempenho do sistema. Competências cognitivas são entendidas como

[...] as modalidades estruturais da inteligência, operações que o sujeito utiliza para estabelecer relações com e entre os objetos, situações, fenômenos e pessoas (observar, representar, imaginar, reconstruir, comparar, classificar, ordenar, memorizar, interpretar, inferir, criticar, supor, levantar hipóteses, escolher, decidir, etc.) (BRASIL, 1998, p. 7).

As habilidades instrumentais, por sua vez, “[...] referem-se especificamente ao plano do saber fazer e decorrem diretamente do nível estrutural das competências adquiridas e que se transformam em habilidades” (BRASIL, 1998, p.7). Com efeito, a partir da descrição das competências e habilidades que os estudantes deveriam dominar em cada etapa avaliada, foram desenvolvidos os respectivos descritores. Esses associam o objeto de conhecimento (conteúdos curriculares) às operações desenvolvidas pelos estudantes (competências e habilidades) (PESTANA, 1998; BRASIL, 2002; RABELO, 2013).

De acordo com Araújo e Luzio (2005, p. 15), “[...] cada descritor dá origem a diferentes itens e, a partir das respostas dadas a eles, verifica-se o que os alunos sabem e conseguem fazer com os conhecimentos adquiridos”. Assim, conforme Rabelo (2013), os descritores são capazes de expressar quais foram os saberes significativos desenvolvidos pelos alunos durante o processo de ensino-aprendizagem.

Desse modo, a primeira versão das matrizes de referência do Saeb, criada em 1997, englobava as disciplinas Língua Portuguesa, Matemática, Ciências, Física, Química e Biologia, avaliadas nesta edição. Já a segunda versão, em 2001, passou a enfatizar somente Língua Portuguesa e

Matemática, devido ao “novo foco” dado pelo sistema. Por essa razão, vale a pena destacar as diferenças e/ou similaridades das versões com referência a essas duas disciplinas.

3.4.1 Língua Portuguesa

Na primeira versão, com relação à disciplina Língua Portuguesa, os conteúdos foram divididos nos seguintes temas: a) práticas de leitura de textos; b) análise e reflexão linguística; e c) práticas de produção de textos (BRASIL, 1998). Nos dois primeiros itens, a distribuição dos descritores foi feita de acordo com os temas, que, por sua vez, foram subdivididos em tópicos. No terceiro, por se tratar de uma produção textual, não havia descritores. A seguir, o Quadro 3 ilustra a distribuição desses descritores.

QUADRO 3 – Distribuição dos descritores da Matriz de Referência de Língua Portuguesa do Saeb 1997

Tema	Tópicos	Descritores
4ª série do ensino fundamental		
Práticas de leitura de textos	Estratégias de leitura determinadas pelos diferentes objetivos da própria leitura	D1 ao D15
	Articulação texto e contexto	D16 ao D18
	Utilização de mecanismos básicos de coesão no processamento do texto	D19 ao D24
	Relações na progressão temática do texto	D25 ao D28
	Relações entre recursos expressivos e efeitos de sentido	D29 ao D36
Análise e reflexão linguística	Texto e gênero	D37 ao D43
	Variação linguística	D44 ao D47
	Operações linguísticas de estabelecimento de relações entre forma e sentido	D48 ao D53
	Paradigmas linguísticos (regras, classes)	D54 ao D59
8ª série do ensino fundamental		
Práticas de leitura de textos	O gênero e suas características	D1 ao D15
	Sentidos constituídos em função das características do portador, do gênero e/ou do enunciador	D16 ao D23
	Estratégias de leitura	D24 ao D43
	Relação entre textos	D44 ao D48
	Mecanismos de coesão referencial no processamento do texto.	D49 ao D53
	Progressão temática e um texto	D54 ao D61
Análise e reflexão linguística	Recursos expressivos recorrentes em função da estratégia de construção textual ou do gênero em que o texto se insere (adjetivação; nominalização; pronominalização; apassivação; omissão do agente; ordem direta ou indireta; impessoalidade, coloquialidade)	D62 ao D64
	Relações texto, gênero e contexto	D68 ao D70
	Operações linguísticas de estabelecimento de relações entre forma e sentido	D71 ao 76
	Variação linguística	D77 ao D85
3ª série do ensino médio		
Práticas de leitura de textos	Características e estratégias próprias de cada gênero	D1 ao D14
	Valores e sentidos constituídos pelo texto em função das características do portador e/ou enunciador	D15 ao D22
	Estratégias de leitura	D23 ao D47
	Relação entre texto e contexto, e entre texto e outros textos implicados pelo primeiro	D48 ao D58
	Mecanismos de coesão referencial no processamento do texto	D59 ao D64
	Índices formais de estabelecimento de relação sintático-semântica na progressão Temática	D65 ao D91
	Progressão temática e estratégia argumentativa	D92 ao D98
	Contexto histórico-social na análise textual	D99 ao D101
	Estratégias de construção da verossimilhança do universo ficcional	D102 ao D106
	Recursos expressivos utilizados na produção de um texto	D107 ao D121
Análise e reflexão linguística	Recursos expressivos recorrentes em função da estratégia de construção textual ou do gênero em que o texto se insere (adjetivação; nominalização; apassivação; modalização; omissão do agente; ordem direta ou indireta; paráfrase; impessoalidade)	D122 ao D124
	Construção da imagem do locutor e do alocutário	D125 ao D127
	Relações de intertextualidade entre dois enunciados (citação, paródia, articulação temática, associação lexical ou de estilo)	D128 ao D130
	Operações linguísticas de estabelecimento de relações entre forma e construção de sentido	D131 ao D153
	Variação linguística	D154 ao D163
	O texto enquanto objeto sociohistoricamente construído	D164 ao D171
	Funções da linguagem e atos de fala	D172 ao D175

Fonte: Elaborado pelos autores a partir de Brasil (1998).

Como visto, ao mudar de série, a contagem dos descritores se iniciava novamente. O descritor³⁸ D1 da 8ª série do ensino fundamental, por exemplo, era diferente do descritor D1 da 3ª série do ensino médio. Assim, para a 4ª e 8ª séries do ensino fundamental e 3ª série do ensino médio, foram encontrados 59, 85 e 175 descritores distribuídos, respectivamente, totalizando, portanto, 243 descritores para a primeira versão da matriz de referência de Língua Portuguesa do Saeb.

Ressalta-se, ainda, que, apesar de os temas serem comuns a todas as séries avaliadas, os tópicos eram diferentes. Conseqüentemente, os descritores também foram distintos, uma vez que contemplam as habilidades esperadas pelo estudante em determinado conteúdo de acordo com a série cursada. Assim, na medida em que eles avançavam nas etapas, os descritores também eram aprofundados conforme as especificidades dos ciclos.

Na segunda versão,³⁹ extinguiu-se a organização por temas/áreas. A distribuição dos descritores passou a ser realizada de acordo com os tópicos abordados que, desde então, são comuns para todas as séries avaliadas, conforme disposto no Quadro 4.

QUADRO 4 – Distribuição dos descritores da Matriz de Referência de Língua Portuguesa do Saeb 2001

Tópicos	Descritores		
	4ª série do ensino fundamental	8ª série do ensino fundamental	3ª série do ensino médio
Procedimentos de leitura	D1, D3, D4, D6, D11	D1, D3, D4, D6, D14	D1, D3, D4, D6, D14
Implicações do suporte, do gênero e/ou do enunciador na compreensão do texto	D5, D9	D5, D12	D5, D12
Relação entre textos	D15	D20, D21	D20, D21
Coerência e coesão no processamento do texto	D2, D7, D8, D12	D2, D7, D8, D9, D10, D11, D15	D2, D7, D8, D9, D10, D11, D15
Relações entre recursos expressivos e efeitos de sentido	D13, D14	D16, D17, D18, D19	D16, D17, D18, D19
Variação linguística	D10	D13	D13

Fonte: Elaborado pelos autores a partir de Brasil (2002).

Ao comparar as duas versões da matriz de referência de Língua Portuguesa, verifica-se, na última, não só um número menor de descritores, com um total de 57, como também aqueles que são comuns as três séries avaliadas. Entretanto, há diferenças quanto ao grau de complexidade desses descritores. Por isso, de um mesmo descritor podem derivar itens de

³⁸ Os descritores estão disponíveis no *site*: <http://www.dominiopublico.gov.br/download/texto/me002747.pdf>

³⁹ Disponível em: https://download.inep.gov.br/download/saeb/2001/Miolo_Novas_Perspectivas2001.pdf

graus de complexidades distintos, seja do ponto de vista do objeto analisado (texto), seja da tarefa (BRASIL, 2002).

Além do mais, desde 2001, o Saeb optou por ter a leitura como foco de avaliação, justificativa que foi dada por motivos “meramente operacionais” (BRASIL, 2002). Diante disso, extinguiram-se os conteúdos referentes a práticas de produções de textos, abordados na primeira versão. Essa configuração se mantém na versão atual da matriz de referência de Língua Portuguesa do Saeb.⁴⁰

3.4.2 Matemática

Com relação à matriz de referência de Matemática do Saeb, em sua primeira versão, propôs-se a proposição e solução de situações-problema como forma primordial para avaliar, de maneira satisfatória, as competências que evidenciavam a aprendizagem dos conteúdos matemáticos considerados básicos para o ensino fundamental e médio (BRASIL, 1998). Nesse caso, diferentemente do que ocorreu em Língua Portuguesa, os temas eram distintos para cada série avaliada e subdivididos em tópicos, como ilustra o Quadro 5.

⁴⁰A versão atual da Matriz de Referência do Saeb está disponível no *site*: https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_exames_da_educacao_basica/matriz_de_referencia_de_lingua_portuguesa_e_matematica_do_saeb.pdf

QUADRO 5 – Distribuição dos descritores da Matriz de Referência de Matemática do Saeb 1997

Tema	Tópicos	Descritores
4ª série do ensino fundamental		
Geometria	Espaço - localização, movimentação e representação (pontos de referência)	D1 ao D12
	Formas - bidimensionais e tridimensionais (elementos e propriedades)	D13 ao D31
	Curvas - região interior e exterior	D32 ao D37
Medidas	Significado e unidades de medida - comprimento, superfície, capacidade e massa	D38 ao D45
	Unidades de medida de tempo	D46 ao D48
	Unidades de medida de temperatura	D49 ao D50
	Unidades de medida do sistema monetário	D51 ao D54
Números	Números naturais	D55 ao D58
	Sistemas de numeração	D59 ao D60
	Números racionais positivos: representação decimal	D61 ao D64
Operações	Significado das operações: adição, subtração, multiplicação, divisão	D65 ao D70
	Propriedades das operações: adição, subtração, multiplicação, divisão	D71 ao D81
	Cálculo mental	D82 ao D84
Estatística	Lista, tabela e gráfico	D85 ao D87
8ª série do ensino fundamental		
Números	Números naturais e operações - inteiros, racionais e reais	D1 ao D3
	Números inteiros e operações	D4 ao D11
	Números racionais e operações	D12 ao D21
	Operações algébricas	D22 ao D26
Geometria e medidas	Retas	D27 ao D30
	Ângulos	D31 ao D35
	Figuras planas	D36 ao D48
	Figuras espaciais	D49 ao D56
Estatística	Noções de proporcionalidade, porcentagem e juros	D57 ao D67
3ª série do ensino médio		
Geometria euclidiana	Retas e planos	D1 ao D2
	Áreas de figuras planas	D3 ao D4
	Sólidos	D5 ao D8
Geometria analítica plana	Ponto e reta	D9 ao D16
	Circunferência	D17 ao D21
	Cônicas	D22 ao D23
Trigonometria	Não tem	D24 ao D32
Relações e funções	Conceituação e propriedades.	D33 ao D38
	Funções, equações, inequações de 1º grau	D39 ao D46
	Módulo	D47 ao D48
	Exponencial e logaritmo	D49 ao D53
	Funções trigonométricas	D54
Matrizes, determinantes e sistemas lineares	Matrizes	D55 ao D57
	Determinantes	D58 ao D59
	Sistemas lineares	D60 ao D61
Análise combinatória e probabilidade	Não tem	D62 ao D66
Estatística	Não tem	D67 ao D68
Números Complexos	Não tem	D69 ao D77
Polinômios e equações polinomiais	Não tem	D78 ao D85

Fonte: Elaborado pelos autores a partir de Brasil (1998).

Como visto, para a 4ª e 8ª séries do ensino fundamental e 3ª série do ensino médio, foram identificados 87, 67 e 85 descritores, respectivamente, somando 239 descritores para a primeira versão da matriz de referência de Matemática do Saeb.

Na segunda versão, por sua vez, similarmente ao que aconteceu com Língua Portuguesa, optou-se por uma organização que consiste no uso dos tópicos e seus respectivos descritores. Diferentemente da anterior, esses tópicos passaram a ser comuns a todas as séries avaliadas. São eles: a) espaço e forma; b) grandezas e medidas; c) números, operações/álgebra e funções; e d) tratamento da informação (BRASIL, 2002), conforme ilustra o Quadro 6.

QUADRO 6 – Distribuição dos descritores da Matriz de Referência de Matemática do Saeb 2001

Tópicos	Descritores		
	4ª série do ensino fundamental	8ª série do ensino fundamental	3ª série do ensino médio
Espaço e forma	D1, D2, D3, D4, D5	D1, D2, D3, D4, D5, D6, D7, D8, D9, D10, D11	D1, D2, D3, D4, D5, D6, D7, D8, D9, D10
Grandezas e medidas	D6, D7, D8, D9, D10, D11, D12	D12, D13, D14, D15	D11, D12, D13
Números e operações/álgebra e funções	D13, D14, D15, D16, D17, D18, D19, D20, D21, D22, D23, D24, D25, D26	D16, D17, D18, D19, D20, D21, D22, D23, D24, D25, D26, D27, D28, D29, D30, D31, D32, D33, D34, D35	D14, D15, D16, D17, D18, D19, D20, D21, D22, D23, D24, D25, D26, D27, D28, D29, D30, D31, D32, D33
Tratamento da informação	D27, D28	D36, D37	D34, D35

Fonte: Elaborado pelos autores a partir de Brasil (2002).

Ressalta-se, portanto, que, ao comparar as duas versões das Matrizes de Referência de Matemática do Saeb, verifica-se, na última, uma quantidade menor de descritores, totalizando 100, bem como a existência de descritores comuns as três séries avaliadas. Além do mais, observa-se que essa configuração também foi mantida para a versão atual da matriz de referência da disciplina.

Em conclusão, é preciso destacar que as matrizes de referência são, ainda, as responsáveis por nortear a elaboração dos itens que irão compor os testes de desempenho escolar do Saeb. Entretanto, a partir do ciclo de 2019, verificou-se, novamente, uma reestruturação no sistema para se adequar à Base Nacional Comum Curricular (BNCC) (BRASIL, 2018a). Essa base se tornará referência para formulação dos itens do 2º (Língua Portuguesa e Matemática) e do 9º anos do ano ensino fundamental, para os testes de Ciências da Natureza e Ciências Humanas, que passaram a ser avaliados a partir dessa edição de forma amostral.

3.5 CONSIDERAÇÕES PARCIAIS

Levando em consideração a trajetória do Saeb, este capítulo empenhou-se em discutir as principais modificações sofridas pelo sistema que contribuíram para o seu processo de desenvolvimento e institucionalização. Para isso, realizou-se um estudo das alterações mais importantes, desenvolvidas no período compreendido entre os anos de 1990 e 2001, que auxiliaram na efetivação de uma Política Nacional de Avaliação: a adoção da TRI e das Matrizes de Referência.

Nota-se que uma das maiores contribuições da preferência pelo uso da metodologia da TRI em relação à TCT consiste na possibilidade de comparação dos resultados ao longo do tempo. Dessa alteração surgiu uma nova utilidade significativa para o sistema que, além de fornecer um diagnóstico da educação básica, também é indutor de políticas públicas que exercem funções importantes para a avaliação, como o juízo de valor e a tomada de decisão.

Além do mais, no que se refere à adoção dessas políticas, o ideal seria que fossem disponibilizadas unidades comparativas, já que assim seria viável avaliar se está ou não havendo melhorias em determinado indicativo de desempenho por meio de sua implementação. Do contrário, não seria possível promover discussões que permitissem corrigir a própria política.

Compreende-se, portanto, que, apesar dos avanços com a adoção da TRI e do caráter censitário da Prova Brasil, a comparação dos resultados pode levar ao uso exacerbado do sistema como política de responsabilização (do inglês, *accountability*), apequenando-se, assim, o que o Saeb se propõe fazer. Dessa forma, aponta-se para a necessidade de estudos que investiguem os *usos e consumos* (GINZBURG, 1989) que se têm feito a partir das médias de desempenho escolar provenientes da Prova Brasil.

Percebeu-se, ainda, que aos poucos a participação das Secretarias Estaduais e Municipais de Educação foi se esvaindo, uma vez que, com a adoção das Matrizes de Referência do Saeb, não foi mais necessária a realização de consulta às Secretarias. Aliás, todos os serviços passaram a ser coordenados pelo Inep/MEC e/ou pelas fundações responsáveis por sua operacionalização.

As Matrizes de Referência do Saeb, apesar de não englobar a totalidade dos currículos (BRASIL, 2002), são responsáveis pelo estabelecimento das competências e habilidades em

cada disciplina e ciclo escolar avaliado. É importante destacar que, desde a atualização de 2001, manteve-se a configuração para as disciplinas de Língua Portuguesa e Matemática, que nunca deixaram de ser avaliadas pelo sistema e também compuseram o quadro de outras avaliações que surgiram posteriormente.

Ademais, há *indícios* (GINZBURG, 1989) de que a notoriedade das Matrizes de Referência do Saeb colaborou, inclusive, para a adequação dos sistemas de ensino à BNCC, que irá definir o conjunto de aprendizagens “essenciais” não só para a avaliação, mas também ao longo de todas as etapas e modalidades da educação básica (BRASIL, 2018a). Inicialmente, ela foi utilizada pelo Saeb como referência para a elaboração dos itens do 2º ano do ensino fundamental (Língua Portuguesa e Matemática) e do 9º ano para os testes de Ciências da Natureza e Ciências Humanas, aplicados de forma amostral.

