

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE CAMPINAS  
CENTRO DE CIÊNCIAS HUMANAS E SOCIAIS APLICADAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO**

**THALITA REGINA GARCIA DA SILVA**

**META-AVALIAÇÃO DO PROGRAMA DE AVALIAÇÃO EXTERNA EM LARGA  
ESCALA IMPLEMENTADO POR UMA REDE PRIVADA DE EDUCAÇÃO BÁSICA  
DE ABRANGÊNCIA NACIONAL**

**Campinas**

**2022**

Ficha catalográfica elaborada por Adriane Elane Borges de Carvalho CRB 8/9313  
Sistema de Bibliotecas e Informação - SBI - PUC-Campinas

372 R586m	<p data-bbox="526 940 718 974">Silva, Thalita Regina da</p> <p data-bbox="526 985 1228 1052">Meta-avaliação do programa de avaliação externa em larga escala implemetado por uma rede privada de Educação Básica de abrangência nacional / Thalita Regina da Silva. - Campinas: PUC-Campinas, 2021.</p> <p data-bbox="526 1064 638 1097">275 f.: il.</p> <p data-bbox="526 1097 877 1131">Orientador: Mônica Piccione Gomes Rios.</p> <p data-bbox="526 1142 1228 1209">Tese (Doutorado em Educação) - Programa de Pós-Graduação em Educação, Centro de Ciências Humanas e Sociais Aplicadas, Pontifícia Universidade Católica de Campinas, Campinas, 2021.</p> <p data-bbox="526 1209 718 1243">Inclui bibliografia.</p> <p data-bbox="526 1254 1228 1355">1. Educação básica . 2. Avaliação . 3. Políticas públicas - Educação . I. Rios, Mônica Piccione Gomes. II. Pontifícia Universidade Católica de Campinas. Centro de Ciências Humanas e Sociais Aplicadas. Programa de Pós-Graduação em Educação. III. Título.</p> <p data-bbox="1069 1355 1228 1377">CDD - 22. ed. 372</p>
--------------	---

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE CAMPINAS  
CENTRO DE CIÊNCIAS HUMANAS E SOCIAIS APLICADAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO**

**THALITA REGINA GARCIA DA SILVA**

**META-AVALIAÇÃO DO PROGRAMA DE AVALIAÇÃO EXTERNA EM LARGA  
ESCALA IMPLEMENTADO POR UMA REDE PRIVADA DE EDUCAÇÃO BÁSICA  
DE ABRANGÊNCIA NACIONAL**

Tese apresentada ao Programa de Pós-graduação em Educação da PUC-Campinas como parte dos requisitos para obtenção do título de doutor em Educação. Orientadora: Prof<sup>a</sup> Dr<sup>a</sup> Mônica Piccione Gomes Rios. Linha de pesquisa: Políticas Públicas em Educação. Grupo de Pesquisa: Avaliação Políticas e Sistemas Educacionais – GRAPSE.

**Campinas**

**2022**

**Thalita Regina Garcia da Silva**

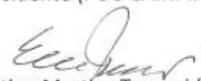
**META-AVALIAÇÃO DO PROGRAMA DE AVALIAÇÃO EXTERNA  
EM LARGA ESCALA IMPLEMENTADO POR UMA REDE  
PRIVADA DE EDUCAÇÃO BÁSICA DE ABRANGÊNCIA  
NACIONAL**

Este exemplar corresponde à  
redação final da Tese de Doutorado em Educação da  
PUC-Campinas, e aprovada pela Banca Examinadora.

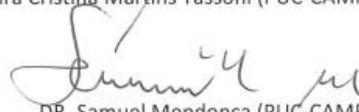
APROVADA: 22 de fevereiro de 2022.



DRA. Monica Piccione Gomes Rios  
Presidente (PUC-CAMPINAS)



DRA. Elvira Cristina Martins Tassoni (PUC-CAMPINAS)



DR. Samuel Mendonça (PUC-CAMPINAS)



DR. Elton Luiz Nardi (UNOESC)



DR. João Luiz Horta Neto (UNB)

Dedico a meu pai (*in memoriam*) que despertou em mim o amor pelas palavras e pelo estudo, tendo estudado tão pouco nas escolas formais, sendo um mestre na escola da vida.

## AGRADECIMENTOS

Exercitar a gratidão num momento de fechamento de ciclo parece ser mais fácil do que manifestá-la como um estilo de vida ao longo do percurso, quando o reconhecimento constante da dependência que temos dos outros requer de nós a humildade de pedir ajuda; de dizer que não sabemos; de pedir perdão; e de aceitar o oferecimento da correção, quando necessitamos. Porém, este espaço se transforma num momento privilegiado de reflexão e de recordação, ao dedicar um tempo para tentar expressar em palavras o sentimento de gratidão que me move neste final de percurso.

Gratidão a Deus, a quem tenho o privilégio de chamar de Pai, pois em todos os momentos que clamei, Ele se fez presente e real. Por vezes, no apoio de um amigo (Armando Gonçalves), na oração de uma amiga (Magali Arnais), ou no sussurrar ao meu ouvido: “Este é o caminho”.

Gratidão a meu pai Aduino Garcia que me ensinou a amar o mundo das palavras, tendo estudado apenas até o quarto ano primário, e que faleceu este ano sem ver mais uma filha “doutora”.

Gratidão a minha mãe Esmeralda Garcia, cuja fé inabalável sobrepuja o mal de Alzheimer e lhe faz lembrar daquilo que me ensinou e dá sentido à minha existência: servir a Deus.

Gratidão a meu amado esposo Jonas Quintela que ofereceu todas as condições para que pudesse estudar, sendo que a mais significativa de todas foi seu amor incondicional.

Gratidão a meu querido irmão Dr. Aduino Garcia de Jesus Jr que me motivou, apoiou, ajudou com ideias e sempre se fez presente nesta empreitada. Minhas irmãs Tania, Telma, Thais e Triana, que são grande inspiração como mulheres, mães e profissionais de excelência.

Gratidão aos docentes e servidores do Programa de Pós-Graduação em Educação da PUC de Campinas, com um carinho especial ao Doutor Adolfo Ignacio Calderón, que me apresentou o Programa, torceu por mim e me fez crescer e desejar voos mais altos.

Gratidão à minha querida orientadora Doutora Mônica Piccione Gomes Rios, que pacientemente me conduziu e orientou, conseguindo enxergar virtudes onde eu só via fraquezas. Abriu os horizontes do conhecimento científico partilhando suas crenças e valores que, graças ao bom Deus, são tão semelhantes aos meus. Fez o difícil tornar-se mais fácil e conseguiu que eu acreditasse que sou capaz. Obrigada!!!