Vale a pena destacar que, conforme disposto na Portaria n.º. 10, de 8 de janeiro de 2021 (BRASIL, 2021a), o Saeb passará por ajustes técnico-pedagógicos, tendo em vista a implementação da BNCC. Além disso, a portaria prevê que todos os anos/séries da educação básica, das escolas públicas e privadas, serão avaliados anualmente em todas as áreas de conhecimento (BRASIL, 2021c).

Verificou-se, também, a influência da legislação brasileira que estabelece o Saeb como ação prioritária (BRASIL, 1993, 2001), cuja avaliação tem o objetivo de monitorar a garantia do padrão de qualidade da educação básica (BRASIL, 1988, 1996, 2001). Atenta-se, portanto, para a uma crescente utilização dos seus dados como um forte auxílio na tomada de decisões e implementação de políticas, inclusive, para a distribuição de recursos educacionais.

Nesse sentido, o Plano Nacional de Educação (PNE) se caracteriza como um desdobramento do Saeb como política de avaliação. A partir do PNE 2001-2011, instituído pela Lei n.º. 10.172, de 9 de janeiro de 2001 (BRASIL, 2001), os aspectos que dizem respeito à avaliação, como os processos de coleta e difusão dos dados, são considerados “[...] instrumentos indispensáveis para a gestão do sistema educacional e melhoria do ensino” (BRASIL, 2001, p. 4).

Assim, a avaliação é, portanto, uma forma de garantir a supervisão e inspeção da qualidade do ensino ofertado pelas escolas (BRASIL, 2001). Configura-se como um elemento importante capaz de direcionar outras ações do plano, como objetivos e metas. Os resultados produzidos

pelo Saeb são utilizados para identificar o cumprimento desses objetivos e metas, o que revela o impacto do sistema e seus desdobramentos na política educacional brasileira.

Esse aspecto também é evidenciado pelo PNE 2014-2024, aprovado por meio da Lei n°. 13.005, de 25 de junho de 2014 (BRASIL, 2014), atualmente em vigor. O Saeb, além de acompanhar as metas, também é utilizado para fixá-las, juntamente com o Ideb, no qual também se utilizam as médias de desempenho escolar do sistema (Prova Brasil).

Dessa forma, aponta-se para a necessidade de estudos que possam investigar como o Saeb se torna um instrumento de monitoramento da qualidade da educação básica. Além disso, surgem alguns questionamentos, sobretudo com a intenção de compreender quais foram e como se deram as primeiras iniciativas de responsabilização e suas implicações para a educação no Brasil.

Finalmente, mais uma vez tornou-se evidente a importância das entidades brasileiras da área de avaliação, como a FCC e a Cesgranrio. Ademais, nota-se que o Saeb contribuiu, gradativamente, para uma valorização da avaliação que aconteceu por vários órgãos durante o período estudado, por exemplo, pelo Consed e pela Undime. Do mesmo modo, foram incentivadas iniciativas estaduais e municipais de avaliação externa, além de impulsionar o aparecimento de outras avaliações nacionais, como o Enem.

Espera-se, portanto, que este capítulo venha esclarecer como se deu o processo de desenvolvimento e institucionalização do Saeb e de que maneira a adoção da TRI e das Matrizes de Referência, além da legislação brasileira, contribuíram para a sua consolidação e aperfeiçoamento e, conseqüentemente, a efetivação de uma Política Nacional de Avaliação.

CAPÍTULO 4

O SAEB CENSITÁRIO E A EFETIVAÇÃO DE UMA CULTURA POLÍTICA DE AVALIAÇÃO PAUTADA NA RESPONSABILIZAÇÃO DA EDUCAÇÃO BÁSICA BRASILEIRA

4.1 INTRODUÇÃO

Conforme aludido nos capítulos anteriores, a partir da década de 1990, as avaliações externas aplicadas em larga escala ganharam destaque nas discussões acadêmicas do Brasil. Alguns estudos atribuem esse fato à influência de uma onda mundial de reformas neoliberais que propuseram uma associação entre a qualidade da educação e os índices de produtividade do ensino em todo o mundo (MENDONÇA, 2014; PASCHOALIN, 2019; FREITAS, 2007, 2016).

Essas avaliações são vistas pelo Estado como um instrumento de “prestação de contas” à sociedade (do inglês, *accountability*), em busca de uma maior transparência na distribuição dos recursos públicos destinados à educação. Em teoria, é esperado que os seus resultados auxiliem na autoavaliação das escolas, bem como dos sistemas de ensino, no aperfeiçoamento dos projetos educacionais desenvolvidos, na realização de diagnósticos do ensino que é ofertado, na elaboração de políticas educacionais etc.

Seguindo esse raciocínio, a avaliação externa, conforme Freitas (2005), tornou-se uma condição necessária ao planejamento e à administração educacional, a fim de garantir uma gestão menos governamental (ou estatal) e mais “pública” e democrática. Dessa forma, coube à avaliação aferir um padrão mínimo de qualidade ofertado pelas instituições escolares em todas as instâncias governamentais (municipal, estadual e federal).

Em consoante com Paschoalin (2019), na segunda metade da década, início do governo Fernando Henrique Cardoso (FHC), o Estado passou uma por uma reforma cujo intuito era minimizar sua atuação como prestador de serviços públicos. Desde então, esse passou a “compartilhar” suas responsabilidades.

Entende-se que diante dessa nova configuração de Estado há uma desresponsabilização deste quanto aos resultados e, ao mesmo tempo, à regulação. A avaliação serve, dessa forma, como instrumento para que o Estado tenha controle sobre as políticas educacionais, visando a resultados, mas responsabilizando as instituições de ensino por eles (PASCHOALIN, 2019, p. 48).

A partir daí, iniciou-se no país uma nova concepção de avaliação que, além do seu uso como instrumento de medida da qualidade da educação, fez-se também de regulação. Desse modo, por meio da avaliação externa aplicada em larga escala, recomendada por organizações multilaterais internacionais, como o Banco Mundial, o Estado adquiriu certo controle sobre as políticas educacionais, passando de “executor” a “regulador” (PASCHOALIN, 2019).

O contexto nacional era de intensificação e internacionalização do capitalismo. Dentre os princípios declarados, para Sousa (1998, p. 167), destacam-se a equidade e a responsabilização dos agentes educativos:

Avaliar com equidade significa analisar se a qualidade da educação que está sendo oferecida atende igualmente a todos os setores sociais. E a responsabilidade para garantir a equidade é de toda a sociedade, sobretudo dos governos, e exige o compromisso expresso e definitivo dos agentes do processo educacional.

Além desses, Paschoalin (2019) ressalta outros princípios norteadores, como produtividade, desempenho, incentivo ao mérito e atendimento às necessidades do mercado de trabalho. Ademais, a autora chama a atenção para o aumento das avaliações de programas e ações do governo que ocorreram durante essa década, inclusive, o próprio Sistema de Avaliação da Educação Básica (Saeb).

Nessa perspectiva, este capítulo tem como objetivo realizar um levantamento sobre as primeiras iniciativas de uma política de avaliação nacional pautada nos processos de prestação de contas e de responsabilização, a partir dos resultados do Saeb. Para isso, considerou-se o período compreendido entre os anos de 2005 e 2019, desde a primeira aplicação da avaliação nacional censitária até o último ciclo de avaliação, realizada no momento em que esta pesquisa foi desenvolvida. O interesse, assim, é responder à seguinte questão: qual a contribuição do Saeb para as primeiras iniciativas de responsabilização na educação básica brasileira?

Para a exposição, além desta introdução, este capítulo se estrutura em outras quatro seções. Na primeira, abarca os instrumentos atuais do Saeb. Na segunda, aborda a incorporação do modelo de prestação de contas na educação básica do país. Na terceira, discute acerca da

Prova Brasil e das primeiras iniciativas de responsabilização escolar. Na quarta e última seção, apresentam-se os apontamentos finais específicos, encerrando o capítulo.

4.2 OS INSTRUMENTOS DO SAEB

Desde a sua implementação, em 1990, o Saeb é aplicado de maneira amostral, ou seja, apenas parte representativa da população é selecionada para a realização das provas. Em sua primeira edição, o sistema avaliou o desempenho escolar dos estudantes das escolas públicas que cursaram as 1ª, 3ª, 5ª e 7ª séries do ensino fundamental. As disciplinas avaliadas foram Língua Portuguesa, Matemática, Ciências Naturais e Redação, cuja formulação dos itens era de acordo com os currículos dos sistemas estaduais. Esse formato foi mantido em 1993.

Em 1995, o Saeb passou a avaliar a terminalidade dos ciclos escolares (4ª e 8ª séries do ensino fundamental e 3ª série do ensino médio) e a realizar levantamento de dados sobre os estudantes e as escolas por meio da aplicação de questionários contextuais. Além disso, optou por proceder à avaliação somente com as disciplinas Língua Portuguesa e Matemática.

A partir de 1997, a análise de desempenho dos estudantes passou a ser realizada de acordo com o seu nível nas escalas de proficiência, construídas para cada área do conhecimento, em conformidade com a Matriz de Referência. Nesse ciclo, em especial, compuseram o quadro de avaliação: Ciências, Física, Química e Biologia. No ano seguinte, em 1999, também foram realizados testes de História e Geografia.

Já em 2001, o ciclo ficou conhecido pelo “novo foco” dado às disciplinas Língua Portuguesa e Matemática que, como discutido, nunca deixaram de ser avaliadas. Em 2003, o sistema manteve o formato da edição anterior.

No ano de 2005, o Saeb passou a ser censitário. O objetivo dessa mudança era a atribuição de um caráter mais universal à avaliação da educação básica no Brasil (RABELO, 2013). Além disso, anos mais tarde, o sistema introduziu a avaliação da educação infantil.

4.2.1 A Avaliação Nacional da Educação Básica (Aneb) e a Avaliação Nacional do Rendimento Escolar (Anresc/Prova Brasil)

Em 2005, o Saeb teve sua avaliação reestruturada por meio da Portaria Ministerial nº. 931, de 21 de março de 2005 (BRASIL, 2005a), quando passou a ser composto pela Avaliação Nacional da Educação Básica (Aneb) e pela Avaliação Nacional do Rendimento Escolar (Anresc). Ambas são avaliações externas aplicadas em larga escala, cujo órgão responsável é o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) – autarquia federal do Ministério da Educação (MEC). No entanto, existem diferenças importantes entre elas, principalmente no que se refere à aplicação, público-alvo e finalidades.

A Aneb é aplicada bianualmente nas escolas públicas e privadas, desde que atendam ao mínimo de dez estudantes matriculados por turma. Por meio dessa avaliação, mantiveram-se os objetivos, características e procedimentos por amostragem, realizados até então pelo sistema. Seus resultados são denominados estratos amostrais do Saeb. Conforme disposto na Portaria nº. 89, de 25 de maio de 2005 (BRASIL, 2005c), estabeleceram-se os seguintes objetivos específicos para a Aneb:

I - Aplicar instrumentos (provas e questionários) em uma amostra representativa de alunos de 4ª e 8ª séries do Ensino Fundamental e de 3ª série do Ensino Médio das escolas das redes pública e privada, localizadas nas zonas rural e urbana, distribuídas nas 27 Unidades da Federação; II - Aplicar provas de Língua Portuguesa e de Matemática, construídas a partir da Matriz de Referência do Sistema Nacional de Avaliação da Educação Básica; III - Produzir informações sobre o desempenho dos alunos, assim como sobre as condições intra e extraescolares que incidem sobre o processo de ensino e aprendizagem; IV - Avaliar a qualidade, a equidade e a eficiência dos sistemas e redes de ensino brasileiros; e V - Manter a construção de séries históricas, permitindo comparabilidade entre anos e entre séries escolares (BRASIL, 2005c, p. 29).

Como visto, as informações produzidas pela Aneb não serão utilizadas para identificar escolas, turmas, alunos, professores ou diretores (BRASIL, 2005a). Pelo contrário, o interesse é produzir informações sobre o desempenho dos estudantes e das condições intra e extraescolares que incidem sobre o processo de ensino e aprendizagem, o que permite a construção de períodos históricos (BRASIL, 2005c). Assim, a Aneb é aplicada a todas as etapas escolares avaliadas pelo Saeb.

A Anresc, mais conhecida como Prova Brasil, inicialmente, passou a avaliar de forma censitária apenas as escolas públicas. Seus resultados são denominados estratos censitários do Saeb. De acordo com a Portaria Ministerial nº. 931 (BRASIL, 2005a), foram estabelecidos os seguintes objetivos gerais:

[...] a) avaliar a qualidade do ensino ministrado nas escolas, de forma que cada unidade escolar receba o resultado global; b) contribuir para o desenvolvimento, em todos os níveis educativos, de uma cultura avaliativa que estimule a melhoria dos padrões de qualidade e equidade da educação brasileira e adequados controles sociais de seus resultados; c) concorrer para a melhoria da qualidade de ensino, redução das desigualdades e a democratização da gestão do ensino público nos estabelecimentos oficiais, em consonância com as metas e políticas estabelecidas pelas diretrizes da educação nacional; e d) oportunizar informações sistemáticas sobre as unidades escolares. Tais informações serão úteis para a escolha dos gestores da rede a qual pertençam (BRASIL, 2005a, p. 17).

Como discutido, ao contrário da Aneb, a Prova Brasil consiste em gerar os resultados por escola, o que, em teoria, estimula a “[...] melhoria dos padrões de qualidade e equidade da educação brasileira” (BRASIL, 2005a, p. 17). Assim, as instituições escolares que participarem dessa avaliação receberão os resultados da média geral da escola e do percentual de estudantes por nível da escala de proficiência e habilidades do Saeb (BRASIL, 2005b).

Ainda nesse ano, por meio da Portaria n.º. 69, de 4 de maio de 2005 (BRASIL, 2005b), estabeleceu-se a sistemática para a realização da Prova Brasil com os seguintes objetivos específicos:

I - aplicar a avaliação nas escolas públicas, localizadas em zona urbana, que possuam pelo menos 30 alunos matriculados em cada uma das séries avaliadas; II - a aplicação nas escolas definidas no inciso I irá ocorrer nas turmas de 4ª e 8ª séries do Ensino Fundamental Regular de 08 anos e nas turmas de 5º e 9º ano em escolas que estejam organizadas no regime de 09 anos para o Ensino Fundamental; III - serão aplicados testes de Língua Portuguesa com foco nas competências e habilidades de leitura definidas na Matriz de Especificações do Sistema de Avaliação da Educação Básica; e IV - oportunizar informações sistemáticas sobre as unidades escolares. Tais informações serão úteis para os gestores da rede a qual pertençam às escolas avaliadas (BRASIL, 2005b, p. 13).

Dessa forma, a Prova Brasil é aplicada somente para a escola pública e restrita aos 5º e 9º anos do ensino fundamental. A princípio, conforme a referida portaria, o critério mínimo de participação foi de 30 estudantes matriculados por ano. Em 2007, alterou-se para pelo menos 20 alunos matriculados em cada uma das turmas de cada escola (BRASIL, 2007b). Por último, desde 2017, adotou-se como critério o mínimo de dez estudantes por turma. Vale ressaltar que, nesse ano, também foi atribuído o caráter censitário para o ensino médio⁴¹ (BRASIL, 2017a, 2017b).

⁴¹Embora a avaliação tenha passado a ser censitária para o ensino médio, ela não deveria ser confundida com a Prova Brasil, que continuou restrita aos anos avaliados pelo sistema no ensino fundamental. Essa etapa foi denominada Saeb censitário.

Esse novo formato de avaliação, amostral e censitário, permitiu ao Inep combinar as médias de desempenho dos estudantes no Saeb com as taxas de promoção, retenção e evasão apuradas pelo Censo Escolar da Educação Básica. Em consoante com a Portaria n.º 316, de 4 de abril de 2007 (BRASIL, 2007b, p. 1), o Censo é “[...] realizado em regime de colaboração entre a União, os Estados e os municípios, a partir de um processo descentralizado de coleta de dados individualizados de alunos, turmas, profissionais de educação e de escolas”. É considerado a pesquisa estatística educacional mais importante do país (BRASIL, 2021e) e também está sob a coordenação do Inep.

Segundo o site oficial,⁴² sua finalidade é disponibilizar as informações para que os atores educacionais possam compreender a situação do ensino do país e de outras instâncias governamentais. Os dados são apurados anualmente e também servem de base para a distribuição dos recursos do Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (Fundeb) (BRASIL, 2007b). É preciso destacar que todo recurso gerado pelo fundo, independentemente da fonte de origem, é redistribuído para aplicação de forma exclusiva na educação básica pública.

O Censo Escolar conta com a participação de todas as escolas brasileiras, tanto da rede de ensino pública como privada. Além disso, abrange as diferentes etapas e modalidades da educação básica e profissional, como: o ensino regular (educação infantil, ensino fundamental e médio); a educação especial; a Educação de Jovens e Adultos (EJA); e a educação profissional (cursos técnicos e cursos de formação inicial continuada ou qualificação profissional) (BRASIL, 2021e).

Ademais, ressalta-se que, desde 2019, extinguiu-se o uso das siglas Aneb e Anresc/Prova Brasil. Desse modo, todas as avaliações passaram a ser identificadas pelo nome Saeb, acompanhado das etapas, áreas de conhecimento e tipos de instrumentos envolvidos (BRASIL, 2020a).

⁴² Site oficial. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-escolar>. Acesso em: 9 fev. 2022.