## RESUMO

**Resumo:** As avaliações em larga escala que integram as políticas públicas de avaliação no Brasil avançam para a terceira década de atuação, desde as primeiras iniciativas nos anos 1990. As diferentes esferas do poder público, nos níveis federal, estadual e municipal, têm desenvolvido, para diferentes níveis de ensino, sistemas de avaliação em larga escala cujo principal propósito declarado é a promoção da qualidade educacional. Em nível nacional, a avaliação conhecida como SAEB se sobressaiu, sobretudo em função de compor o Índice de Desenvolvimento da Educação Básica (IDEB). Neste cenário, também há sistemas de avaliação sendo implementados em redes privadas de ensino. O problema gerador da pesquisa questiona ganhos (aspectos desenvolvidos) e dificuldades (aspectos a desenvolver) de um programa de avaliação da educação básica de uma rede privada de extensão nacional, que desenvolveu um programa de avaliação em larga escala com vistas a assegurar a qualidade da educação oferecida, pautado nas mesmas premissas e metodologia adotadas pela então denominada Prova Brasil. Derivado do problema, estabeleceu-se como objetivo geral investigar os ganhos e as dificuldades que podem ser abstraídos da experiência do referido programa de avaliação. A pesquisa adotou uma abordagem qualitativa, sem desprezar os dados quantificáveis. O material empírico foi produzido por meio de análise documental, entrevista semiestruturada com o idealizador do programa e gestores da rede e questionários mistos on-line para gestores e coordenadoras pedagógicas dos níveis de gestão regionais e sub-regionais da rede. Os principais autores que balizam este estudo são Domingos Fernandes, Luiz Fernandes Dourado, Luiz Carlos de Freitas e José Dias Sobrinho. A análise foi realizada em um esforço de diálogo com os autores e os dados levantados pela pesquisa. Por meio de revisão de literatura em bases de dissertações e teses e de periódicos – Biblioteca Digital de Teses e Dissertações (BDTD), Catálogo de Teses e Dissertações da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e na *Scientific Electronic Library Online (SciELO)* –, constatou-se que há várias pesquisas desenvolvidas sobre a temática da meta-avaliação, porém, não há publicação sobre avaliação em larga escala em sistemas privados de ensino. A pesquisa oportunizou a realização de uma meta-avaliação, indicando ganhos e dificuldades do programa investigado. Dentre os ganhos abstraídos destaca-se: a oportunidade de a avaliação externa contribuir para desvelar a realidade; indicar pontos fortes e fracos; permitir a reflexão sobre a prática pedagógica e ser propulsora de mudanças. Dentre as dificuldades destaca-se: a necessidade de formação de professores e gestores para o uso dos resultados; o desenvolvimento de uma cultura avaliativa; o conhecimento dos objetivos da avaliação externa pela comunidade interna e externa das escolas. Considera-se que a investigação contribuiu para a reflexão e produção de conhecimento neste campo de estudos.

**Palavras-chave:** avaliação em larga escala; educação básica; políticas públicas de avaliação.

## ABSTRAT

**Abstrat:** Large-scale evaluations that are part of public evaluation policies in Brazil are advancing into the third decade of operation, since the first initiatives in the 1990s. The different spheres of public power, at the federal, state and municipal levels, have developed, for different levels of education, large-scale assessment systems whose main stated purpose is the promotion of educational quality. At the national level, the assessment known as SAEB stood out, mainly due to its inclusion in the Basic Education Development Index (IDEB). In this scenario, there are also assessment systems being implemented in private education networks. The problem that generates the research questions the gains (aspects developed) and difficulties (aspects to be developed) of a basic education evaluation program of a private network of national extension, which developed a large-scale evaluation program with a view to ensuring the quality of education offered, based on the same premises and methodology adopted by the so-called Prova Brasil. Derived from the problem, it was established as a general objective to investigate the gains and difficulties that can be abstracted from the experience of the aforementioned evaluation program. The research adopted a qualitative approach, without neglecting quantifiable data. The empirical material was produced through document analysis, semi-structured interviews with the program's creator and network managers, and mixed online questionnaires for managers and pedagogical coordinators at the network's regional and sub-regional management levels. The main authors that guide this study are Domingos Fernandes, Luiz Fernandes Dourado, Luiz Carlos de Freitas and José Dias Sobrinho. The analysis was carried out in an effort to dialogue with the authors and the data collected by the research. By means of a literature review in databases of dissertations and theses and journals - Digital Library of Theses and Dissertations (BDTD), Catalog of Theses and Dissertations of the Coordination for the Improvement of Higher Education Personnel (CAPES) and the Scientific Electronic Library Online (SciELO) -, it was found that there are several studies developed on the subject of metaevaluation, however, there is no publication on large-scale evaluation in private education systems. The research provided an opportunity to carry out a metaevaluation, indicating gains and difficulties of the investigated program. Among the abstracted gains, the following stand out: the opportunity for external evaluation to contribute to unveiling the reality; indicate strengths and weaknesses; allow reflection on pedagogical practice and be a driver of changes. Among the difficulties, the following stand out: the need to train teachers and managers to use the results; the development of an evaluative culture; the knowledge of the objectives of the external evaluation by the internal and external community of the schools. It is considered that the investigation contributed to the reflection and production of knowledge in this field of study.

**Keywords:** large-scale assessment; basic education; public evaluation policies.

## LISTA DE FIGURAS

Figura 1 - Organograma da estrutura administrativa da rede .....	121
Figura 2 - Exemplo de distribuição percentual dos alunos na escala de proficiência .....	128
Figura 3 - Comparativo de desempenho dos alunos e participantes de uma unidade escolar e os demais níveis administrativos da rede .....	128
Figura 4 - Comparativo dos níveis de desenvolvimento da Unidade Escolar .....	130
Figura 5 - Resultado disponibilizado para alunos do 5º ano (2013).....	132
Figura 6 - Desempenho apresentado no boletim de aluno de ensino médio .....	132
Figura 7 - Instrumentos e participantes representados na estrutura organizacional .....	173