4.2.2 A Avaliação Nacional da Alfabetização (ANA)

Um dos componentes atuais do Saeb é a Avaliação Nacional de Alfabetização (ANA), instituída pela Portaria nº. 482, de 7 de junho de 2013 (BRASIL, 2013), que está vinculada ao Pacto Nacional pela Alfabetização na Idade Certa (Pnaic) – lançado em 2012 pelo Governo Federal. Dentre as ações do Pnaic, uma consiste em “[...] garantir que todos os estudantes dos sistemas públicos de ensino estejam alfabetizados, em Língua Portuguesa e em Matemática, até o final do 3º ano do ensino fundamental” (BRASIL, 2012, p. 2). Desse modo, a criação da ANA teve como objetivos:

I - avaliar a qualidade, a equidade e a eficiência (incluindo as condições de oferta) do Ciclo de Alfabetização das redes públicas; e II - produzir informações sistemáticas sobre as unidades escolares, de forma que cada unidade receba o resultado global (BRASIL, 2013, p. 17).

A ANA teve sua aplicação regulamentada pela Portaria nº. 410, de 22 de julho de 2016 (BRASIL, 2016), na qual foram estabelecidos os seguintes objetivos específicos:

I - estimular a melhoria dos padrões de qualidade e equidade da educação brasileira; II - subsidiar a elaboração de políticas educacionais para o ciclo de alfabetização; III - aferir o nível de alfabetização e letramento em Língua Portuguesa e alfabetização em Matemática dos estudantes ao final do 3º ano do ensino fundamental, por meio de testes de leitura, escrita e matemática; IV - produzir indicadores sobre o contexto em que se realiza o trabalho escolar (BRASIL, 2016, p. 393).

A participação para a ANA é a mesma das demais avaliações do Saeb, cujo critério mínimo estabelecido é de dez estudantes matriculados em turmas regulares do 3º ano do ensino fundamental, organizado no regime de nove anos (BRASIL, 2016). Contudo, a partir de 2019, a avaliação passou a ser realizada no 2º ano do ensino fundamental, conforme disposto na Portaria nº. 366, de 29 de abril de 2019 (BRASIL, 2019).

Para realizar esse diagnóstico ao final do ciclo de alfabetização, a formulação dos itens, desde 2019, fundamenta-se na Base Nacional Comum Curricular (BRASIL, 2019). Nos anos anteriores, foram utilizadas Matrizes de Referência de Língua Portuguesa (Leitura e Escrita) e Matemática próprias para ANA.⁴³

A técnica de montagem adotada para os cadernos de prova da ANA também é a mesma empregada pelas outras avaliações do Saeb, denominada Blocos Incompletos Balanceados (BIB). Em síntese, essa técnica permite a elaboração de vários cadernos de prova que, embora

⁴³As matrizes utilizadas na edição da ANA 2016 são as mesmas das edições dos anos de 2013 e 2014.

distintos, apresentam o mesmo grau de dificuldade. Isso permite avaliar um número maior de itens e, assim, uma maior variedade de habilidades (RABELO, 2013). Além das provas, a avaliação também aplica os questionários contextuais, cujo foco é a gestão escolar (BRASIL, 2018b).

Para aplicação dos cadernos de prova da ANA, de acordo com Brasil (2018b), em um primeiro dia, os estudantes são avaliados por meio de 20 itens de múltipla escolha nos testes de Leitura e três itens de resposta construídos nos testes de Escrita. No dia seguinte, serão aplicados mais 20 itens de múltipla escolha referentes à disciplina Matemática.

É importante destacar que a aplicação da ANA 2016, última edição com os resultados divulgados até o momento em que esta pesquisa foi realizada, deu-se por meio de um consórcio aplicador formado por três instituições externas ao Inep: Fundação Cesgranrio; Centro de Políticas Públicas e Avaliação da Educação (CAEd); e Fundação Getúlio Vargas (FGV) (BRASIL, 2016). Além da aplicação, “[...] o consórcio ficou responsável pela leitura óptica das folhas de resposta, consistência das bases de dados e pelo processamento de todos os testes aplicados na avaliação” (BRASIL, 2018b, p. 27).

Para uma análise de caráter mais qualitativo da ANA, a proficiência dos estudantes é posicionada em três escalas (Leitura, Escrita e Matemática),⁴⁴ nas quais o indicador utiliza 500 como média de desempenho e 100 de desvio padrão. Entretanto, apesar de apresentarem os mesmos valores para essas estatísticas, é preciso considerar que uma proficiência igual a 250 em Leitura é diferente de outra proficiência igual a 250 pontos em Matemática, por exemplo. Por isso é necessário estar atento aos padrões de desempenho, delimitados mediante a participação da escala. Por meio deles, pode-se desenvolver uma interpretação pedagógica das habilidades e competências alcançadas pelos estudantes ao longo dos anos.

Ademais, a ANA também traz um elemento novo que a distingue das demais modalidades de avaliação integradas ao Saeb (DICKEL, 2016), uma vez que se encontra acompanhada de um projeto de formação de professores. Além dos indicadores de nível socioeconômico (NSE) e da complexidade de gestão das escolas, são coletadas, ainda, informações sobre o grau de instrução do professorado, que resultam no Indicador de Adequação da Formação Docente (IAFD). Esse indicador foi “[...] criado pelo Inep para classificar os docentes em exercício na

⁴⁴Escala da ANA. Disponível em:

https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_exames_da_educacao_basica/escala_de_proficiencia_do_saeb_ensino_fundamental.pdf. Acesso em: 9 fev. 2022.

educação básica, considerando a etapa da educação básica em que lecionam e disciplina que ministram” (BRASIL, 2018b, p. 62).

4.2.3 Síntese das avaliações do Saeb

Conforme supraexposto, atualmente as avaliações do Saeb são aplicadas aos anos finais de cada ciclo escolar: 2º, 5º e 9º anos do ensino fundamental e 3º ano do ensino médio. Quanto à sua natureza, a avaliação é censitária para as escolas públicas (com exceção do 2º ano do ensino fundamental) e amostral para as escolas privadas. Além disso, em 2019, o sistema também começou a avaliar, em um estudo piloto, a educação infantil por meio de questionários eletrônicos aplicados exclusivamente para uma amostra de professores, diretores e secretários municipais de Educação (BRASIL, 2019, 2020a). A seguir, no Quadro 7, apresenta-se uma síntese das avaliações do Saeb, compreendendo os anos avaliados, os instrumentos e a natureza da aplicação.

QUADRO 7 – Síntese das avaliações do Saeb

Anos avaliados	Instrumentos	Aplicação
Educação infantil	Questionários (secretários municipais de educação, diretores e professores)	Amostral
	Não há testes	
2º ano do ensino fundamental	Não há questionários	Amostral
	Testes de Língua Portuguesa e Matemática	
5º ano do ensino fundamental	Questionários (secretários municipais de educação, diretores, professores e estudantes)	Censitária (públicas) Amostral (privadas)
	Testes de Língua Portuguesa e Matemática	
9º ano do ensino fundamental	Questionários (secretários municipais de educação, diretores, professores e estudantes)	Censitária (públicas) Amostral (privadas)
	Testes de Língua Portuguesa e Matemática e Testes Ciências da Natureza e Ciências Humanas (amostral)	
3ª e 4ª séries do ensino médio	Questionários (secretários municipais de educação, diretores, professores e estudantes)	Censitária (públicas) Amostral (privadas)
	Testes de Língua Portuguesa e Matemática	

Fonte: Elaboração dos autores.

Como visto, os estudantes são avaliados pelo seu desempenho nas disciplinas Língua Portuguesa e Matemática. Somente para o 9º ano do ensino fundamental são aplicados, de maneira amostral, testes em outras áreas do conhecimento, como Ciências da Natureza e Ciências Humanas.

Também são aplicados questionários contextuais aos secretários municipais de Educação, diretores, professores e estudantes. Essas informações, combinadas com os resultados dos testes, oferecem subsídios para a produção de indicadores de qualidade da educação básica (BRASIL, 2021a).

Desse modo, a partir de 2019, o Saeb passou a avaliar todas as etapas escolares da educação básica brasileira (educação infantil, ensino fundamental e ensino médio), tornando-se um importante instrumento norteador para as políticas de responsabilização implementadas no país a partir de 2005, com a inserção da avaliação censitária, o que será tratado com maior detalhamento nas seções seguintes.

4.3 A INCORPORAÇÃO DO MODELO DE PRESTAÇÃO DE CONTAS NA EDUCAÇÃO BÁSICA BRASILEIRA

Conforme discutido na seção anterior, em 2005, o Saeb passou por uma reestruturação que o dividiu em dois processos de avaliação: Aneb e Anresc/Prova Brasil. Enquanto a Aneb manteve as características e objetivos do sistema quanto ao delineamento amostral, a Prova Brasil assumiu a aplicação censitária, com abrangência universal aos estudantes de escolas públicas que cursam os anos finais de cada ciclo do ensino fundamental.

A criação da Prova Brasil possibilitou que fossem apresentadas “[...] informações que discriminam os resultados para cada município e escola participante” (WERLE, 2011, p. 786). Dessa forma, direcionou-se, ainda mais, o processo de avaliação para a obtenção de resultados em todos os níveis, etapas e modalidades de ensino.

Desde o seu primeiro ciclo de avaliação, houve uma efetiva participação das Secretarias Municipais e Estaduais de Educação (PESTANA, 1998) que, posteriormente, também se configuraram na Prova Brasil (FERNANDES, 2016). O Inep, responsável pelo planejamento e operacionalização das avaliações, ressalta que as informações coletadas pelo Saeb permitem

não só a realização de um diagnóstico da educação básica brasileira como também de outros fatores que podem interferir no desempenho do estudante (BRASIL, 2020c).

Com vistas à realização de tal diagnóstico, efetivou-se no país a instauração de uma política educacional baseada no modelo *accountability*, que utiliza a avaliação como instrumento de monitoramento. Essa política, pautada na responsabilização, materializa-se no Brasil por meio do Plano de Desenvolvimento da Educação (PDE). Trata-se de um conjunto de ações e programas governamentais voltados para a melhoria da qualidade da educação, por meio do qual se estabelecem “[...] inéditas conexões entre avaliação, financiamento e gestão” (BRASIL, 2008, p. 11).

O PDE introduziu os conceitos de responsabilização e mobilização social que, de acordo com o documento, tornam a escola “menos estatal” e mais “pública” (BRASIL, 2008). Além disso, trouxe um sentido mais operacional às avaliações, a partir da consolidação de metas de qualidade para todas as modalidades de ensino e em todas as esferas administrativas (estadual, municipal e federal).

Dessa preocupação, para o acompanhamento e monitoramento de tais metas, criou-se, em 2007, o Ideb, a principal referência da qualidade do ensino no Brasil (SOARES, D.; SOARES, T.; DOS SANTOS, 2020). O Ideb⁴⁵ é um indicador de qualidade educacional, expresso em uma escala de zero a dez, que combina em um só indicador as médias de desempenho da Prova Brasil com os indicadores de fluxo (promoção, retenção e evasão) obtidos pelo Censo Escolar. Em termos de diagnóstico, seu objetivo consiste em: “[...] (a) detectar escolas e/ou redes de ensino cujos alunos apresentem baixa performance em termos de rendimento e proficiência; (b) monitorar a evolução temporal do desempenho dos alunos dessas escolas e/ou redes de ensino” (BRASIL, 2007b, p. 2).

Calculado por escola, por rede de ensino e pelo país, foi possível fixar metas de desenvolvimento educacional de médio e curto prazo para cada uma dessas instâncias (BRASIL, 2008). Assim, com o intuito de obter um maior “comprometimento” das redes de ensino e das escolas com o objetivo de melhoria do Ideb (FERNANDES, 2016), estabeleceu-

⁴⁵Segundo Fernandes (2016), para o cálculo da nota do Ideb em um determinado ano, multiplica-se a proficiência média dos alunos da escola no Saeb daquele ano (N), padronizada para um valor entre zero e dez, pela taxa média de aprovação dos estudantes na etapa de ensino avaliada (P): $Ideb = NP$. Por exemplo, caso uma escola tenha obtido nota padronizada igual a sete no Saeb e taxa de aprovação de 90%, o Ideb dessa escola será $7 \times 0,9 = 6,3$.

se o Plano de Metas Compromisso Todos pela Educação, por meio do Decreto nº. 6.094, de 24 de abril de 2007 (BRASIL, 2007a).

O sistema de metas, inaugurado pelo Ideb⁴⁶ e pactuado entre o MEC e as Secretarias Municipais e Estaduais de Educação, em teoria, serviria para aumentar a mobilização da sociedade em favor da qualidade da educação (FERNANDES; GREMAUD, 2009). Para isso, uma das diretrizes do plano consistia em divulgar na escola e na comunidade os dados educacionais com ênfase no Ideb. Desse modo, ao “traduzir” o quanto as escolas, municípios e Estados avançaram em relação às suas metas, segundo Schneider e Nardi (2014), o índice passou a configurar-se como uma importante “prestação de contas” dos atores que permeiam o cotidiano das escolas, acerca das metas de qualidade preestabelecidas.

Como resultado, no ciclo de avaliação seguinte, em 2009, fornecer dados para o cálculo do Ideb passou a ser um dos objetivos do Saeb, por meio da publicação da Portaria nº. 87, de 7 de maio de 2009 (BRASIL, 2009b). Consequentemente, é creditado à Prova Brasil maior potencial em termos de monitoramento de resultados educacionais (SCHNEIDER; NARDI, 2014), tanto por meio do Ideb como pelo próprio sistema.

Para Fernandes e Gremaud (2009), essa divulgação, além de fornecer parâmetros que podem auxiliar as escolas, professores e diretores, também passou a contribuir para desfazer certas percepções que a escola ou a rede de ensino têm de si mesmas. Além disso, os autores acrescentam, ainda, que a ausência de resultados comparáveis longitudinalmente pode gerar uma percepção equivocada, visto que leva os profissionais da educação a acreditar estar fazendo um bom trabalho quando, na verdade, não estão. Para eles, a divulgação funciona como um elemento de pressão pela melhoria da qualidade da educação por parte da sociedade, como as famílias responsáveis pelos estudantes.

Esse tipo de mecanismo de responsabilização tem como pressuposto que o conhecimento dos resultados favorece a mobilização das equipes escolares para a melhoria da educação, bem como a pressão dos pais e da comunidade sobre a escola (BONAMINO; SOUSA, 2012, p. 375).

Sob essa perspectiva neoliberal, ao expor os resultados da avaliação externa à sociedade, acredita-se que a ação do mercado contribuiria para a elevação da qualidade do ensino. No entanto, o ambiente educativo é diferente do mercado. Conforme discutido por Freitas (2016,

⁴⁶As metas do Ideb foram estipuladas para o ano de 2021 com previsão de divulgação em 2022.

p. 146), “[...] o mercado é uma área competitiva onde ganhar e perder são consequências naturais do lance dado, mas na educação só deve haver ganhadores”.

Transferir relações de mercado para dentro dos processos educacionais apenas encobre o que realmente acontece no interior das escolas, força relações artificiais, produz disputas e mina a confiança e a colaboração entre os agentes educativos. Desagrega mais do que organiza (FREITAS, 2016, p. 146).

De acordo com Freitas (2007), essa ideologia tem como base a meritocracia e o empreendedorismo pessoal. Apesar de admitirem a “igualdade de acesso”, os neoliberais não podem conviver com a “igualdade de resultados” sem competição. Ou seja, falam de igualdade de oportunidades, não de resultados (FREITAS, 2007, 2012, 2014, 2016). Dadas as oportunidades, o que faz a diferença é o “esforço pessoal” do estudante, como se o Brasil não apresentasse acentuados contrastes socioeconômicos em sua população.

Dessa maneira, criou-se o conceito de equidade, entendido como a responsabilidade da escola de ensinar a todos, independentemente do NSE (FREITAS, 2014). Apesar de parecer tentadora, essa “equidade neoliberal” é limitada ao acesso, pois tê-la não significa que todos estão em condições iguais para uma educação de qualidade (ALTMANN, 2002).

Nesse sentido, Freitas (2007) chama a atenção para os estudantes das camadas populares, que foram incluídos, porém não se sabe como ensiná-los. Na verdade, para o autor, a pobreza só se torna alvo de reflexões quando as médias de desempenho começam a cair, inclusive, na Prova Brasil.

Essa postura de postergar os problemas políticos e sociais que o próprio neoliberalismo enfrenta com a sua política econômico-social, ainda segundo com Freitas (2007), não resolve a “universalização” da qualidade da educação básica. Pelo contrário, faz com que os estudantes mais vulneráveis aguardem a sua “eliminação adiada”, termo empregado pelo autor para se referir aos que

[...] vivem o seu ocaso no interior das escolas, desacreditados nas salas de aula ou relegados a programas de recuperação, aceleração, progressão continuada e/ou automática, educação de jovens e adultos, pseudo-escolas de tempo integral, cuja eliminação da escola foi suspensa ou adiada e aguardam sua eliminação definitiva na passagem entre ciclos ou conjunto de séries, quando então saem das estatísticas de reprovação, ou em algum momento de sua vida escolar onde a estatística seja mais confortável (FREITAS, 2007, p. 968).

No que se refere à Prova Brasil e aos usos previstos para ela, conforme disposto na Portaria Ministerial n.º. 931 (BRASIL, 2005a, p. 17), um dos objetivos é “[...] contribuir para o

desenvolvimento [...] de uma cultura avaliativa que estimule a melhoria dos padrões de qualidade e equidade da educação brasileira e adequados controles sociais de resultados”. Como visto, está enraizada na avaliação a ideia de que, ao “responsabilizar a escola”, expondo os seus resultados à sociedade, a qualidade do ensino irá melhorar. Isso significa que a Prova Brasil poderia fazer parte de qualquer programa neoliberal (FREITAS, 2007).

Conseqüentemente, para Freitas (2007), há uma redução da ênfase na avaliação do estudante, aplicada pelo professor em sala de aula, em razão do fortalecimento do monitoramento via avaliação externa. Em outras palavras, o desempenho individual é deixado de lado em frente ao desempenho geral, o que é contraditório, pois esse não é capaz de monitorar se todos realmente estão aprendendo.

Em termos matemáticos, um aumento significativo na média de desempenho escolar não quer dizer que, necessariamente, todos os estudantes seguiram esse comportamento. Existe a possibilidade de determinado grupo de estudantes, que atingiram as médias mais altas, compensar o desempenho daqueles que não obtiveram tal resultado.