## LISTA DE GRÁFICOS

Gráfico 1 - Distribuição de trabalhos por ano e por IES .....	322
Gráfico 2 - Distribuição de Trabalhos por ano e IES no Banco de Teses da Capes.....	36
Gráfico 3 - Questão 1: Sexo dos participantes .....	174
Gráfico 4 - Questão 2: Idade .....	175
Gráfico 5 - Questão 3: Tempo de atuação na rede .....	175
Gráfico 6 - Questão 4: Tempo de atividade na função atual .....	176
Gráfico 7 - Questão 5: Tempo de magistério na Educação Básica .....	176
Gráfico 8 - Questão 6: Edições do PAER de que já participou na função atual .....	177
Gráfico 9 - Conheço os objetivos do PAER.....	178
Gráfico 10 - Os diretores conhecem os objetivos do PAER .....	178
Gráfico 11 - Os professores conhecem os objetivos do PAER .....	179
Gráfico 12 - Os pais conhecem os objetivos do PAER .....	181
Gráfico 13 - Os alunos conhecem os objetivos do PAER .....	181
Gráfico 14 - Considero que o PAER alcançou seus objetivos .....	182
Gráfico 15 - Analisei os dados das avaliações do PAER nos últimos anos .....	184
Gráfico 16 - Os dados do PAER são analisados pelas equipes administrativas.....	185
Gráfico 17 - Os resultados do PAER são analisados pelos professores .....	185
Gráfico 18 - Melhoria da qualidade do ensino após a introdução do PAER na rede .....	187
Gráfico 19 - Houve mudanças significativas na qualidade do ensino.....	187
Gráfico 20 - Em meu planejamento há uma meta sobre o PAER .....	189

## LISTA DE TABELAS

Tabela 1 - Primeira aplicação do Saeb -1990 .....	77
Tabela 2 - Evolução dos ciclos do Saeb – 1990 a 2003 .....	88
Tabela 3 - Educação Básica no Brasil – Estabelecimentos e Matrículas .....	118
Tabela 4 - Provas do PAER – edição de 2019.....	125
Tabela 5 - Média de Proficiência em Língua Portuguesa e Matemática .....	133
Tabela 6 - Pergunta 8: Categorias de análise e incidência .....	191
Tabela 7 - Resposta 9: Categorias de análise e Incidência .....	193
Tabela 8 - Resposta 10 – Categorias de análise e incidência .....	197
Tabela 9 - Resposta 11: Cenário de 2020 e implicações para o PAER.....	200
Tabela 10 - Perfil dos participantes da entrevista.....	204

## LISTA DE QUADROS

Quadro 1 - Dissertações e Teses na BDTD tendo como referência a meta-avaliação.....	31
Quadro 2 - Dissertações e Teses do Banco de Teses da Capes .....	34
Quadro 3 - Artigos disponibilizados na Scientific Eletronic Library Online (SciELO) .....	38
Quadro 4 - Lista de Verificação dos Padrões de Avaliação do Joint Committee.....	55
Quadro 5 - Lista Chave de Verificação de Avaliação de Scriven – KEC .....	57
Quadro 6 - Critérios Fundamentais de Davidson para a meta-avaliação .....	58
Quadro 7 - Critérios Paralelos e de Autenticidade de Guba e Lincoln.....	60
Quadro 8 - Eixos, dimensões e indicadores da avaliação da educação básica .....	76
Quadro 9 - Diretrizes e dimensões do Sistema Nacional de Avaliação da Educação Básica ..	99
Quadro 10 - Melhor rendimento mensal domiciliar per capita (2019).....	120
Quadro 11 - Pior rendimento mensal domiciliar per capita (2019).....	120
Quadro 12 - Número de participantes nas edições do PAER .....	125
Quadro 13 - Formas de divulgação dos resultados do PAER .....	126
Quadro 14 - Periodização, segundo Stufflebeam e Skinkfield.....	149
Quadro 15 - Periodização, segundo Guba e Lincoln.....	150
Quadro 16 - Comparação entre o sistema de crenças convencional e o construtivista .....	158
Quadro 17 - Dimensões mínimas comuns da qualidade da educação.....	166
Quadro 18 - Questões da entrevista semiestruturada.....	204
Quadro 19 - Categorias e frequência das respostas do Eixo Visão .....	206
Quadro 20 - Categorias de análise do Eixo Criação/ Implementação .....	213
Quadro 21 - Categorias de análise do Eixo Aplicação .....	216
Quadro 22 - Categorias de análise do Eixo Articulação.....	220
Quadro 23 - Categorias de análise do Eixo Objetivo .....	222
Quadro 24 - Categorias de análise do Eixo Uso dos resultados .....	225
Quadro 25 - Categorias de análise do Eixo Ganhos .....	229
Quadro 26 - Categorias de análise do Eixo Dificuldades .....	236
Quadro 27 - Categorias de análise Eixo Implicações .....	242
Quadro 28 - Ganhos e dificuldades do PAER .....	246

## LISTA DE ABREVIATURAS E SIGLAS

<b>AEA</b>	American Evaluation Association
<b>ANA</b>	Avaliação Nacional da Alfabetização
<b>ANEB</b>	Avaliação Nacional da Educação Básica
<b>ANRESC</b>	Avaliação Nacional do Rendimento Escolar
<b>BDTD</b>	Biblioteca Digital Brasileira de Teses e Dissertações
<b>BM</b>	Banco Mundial
<b>BNCC</b>	Base Nacional Comum Curricular
<b>CAPES</b>	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
<b>CNE</b>	Conselho Nacional de Educação
<b>CPA</b>	Comissão Própria de Avaliação
<b>ECIEL</b>	Programa de Estudos Conjuntos de Integração Econômica Latino-americana
<b>EDURURAL</b>	Programa de Expansão e Melhoria do Ensino no Meio Rural do Nordeste Brasileiro
<b>ENADE</b>	Exame Nacional de Desempenho dos Estudantes
<b>ENEM</b>	Exame Nacional do Ensino Médio
<b>ERS</b>	Evaluation Research Society
<b>GESTEC</b>	Programa de Pós-graduação em Gestão e Tecnologias Aplicadas à Educação
<b>IBGE</b>	Instituto Brasileiro de Geografia e Estatística
<b>IBM</b>	International Business Machine Corporation
<b>IDD</b>	Índice de Diferença do Desempenho
<b>IDEB</b>	Índice de Desenvolvimento da Educação Básica
<b>IDH</b>	Índice de Desenvolvimento Humano
<b>INEP</b>	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

<b>JCEE</b>	Joint Committee on Standards for Educational Evaluation
<b>KEC</b>	Key Evaluation Checklist
<b>LDB</b>	Lei de Diretrizes e Bases da Educação Nacional
<b>MEC</b>	Ministério da Educação
<b>OCDE</b>	Organização para Cooperação e Desenvolvimento Econômico
<b>ONU</b>	Organização das Nações Unidas
<b>PAER</b>	Programa de Avaliação Externa da Rede
<b>PNAIC</b>	Pacto Nacional pela Alfabetização na Idade Certa
<b>PNE</b>	Plano Nacional da Educação
<b>PISA</b>	Programme for International Student Assessment
<b>SAEB</b>	Sistema de Avaliação da Educação Básica
<b>SAEP</b>	Sistema Nacional de Avaliação das Escolas Públicas de 1º Grau
<b>SICEVAES</b>	Sistema Centroamericano de Evaluación y Armonización de la Educación Superior
<b>SINAEB</b>	Sistema Nacional de Avaliação da Educação Básica
<b>SINAES</b>	Sistema Nacional de Avaliação da Educação Superior
<b>SPSS</b>	Statistic Program for Social Sciencies
<b>TRI</b>	Teoria de Resposta ao Item
<b>UNESCO</b>	Organização das Nações Unidas para a Educação, a Ciência e a Cultura

## SUMÁRIO

INTRODUÇÃO .....	16
CAPÍTULO 1 - CAMINHOS DA PESQUISA.....	29
1.1 Revisão da Literatura .....	30
1.2 Aspectos Metodológicos .....	39
1.2.1 Técnicas e Instrumentos de produção do material empírico .....	42
1.3 O locus da Pesquisa.....	44
1.4 Os participantes da pesquisa .....	45
1.5 Análise dos dados.....	46
CAPÍTULO 2 - FUNDAMENTOS DA META-AVALIAÇÃO.....	49
2.1 Tipos, modelos e padrões de Meta-avaliação.....	53
2.1.1 Tipos de Meta-avaliação .....	53
2.1.2 Abordagens somativas e formativas.....	54
2.1.3 Modelo de critérios paralelos e critérios de autenticidade de Guba e Lincoln.....	60
2.2 O meta-avaliador .....	61
2.3 A meta-avaliação nesta pesquisa.....	63
CAPÍTULO 3 - TRAJETÓRIA DAS AVALIAÇÕES EM LARGA ESCALA DA EDUCAÇÃO BÁSICA .....	66
3.1 Iniciativas embrionárias do Sistema de Avaliação da Educação Básica .....	66
3.2 Os oito primeiros ciclos do Saeb – 1990 a 2003 .....	74
3.3 O Saeb de 2005 à segunda década do século XXI .....	89
CAPÍTULO 4 – PROGRAMA DE AVALIAÇÃO EM LARGA ESCALA DA REDE PRIVADA DE ENSINO: TRAJETÓRIA E CARACTERÍSTICAS .....	104
4.1 Uma breve história do ensino privado no Brasil .....	104
4.2 O ordenamento jurídico constitucional e na LDB para o ensino privado .....	114
4.3 Cenário do ensino privado no final da segunda década do século XXI .....	118
4.4 O Programa de Avaliação Externa da Rede (PAER) .....	121
CAPÍTULO 5 – CONCEPÇÕES DE AVALIAÇÃO E QUALIDADE EDUCACIONAL .....	136
5.1 Primórdios da avaliação educacional .....	137
5.2 A avaliação educacional no início do século XX.....	142
5.3 A avaliação educacional como campo de estudos.....	147
5.4 Os paradigmas avaliativos.....	151
5.4.1 O paradigma dominante .....	153
5.4.2 O Paradigma Emergente.....	157
5.5 Concepções de qualidade educacional .....	161
CAPÍTULO 6 – AVALIAÇÃO DO PROGRAMA DE AVALIAÇÃO EXTERNA DA REDE .....	172
6.1 Análise dos Questionários .....	173

6.1.1 Perfil dos participantes do questionário .....	174
6.1.2 Questões objetivas.....	177
6.1.3 Questões abertas.....	190
6.2 Análise das Entrevistas.....	203
6.2.1 – 1º Foco: Percepção sobre o PAER .....	205
6.2.2 – 2º Foco: Processo .....	213
6.2.3 – 3º Foco: Resultados.....	222
6.3 Síntese integradora dos ganhos e das dificuldades.....	245
CONSIDERAÇÕES FINAIS.....	252
REFERÊNCIAS.....	258
APÊNDICE.....	275

## INTRODUÇÃO

Este estudo foi desenvolvido como requisito do Programa de Pós-Graduação em Educação da Pontifícia Universidade Católica de Campinas, encontra-se vinculado à linha de pesquisas de Políticas Públicas em Educação, cuja ênfase é voltada para a atuação do Estado brasileiro na perspectiva da garantia da educação para todos e nos processos de formulação, planejamento, implantação e avaliação dessas políticas públicas relacionadas à educação. A pertinência do tema deve-se ao fato de que, dentre as políticas educacionais criadas e implementadas nos últimos anos, a avaliação de sistemas tem sido a mais estável e permanente, preservada por governos de diferentes espectros políticos, presente em todos os níveis de ensino e capilarizando-se do sistema federal aos estaduais e municipais. Assim, a pesquisa tem como foco a avaliação externa e em larga escala<sup>1</sup>, porém implementada por uma rede privada de educação básica, que possui cerca de 350 escolas distribuídas em todos os estados brasileiros, e que tem sido desenvolvida com base no sistema de avaliação da educação básica vigente no Brasil nas últimas três décadas.

Sabidamente, a temática da avaliação vem assumindo centralidade na formulação e implementação de políticas no setor educacional nos últimos anos e, concomitantemente, a produção científica crescente sobre o tema o tem caracterizado como campo de estudos em expansão, com uma abrangência significativa de subcampos de diferentes características que transitam desde a avaliação da aprendizagem, em sala de aula; avaliação de docentes; avaliação institucional; até a avaliação de sistemas educacionais e de programas nos diferentes níveis de ensino.

No último quarto de século, a avaliação de sistemas educacionais tem se consolidado como objeto de estudos, pesquisas, discussões e interesse dentro deste amplo campo. Para Gatti (2015), no Brasil, um campo ainda sujeito a fortes críticas ideológicas com uma massa crítica intelectual relativamente pouco expressiva a debruçar-se sobre as questões afetas ao tema.

---

<sup>1</sup> Neste estudo, a concepção adotada para avaliação externa e avaliação em larga escala é a elaborada por Santana (2018), que considera como avaliação externa toda avaliação padronizada que é realizada ou idealizada por órgãos externos à escola. Ela é pensada e montada fora da escola, mesmo que seja aplicada pelos professores da própria instituição de ensino, podendo ser de larga escala, ou não. Ela receberá o nome de avaliação de larga escala se for aplicada num sistema de ensino inteiro [...]. Ela ainda pode ser chamada de amostral – caso avalie somente alguns alunos desse sistema – ou censitária –, caso avalie todos os alunos matriculados numa mesma série de determinado sistema de ensino (SANTANA, 2018, p. 38).