De todo o modo, a avaliação externa como prestação de contas não assegura um diálogo entre quem presta contas e quem as solicita, ao contrário, prevalece uma relação unilateral (SCHNEIDER; NARDI, 2014). Com efeito, recai sobre a escola a obrigação de justificar os resultados alcançados pela comunidade.

Talvez por falta de tempo ou, ainda, por dificuldade de interpretar o mecanismo de composição do Ideb, especialmente no tocante às notas da Prova Brasil, em algumas situações, escolas aceitam passivamente os resultados e, se chamadas a justificar-se, apresentam argumentos pouco consistentes acerca de uma pontuação considerada indesejável (abaixo da meta projetada para o biênio). Se a pontuação do Ideb for favorável (compatível com a meta projetada), não raras vezes essas mesmas escolas esquecem-se de promover uma reflexão mais ampla sobre o significado real daquela pontuação (SCHNEIDER; NARDI, 2014, p. 22).

Em conclusão, como a educação é uma atividade socialmente determinada (COTTA, 2001), o “fracasso” da escola evidenciado pelo sistema de avaliação pode ser, na verdade, um fracasso na política pública que orienta esse contexto. Por isso a avaliação externa deveria ser destinada, preferencialmente, à avaliação das próprias políticas públicas, e não das escolas (FREITAS, 2007).

4.4 O SAEB E A EFETIVAÇÃO DE UMA CULTURA POLÍTICA DE AVALIAÇÃO PAUTADA NA RESPONSABILIZAÇÃO ESCOLAR

Em um estudo realizado por Bonamino e Sousa (2012), são identificadas três gerações de avaliação da educação básica no país. A primeira está voltada para as avaliações cujo objetivo é acompanhar a evolução da qualidade da educação. Em geral, essas avaliações “[...] divulgam seus resultados na internet, para consulta pública, ou utilizam-se da mídia ou de outras formas de disseminação, sem que os resultados da avaliação sejam devolvidos para as escolas” (BONAMINO; SOUSA, 2012, p. 375).

As avaliações de segunda geração, por sua vez, contemplam não só a divulgação pública para a sociedade, como também a devolutiva dos resultados para as escolas. Porém, assim como a geração anterior, essas avaliações não estabelecem consequências diretas, são apenas simbólicas, a depender da apropriação dessas informações pelas instituições escolares e pela sociedade. São chamadas *low stakes* ou de *responsabilidade branda* (BONAMINO; SOUSA, 2012).

As avaliações de terceira geração “[...] são aquelas que referenciam políticas de *responsabilização forte* ou *high stakes*, contemplando sanções ou recompensas em decorrência dos resultados de alunos e escolas” (BONAMINO; SOUSA, 2012, p. 375). Nesse caso, os resultados possuem implicações importantes para os avaliados, como o impedimento de acesso aos certificados e graus acadêmicos e, no caso das escolas, podem dar origem a processos que vão desde a advertência pelos maus resultados até o fechamento da instituição de ensino quando não atinge determinados níveis ou padrões de desempenho (AFONSO, 2012).

Dessarte, é possível observar que as três gerações de avaliação da educação básica, discutidas por Bonamino e Sousa (2012), estão associadas a uma cultura política (BERSTEIN, 1998) de avaliação educacional, organizada em três dimensões, conforme indicado por Castillo Arredondo e Diago (2009). Resumidamente, as avaliações da primeira, segunda e terceira gerações dizem respeito à obtenção de informações, formulação de juízos de valor e tomada de decisões, respectivamente.

Durante os primeiros 15 anos, o Saeb era voltado para a coleta de dados sobre a educação básica, pois as informações educacionais eram consideravelmente escassas no país. Até então,

a instituição responsável por esse tipo de trabalho era o Censo Escolar, que reunia informações meramente quantitativas, como o número de escolas, matrículas, profissionais etc.

Como discutido, o sistema era aplicado apenas de maneira amostral e por isso apresentava baixo nível de interferência no cotidiano das escolas (BONAMINO; SOUSA, 2012). Desse modo, é plausível dizer que, entre 1990 e 2005, o Saeb pertenceu à primeira geração de avaliação da educação básica, porém, com a publicação dos resultados por instituição escolar, a partir da Prova Brasil, observa-se a transição do sistema para a segunda geração.

Atualmente, além de um relatório físico e individualizado por escola, o Saeb também disponibiliza login e senha para a escola acessar as informações restritas de forma on-line. Pode-se, também, realizar uma consulta pública dos dados⁴⁷ por meio dos seguintes recursos: a) Relatório Saeb, publicação que consiste na divulgação dos resultados gerais da avaliação aplicada em determinada edição; b) Painel Educacional,⁴⁸ em que são apresentadas informações agregadas sobre o cenário educacional estadual e municipal; c) Boletim da Escola,⁴⁹ no qual podem ser acessadas as informações gerais sobre as escolas; d) Resultados em tabelas,⁵⁰ estratificados de acordo com os resultados do país, estados e municípios; e e) Seminários promovidos pelo Inep, em que são divulgados os resultados para a sociedade.

Todavia, existe uma crítica quanto à linguagem utilizada pelo sistema na divulgação dos resultados, pois não são facilmente compreendidos por grande parte dos profissionais da educação. Por outro lado, ao considerar a educação como um fenômeno social influenciado por múltiplos fatores, internos ou externos à escola e ao sistema educacional, para Cotta (2001), essa dificuldade do Saeb em produzir informações acessíveis ao grande público é compreensível. Nas palavras da autora,

[...] deve-se admitir que, em níveis mais altos de agregação, é quase impossível explicar o desempenho sem lançar mão de procedimentos estatísticos sofisticados. Este fato torna a interpretação dos resultados de avaliações como o Saeb um exercício analítico complexo, dado que não há uma explicação única para as estimativas e as tendências de evolução do desempenho (COTTA, 2001, p. 106).

⁴⁷Microdados do Saeb e demais informações sobre os resultados. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/saeb/resultados>. Acesso em: 11 fev. 2022.

⁴⁸Disponibilizam-se informações sobre os resultados das avaliações e a série histórica com informações sobre as médias e resultados por nível de proficiência das redes estaduais e municipais.

⁴⁹Disponibilizam-se informações sobre: o nível socioeconômico da escola; a formação docente; a participação na avaliação; e a distribuição percentual dos estudantes da escola por nível de proficiência.

⁵⁰São apresentados os dados numéricos brutos obtidos pelo Saeb.

De qualquer forma, não existe um único meio para que os interessados obtenham informações sobre a instituição escolar e/ou contexto em que atuam. Além do mais, em alguns casos, a realidade pode ser ainda mais agravante, dado que receber o relatório não significa que a escola realmente irá considerá-lo.

Conforme discutido por Bonamino e Sousa (2012), verifica-se, portanto, uma mudança da ênfase no uso dos resultados da avaliação. Inicialmente, a avaliação do Saeb possuía uma finalidade diagnóstica da educação, porém, após a implementação da Prova Brasil, passou a perder força mediante a tendência de focalizar esse uso como subsídio para a tomada de decisões. Inclusive, de acordo com a Portaria n.º. 250, de 5 de julho de 2021 (BRASIL, 2021b), esse é um dos objetivos do sistema.

I - Produzir indicadores educacionais para o Brasil, suas Regiões e Unidades da Federação e, quando possível, para os Municípios e as Instituições Escolares, tendo em vista a manutenção da comparabilidade dos dados, permitindo, assim, o incremento das séries históricas; II - Avaliar a qualidade, a equidade e a eficiência da educação praticada no país em seus diversos níveis governamentais; III - Subsidiar a elaboração, o monitoramento e o aprimoramento de políticas públicas em educação baseadas em evidências, com vistas ao desenvolvimento social e econômico do Brasil; e IV - Desenvolver competência técnica e científica na área de avaliação educacional, ativando o intercâmbio entre instituições de ensino e pesquisa (BRASIL, 2021b, p. 39).

Ao analisar o Saeb, Cotta (2001) aponta que se encarregar da qualidade técnica e metodológica não é o bastante. Da mesma forma, necessita-se de um esforço deliberado em promover a incorporação das informações oferecidas pela avaliação no ciclo das políticas públicas educacionais. Para isso, segundo a autora,

As expectativas dos principais usuários são determinantes para o aproveitamento das informações de avaliação no processo decisório. É preciso que eles saibam exatamente para que serve um estudo desta natureza e que estejam conscientes das limitações impostas pelo desenho da investigação (COTTA, 2001, p. 108).

Cotta (2001) acrescenta, ainda, que até mesmo os gestores educacionais, em alguns contextos, desconhecem a potencialidade da avaliação como instrumento no processo de tomada de decisões, o que foi evidenciado na pesquisa realizada por Horta Neto (2006). Nesses casos, “[...] a formulação de políticas tende a acontecer em um ambiente de baixa racionalidade, em que critérios históricos de alocação de recursos e cálculos políticos ainda têm peso considerável” (COTTA, 2001, p. 108).

Para que isso não aconteça, a autora apresenta algumas observações relevantes para que a avaliação realmente seja capaz de direcionar o processo decisório, como: disponibilizar informações de boa qualidade, que sejam relevantes, acuradas, válidas e confiáveis; estabelecer canais de comunicação permanentes com os principais usuários; e ser capaz de se adaptar ao ambiente político e burocrático (COTTA, 2001). Assim, é possível afirmar que o Saeb se consolidou acerca dessas ações, pois, além de produzir e divulgar informações educacionais sobre a educação básica brasileira, as alterações no sistema encaminham para sua adaptação de acordo com a política de interesse, por exemplo, a responsabilização.

Apesar de a Prova Brasil ser considerada a primeira avaliação externa aplicada em larga escala com objetivos de *accountability* na educação básica do país (SCHNEIDER; NARDI, 2014), ela ainda não pertence às avaliações de terceira geração. Até o momento, não foram encontradas iniciativas da avaliação atreladas a prêmios ou sanções aos seus resultados, o que é característico das políticas de *high stakes*.

Em contrapartida, os resultados da Prova Brasil são aplicados para a composição do Ideb, “[...] utilizado para medir a qualidade do aprendizado nacional e estabelecer metas para a melhoria da educação básica no Brasil e, também, monitorar as políticas públicas educacionais postas em prática no país” (SOARES, D.; SOARES, T.; DOS SANTOS, 2020). Embora, também, não exista nenhuma alusão a políticas de bonificação na legislação federal, o Plano Nacional de Educação (PNE) prevê como estratégia⁵¹ a instauração de “[...] políticas de estímulo às escolas que melhorarem o desempenho no Ideb, de modo a valorizar o mérito do corpo docente, da direção e da comunidade escolar” (BRASIL, 2014, p. 11).

Nesse sentido, a Emenda Constitucional n.º. 108, de 26 de agosto de 2020 (BRASIL, 2020d), sobre o novo Fundeb, prevê, como complementação dos recursos da União, que 2,5 pontos percentuais serão repassados para as redes públicas que cumprirem as condicionalidades previstas em lei, que dizem respeito à melhoria de gestão, evolução de indicadores a serem definidos e de atendimento e melhoria da aprendizagem com redução das desigualdades.

Como exposto no documento oficial, esse quantitativo será distribuído às redes de ensino públicas do país que cumprirem condicionalidades de melhoria de gestão e alcançarem evolução nos indicadores de atendimento escolar e aprendizagem, com redução das desigualdades (BRASIL, 2020e, 2020f). É preciso destacar que esses critérios ainda serão

⁵¹Estratégia 7.36 do PNE aprovado pela Lei n.º. 13.005, de 25 de junho de 2014.

regulamentados por lei (BRASIL, 2020e) e estão previstos para serem implementados paulatinamente a partir de 2023 (TODOS PELA EDUCAÇÃO, 2020).

Ademais, esses “critérios de qualidade” (BRASIL, 2020e) do Fundeb serão estabelecidos nos termos do Sistema Nacional de Avaliação da Educação Básica (Sinaeb),⁵² que inclui a realização das avaliações do Saeb e o cálculo do Ideb. Quanto à elaboração das avaliações, foi anunciado que os referenciais curriculares serão alinhados à Base Nacional Comum Curricular (BNCC) (BRASIL, 2018a; G1, 2020).

No âmbito nacional, tem-se o prêmio “Destques da Educação Fundamental”, organizado pelo MEC. Conforme Menezes (2019), são valorizadas as instituições escolares que possuírem maior Ideb e maior aumento no Ideb. Ao todo, são contempladas quatro escolas por região e edição, premiando um diretor e três alunos. No entanto, esse prêmio se restringe ao reconhecimento da escola por meio da emissão de certificados e divulgação na mídia.

Por outro lado, para Sousa (2014), é possível observar que Estados e municípios brasileiros, gradualmente, vêm considerando os resultados do Saeb, que são utilizados como referência para iniciativas de gestão escolar; acompanhamento das metas estipuladas pelo PNE e/ou governo local; e para a adoção de políticas de *accountability*, principalmente por meio do Ideb.

No Estado de São Paulo, professores e diretores receberam um complemento aos salários no final do ano letivo, entre 2001 e 2005. Segundo Brooke (2008, p. 99), “[...] esse bônus foi calculado de acordo com a assiduidade média da equipe, o desempenho global da escola medido durante o ano escolar através do sistema estadual de avaliação”, o Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo (Saresp).⁵³ Dentre as metas estaduais, destaca-se a comparação dos resultados do Saresp com a Prova Brasil, que serviria de acompanhamento das metas a serem atingidas pelas escolas (BONAMINO; SOUSA, 2012). Em função disso, na atualidade, mantêm-se os mesmos critérios e parâmetros do Saeb (MOREIRA; FELICORI; OSTI, 2021).

⁵²De acordo com Brasil (2016, p. 26), “[...] o Sinaeb tem por finalidade subsidiar o desenvolvimento e o aprimoramento de políticas públicas voltadas para a educação básica por meio da realização de estudos, pesquisas e avaliações envolvendo instituições, contextos, recursos, processos, instrumentos, indicadores e resultados educacionais”.

⁵³O Saresp, segundo Brooke (2008, p. 99), “[...] diferencia-se dos outros sistemas estaduais na medida em que seus testes são elaborados, aplicados e corrigidos com a ajuda dos professores”.

Goiás estabeleceu o “Prêmio Escola”, em 2011, como forma de reconhecimento ao “[...] bom desempenho das unidades educacionais da rede estadual que, por realizar um trabalho de excelência” (GOIÁS, 2011, p. 4), receberiam uma quantia para aplicação na própria unidade. O prêmio era destinado às instituições de ensino que alcançassem as metas estabelecidas para o Ideb, por isso obrigava as escolas a divulgar o índice na porta da unidade (BRASIL, 2011), o que repercutiu em todo o país.

No Maranhão, atualmente, está em vigor o “Prêmio Mais Ideb”, instituído pela Portaria nº. 814, de 21 de maio de 2019 (MARANHÃO, 2019). De acordo com o art. 6º, o objetivo é “[...] mobilizar estudantes, professores e gestores escolares para implementação de ações didático-pedagógicas focadas na aprendizagem” (MARANHÃO, 2019, p. 2), com vistas ao alcance das metas do Ideb estipuladas para o Estado (MARANHÃO, 2014).

O “Prêmio Mais Ideb” é calculado com base no resultado do Índice de Desempenho Escolar do Maranhão, realizado pelo Sistema Estadual de Avaliação do Maranhão (Seama) e/ou simulados realizados no âmbito do programa. Além disso, consideram-se a participação das escolas na Prova Brasil e o seu resultado no Ideb, em que são premiadas as dez com maiores percentuais de crescimento e as dez escolas com melhores resultados, similarmente à premiação do “Destaque da Educação Fundamental”. As premiações variam de acordo com a categoria, mas, para os dez estudantes mais bem classificados no ensino médio, poderão ser ofertadas bolsas, com duração de quatro anos, para o seu ingresso no ensino superior.

Além disso, existem outras iniciativas estaduais que, apesar de não utilizarem os dados da Prova Brasil e/ou do Ideb, possuem políticas de responsabilização escolar, por exemplo, Rio de Janeiro, Ceará, Acre e Espírito Santo (COSTA; VIDAL, 2021; FERREIRA FILHO; SOUSA; PONTES JUNIOR, 2021; SANTOS et al., 2021; CARVALHO; MORAIS, 2021). Com base nos resultados dos sistemas estaduais de avaliação e seus respectivos índices, os Estados promovem o pagamento de prêmios e/ou bônus salarial aos profissionais que atingem os requisitos preestabelecidos.

No Rio de Janeiro, destaca-se o “Programa Nova Escola”, criado pelo Decreto nº. 25.959, de 12 de janeiro de 2000 (RIO DE JANEIRO, 2000). Considerado o modelo pioneiro de responsabilização, de acordo com Brooke (2006, p. 387), o programa era “[...] uma tentativa de melhorar a gestão das escolas estaduais e de responsabilizar seus diretores e funcionários pelos resultados dos alunos”. Para isso oferecia uma gratificação às escolas, somada ao

vencimento básico dos servidores, que variava de acordo com o desempenho obtido nos testes do sistema estadual de avaliação, baseando-se em alguns itens e na escala do Saeb.

No Ceará, conforme discutido por Costa e Vidal (2021), o Estado já implementa políticas de *accountability* desde os anos 2000, anteriormente à Prova Brasil e ao Ideb. O Estado é conhecido pela variedade de prêmios ao longo da implementação do Sistema Permanente de Avaliação da Educação Básica do Ceará (Spaece), como o “Prêmio Escola do Novo Milênio”, “Prêmio Escola Nota Dez”, “Prêmio Aprender pra Valer”, “Prêmio Foco na Aprendizagem” etc. (COSTA; VIDAL, 2021; FERREIRA FILHO; SOUSA; PONTES JUNIOR, 2021).

No Acre, por sua vez, o abono financeiro também é realizado por meio de premiações com base no Sistema Estadual de Avaliação da Aprendizagem Escolar (Seape). Destacam-se o “Prêmio Anual de Valorização e Desempenho Profissional” e o “Prêmio Anual de Valorização de Desenvolvimento da Gestão” destinados aos professores e às equipes gestoras, respectivamente (CARVALHO; MORAIS, 2021).