Esta é uma temática complexa porque envolve concepções, crenças, valores, teorias, princípios muito específicos e, quando a avaliação tem como foco a educação, os desafios tornam-se potencialmente mais complexos. Freitas *et al.* (2017) afirmam que a avaliação é uma categoria pedagógica complexa e “um campo de pesquisa de difícil acesso e compreensão” (p.8). Afonso (2000) considera a avaliação como um dos eixos estruturantes das políticas educativas contemporâneas. O autor a pressupõe, ainda, como uma atividade política que precisa ser compreendida no contexto das mudanças educacionais e das mudanças econômicas e políticas mais amplas. Sendo assim, configura-se um campo de tensão entre o que ela deveria ser, como atividade política, e o que de fato ela é, como prática avaliativa. Há que se levarem em conta essas dimensões, quando escolhermos como objeto de estudo um processo avaliativo que ocorre em um *locus* educacional que não é isolado nem alheio às influências do contexto social e político. Ao contrário, é por este determinado.

A despeito de existirem várias concepções de políticas públicas, neste estudo, elegeu-se a definição de Domingos Fernandes, como “processos mais ou menos complexos, participados e articulados, que normalmente se traduzem em ações destinadas a resolver problemas dos cidadãos e das sociedades num determinado domínio.” (FERNANDES, 2018, p. 20), complementada pela ideia de Höfling acerca do “Estado implantando um projeto de governo, através de programas, de ações voltadas para setores específicos da sociedade.” (HÖFLING, 2001, p. 31).

A iniciativa do Estado de elaborar políticas públicas surge a partir de um “problema” ou uma “necessidade” oriunda da sociedade que, ao adquirir o grau de questão pública, requer a implantação de um projeto ou programa de governo que atenda às demandas e aos interesses sociais. Portanto, nessa ótica, as políticas públicas voltadas para o domínio educacional são formuladas para “resolver problemas dos cidadãos” que estão relacionados a um espaço muito específico: os sistemas de ensino. São também denominadas políticas educacionais e se consolidam em projetos e programas que necessitam da estrutura do Estado e das ações do governo para a sua implementação, mas dependem significativamente dos atores sociais envolvidos na educação: dirigentes educacionais, professores, pais, alunos e comunidade em geral.

Em uma análise sobre a relação entre avaliação e política pública, Domingos Fernandes considera que avaliação é uma atividade política e, como tal, interfere e influencia a vida e a política pública (FERNANDES, 2018). Todavia, a avaliação não está integrada na política “porque deve gerar e difundir informação credível. Isto é, informação que seja relevante, significativa e imparcial, que decorra de uma efetiva participação dos intervenientes

e interessados no processo de avaliação” (FERNANDES, 2018, p.33). Nessa perspectiva, a avaliação está comprometida com a política por ter uma agenda orientada para a melhoria da sociedade e sua democratização. O autor conclui que “quando o entendimento da política é o de que ela está presente em todos os setores da vida social, pública e mesmo privada, fazendo parte do dia a dia das atividades de cada cidadão [...] não há separação possível com a avaliação.” (FERNANDES, 2018, p. 33).

Contudo, Horta Neto (2018) alerta para o fato de que “nos últimos tempos, a avaliação se tornou sinônimo de política educacional. E isso é um equívoco tremendo” (p.51). Para o autor, a avaliação é um instrumento para aprimorar os processos educacionais e as políticas públicas relacionadas. Um importantíssimo instrumento, mas, apenas um instrumento. Uma política educacional vai muito mais além da busca por resultados ou metas de aprendizado, como o que tem sido visto nos últimos anos.

Tendo por base essas ideias a respeito da criação e manutenção de uma política pública a partir de um problema ou necessidade demandada pelos cidadãos, a avaliação de sistemas educacionais buscaria atender ao problema da baixa qualidade do ensino percebida após a expansão da ampliação de oferta de vagas tanto na educação básica como na educação superior, expansão esta, iniciada ao final da década de 1970 (CASTRO, 2009). Contudo, as iniciativas de avaliação de sistemas nas últimas décadas não compõem uma agenda exclusiva da realidade brasileira. Ao contrário, representam uma tendência internacional. Em um estudo sobre o crescimento do número de sistemas próprios de avaliação em larga escala nas redes de ensino no Brasil, nos últimos anos, ao comentarem as justificativas desse crescimento, alguns autores consideram que acompanha “uma agenda e tendências internacionais já consolidadas que consideram nesse tipo de ação, uma forma de se aferir a qualidade e a efetividade dos sistemas educacionais.” (GIMENES *et al.*, 2013, p. 16).

O crescimento de políticas de avaliação educacional em diversos níveis e esferas não é um fenômeno restrito à realidade brasileira. É sim, um movimento amplo, principalmente nos países do eixo central, e desencadeado por inúmeros fatores.

Alguns autores sinalizam como marco temporal para a ênfase em avaliações em larga escala por meio de provas padronizadas, a publicação do relatório *Coleman*, em 1966, nos Estados Unidos, considerado o mais importante estudo de política educacional a relacionar qualidade da educação a resultados de testes cognitivos aplicados em alunos (GIMENES *et al.*, 2013; FERNANDES; GREMAUD, 2009). A importância atribuída ao relatório é assim explicada:

Com o crescimento dos exames em larga escala, a partir da década de 1950, as avaliações ganharam um papel adicional: o de diagnosticar e monitorar o sistema educacional. Um marco nas avaliações diagnósticas é dado pelo Relatório Coleman, que tinha como objetivo estudar a segregação racial no sistema educacional nos EUA. Coleman e seus colaboradores examinaram diferenças nos resultados dos exames padronizados e procuraram explicá-las com bases nas diferenças de recursos das escolas e das características socioeconômicas dos alunos. Este trabalho é importante por uma série de razões, entre as quais a mudança de foco na forma de se identificar a qualidade das escolas: de insumos e processos para resultados. (FERNANDES; GREMAUD, 2009, p.4)

Já Rothen (2018) apresenta, como marco das políticas públicas voltadas para a avaliação de sistemas de ensino, a década de 1980 com as iniciativas dos Estados Unidos e Inglaterra de diminuir a participação do Estado na execução dos serviços públicos a fim de reduzir os gastos estatais, já que havia uma crítica à ineficiência na oferta desses serviços pelo Estado. Consequentemente, defendiam que, diminuindo a regulação pública das escolas, os pais teriam a possibilidade de escolha, gerando uma competição entre elas e que, dessa forma, haveria um aumento da qualidade da educação. As políticas desenvolvidas nesse período transferem as leis do mercado para a educação e, com a defesa do Estado Mínimo<sup>2</sup>, a presença regulatória cresceu. Segundo Barroso (2000), as políticas de descentralização aumentaram o controle do Estado ao invés de transferir efetivamente o poder para a escola e para as comunidades locais, e um dos dispositivos usados para esse aumento de controle foi a criação dos sistemas avaliativos.