No Espírito Santo, de acordo com Santos et al. (2021), a gratificação dos professores é anual, podendo chegar até 150% do valor do seu rendimento mensal. Como critério, utiliza-se o desempenho dos estudantes na Avaliação Interna Trimestral Diagnóstica da Aprendizagem (Paebes TRI).⁵⁴ Outro aspecto que o diferencia dos demais Estados é o fato de o Paebes TRI compor parte da nota trimestral do aluno.

Em nível municipal, Bauer et al. (2017) realizaram uma pesquisa, com abrangência nacional, sobre as iniciativas de avaliação do ensino em municípios brasileiros. Segundo os autores, até 2004, foram registrados 103 municípios que possuíam iniciativas próprias de avaliação. Entre 2005 e 2014, as avaliações se intensificaram e registraram 1.280 novas iniciativas, o que evidencia a influência da implementação da Prova Brasil.

No que se refere à avaliação dos estudantes, realizada por 82,8% dos municípios, para Bauer et al. (2017), a principal motivação declarada é a necessidade de fazer o diagnóstico e o monitoramento do ensino e das aprendizagens. Em seguida, estão o apoio ao gerenciamento educacional e a melhoria dos índices educacionais, como o Ideb. Existem referências explícitas ao aumento desse índice. Para a elaboração das provas, conforme os autores, a maioria dos municípios declarou utilizar as Matrizes de Referência da Prova Brasil e/ou

⁵⁴Essa avaliação compõe o Programa de Avaliação da Educação Básica do Espírito Santo (Paebes).

documentos curriculares, por exemplo, os Parâmetros Curriculares Nacionais (PCN) e as propostas da rede municipal de ensino.

Quanto ao uso dos resultados da avaliação municipal para a distribuição de incentivos, percebeu-se que tais práticas não são recorrentes. Na realidade, “[...] os depoimentos ressaltam o incentivo ao alcance dos melhores resultados possíveis pelas escolas sem que, necessariamente, sejam atreladas a eles benesses de algum tipo” (BAUER et al., 2017, p. 16). Dos 13,3% que afirmaram fazer algum tipo de premiação, nota-se que as ações são destinadas prioritariamente aos alunos.

A título de exemplo, em 2006, foi criado o “Informativo da Escola Municipal” de Belo Horizonte. Para os indicadores pedagógicos, eram considerados os resultados do Sistema Mineiro de Avaliação e Equidade da Educação Pública (Simave) e da Prova Brasil. O informativo apresentava o percentual dos estudantes em cada um dos quatro níveis de desempenho, como no Saeb, cujas comparações eram realizadas somente com escolas que tivessem condições socioeconômicas similares (BROOKE, 2008). Apesar disso, essa iniciativa não efetuava o pagamento de recompensas. O seu uso era restrito à divulgação das informações para as famílias e para a comunidade local.

Desse modo, é possível afirmar que as iniciativas estaduais de responsabilização, via avaliação externa, são do tipo *high stakes*, cujos resultados possuem implicação direta na vida dos atores educacionais, por exemplo, políticas de bonificação. As iniciativas municipais mostraram ser do tipo *low stakes*, restritas às dimensões da avaliação em obter informações e formular juízos de valor, sem consequências materiais para os envolvidos.

Em todo caso, utilizam-se como critérios os resultados alcançados pelas instituições escolares nas avaliações externas aplicadas em larga escala, seja por meio da Prova Brasil e/ou Ideb, seja por indicadores próprios, para fundamentar ações que estimulem a competitividade com vistas à melhoria da qualidade da educação básica.

4.5 CONSIDERAÇÕES PARCIAIS

Este capítulo foi desenvolvido com o objetivo de realizar um levantamento sobre as primeiras iniciativas de uma política de avaliação nacional pautada nos processos de prestação de contas

e de responsabilização, a partir dos resultados do Saeb. Com base nas discussões acerca do objeto, evidenciou-se que, por meio do discurso declarado de “aumentar a transparência da destinação dos recursos públicos para a educação” no Brasil, datado de 2005, foi criada uma *cultura política* (BERSTEIN, 1998) de avaliação pautada no modelo *accountability*.

Como discutido, um fator essencial para a sua implementação foi o uso da avaliação externa aplicada em larga escala como “condição” ao planejamento e a gestão educacional. Esse tipo de avaliação, em geral, é responsável por aferir um padrão mínimo de qualidade da educação, como as avaliações do Saeb.

Com relação à Prova Brasil, compreendeu-se que a sua implementação foi o primeiro passo para a chegada das políticas de responsabilização no país. Nessa vertente, percebeu-se que a concepção neoliberal de que a divulgação dos resultados da avaliação por instituição escolar contribui para a melhoria da qualidade da educação está enraizada nas políticas educacionais implementadas após a sua criação. Inclusive, na própria legislação federal, que regulariza a criação dessa avaliação, são admitidas práticas que poderiam ocasionar “[...] adequados controles sociais de seus resultados” (BRASIL, 2005a, p. 17).

Desse modo, ao assegurar o gerenciamento e a aplicação do Saeb, a importância dada à Prova Brasil favoreceu a conversão do Inep em uma agência nacional de avaliação (FREITAS, 2014). Atualmente, a instituição é responsável pela coordenação das avaliações externas e pela produção dos indicadores educacionais em nível nacional.

Outro fator que auxiliou na incorporação do modelo de prestação de contas no país foi o PDE que, além de estabelecer metas educacionais, passou a relacionar os seguintes aspectos: avaliação, financiamento e gestão. Para monitorar o cumprimento dessas metas, desde 2007, passou-se a calcular o Ideb, que utiliza informações fornecidas pela Prova Brasil e pelo Censo Escolar. Esse indicador reforça, ainda mais, a intenção do *uso e consumo* (GINZBURG, 1989) regulatórios da avaliação por meio do Saeb.

Como consequência, apesar de ter sido encontrada apenas uma iniciativa atrelada a prêmios no âmbito nacional, ela se restringe à emissão de certificados. No entanto, verificou-se uma crescente utilização dos resultados do Saeb e/ou de indicadores próprios em Estados e municípios brasileiros. De maneira geral, percebeu-se como objetivo comum o uso dos dados como referência para a melhoria dos índices educacionais (municipal, estadual e nacional) e da gestão educacional.

Constatou-se, portanto, uma aproximação dos sistemas estaduais e municipais de avaliação com o Saeb, principalmente no que se refere à Prova Brasil, por exemplo, as escalas de proficiência, Matrizes de Referência e disciplinas avaliadas etc. Esta última revelou *sinais* (GINZBURG, 1989) da ênfase dada à Língua Portuguesa e à Matemática.

Quanto às políticas de “incentivo”, observou-se que as avaliações externas dos sistemas estaduais pertencem à terceira geração de avaliação da educação básica. Por meio delas, são atribuídos prêmios e/ou bonificações aos profissionais e/ou estudantes que atingiram as metas preestabelecidas. Esse tipo de responsabilização possui implicações diretas na vida dos envolvidos, por isso se configura como *responsabilização forte (high stakes)*.

Por outro lado, notou-se que tais práticas não são tão frequentes como se esperava pelo que vem sendo divulgado na literatura acadêmica (BAUER et al., 2017). Quando existem, mostraram ser restritas ao reconhecimento por meio da emissão de certificados, sobretudo dos estudantes. Assim, por serem simbólicas e não apresentarem consequências diretas, caracterizam-se como *responsabilidade branda (low stakes)*.

Em síntese, ao atribuir o caráter censitário à avaliação nacional, pôde-se perceber que essa mudança alterou, também, a ênfase dada aos resultados do Saeb. A partir da Prova Brasil, o sistema passou a pertencer à segunda geração de avaliação, ou seja, configura-se como uma política *low stakes*.

Em contrapartida, além do Novo Fundeb, que prevê o repasse com base em “critérios de qualidade” via avaliação externa, a Portaria n.º. 458, de 5 maio de 2020 (Brasil, 2020a), anuncia que, futuramente, os resultados do Saeb deverão possibilitar o acesso à educação superior. Diante disso, há *indícios* (GINZBURG, 1989) de que o sistema dá os primeiros passos em direção às avaliações de terceira geração, implementando políticas educacionais fundamentadas no modelo *accountability* e do tipo *high stakes*.

CONSIDERAÇÕES FINAIS

Esta pesquisa teve como objetivo geral analisar a trajetória histórica do Saeb ao longo dos mais de trinta anos de implementação, identificando os princípios e as concepções que fundamentam o sistema. Para tanto, buscando compreender a motivação das várias mudanças no seu desenvolvimento desde a sua criação, em 1990, discutimos os principais acontecimentos durante a sua trajetória histórica que contribuíram para o seu uso como política de responsabilização.

Conforme evidenciamos no decorrer do estudo, dentre as primeiras experiências de avaliação que influenciaram o desenho e a montagem do Saeb, destacam-se: o Projeto Edurural; o Projeto Nordeste; a Avaliação do Rendimento Escolar de Alunos de Escolas de 1º Grau da Rede Pública; e o Saep. Tal como sugerem Horta Neto (2007), Freitas (2013) e Pestana (2016), este último é assumido como estudo piloto, constituindo-se como a primeira designação do sistema.

Observamos que as experiências supracitadas se mostraram precursoras para o Saeb, principalmente para a escolha das séries, disciplinas e conteúdos curriculares a serem avaliados. Além disso, elas revelaram a inexistência de dados educacionais naquele período (entre 1980 e 1990), sobretudo em nível de sistema. Diante de tal quadro, avançou-se em direção à criação de um Sistema Nacional de Avaliação, cujo objetivo era obter informações sobre a realidade da educação básica brasileira, a primeira dimensão da avaliação proposta por Castillo Arredondo e Diago (2009).

Durante esses mais de trinta anos de existência, o Saeb passou por diversos aprimoramentos, adaptações e alterações metodológicas. Destaca-se o ciclo de avaliação de 1995, que consideramos como um marco significativo nas alterações do sistema. A partir daí, ele passou a avaliar a terminalidade dos ciclos escolares e a realizar um levantamento de dados contextuais por meio de questionários aplicados aos estudantes e aos professores, primeiro movimento em prol da análise dos processos que ocorrem dentro da escola e que podem oferecer significativa influência no desempenho dos estudantes.

Além disso, para a construção das avaliações e análises dos resultados, o Saeb adotou como metodologia a TRI, que foi o seu uso pioneiro no território nacional. Essa metodologia permite a comparação dos resultados da avaliação ao longo do tempo. Dessa alteração, surge,

portanto, um novo *uso e consumo* (GINZBURG, 1989) para o sistema que, além de fornecer um diagnóstico da educação básica, também se torna um instrumento indutor de políticas públicas que exercem funções importantes para a avaliação, como o juízo de valor e a tomada de decisão, que correspondem às demais dimensões da avaliação propostas por Castillo Arredondo e Diago (2009).

Nessa perspectiva, observamos que há *sinais* (GINZBURG, 1989) de que a própria legislação federal contribuiu, paulatinamente, para tal prática. Em 1993, o Saeb recebeu o apoio institucional do Plano Decenal de Educação para Todos 1993-2003, que o definiu como ação prioritária nas diversas instâncias governamentais (estadual, municipal e federal). Dentre as finalidades, destacamos o uso do sistema para aferir o desempenho das escolas e prover informações para avaliação e revisão de planos e programas educacionais.

No ano seguinte, o Saeb foi institucionalizado por meio da Portaria nº. 1.795, de 27 de dezembro de 1994, documento governamental que oficializa a sua criação. Essa implementação tardia e seguida do plano anterior são *indícios* (GINZBURG, 1989) de que já se pretendia a utilização do sistema como subsídio para a formulação de políticas públicas educacionais. Inclusive, a referida portaria o define como responsável pela realização dos processos de monitoramento e avaliação dos resultados dessas políticas.

Em 1996, houve a promulgação da LDB, um marco na agenda de reformas do governo Fernando Henrique Cardoso. Como observa Horta Neto (2007), essa foi a primeira vez em que a legislação federal se referiu à coleta de informações e à avaliação, cujo objetivo era a garantia do padrão de qualidade da educação básica. Da mesma forma, coube à avaliação “monitorar” a qualidade, principalmente por meio do Saeb.

Posteriormente, um desdobramento do Saeb como política de avaliação foi a implementação do Plano Nacional de Educação (PNE) de 2001, em que a avaliação aparece como instrumento indispensável para a gestão do sistema educativo e melhoria da qualidade da educação. Com o PNE reforçou-se ainda mais o *uso e consumo* (GINZBURG, 1989) da avaliação como forma de supervisão e inspeção da qualidade do ensino ofertado nas escolas.

Outra mudança significativa ocorreu em 1997, a partir da adoção de Matrizes de Referência para a elaboração dos itens das provas do Saeb, responsáveis por estabelecer o “conteúdo mínimo” que os estudantes devem aprender em cada etapa e disciplina avaliada. Assim, por

meio da avaliação externa, é possível aferir o quão distante/próximo os estudantes estão desse mínimo estabelecido.

Para a elaboração da primeira versão, foi realizada uma ampla consulta nacional às Secretarias Estaduais e Municipais de Educação, processo que, anteriormente, era utilizado para seleção dos conteúdos das provas. Porém, para a “atualização” dessas matrizes, em 2001, contou-se apenas com a participação de especialistas da área do Inep/MEC e das fundações de pesquisa. Temos, portanto, uma *pista* (GINZBURG, 1989) de que a adoção das Matrizes de Referência pelo Saeb contribuiu para que a participação das Secretarias de Educação, aos poucos, fosse se esvaindo.

Nesse sentido, é preciso destacar a influência de organizações multilaterais internacionais, sobretudo o Banco Mundial. Esse não acreditava na existência do sistema de avaliação, relativamente, em parcerias com as Secretarias de Educação. Em razão disso, recusou-se a financiar a primeira avaliação do Saeb, como discute Altmann (2002). Para o banco, o MEC deveria estar à frente das negociações e repassá-las às demais instâncias governamentais.

Já que para a concessão de empréstimos era necessário “estar de acordo” com as condicionantes do Banco Mundial, como consequência, o financiamento só foi retomado quando o Inep/MEC passou a conduzir as avaliações no país. Dessa forma, ao diminuir a participação das Secretarias de Educação, acreditamos que isso tenha favorecido a conversão do Inep em uma agência nacional de avaliação, como argumenta Freitas (2014). Atualmente, a instituição é a responsável pela coordenação das avaliações externas e pela produção dos indicadores educacionais em nível nacional.

Ademais, de acordo com as discussões apresentadas ao longo desta pesquisa, há *sinais* (GINZBURG, 1989) de que o Banco Mundial também influenciou sobre outros aspectos considerados pelo Saeb, como a aplicação de avaliações amostrais e bianuais. Do mesmo modo, é atribuída a ele a preferência por avaliar o desempenho dos estudantes nos anos finais de cada ciclo escolar e a ênfase dada ao ensino da Língua Vernácula e de Matemática. Com exceção da Avaliação do Rendimento Escolar de Alunos de Escolas de 1º Grau da Rede Pública, o banco financiou as experiências que serviram como base para o sistema e por isso elas apresentavam essas mesmas características.

Dando continuidade à trajetória do Saeb, em 2005, inicia-se a aplicação da avaliação nacional censitária, via Prova Brasil. Essa avaliação possibilitou a cobertura geográfica da totalidade

das unidades da Federação, visto que, anteriormente, era realizada somente de maneira amostral. Com isso, tornou-se possível a divulgação dos resultados por escola e por redes de ensino.

Ao analisarmos os documentos oficiais que tratam da realização da Prova Brasil, verificamos a presença de *indícios* (GINZBURG, 1989) de que essa avaliação está fundamentada em uma concepção neoliberal. Por exemplo, a ideia de que a publicação dos resultados da avaliação por instituição escolar possa contribuir para a melhoria da qualidade da educação, pois favoreceria a mobilização das equipes escolares assim como incentivaria a pressão dos pais e da comunidade sobre a escola.

Compreendemos, portanto, que a implementação da Prova Brasil foi o primeiro passo para a chegada das políticas de responsabilização no país. Como resultado, iniciou-se a incorporação dos resultados do Saeb na formulação de políticas públicas e no planejamento do ensino, bem como na elaboração de propostas curriculares desenvolvidas pelas Secretarias Estaduais e Municipais de Educação.

Outro fator que auxiliou na incorporação do modelo de prestação de contas foi o Plano de Desenvolvimento da Educação (PDE), que passou a relacionar os seguintes aspectos: avaliação, financiamento e gestão. Também introduziu os conceitos de responsabilização e mobilização social que, segundo o documento, tornam a escola “menos estatal” e mais “pública”. Com isso o programa trouxe um sentido mais operacional às avaliações, a partir da consolidação de metas de qualidade para todas as modalidades de ensino e em todas as esferas administrativas.

Para monitorar o “acompanhamento” das metas, desde 2007, passou-se a calcular o Ideb, que utiliza informações fornecidas pelo Saeb e pelo Censo Escolar. Assim, o Ideb se tornou a principal referência de qualidade em educação, inclusive, para o repasse e a distribuição de recursos públicos. Como efeito, a partir do ciclo de avaliação seguinte, em 2009, fornecer dados para o cálculo do Ideb passou a ser um dos objetivos do Saeb.

Tendo em vista um maior “comprometimento” das escolas e das redes de ensino com o objetivo de melhoria do Ideb, como discute Fernandes (2016), foi estabelecido o Plano de Metas Compromisso Todos pela Educação, em 2007. Uma das diretrizes do plano consiste em divulgar, na escola e na comunidade, os dados educacionais com ênfase no Ideb.

Desse modo, ao “traduzir” o quanto as escolas, municípios e Estados estavam avançando em relação às suas metas, em concordância com Schneider e Nardi (2014), o Ideb passou a se configurar como uma importante “prestação de contas” dos atores que permeiam o cotidiano das escolas. Diante disso, entendemos que esse indicador reforça ainda mais o *uso e consumo* (GINZBURG, 1989) regulatório da avaliação, ampliando o impacto do Saeb nas políticas educacionais brasileiras.

Em 2013, a Avaliação Nacional da Alfabetização (ANA), que estava prevista no Pacto Nacional pela Alfabetização na Idade Certa (Pnaic), passou a compor o quadro de avaliações do Saeb. Como destaca Dickel (2016), com ela surge um elemento novo que a distingue das demais modalidades de avaliação integradas ao sistema, pois se encontra acompanhada de um projeto de formação de professores.

Além dos indicadores de nível socioeconômico (NSE) e da complexidade de gestão das escolas, por meio da ANA, são coletadas, ainda, informações sobre o grau de instrução do professorado, que resultam no Indicador de Adequação da Formação Docente (IAFD). Esse indicador foi criado para classificar os docentes da educação básica, considerando a etapa em que lecionam e a disciplina que ministram. Por esse motivo, acreditamos que a ANA seja mais uma *pista* (GINZBURG, 1989) sobre o fortalecimento do uso regulativo e punitivo do Saeb, agora voltado especificamente aos professores.

É importante ressaltar que, desde 2019, se extinguiu o uso das siglas Aneb, Anresc/Prova Brasil e ANA. Apesar de na literatura acadêmica ainda fazerem seu uso, todas as avaliações passaram a ser identificadas pelo nome Saeb, acompanhado das etapas, áreas de conhecimento e tipos de instrumentos envolvidos.

Atualmente, as avaliações do Saeb são aplicadas aos anos finais de cada ciclo escolar: 2º, 5º e 9º anos do ensino fundamental e 3º ano do ensino médio. Quanto à sua natureza, a avaliação é censitária para as escolas públicas (com exceção do 2º ano do ensino fundamental) e amostral para as escolas privadas. Além disso, ainda em 2019, o sistema também começou a avaliar, em caráter de estudo piloto, a educação infantil por meio de questionários eletrônicos aplicados exclusivamente para uma amostra de professores, diretores e secretários municipais de Educação.

Dessarte, o Saeb passou a avaliar todas as etapas escolares da educação básica brasileira (educação infantil, ensino fundamental e ensino médio), cujos *sinais* (GINZBURG, 1989) o

apontam como um importante instrumento norteador para as políticas de responsabilização implementadas.

Com base nas discussões acerca do objeto, evidenciamos que a busca por “aumentar a transparência da destinação dos recursos públicos para a educação” no Brasil, em 2005, foi criada uma *cultura política* (BERSTEIN, 1998) de avaliação pautada no modelo *accountability*. Ao atribuir o caráter censitário à avaliação nacional, percebemos que essa mudança alterou também a ênfase dada aos resultados do Saeb. A partir da Prova Brasil, o sistema passou a pertencer ao que Bonamino e Souza (2012) denominam de segunda geração de avaliação da educação básica, configurando-se como uma política *low stakes* ou *responsabilização branda*.

Em contrapartida, em 2020, o Saeb prevê uma nova alteração, cuja avaliação passará a ser aplicada anualmente abrangendo todas as etapas a partir do 2º ano do ensino fundamental. Além do Novo Fundeb, que prevê o repasse de recursos financeiros com base em “critérios de qualidade” via avaliação externa, foi anunciado que, futuramente, os resultados do Saeb deverão possibilitar o acesso à educação superior. Diante disso, há *indícios* (GINZBURG, 1989) de que o sistema dá os primeiros passos em direção às avaliações de terceira geração, implementando políticas educacionais fundamentadas no modelo *accountability* e do tipo *high stakes* ou *responsabilização forte*, o que confirma nossa hipótese.

A experiência desta pesquisa nos ajudou a compreender que as avaliações externas devem ser instrumentos de monitoramento de políticas públicas e não de pressão das escolas. Não dá para colocar uma política em prática e só depois ir atrás de evidências, como argumenta Freitas (2013). Pelo contrário, as políticas educacionais devem estar baseadas em dados consistentes da avaliação que sustentem a sua necessidade de implementação, visto que podem interferir na vida de alunos, professores, pais e gestores.

Esperamos que as discussões e análises empreendidas neste estudo contribuam para as investigações acerca do uso regulativo e punitivo do Saeb em detrimento de uma perspectiva avaliativa do próprio papel desempenhado pelo Estado como condutor das políticas educacionais. Além disso, pretendemos instigar novos assuntos sobre a temática para futuras investigações.

REFERÊNCIAS

ABRÃO, Jorge Luís Ferreira. As influências da psicanálise na educação brasileira no início do século XX. **Psicologia: Teoria e Pesquisa**, Brasília, v. 22, n. 2, p. 233-240, maio/ago. 2006.

ABRAMOWICZ, Mere. Avaliação, tomada de decisão e políticas públicas subsídios para um repensar. **Estudos em Avaliação Educacional**, São Paulo, n. 10, p. 81-101. 1994. DOI: <https://doi.org/10.18222/ae01019942321>

AFONSO, Almerindo Janela. Para uma concetualização alternativa de accountability em educação. **Educação & Sociedade**, Campinas, v. 33, n. 119, p. 471-484, abr./jun. 2012. DOI: <https://doi.org/10.1590/S0101-73302012000200008>

ALTMANN, Helena. Influências do Banco Mundial no projeto educacional brasileiro. **Educação e Pesquisa**, São Paulo, v. 28, n. 1, p. 77-89, jun./set. 2002. DOI: <https://doi.org/10.1590/S1517-97022002000100005>

ALVES, Fátima. Políticas educacionais e desempenho escolar nas capitais brasileiras. **Cadernos de Pesquisa**, São Paulo, v. 38, n. 134, p. 413-440, ago./out. 2008. DOI: <http://dx.doi.org/10.1590/S0100-15742008000200008>

ANDRADE, Dalton Francisco; TAVARES, Heliton Ribeiro; VALLE, Raquel da Cunha. Teoria da resposta ao item: conceitos e aplicações. In: SIMPÓSIO NACIONAL DE PROBABILIDADE E ESTATÍSTICA (SINAPE), 14. **Anais**. Caxambu: Associação Brasileira de Estatística, 2000. p. 164. Disponível em: https://www.researchgate.net/profile/Dalton-Andrade/publication/228593873_Teoria_da_Resposta_ao_Item_conceitos_e_aplicacoes/links/54183ff70cf25ebee988107c/Teoria-da-Resposta-ao-Item-conceitos-e-aplicacoes.pdf. Acesso em: 8 fev. 2021.

ARAÚJO, Carlos Henrique; LUZIO, Nildo. **Avaliação da educação básica: em busca da qualidade e equidade no Brasil**. Brasília: Inep, 2005.

BANCO MUNDIAL. **Second Northeast Basic Education Project Brazil**. Banco Mundial, 1993. Disponível em: <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/168121468228272836/brazil-second-northeast-basic-education-project>. Acesso em: 10 jan. 2022.

BANCO MUNDIAL. **Prioridades y estrategias para la educación**. Banco Mundial, 1996. Disponível em: <http://documents1.worldbank.org/curated/pt/715681468329483128/pdf/14948010spanish.pdf>. Acesso em: 19 jun. 2020.

BAUER, Adriana. Usos dos resultados das avaliações de sistemas educacionais: iniciativas em curso em alguns países da América. **Revista Brasileira de Estudos Pedagógicos**, Brasília, v. 91, n. 228, p. 315-344, maio/ago. 2010. DOI: <http://doi.org/10.24109/2176-6681.rbep.91i228.576>

BAUER, Adriana; ALAVARSE, Ocimar Munhoz; OLIVEIRA, Romualdo Portela. Avaliação em larga escala: uma sistematização do debate. **Educação e Pesquisa**, São Paulo, v. 41, n. especial, p. 1367-1382, mar./dez. 2015. DOI: <http://dx.doi.org/10.1590/S1517-9702201508144607>

- BAUER, Adriana. Avaliação de Redes de Ensino e Gestão Educacional: aportes teóricos. **Educação & Realidade**, Porto Alegre, v. 44, n. 1, dez./fev. 2019. DOI: <https://doi.org/10.1590/2175-623677006>
- BEATON, Albert E.; ALLEN, Nancy L. Chapter 6: Interpreting Scales Through Scale Anchoring. **Journal of Educational Statistics**, v. 17, n. 2, p. 191-204, 1992. DOI: <https://doi.org/10.2307/1165169>
- BERSTEIN, S. Culturas políticas e historiografia. In: AZEVEDO, Cecília *et al.* **Cultura política, memória e historiografia**. Rio de Janeiro: Editora FGV, 2009. p. 29-46.
- BIRNBAUM, Allan. Some latent trait models and their use in inferring an examinee's ability. **Statistical Theories of Mental Test Scores**, Boston, p. 397-479, 1968.
- BLOCH, Marc. **Apologia da história ou o ofício de historiador**. Rio de Janeiro: Jorge Zahar, 2001.
- BONAMINO, Alícia; FRANCO, Creso. Avaliação e política educacional: o processo de institucionalização do Saeb. **Cadernos de Pesquisa**, São Paulo, n. 108, p. 101-132, nov. 1999. DOI: <http://doi.org/10.1590/S0100-15741999000300005>
- BRASIL. Lei nº 3.692, de 15 de dezembro de 1959. Institui a Superintendência do Desenvolvimento do Nordeste e dá outras providências. **Diário Oficial da União**, Brasília, DF: Presidência da República, 1959. Disponível em: https://www.planalto.gov.br/ccivil_03/LEIS/1950-1969/L3692.htm. Acesso em: 10 jan. 2022.
- BRASIL. Decreto nº 85.287, de 23 de outubro de 1980. Cria o Programa de Expansão e Melhoria da Educação no Meio Rural do Nordeste (Edurural – NE) e dá outras providências. **Diário Oficial da União**, Brasília, DF: Presidência da República, 1980. Disponível em: <https://legis.senado.leg.br/norma/507927>. Acesso em: 9 jan. 2022.
- BRASIL. Ministério da Educação. **Programa de Expansão e Melhoria da Educação no Meio Rural do Nordeste**. Brasília: MEC, 1982. Disponível em: http://www.dominiopublico.gov.br/pesquisa/DetalheObraForm.do?select_action=&co_obra=27547. Acesso em: 29 jun. 2020.
- BRASIL. Decreto nº 92.435, 3 de março de 1986. Dispõe sobre a Estrutura Básica da Secretaria Executiva da Superintendência do Desenvolvimento do Nordeste (Sudene) e dá outras providências. **Diário Oficial da União**, Brasília, DF: Presidência da República, 1986a. Disponível em: https://www.planalto.gov.br/ccivil_03/decreto/1980-1989/1985-1987/d92435.htm. Acesso em: 23 dez. 2021.
- BRASIL. Lei nº 7.499, de 25 de junho de 1986. Aprova as diretrizes do Primeiro Plano de Desenvolvimento do Nordeste da Nova República. **Diário Oficial da União**, Brasília, DF: Presidência da República, 1986b. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/1980-1988/17499.htm. Acesso em: 23 dez. 2020.
- BRASIL. Ministério da Educação. **Alcance e funcionamento operacional do Projeto BRA-86-002**. Brasília: MEC, 1987.
- BRASIL. Constituição da República Federativa do Brasil de 1988. **Diário Oficial da União**, Brasília, DF: Presidência da República, 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 15 abr. 2020.

BRASIL. Ministério da Educação. **Plano Decenal de Educação para Todos 1993-2003.**

Brasília, DF: MEC, 1993. Disponível em:

<http://www.dominiopublico.gov.br/download/texto/me001523.pdf>. Acesso em: 18 fev. 2021.

BRASIL. Ministério da Educação. **Sistema nacional de avaliação da educação básica:**

objetivos, diretrizes, produtos e resultados. Brasília, DF: MEC/Inep, 1994a.

BRASIL. Portaria n° 1.795, de 27 de dezembro de 1994. **Diário Oficial da União**, Brasília, DF: Presidência da República, 1994b. Disponível em:

<https://pesquisa.in.gov.br/imprensa/servlet/INPDFViewer?jornal=1&pagina=91&data=28/12/1994&captchafield=firstAccess>. Acesso em: 15 abr. 2020.

BRASIL. Lei n° 9.934, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. **Diário Oficial da União**, Brasília, DF: Presidência da República, 1996.

Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l9394.htm. Acesso em: 13 mar. 2021.

BRASIL. Ministério da Educação. **Matrizes Curriculares de Referência para o SAEB/97.**

Brasília, DF: MEC/Inep, 1998.

BRASIL. Ministério da Educação. **Relatório:** resultados do Saeb 99. Brasília, DF: Inep, 1999. Disponível em:

<http://portal.inep.gov.br/documents/186968/484421/Resultados+do+SAEB-9+relat%C3%B3rio/83240ffc-02b9-49f6-bf81-6af548c17ace?version=1.2>. Acesso em: 6 maio 2020.

BRASIL. Lei n° 10.172, de 9 de janeiro de 2001. Aprova o Plano Nacional de Educação e dá outras providências. **Diário Oficial da União**, Brasília, DF: Presidência da República, 2001.

Disponível em: http://www.planalto.gov.br/ccivil_03/leis/leis_2001/l10172.htm. Acesso em: 6 maio 2020.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Saeb 2001:** novas perspectivas. Brasília, DF: Inep, 2002. Disponível em:

http://portal.inep.gov.br/informacao-da-publicacao/-/asset_publisher/6JYIsGMAMkW1/document/id/484540. Acesso em: 10 fev. 2021.

BRASIL. Portaria n° 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). **Diário Oficial da União**, Brasília, DF: Presidência da República, 2005a.

Disponível em:

http://download.inep.gov.br/educacao_basica/prova_brasil_saeb/legislacao/Portaria931_NovoSaeb.pdf. Acesso em: 22 fev. 2022.

BRASIL. Portaria n° 69, de 4 de maio de 2005. **Diário Oficial da União**, Brasília, DF: Presidência da República, 2005b. Disponível em:

http://download.inep.gov.br/educacao_basica/prova_brasil_saeb/legislacao/Portaria69_Anresc.pdf. Acesso em: 22 fev. 2022.

BRASIL. Portaria n° 89, de 25 de maio de 2005. **Diário Oficial da União**, Brasília, DF: Presidência da República, 2005c. Disponível em:

https://download.inep.gov.br/educacao_basica/prova_brasil_saeb/legislacao/Portaria89_Aneb.pdf. Acesso em: 22 fev. 2022.

BRASIL. Decreto n.º 6.094, de 24 de abril de 2007. Dispõe sobre a implementação do Plano de Metas Compromisso Todos pela Educação, pela União Federal, em regime de colaboração com Municípios, Distrito Federal e Estados, e a participação das famílias e da comunidade, mediante programas e ações de assistência técnica e financeira, visando a mobilização social pela melhoria da qualidade da educação básica. **Diário Oficial da União**, Brasília, DF: Presidência da República, 2007a. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2007/decreto/d6094.htm. Acesso em: 22 fev. 2022.

BRASIL. Portaria n.º 47, de 3 de maio de 2007. **Diário Oficial da União**, Brasília, DF: Presidência da República, 2007b. Disponível em: https://download.inep.gov.br/educacao_basica/prova_brasil_saeb/legislacao/Port47_3MAI07.pdf. Acesso em: 22 fev. 2022.

BRASIL. Ministério da Educação. **O Plano de Desenvolvimento da Educação: Razões, Princípios e Programas**. Brasília, DF: Inep/MEC, 2008. Disponível em: <http://portal.inep.gov.br/documents/186968/485287/O+Plano+de+Desenvolvimento+da+Educa%C3%A7%C3%A3o+raz%C3%B5es%2C+princ%C3%ADpios+e+programas/3c6adb19-4c2e-4c60-9ccb-3b476bed9358?version=1.6>. Acesso em: 22 fev. 2022.

BRASIL. Ministério da Educação. **Nota Técnica: Índice de Desenvolvimento da Educação Básica (Ideb)**. Brasília, DF: MEC/SEB/Inep, 2009a. 10p. Disponível em: http://download.inep.gov.br/educacao_basica/portal_ideb/o_que_e_o_ideb/Nota_Tecnica_n1_concepcaoIDEB.pdf. Acesso em: 17 dez. 2020.

BRASIL. Portaria n.º 87, de 7 de maio de 2009. Estabelece a sistemática para a realização da Avaliação Nacional do Rendimento Escolar – Anresc (Prova Brasil) e da Avaliação Nacional da Educação Básica – Aneb, no ano de 2009. **Diário Oficial da União**, Brasília, DF: Presidência da República, 2009b. Disponível em: https://www.normasbrasil.com.br/norma/portaria-87-2009_213756.html. Acesso em: 22 fev. 2022.

BRASIL. PL n.º 1.530/2011 de 07 de junho de 2011. Obriga os estabelecimentos de ensino básico do país a divulgarem o índice Ideb. **Câmara dos Deputados**, 2011.

BRASIL. Portaria n.º 867, de 4 de julho de 2012. Institui o Pacto Nacional pela Alfabetização na Idade Certa e as ações do pacto e define suas diretrizes gerais. **Diário Oficial da União**, Brasília, DF: Presidência da República, 2012. Disponível em: <https://abmes.org.br/arquivos/legislacoes/Port-867-2012-07-04.pdf>. Acesso em: 22 fev. 2022.

BRASIL. Portaria n.º 482, de 7 de junho de 2013. Dispõe sobre o Sistema Nacional de Avaliação da Educação Básica (Saeb). **Diário Oficial da União**, Brasília, DF: Presidência da República, 2013. Disponível em: https://download.inep.gov.br/educacao_basica/prova_brasil_saeb/legislacao/2013/portaria_n_482_07062013_mec_inep_saeb.pdf. Acesso em: 22 fev. 2022.

BRASIL. Lei n.º 13.005, de 25 de junho de 2014. Aprova o Plano Nacional de Educação (PNE) e dá outras providências. **Diário Oficial da União**, Brasília, DF: Presidência da República, 2014. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/113005.htm. Acesso em: 22 fev. 2022.