Sobre esse período da década de 1980, Afonso (2000) considera que o interesse demonstrado pela avaliação, sobretudo por governos neoconservadores e neoliberais, pode ser traduzido pela expressão “Estado Avaliador”<sup>3</sup>, que quer significar a adoção, pelo Estado, de um *ethos* competitivo, da lógica do mercado, trazendo para o domínio público o modelo da gestão privada com “ênfase nos *resultados* ou *produtos* dos sistemas educativos” (AFONSO, 2018, p. 49, itálicos do autor).

A conjuntura do final do século XX tornou-se propícia para o fortalecimento dessas ideias. O cenário que favoreceu a expansão dessas reformas foi constituído por diversos

---

<sup>2</sup> Estado Mínimo, também conhecido por Estado liberal, aborda a ideia de que numa economia é preciso que o peso do setor público seja mínimo em comparação ao do setor privado. Pela ótica do liberalismo, o Estado em seu tamanho reduzido deve funcionar para manter bens públicos que auxiliam o bom funcionamento dos mercados. Além disso, a quantidade mínima de serviços prestados deve contribuir para a organização da ordem social, através de pilares como educação, saúde e segurança. (www.dicionariofinanceiro.com, visitado 25.06.2020, às 17h06min).

<sup>3</sup> A expressão foi proposta por Guy Neave (1988), referindo-se ao estado que se afasta, mas ao mesmo tempo aumenta seu controle sobre a educação por meio da avaliação e a regulação pelo mercado (ROTHEN, 2018, p. 20).

fatores externos à escola e com um fundo comum: a perda de competitividade norte-americana desde o final da década de 1970, diante do forte crescimento do poder produtivo do Japão; a crise econômica gerada pelo choque do petróleo que produziu uma recessão mundial e a estagnação das taxas de crescimento econômico e mudanças no sistema produtivo a partir da introdução de novas tecnologias de informação. Todos os fatores criaram um ambiente fértil para o interesse comum em diversos países de ideologia e de sistemas políticos distintos, em associar o papel da qualidade da educação com a retomada da capacidade de competitividade econômica. Nesse contexto, a avaliação de sistemas educacionais passou a ser um ponto de destaque nas propostas de políticas públicas em vários países.

Aliado a esses fatores, inúmeros documentos publicados por organismos internacionais multilaterais, divulgando as ideias de políticas educacionais desenvolvidas nos Estados Unidos e na Inglaterra, repercutiram em outros países e fortaleceram-nas. Rothen (2018) listou 18 documentos produzidos no período de 1989 a 2007 por agências como o Banco Mundial, a Organização para Cooperação e Desenvolvimento Econômico (OCDE) e a Organização das Nações Unidas para a Educação a Ciência e a Cultura (UNESCO), que divulgavam essas ideias. O autor entende que a tônica dos discursos desses órgãos é em defesa de ideias que possuem enfoques sociais, como a visão de que educação é um bem que deve ser oferecido e assegurado a todos. Mas, afirma que “em linhas gerais, as soluções para a ‘oferta de educação para todos’ têm uma perspectiva gerencial e alinhada à visão economicista” (ROTHEN, 2018, p. 22).

Portanto, é nesse cenário mundial, com processos avaliativos implantados com características peculiares em diferentes países, que surgem e são desenvolvidas, no Brasil, avaliações “que têm como traço comum a utilização de provas padronizadas, aplicadas em larga escala, configurando nitidamente um processo de avaliações externas, com crescente importância no desenho das políticas educacionais de todos os entes federados” (BAUER; ALAVARSE; OLIVEIRA, 2015, p. 1369). E, na percepção dos autores:

A aceitação de tais avaliações, entre nós, longe está de ser consensual. Ao contrário, ainda encontramos parcelas significativas da comunidade acadêmica e educacional que, de maneira geral, desconsideram as eventuais contribuições que podem se originar das medidas educacionais em larga escala. (BAUER; ALAVARSE; OLIVEIRA; 2015, p. 1369)

No cenário nacional, há autores que têm se dedicado à discussão dos propósitos, efeitos, usos, modelos e finalidades das avaliações em larga escala (BONAMINO; SOUZA, 2012; HORTA NETO, 2010; BAUER, 2010; SOUZA, S., 2014; NARDI; SCHNEIDER; RIOS, 2014; COELHO, 2008; FREITAS, D., 2013; OLIVEIRA *et al.*, 2013; DIAS

SOBRINHO, 2008). O debate abriga posições que vão desde uma “contraposição extremada ao reconhecimento da contribuição das avaliações em larga escala e das medidas educacionais como norteadoras de políticas e programas educacionais, incluindo as restrições ao uso que vem sendo feito dos resultados” (BAUER; ALAVARSE; OLIVEIRA, 2015, p. 1367).

Porém, os dados produzidos por tais avaliações têm sido, historicamente, o meio pelo qual a sociedade se vale para conhecer tendências, responsabilidades, resultados e coerências entre teorias e práticas (RABELO; MARINHO-ARAÚJO, 2015). É, nesse aspecto, um tema de interesse social, pois de fato, as ideias de avaliação e qualidade estão intrinsecamente relacionadas; qualidade esta, que é percebida como determinante para o desenvolvimento econômico e social dos países e da população em geral. De igual modo, o interesse público tem crescido na medida em que a divulgação dos resultados das avaliações ocupa um significativo espaço nos veículos midiáticos, fortalecendo a ideia de que o desempenho dos alunos nas avaliações externas é o indicador da qualidade educacional dos sistemas e, com isso, favorecendo as iniciativas de ranqueamento e comparações, tendo por base as notas ou índices gerados por essas avaliações.

Questionamos, nesse contexto, qual avaliação? Qual qualidade? E foram esses questionamentos que, sem esgotar o assunto, nos moveram, fruto da curiosidade epistemológica (FREIRE, 1996), a discorrer sobre os respectivos conceitos no Capítulo cinco desta tese. A priori, enfatizamos a necessidade de a avaliação estar a serviço das aprendizagens e da melhoria da qualidade<sup>4</sup> educacional, a se considerarem os fatores intra e extraescolares.