BRASIL. Portaria n.º 410, de 22 de julho de 2016. **Diário Oficial da União**, Brasília, DF: Presidência da República, 2016. Disponível em:

https://download.inep.gov.br/educacao_basica/prova_brasil_saeb/legislacao/2016/portaria_n410_22072016_ana.pdf. Acesso em: 22 fev. 2022.

BRASIL. Portaria nº 564, de 19 de abril de 2017. Altera a Portaria MEC nº 482, de 7 de junho de 2013 que dispõe sobre o Sistema de Avaliação da Educação Básica (Saeb) e dá outras providências. **Diário Oficial da União**, Brasília, DF: Presidência da República, 2017a. Disponível em:

http://download.inep.gov.br/educacao_basica/prova_brasil_saeb/legislacao/2017/Portaria_mec_gm_n564_de_19042017_saeb.pdf. Acesso em: 22 fev. 2022.

BRASIL. Portaria nº 447, de 24 de maio de 2017. Estabelece diretrizes para o planejamento e a operacionalização do Sistema de Avaliação da Educação Básica (Saeb) no ano de 2017.

Diário Oficial da União, Brasília, DF: Presidência da República, 2017b. Disponível em: https://download.inep.gov.br/educacao_basica/saeb/2017/legislacao/portaria_n447_24052017.pdf. Acesso em: 22 fev. 2022.

BRASIL. Ministério da Educação. **Base Nacional Comum Curricular**. Brasília, DF: MEC, 2018a. Disponível em:

http://basenacionalcomum.mec.gov.br/images/BNCC_EI_EF_110518_versaofinal_site.pdf. Acesso em: 22 fev. 2022.

BRASIL. **Relatório SAEB/ANA 2016**: panorama do Brasil e dos estados. Brasília, DF: Inep/MEC, 2018b. Disponível em:

https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_exames_da_educacao_basica/relatorio_saeb_ana_2016_panorama_do_brasil_e_dos_estados.pdf. Acesso em: 22 fev. 2022.

BRASIL. Portaria nº 366, de 29 de abril de 2019. Estabelece as diretrizes de realização do Sistema de Avaliação da Educação Básica (Saeb) no ano de 2019. **Diário Oficial da União**, Brasília, DF: Presidência da República, 2019. Disponível em:

https://download.inep.gov.br/educacao_basica/saeb/2019/legislacao/portaria_n366_29042019.pdf. Acesso em: 22 fev. 2022.

BRASIL. Portaria nº 458, de 5 maio de 2020. Institui normas complementares necessárias ao cumprimento da Política Nacional de Avaliação da Educação Básica. **Diário Oficial da União**, Brasília, DF: Presidência da República, 2020a. Disponível em:

<https://www.in.gov.br/web/dou/-/portaria-n-458-de-5-de-maio-de-2020-255378342>. Acesso em: 17 fev. 2022.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Matrizes de Referência de língua portuguesa e matemática do Saeb**: documento de referência do ano de 2001. Brasília, DF: Inep, 2020b. Disponível em:

https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_exames_da_educacao_basica/matriz_de_referencia_de_lingua_portuguesa_e_matematica_do_saeb.pdf. Acesso em: 08 mar. 2021.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Histórico do Saeb atualizado em 2020**. Brasília, DF: Inep, 2020c. Disponível em:

<https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/saeb/historico>. Acesso em: 22 fev. 2022.

BRASIL. Emenda Constitucional nº. 108, de 26 de agosto de 2020. Altera a Constituição Federal para estabelecer critérios de distribuição da cota municipal do Imposto sobre

Operações Relativas à Circulação de Mercadorias e sobre Prestações de Serviços de Transporte Interestadual e Intermunicipal e de Comunicação (ICMS), para disciplinar a disponibilização de dados contábeis pelos entes federados, para tratar do planejamento na ordem social e para dispor sobre o Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (Fundeb); altera o Ato das Disposições Constitucionais Transitórias; e dá outras providências. **Diário Oficial da União**, Brasília, DF: Presidência da República, 2020d. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/emendas/emc/emc108.htm. Acesso em: 22 fev. 2022.

BRASIL. **Conheça o novo Fundeb, que amplia gradualmente os recursos da educação**. Agência Câmara de Notícias, 2020e. Disponível em: <https://www.camara.leg.br/noticias/687499-conheca-o-novo-fundeb-que-amplia-gradualmente-os-recursos-da-educacao/>. Acesso em: 22 fev. 2022.

BRASIL. **Novo Fundeb garante mais recursos da União para a educação básica brasileira**. Fundo Nacional de Desenvolvimento da Educação, 2020f. Disponível em: <https://www.fnde.gov.br/index.php/aceso-a-informacao/institucional/area-de-imprensa/noticias/item/14012-novo-fundeb-garante-mais-recursos-da-uni%C3%A3o-para-a-educac%C3%A7%C3%A3o-b%C3%A1sica-brasileira>. Acesso em: 22 fev. 2022.

BRASIL. Portaria nº 10, de 8 de janeiro de 2021. Estabelece parâmetros e fixa diretrizes gerais para implementação do Sistema de Avaliação da Educação Básica (Saeb) no âmbito da Política Nacional de Avaliação da Educação Básica. **Diário Oficial da União**, Brasília, DF: Presidência da República, 2021a. Disponível em: <https://www.in.gov.br/en/web/dou/-/portaria-n-10-de-8-de-janeiro-de-2021-298322305>. Acesso em: 22 fev. 2022.

BRASIL. Portaria nº 250, de 5 de julho de 2021. Estabelece as diretrizes de realização do Sistema de Avaliação da Educação Básica (Saeb) no ano de 2021. **Diário Oficial da União**, Brasília, DF: Presidência da República, 2021b. Disponível em: <https://www.in.gov.br/en/web/dou/-/portaria-n-250-de-5-de-julho-de-2021-330276260>. Acesso em: 22 fev. 2022.

BRASIL Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Inep se prepara para implementação do Novo Saeb em 2021**. Brasília, DF: Inep/MEC, 2021c. Disponível em: http://portal.inep.gov.br/artigo/-/asset_publisher/B4AQV9zFY7Bv/content/inep-se-prepara-para-implantacao-do-novo-saeb-em-2021/21206. Acesso em: 22 fev. 2021.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Nota de esclarecimento**: Sistema de Avaliação da Educação Básica (Saeb) 2021. Brasília, DF: Inep, 2021d. Disponível em: <https://www.gov.br/inep/pt-br/assuntos/noticias/saeb/nota-de-esclarecimento-sistema-de-avaliacao-da-educacao-basica-saeb-2021>. Acesso em: 27 jul. 2021.

BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). **Censo Escolar**. Brasília, DF: INEP, 2021e. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-escolar>. Acesso em: 22 fev. 2022.

BRAUN, Virginia; CLARKE, Victoria. Using thematic analysis in psychology. **Qualitative Research in Psychology**, v. 3, n. 2, p. 77-101, jul./set. 2006. DOI: <http://doi.org/10.1191/1478088706qp063oa>

- BROOKE, Nigel. Responsabilização educacional no Brasil. **RIEE. Revista Iberoamericana de Evaluación Educativa**, 2008. Disponível em: <https://revistas.uam.es/riee/article/view/4684>. Acesso em: 20 fev. 2022.
- CABRAL NETO, Antônio. **Política Educacional no Projeto Nordeste**: discursos, embates e práticas. Natal: EDUFRN, 1997.
- CARNOY, Martin; KHAVENSON, Tatiana; FONSECA, Izabel; COSTA, Leandro; MAROTTA, Luana. A educação brasileira está melhorando? Evidências do Pisa e Saeb. **Cadernos de Pesquisa**, São Paulo, v. 45, n. 157, p. 450-485, jul./set. 2015. DOI: <http://doi.org/10.1590/198053143331>
- CASTILLO ARREDONDO, Santiago; DIAGO, Jesús Cabrerizo. **Avaliação educacional e promoção escolar**. São Paulo: Unesp, 2009.
- CASTRO, Vanessa Gomes de; TAVARES JÚNIOR, Fernando. Jovens em contextos sociais desfavoráveis e sucesso escolar no ensino médio. **Educação & Realidade**, Porto Alegre, v. 41, n. 1, p. 239-258, jan./mar. 2016. DOI: <http://dx.doi.org/10.1590/2175-623656080>
- CASTRO, Maria Helena Guimarães de. O Saeb e a agenda de reformas educacionais: 1995 a 2002. **Em Aberto**, Brasília, v. 29, n. 96, p. 85-98, maio/ago. 2016. DOI: <https://doi.org/10.24109/2176-6673.emaberto.29i96.%25p>
- CAPES. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. **Relatório do Qualis Periódicos – Área 38: Educação**. Brasília, DF: Capes, 2019. Disponível em: <https://www.gov.br/capes/pt-br/centrais-de-conteudo/relatorio-qualis-educacao-pdf>. Acesso em: 16 abr. 2020.
- CERTEAU, Michel. **A escrita da história**. Rio de Janeiro: Forense Universitária, 2011.
- COELHO, Maria Inês de Matos. Vinte anos de avaliação da educação básica no Brasil: aprendizagens e desafios. **Ensaio: Avaliação e Políticas Públicas em Educação**, Rio de Janeiro, v. 16, n. 59, p. 229-258, jun./jul. 2008. DOI: <http://doi.org/10.1590/S0104-40362008000200005>
- COTTA, Tereza Cristina. Avaliação educacional e políticas públicas: a experiência do Sistema Nacional de Avaliação da Educação Básica (Saeb). **Revista do Serviço Público**, Brasília, v. 52, n. 4, p. 89-111, out./dez. 2001. DOI: <https://doi.org/10.21874/rsp.v52i4.316>
- CRONBACH, Lee J. Course improvement through evaluation. **Evaluation models**. Springer, Dordrecht, 2000. p. 235-247.
- DALBEN, Ângela Imaculada Loureiro de Freitas. **Avaliação educacional**. Belo Horizonte: Editora UFMG, 2008.
- DANTAS, Claudio Rejane da Silva; MASSONI, Neusa Teresinha; SANTOS, Flávia Maria Teixeira. A avaliação no Ensino de Ciências Naturais nos documentos oficiais e na literatura acadêmica: uma temática com muitas questões em aberto. **Ensaio: Avaliação e Políticas Públicas em Educação**. Rio de Janeiro, v. 25, n. 95, p. 440-482, abr./jun. 2017. DOI: <http://doi.org/10.1590/s0104-40362017002500807>
- DICKEL, Adriana. A Avaliação Nacional da Alfabetização no contexto do Sistema de Avaliação da Educação Básica e do Pacto Nacional pela Alfabetização na Idade Certa: responsabilização e controle. **Cadernos Cedes**, Campinas, SP, v. 36, n. 99, p. 193-206, maio/ago. 2016. DOI: <http://dx.doi.org/10.1590/CC0101-32622016162940>

- FERNANDES, Reynaldo; GREMAUD, Amaury Patrick. Qualidade da educação: avaliação, indicadores e metas. Educação básica no Brasil: construindo o país do futuro. **Elsevier**, Rio de Janeiro, v. 1, p. 213-238, 2009. Disponível em: https://www.cps.fgv.br/ibrecps/rede/seminario/reynaldo_paper.pdf. Acesso em: 12 mai. 2021.
- FERNANDES, Reynaldo. A universalização da avaliação e a criação do Ideb: pressupostos e perspectivas. **Em Aberto**, Brasília, v. 29, n. 96, 2016. DOI: <https://doi.org/10.24109/2176-6673.emaberto.29i96.%25p>
- FERRÃO, Maria Eugénia; BARROS, Gabriela Thamara de Freitas; BOF, Alvana Maria; OLIVEIRA, Adolfo Samuel. Estudo Longitudinal sobre Eficácia Educacional no Brasil: Comparação entre Resultados Contextualizados e Valor Acrescentado. **Dados – Revista de Ciências Sociais**, Rio de Janeiro, v. 61, n. 4, p. 265-300, out./dez. 2018. DOI: <http://dx.doi.org/10.1590/001152582018160>
- FINI, Maria Inês. Currículo e avaliação: articulação necessária em favor da aprendizagem dos alunos da rede pública de São Paulo. **São Paulo em Perspectiva**, São Paulo, v. 23, n. 1, p. 57-72, jan./jun. 2009. Disponível em: http://produtos.seade.gov.br/produtos/spp/v23n01/v23n01_05.pdf. Acesso em: 14 dez. 2020.
- FONSECA, Marília. O financiamento do Banco Mundial à educação brasileira: vinte anos de cooperação internacional. In: DE TOMMASI, Lívia; WARDE, Mirian Jorge; HADDAD, Sérgio. (org.). **O Banco Mundial e as políticas educacionais**. São Paulo: Cortez Editora, 1996. p. 229-252.
- FRANCO, Ana Maria Paiva; MENEZES-FILHO, Naércio. Uma análise de rankings de escolas brasileiras com dados do SAEB. **Estudos Econômicos**, São Paulo, v. 42, n. 2, p. 263-283, abr./jun. 2012. DOI: <https://doi.org/10.1590/S0101-41612012000200002>
- FREITAS, Dirce Nei Teixeira. A Avaliação Educacional como objeto de recomendações internacionais. **Estudos em Avaliação Educacional**, São Paulo, v. 16, n. 31, p. 79-100, jan./jun. 2005. DOI: <https://doi.org/10.18222/eae163120052143>
- FREITAS, Dirce Nei Teixeira. Avaliação da educação básica no Brasil: características e pressupostos. In: BAUER, Adriana; GATTI, Bernadette; TAVARES, Marialva R. (org.). **Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origens e pressupostos**. Florianópolis: Insular, 2013. p. 70-96.
- G1. Novo Fundeb: entenda como vai funcionar o fundo que financia a educação básica.** 26 dez. 2020. Disponível em: <https://g1.globo.com/educacao/noticia/2020/12/26/novo-fundeb-entenda-como-vai-funcionar-o-fundo-que-financia-a-educacao-basica.ghtml>. Acesso em: 22 fev. 2022.
- GATTI, Bernadette. A avaliação educacional no Brasil: experiências, problemas, recomendações. **Estudos em Avaliação Educacional**, São Paulo, n. 10, p. 67-80, jul./dez. 1994. DOI: <http://dx.doi.org/10.18222/eae01019942320>
- GATTI, Bernadette. Avaliação educacional no Brasil: pontuando uma história de ações. **EccoS – Revista Científica**, São Paulo, v. 4, n. 1, p. 17-41, abr./jun. 2002. DOI: <https://doi.org/10.5585/eccos.v4i1.291>
- GATTI, Bernadette. Avaliação de sistemas educacionais no Brasil. **Sísifo**, Lisboa, n. 9, p. 7-18, maio/ago. 2009. Disponível em: <http://sisifo.ie.ulisboa.pt/index.php/sisifo/article/view/144>. Acesso em: 26 jun. 2020.

GATTI, Bernadette. Possibilidades e fundamentos de avaliações em larga escala: primórdios e perspectivas. In: BAUER, Adriana; GATTI, Bernadette; TAVARES, Marialva R. (org). **Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origens e pressupostos**. Florianópolis: Insular, 2013. p. 47-69.

GATTI, Bernadete. Avaliação: contexto, história e perspectivas. **Olhares: Revista do Departamento de Educação da Unifesp**, São Paulo, v. 2, n. 1, p. 08-26, maio. 2014.

GATTI, Bernadette; VIANNA, Heraldo Marelim; DAVIS, Cláudia. Problemas e impasses da avaliação de projetos e sistemas educacionais: dois casos brasileiros. **Estudos em Avaliação Educacional**, São Paulo, n. 4, p. 7-26, jul./dez. 1991. DOI: <http://dx.doi.org/10.18222/ea00419912374>

GINZBURG, Carlo. **Mitos, emblemas, sinais: morfologia e história**. São Paulo: Companhia das Letras, 1989.

GOBBI, Beatriz Christo; LACRUZ, Adonai José; AMÉRICO, Bruno Luiz; ZANQUETTE FILHO, Hélio. Uma boa gestão melhora o desempenho da escola, mas o que sabemos acerca do efeito da complexidade da gestão nessa relação? **Ensaio: Avaliação e Políticas Públicas em Educação**, Rio de Janeiro, v. 28, n. 106, p. 198-220, jan./mar. 2020. DOI: <https://doi.org/10.1590/s0104-40362019002701786>

GOIÁS. **Prêmio Escola**. Governo do Estado de Goiás, 2011. Disponível em: <https://www.yumpu.com/pt/document/read/35388547/cartilha-premio-escola-secretaria-da-educacao-do-estado-de-goias>. Acesso em: 22 fev. 2022.

GOSDEN, Chris. **Pré-história**. Porto Alegre: L&PM, 2012.

HARBISON, Ralph W.; HANUSHEK, Eric A. **Educational performance of the poor: lessons from rural Northeast Brazil**. New York: Oxford University Press, 1992. Disponível em: <https://documents.worldbank.org/pt/publication/documents-reports/documentdetail/310981468770471105/educational-performance-of-the-poor-lessons-from-rural-northeast-brazil>. Acesso em: 23 jul. 2020.

HORTA NETO, João Luiz. **Avaliação externa: a utilização dos resultados do Saeb 2003 na gestão do sistema público de ensino fundamental no Distrito Federal**. 2006. 144 f. Dissertação (Mestrado em Educação) - Universidade de Brasília, Brasília, 2006. Disponível em: <https://repositorio.unb.br/handle/10482/5811>. Acesso em: 11 jan. 2022.