Horta Neto (2018) indica um aumento recente no número de estados brasileiros que utilizam sistemas próprios de avaliação, principalmente a partir do lançamento do Índice de Desenvolvimento da Educação Básica (IDEB)<sup>5</sup> e das metas estabelecidas pelo Compromisso Todos pela Educação, criados pelo Decreto 6.094 de 24 de abril de 2007. Em 2007, havia nove estados que realizavam avaliações próprias. Em 2016, esse número passou a 21 estados. De igual modo, demonstrando esse crescimento, em um levantamento amplo por meio de *survey* eletrônico, a Fundação Carlos Chagas e o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) realizaram uma pesquisa com as secretarias municipais

---

<sup>4</sup> Dada a polissemia do termo, o conceito de qualidade é discutido no quinto capítulo.

<sup>5</sup> Índice que reúne em um só indicador os resultados do fluxo escolar e a média do desempenho dos alunos no Sistema de Avaliação da Educação Básica (Saeb). É a ferramenta para acompanhamento das metas de qualidade para a educação básica, que tem estabelecida como meta para 2022 alcançar média 6 – valor que corresponde a um sistema educacional de qualidade comparável ao dos países desenvolvidos. <http://portal.inep.gov.br/web/guest/ideb>

de Educação de 5.532 municípios. Desses, 4.309 responderam ao *survey* proposto. Os dados divulgados apontam que 1.573 municípios possuíam avaliações próprias e 905 declararam pretender ter uma avaliação, ou seja, 60% dos respondentes possuíam ou tinham a intenção de construir avaliação em larga escala legitimada no âmbito da gestão municipal (BAUER; PIMENTA; HORTA NETO; SOUZA, 2015).

Considera-se, portanto, que a avaliação externa de sistemas de ensino tem se mantido como instrumento estável de política educacional em governos de espectros ideológicos distintos; há um aumento do interesse do tema como objeto de estudos e pesquisas; há um interesse social pelos resultados produzidos por essas avaliações; e, como sinaliza Afonso (2007), um crescente protagonismo e exponencial centralidade do tema que ainda propiciam amplo espaço para conhecimento, reflexão, acréscimo de saberes e – por que não – a avaliação das políticas de avaliação ou da própria avaliação. Daí, a pertinência e a relevância do tema ora investigado.

O nosso interesse na qualidade de pesquisadora por este tema deve-se ao fato de haver desempenhado, durante algum tempo, função administrativa em uma instituição de ensino superior no período de implementação do Sistema Nacional de Avaliação da Educação Superior (SINAES), instituído pela Lei nº 10.861 de 24 de abril de 2004, e acompanhado todo o processo de estruturação da instituição para atender aos dispositivos previstos na lei, cujas dimensões envolviam a avaliação institucional interna e externa, a avaliação de cursos e a avaliação do corpo discente, por meio do Exame Nacional de Desempenho dos Estudantes (ENADE). Atuar na organização de uma Comissão Própria de Avaliação (CPA) e de todo o processo de autoavaliação institucional, bem como receber as comissões de avaliação *in loco* e participar do desenvolvimento de uma cultura de avaliação na comunidade acadêmica foram etapas que compuseram essa experiência profissional. Nesse período, também atuei como avaliadora de uma associação de acreditação de instituições de ensino superior para a América do Sul, tendo avaliado três instituições no Brasil, uma na Argentina e uma na Bolívia. Essa associação segue padrões internacionais de acreditação e pertence a uma organização privada, sem fins lucrativos. Ao deixar de atuar no ensino superior e voltar à lida na educação básica de uma rede privada de ensino, no ano de 2010, me deparei com o interesse da mantenedora em desenvolver um sistema de avaliação com as mesmas características do Sistema de Avaliação da Educação Básica (SAEB), especificamente o modelo da Prova Brasil, que avaliava, nesse período, por meio de provas de Língua Portuguesa e Matemática, os alunos do 5º e do 9º ano do ensino fundamental, e nas áreas de Língua Portuguesa, Matemática,

Ciências Humanas e Ciências da Natureza, o 2º ano do ensino médio. O interesse aprofundou-se na medida em que, desde 2012, acompanhei o desenvolvimento e consolidação do programa de avaliação da rede privada de ensino. O que levaria uma rede privada de extensão nacional a investir uma significativa soma de recursos para avaliar o desempenho dos alunos por meio de uma avaliação em larga escala? Quais os ganhos ou as deficiências de um sistema de avaliação com essas características? Essas questões levaram à busca de uma investigação sistemática, por meio de um programa *stricto sensu*, capaz de ensejar que essas e outras questões pudessem ser consideradas com vistas a ampliar o conhecimento na área, mediante o aprofundamento teórico conceitual e o rigor de uma pesquisa científica. Afinal, tornou-se instigante notar que uma rede privada, com um público distintamente fidelizado pela natureza do serviço educacional que oferece, sente a necessidade de investir em um programa tão extenso de avaliação, tendo como referência o modelo desenvolvido pelas políticas públicas de educação, dispendendo de um alto investimento, tendo em vista as características continentais do país, assim como a dimensão técnica e metodológica requerida por uma iniciativa dessa envergadura.

O sistema de ensino privado que desenvolveu o Programa de Avaliação em Larga Escala, objeto desta pesquisa, tem características distintas. Trata-se de uma rede que está presente em mais de 140 países e com cerca de 8.000 instituições que atendem a todos os níveis de ensino. Na América do Sul, está presente em todos os países e, atualmente, a rede tem mais de 340 mil alunos, dos quais, mais de 220 mil estão no Brasil.

Com uma história na educação básica de mais de 100 anos no território brasileiro, a rede possui cerca de 350 unidades de ensino básico distribuídas em todas as unidades federativas e está organizada em regiões administrativas que são subordinadas a uma coordenação central para toda a América do Sul.

No ano de 2011, um grupo de educadores da rede foi convidado a pensar um modelo de avaliação em larga escala que se assemelhasse ao modelo da Avaliação Nacional do Rendimento Escolar (ANRESC), também conhecida como Prova Brasil, que foi criada em 2005 dentro do Sistema de Avaliação da Educação Básica (SAEB). Com o apoio de alguns estudiosos do tema da avaliação educacional ligados a uma universidade pública e uma equipe multidisciplinar ligada à mantenedora da rede, o programa foi criado e implementado a partir do ano 2012. Em consulta ao documento que descreve o referido sistema, o propósito de desenvolver um sistema de avaliação externa de escolas e colégios (educação básica) em todo o território brasileiro, é produzir informação que permita assegurar a qualidade da educação oferecida.