HORTA NETO, João Luiz. Um olhar retrospectivo sobre a avaliação externa no Brasil: das primeiras mediações até o Saeb de 2005. **Revista Iberoamericana de Educación**, Madrid, v. 42, n. 5, jan./abr. 2007. DOI: <https://doi.org/10.35362/rie4252398>

HORTA NETO, João Luiz. Avaliação educacional no Brasil para além dos testes cognitivos/Educational evaluation in Brazil beyond cognitive tests. **Revista de Educação PUC-Campinas**, Campinas, v. 23, n. 1, p. 37-53, mar. 2018. DOI: <https://doi.org/10.24220/2318-0870v23n1a3990>

KAUKO, Jaakko; CENTENO, Vera Gorodski; CANDIDO, Helena; SHIROMA, Eneida; KLUTAS, Anni. The emergence of quality assessment in Brazilian basic education. **European Educational Research Journal**, Thousand Oaks, v. 15, n. 5, p. 558-579, set. 2016. DOI: <http://doi.org/10.1177/1474904116662889>

- KELLEY, T. L. The selection of upper and lower groups for the validation of test items. **Journal of educational psychology**, Warwick & York, v. 30, n. 1, p. 17-24, 1939. Disponível em: <https://psycnet.apa.org/record/1939-03313-001>. Acesso em: 13 de jul. 2020.
- KLEIN, Ruben. Utilização da teoria de resposta ao item no Sistema Nacional de Avaliação da Educação Básica (Saeb). **Meta: Avaliação**, Rio de Janeiro, v. 1, n. 2, p. 125-140, jul./set. 2009. DOI: <http://dx.doi.org/10.22347/2175-2753v1i2.38>
- KLEIN, Ruben. Alguns aspectos da teoria de resposta ao item relativos à estimação das proficiências. **Ensaio: avaliação e políticas públicas em educação**, Rio de Janeiro, v. 21, n. 78, p. 35-56, jan./mar. 2013a. DOI: <https://doi.org/10.1590/S0104-40362013005000003>
- KLEIN, Ruben. Aspectos Metodológicos e Técnicos: delineamentos assumidos nas avaliações, limites e perspectivas de aprimoramento. In: BAUER, Adriana; GATTI, Bernadette; TAVARES, Marialva R. (org). **Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origens e pressupostos**. Florianópolis: Insular, 2013b. p. 109-130.
- LOCATELLI, Iza. Construção de instrumentos para a avaliação de larga escala e indicadores de rendimento: o modelo SAEB. **Estudos em Avaliação Educacional**, São Paulo, n. 25, p. 3-21, jan./jun. 2002. DOI: <https://doi.org/10.18222/eae02520022189>
- MARANHÃO. **Plano Estadual de Educação do Estado do Maranhão**. Disponível em: https://www.educacao.ma.gov.br/files/2016/05/suplemento_lei-10099-11-06-2014-PEE.pdf. Acesso em: 22 fev. 2022.
- MARANHÃO. Portaria nº. 814, de 21 de maio de 2019. Dispõe sobre a operacionalização do artigo 3º, IX e X, Lei nº 10.995, de 11 de março de 2019, e dá outras providências. **Governo do Estado do Maranhão**, 2019.
- MARQUES, Rodrigo. **A Educação Física no Exame Nacional do Ensino Médio: trajetória, conceitos e práticas pedagógicas**. 2020. Tese (Doutorado em Educação Física) – Universidade Federal do Espírito Santo, Vitória, ES, 2020. Disponível em: <https://educacaofisica.ufes.br/pt-br/pos-graduacao/PPGEF/detalhes-da-tese?id=13947>. Acesso em: 4 fev. 2022.
- MARQUES, Rodrigo; STIEG, Ronildo; SANTOS, Wagner dos. Exames standardizados: uma análise dos modelos e das teorias na produção acadêmica. **Meta: avaliação**, Rio de Janeiro, v. 12, n. 34, p. 1-27, jan./mar. 2020. DOI: <http://dx.doi.org/10.22347/2175-2753v12i34.2342>
- MARQUES, Rodrigo; STIEG, Ronildo ; CUNHA DE PAULA, Sayonara; LOPES NEGREIROS, Heitor; DOS SANTOS, Wagner. Educación física en el Examen Nacional de Escuela Secundaria de Brasil: análisis de las preguntas en el periodo 2009-2017. **Revista Calidad en la Educación**, Santiago, v. 53, p. 113-146, dez. 2020. DOI: <http://dx.doi.org/10.31619/caledu.n53.834>
- MATOS, Daniel Abud Seabra; RODRIGUES, Erica Castilho. Indicadores educacionais e contexto escolar: uma análise das metas do Ideb. **Estudos em Avaliação Educacional**, São Paulo, v. 27, n. 66, p. 662-688, set./dez. 2016. DOI: <http://dx.doi.org/10.18222/eae.v27i66.4012>
- MENDONÇA, Liliane de Paula. **A reforma educacional na América Latina e a disseminação de políticas públicas de avaliação de sistemas educacionais**. Dissertação (Mestrado em Ciências Sociais) – Universidade Federal de Juiz de Fora, Juiz de Fora, MG,

2014. Disponível em: <http://repositorio.ufjf.br:8080/jspui/handle/ufjf/1830>. Acesso em: 22 fev. 2022.

MENEZES, Dyelle. MEC premia alunos e diretores de escolas públicas de todas as regiões do país. **Portal MEC**, 2019. Disponível em: <http://portal.mec.gov.br/component/content/article/211-noticias/218175739/79921-mec-premia-alunos-e-diretores-de-escolas-publicas-de-todas-as-regioes-do-pais?Itemid=164>. Acesso em: 15 fev. 2022.

MENEZES-FILHO, Naércio; VASCONCELLOS, Lígia; WERLANG, Sérgio Ribeiro da Costa; BIONDI, Roberta Loboda. Avaliando o impacto da progressão continuada nas taxas de rendimento e desempenho escolar do Brasil. **Relatório de Avaliação Econômica**, São Paulo: Fundação Itaú Social, 2009. Disponível em: <http://pdfs.semanticscholar.org/5de6/6ca73a91c7c7215046c948fa6031d994083c.pdf>. Acesso em: 10 jun. 2020.

MOHER, David; LIBERATI, Alessandro; TETZLAFF, Jennifer; ALTMAN, Douglas. G. Preferred reporting items for systematic reviews and meta-analyses: the Prisma statement. **PLOS Medicine**, Califórnia, v. 6, n. 7, jul. 2009. DOI: <http://doi.org/10.1371/journal.pmed.1000097>

MUÑIZ, J. **Teoría de respuesta a los ítems**: Un nuevo enfoque en la evolución psicológica y educativa. Madri: Ediciones Pirámide, S. A, 1990.

OLIVEIRA, Ana Cristina Prado de; CARVALHO, Cynthia Paes de. Gestão escolar, liderança do diretor e resultados educacionais no Brasil. **Revista Brasileira de Educação**, Rio de Janeiro, v. 23, nov./mar. 2018. DOI: <https://doi.org/10.1590/s1413-24782018230015>

OLIVEIRA, Maria Marly de. **Como fazer pesquisa qualitativa**. Petrópolis: Vozes, 2007.

PASCHOALIN, Pâmela Cristina Botiglieri. **Heraldo Marelim Vianna na história da avaliação educacional no Brasil**. 2019. Dissertação (Mestrado em Educação) - Universidade Federal de São Carlos, São Carlos, 2019. Disponível em: <https://repositorio.ufscar.br/handle/ufscar/11249>. Acesso em: 11 jan. 2022.

PASQUALI, Luiz. **Psicometria**: teoria dos testes na psicologia e na educação. Petrópolis: Vozes, 2003. 397 p.

PASQUALI, Luiz. Psicometria. **Revista da Escola de Enfermagem da USP**, São Paulo, v. 43, n. SPE, p. 992-999, 2009.

PASQUALI, Luiz. **TRI–Teoria de resposta ao item**: Teoria, procedimentos e aplicações. Editora Appris, 2018.

PEREIRA, Sandra Márcia Campos. **Projeto Nordeste de Educação Básica e o Fundescola: uma análise do discurso governamental e do banco mundial sobre a qualidade da educação**. 2007. 170 f. Tese (Doutorado em Educação) – Universidade Estadual Paulista, Araraquara, SP, 2007. Disponível em: <http://hdl.handle.net/11449/101593>. Acesso em: 17 maio 2020.

PESTANA, Maria Inês. O Sistema Nacional de Avaliação da Educação Básica. **Estudos em Avaliação Educacional**, São Paulo, n. 05, p. 81-83, jan./jun. 1992. DOI: <http://dx.doi.org/10.18222/eae00519922366>

PESTANA, Maria Inês. O sistema de avaliação brasileiro. **Revista brasileira de Estudos Pedagógicos**, Brasília, v. 79, n. 191, p. 65-73, jan./abr. 1998. DOI: <https://doi.org/10.24109/2176-6681.rbep.79i191.1044>

PESTANA, Maria Inês. Trajetória do Saeb: criação, amadurecimento e desafios. **Em Aberto**, Brasília, v. 29, n. 96, p. 71-84, maio/ago. 2016. DOI: <http://doi.org/10.24109/2176-6673.emaberto.29i96.%25p>

RABELO, Mauro Luiz. **Avaliação Educacional**: Fundamentos. Metodologia e Aplicações no Contexto Brasileiro. Rio de Janeiro: Coleção Profmat, Sociedade Brasileira de Matemática, 2013.

RIO DE JANEIRO. Decreto nº. 25.959, de 12 de janeiro de 2000. Institui o Programa Estadual de Reestruturação da Educação Pública (Programa Nova Escola) e dá outras providências. **Governo do Estado do Rio de Janeiro**, 2000.

SÁ-SILVA, Jackson Ronie; ALMEIDA, Cristóvão Domingos de; GUINDANI, Joel Felipe. Pesquisa documental: pistas teóricas e metodológicas. **Revista brasileira de história & ciências sociais**, Rio Grande, v. 1, n. 1, p. 1-15, jan./jun. 2009. Disponível em: <https://periodicos.furg.br/rbhcs/article/view/10351>. Acesso em: 10 fev. 2022.

SCHNEIDER, Marilda Pasqual; NARDI, Elton Luiz. O IDEB e a construção de um modelo de accountability na educação básica brasileira. **Revista Portuguesa de Educação**, Braga, v. 27, n. 1, p. 07-28, 2014. DOI: <https://doi.org/10.21814/rpe.4295>

SILVA, Gustavo Felipe C.; FILGUEIRAS, Juliana M.; VIEIRA, Juliana S. O ginásio “Inconfidência” de Alfenas: vestígios da cultura escolar e do ensino de história. **Educação em Revista**, Belo Horizonte, v. 36, p. 1-19, fev./jun. DOI: <https://doi.org/10.1590/0102-4698219654>

SOARES, Denilson Junio Marques. **Teoria clássica dos testes e teoria de resposta ao item aplicadas em uma avaliação de matemática básica**. 2018. 121 f. Dissertação (Mestrado em Estatística) - Universidade Federal de Viçosa, Viçosa, MG, 2018.

SOARES, Denilson Junio Marques; SOARES, Talita Emidio Andrade; SANTOS, Wagner dos. Infraestrutura e desempenho escolar na Prova Brasil: aspectos e conexões. **Olhar de Professor**, Ponta Grossa, v. 23, p. 1-18, jul. 2020. DOI: <https://doi.org/10.5212/10.5212/OlharProfr.v.23.2020.15023.209209227242.0626>

SOARES, Denilson Junio Marques; SOARES, Talita Emidio Andrade; SANTOS, Wagner dos. Análise da qualidade psicométrica da prova de matemática do Exame Nacional do Ensino Médio brasileiro de 2018. **Actualidades Investigativas en Educación**, San José, v. 21, p. 1-28, jan./abr. 2021. DOI: <http://dx.doi.org/10.15517/aie.v21i1.42338>

SOARES, Talita Emidio Andrade; SOARES, Denilson Junio Marques; SANTOS, Wagner dos. Sistema de Avaliação da Educação Básica (Saeb): Revisão Sistemática da Literatura. **Estudos em Avaliação Educacional**, São Paulo, v. 32, p. 1-25, jul. 2021a. DOI: <https://doi.org/10.18222/eae.v32.7839>

SOARES, Talita Emidio Andrade; SOARES, Denilson Junio Marques; SANTOS, Wagner dos. Medidas de Tendência Central: análise da qualidade das questões do ENEM de 2016 a 2018. **Jornal Internacional de Estudos em Educação Matemática**, Londrina, v. 14, n. 4, p. 119-128, abr. 2021b. DOI: <https://doi.org/10.17921/2176-5634.2021v14n1p119-128>

SOARES NETO, Joaquim José; JESUS, Girlene Ribeiro de; KARINO, Camila Akemi; ANDRADE, Dalton Francisco de. Uma escala para medir a infraestrutura escolar. **Estudos em Avaliação Educacional**, São Paulo, v. 24, n. 54, p. 78-99, jan./abr. 2013. DOI: <https://doi.org/10.18222/ea245420131903>

SOUSA, Clarilza Prado. Descrição de uma trajetória na/da avaliação educacional. **Idéias**, São Paulo, n. 30, p. 161-174, 1998. Disponível em: http://www.crmariocovas.sp.gov.br/pdf/ideias_30_p161-174_c.pdf. Acesso em: 10 jan. 2022.

SOUSA, Sandra Zákia. Avaliação da aprendizagem nas pesquisas no Brasil de 1930 a 1980. **Cadernos de Pesquisa**, São Paulo, n. 94, p. 43-49, ago. 1995. Disponível em: <http://publicacoes.fcc.org.br/index.php/cp/article/view/837>. Acesso em: 12 jan. 2022

SOUSA, Sandra Zákia; OLIVEIRA, Romualdo Portela. Sistemas estaduais de avaliação: uso dos resultados, implicações e tendências. **Cadernos de Pesquisa**, São Paulo, v. 40, n. 141, p. 793-822, set./dez. 2010. DOI: <http://doi.org/10.1590/S0100-15742010000300007>

SOUZA, Luciana Karine. Pesquisa com análise qualitativa de dados: conhecendo a Análise Temática. **Arquivos Brasileiros de Psicologia**, Rio de Janeiro, v. 71, n. 2, p. 51-67, jul./dez. 2019. Disponível em http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1809-52672019000200005&lng=pt&nrm=iso. Acesso em: 2 jun. 2020.

STADLER, Jocasta Conceição. **Prova Brasil de Matemática do 5º ano do Ensino Fundamental**: resultados nas plataformas devolutivas pedagógicas e QEdu. 2017. 165f. Dissertação (Mestrado em Educação). Universidade Estadual de Ponta Grossa, Ponta Grossa, PR, 2017.

TAVARES, Antônio Vanderlei. **Avaliação de Larga Escala**: resultados e tomada de decisão. Dissertação (Mestrado em Educação). Pontifícia Universidade Católica de São Paulo, São Paulo, SP, 2012.

TODOS PELA EDUCAÇÃO. **Novo Fundeb**: o que muda se for aprovada a PEC 26/2020 no Senado Federal, 2020. Disponível em: <https://todospelaeducacao.org.br/noticias/novo-fundeb-pec-26-2020-no-senado-federal/>. Acesso em: 17 fev. 2022.

TÔRRES, Frederico Carvalho. **Uma aplicação da teoria de resposta ao item em um simulado de matemática no modelo ENEM**. 2015. 116 p. Dissertação (Mestrado em Matemática) – Programa de Mestrado Profissional em Matemática em Rede Nacional, Universidade de Brasília, Brasília, DF, 2015.

UCZAK, Lucia Hugo. **O PREAL e as políticas de avaliação educacional para a América Latina**. 2014. 211 f. Tese (Doutorado em Educação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, 2014.

VIANNA, Heraldo Marelim; GATTI, Bernadette. Avaliação do rendimento de alunos de escolas de 1º grau da rede pública: uma aplicação experimental em 10 cidades. **Educação e Seleção**, São Paulo, n. 17, p. 5-52, jul./nov. 1988. Disponível em: <http://publicacoes.fcc.org.br/ojs/index.php/edusel/article/view/2626>. Acesso em: 5 jan. 2022.

VIANNA, Heraldo Marelim. Avaliação do rendimento de alunos de escolas de 1º grau da rede pública: um estudo em 20 cidades. **Educação e Seleção**, São Paulo, n. 19, p. 33-98, jul./nov. 1989a. Disponível em: <http://publicacoes.fcc.org.br/ojs/index.php/edusel/article/view/2646>. Acesso em: 28 dez. 2021.

VIANNA, Heraldo Marelim. Avaliação do rendimento de alunos de escolas de 1º grau da rede pública: um estudo em 39 cidades. **Educação e Seleção**, São Paulo, n. 20, p. 5-56, jul./nov. 1989b. Disponível em: <http://publicacoes.fcc.org.br/ojs/index.php/edusel/article/view/2647>. Acesso em: 28 dez. 2021.

VIANNA, Heraldo Marelim. Avaliação educacional: uma perspectiva histórica. **Estudos em Avaliação Educacional**, São Paulo, n. 12, p. 7-24, jul./dez. 1995. DOI: <http://dx.doi.org/10.18222/eae01219952297>

VIANNA, Heraldo Marelim. Estudos em avaliação educacional: Tyler, Cronbach e Scriven. **Estudos em Avaliação Educacional**, São Paulo, n. 17, p. 73-135, jan./jun. 1998. DOI: <https://doi.org/10.18222/eae01719982263>

VIANNA, Heraldo Marelim. **Avaliação em debate: Saeb, Enem, Provão**. Brasília: Liber Livro, 2010.

WASELFISZ, Júlio Jacobo; HORTA NETO, João Luiz. As origens do Saeb. **Em Aberto**, Brasília, v. 29, n. 96, p. 177-193, maio/ago. 2016. DOI: <https://doi.org/10.24109/2176-6673.emaberto.29i96.2705>

WASELFISZ, Júlio Jacobo. O Sistema Nacional de Avaliação do Ensino Público de 1º Grau. **Estudos em Avaliação Educacional**, São Paulo, n. 4, p. 65-73, jul./nov. 1991. DOI: <http://dx.doi.org/10.18222/eae00419912376>

WERLE, Flávia Obino Corrêa. Sistema de avaliação da educação básica no Brasil: abordagem por níveis de segmentação. In: Flávia Obino Corrêa Werle. (org.). **Avaliação em larga escala: foco na escola**. 1ed. São Leopoldo; Brasília: Oikos; Liberlivro, 2010, v.1, p. 21-36.