Diante do exposto, busca-se a resposta para o seguinte problema: quais ganhos<sup>6</sup> (aspectos desenvolvidos) e quais dificuldades (aspectos a desenvolver) podem ser identificados com a implementação do programa de avaliação externa em larga escala por uma rede privada de ensino de abrangência nacional, com vistas a assegurar qualidade à educação oferecida?

O problema desdobra-se nas seguintes questões de pesquisa:

- Qual a trajetória das políticas públicas de avaliação da educação básica brasileira?
- Qual a trajetória e caracterização do Programa de Avaliação Externa e em Larga Escala objeto da pesquisa?
- Que concepções de “avaliação” e de “qualidade educacional” estão expressas no Programa de Avaliação Externa e em Larga Escala da educação básica a ser pesquisado?
- Quais os ganhos e as dificuldades podem ser abstraídos da experiência com o Programa de Avaliação em Larga Escala pesquisado?

Estas questões nos apresentam objetivos a serem alcançados, destacando-se, como objetivo geral: investigar os ganhos (aspectos desenvolvidos) e as dificuldades (aspectos a desenvolver) do Programa de Avaliação em Larga Escala implementado em uma rede de ensino privada de extensão nacional. Este objetivo desdobra-se em objetivos específicos, assim descritos: i) Reconhecer a trajetória das políticas públicas de avaliação da educação básica brasileira; ii) Reconhecer a trajetória do Programa de Avaliação em Larga Escala da rede de ensino privado em estudo; iii) Identificar concepções de “avaliação” e “qualidade educacional” expressas no Programa de Avaliação em Larga Escala investigado; iv) Analisar ganhos e dificuldades abstraídos do Programa de Avaliação em Larga Escala.

A pesquisa em pauta, que visa contribuir com o apontamento de possíveis caminhos para a promoção da qualidade desejada, é considerada uma meta-avaliação, termo criado por Scriven (2009) para designar a avaliação de uma avaliação. Ou, como Stufflebeam (2001), que procurou ampliar essa definição de meta-avaliação, como

*[...]the process of delineating, obtaining, and applying descriptive information and judgmental information—about the utility, feasibility, propriety, and accuracy of an evaluation and its systematic nature, competent conduct, integrity/honesty,*

---

<sup>6</sup> Para designar essa síntese dos resultados e conclusões que uma avaliação pode oferecer, Stufflebeam (2001) utiliza as palavras “fraquezas” e “forças”; Dias Sobrinho (2003) lança mão dos termos “fortalezas” e “problemas”. Neste estudo, elegemos a designação de Cappelletti (2015), que nomeia “ganhos” e “dificuldades” os indicativos que uma avaliação pode oferecer ao gestor ou interessados pelos resultados do processo avaliativo.

*respectfulness, and social responsibility—to guide the evaluation and/or report its strengths and weaknesses. (STUFFLEBEAM, 2001, p. 185)<sup>7</sup>*

Conforme será demonstrado no segundo capítulo, no Brasil, ainda são recentes os estudos que se valem de processos meta-avaliativos. Encontramos autores que se pautam, principalmente, pelos padrões e publicações de outros países, com destaque para os Estados Unidos, onde o tema tem cerca de cinco décadas de construção teórica e prática. Tanto é, que podemos identificar algumas formas distintas de substantivação para a meta-avaliação: Elliot (2011) a denomina como uma “metodologia” (p. 960); Penna Firme e Letichevsky (2010), como uma “abordagem” (p.186); Gimenes (2006), como “atividade” e “mecanismo” (p.38); Bordini (2017), como “modalidade avaliativa” e “instrumento” (p. 29). Seja qual for a designação atribuída ao termo, há um consenso acerca de que a concepção de meta-avaliação está vinculada à própria ideia que se tem da avaliação. Portanto, conhecer as diferentes concepções de avaliação e de qualidade – termo este inerente ao processo de avaliar ou de meta-avaliar, pois sempre implica num juízo de valor ou mérito do objeto avaliado – é fundamental para a realização de uma meta-avaliação que também promova reflexão sobre as concepções teóricas subjacentes ao processo avaliativo em questão.

Contudo, como pressuposto para esta tese, nos valem das ideias de Guba e Lincoln (2011) que, ao introduzirem seus estudos sobre os modelos de avaliação ao longo da história, afirmam que “não existe *nenhuma* forma ‘correta’ de definir *avaliação*, pois, se fosse possível encontrar esse sentido, isso poria fim [...] à discussão acerca de como avaliação deve ser conduzida e sobre quais são os seus propósitos” (p. 27, *itálicos dos autores*). Consideram as definições de avaliação como interpretações mentais humanas e que correspondem a uma determinada “realidade” que reflete um momento histórico, as suposições filosóficas dos autores e que, como construções humanas, estão em processo de refinamento, de adequação e contínua reconstrução. Sendo assim, não consideramos uma concepção de avaliação certa ou errada. Porém, serão analisadas as ideias acerca de avaliação sob a perspectiva dos paradigmas que as fundamentam e à luz de autores referenciados da área.

A proposição que se enuncia para discussão é formulada sob a tese de que, de fato, para que o Programa de Avaliação em Larga Escala da rede pesquisada promova a melhoria da qualidade ensejada é necessário lançar atenção ao uso dos resultados, em uma perspectiva

---

<sup>7</sup> [...]” o processo de delinear, obter e aplicar informações descritivas e informações de julgamento - sobre a utilidade, viabilidade, propriedade e precisão de uma avaliação e sua natureza sistemática, conduta competente, integridade / honestidade, respeito e responsabilidade social - para orientar avaliação e/ ou relatar seus pontos fortes e fracos (STUFFLEBEAM, 2001, p.185, tradução nossa).

emancipatória. De igual modo, os possíveis ganhos ou dificuldades abstraídos da experiência com o programa, necessitam ser considerados a partir desta perspectiva. Há necessidade de articulação entre a referida avaliação externa, foco do presente estudo, e a avaliação da aprendizagem, que oportunize e promova a reflexão para transformação. Considerar com exclusividade os resultados da avaliação externa, centrada em avaliação de desempenho, torna-se insuficiente para expressar a qualidade educacional.

Ainda que o programa de avaliação em análise seja desenvolvido em uma rede de natureza privada, os princípios da avaliação emancipatória podem ser aplicados não somente ao longo da execução do programa, mas, principalmente, na análise dos resultados, levando em conta os pressupostos de sua criação detalhados no capítulo quatro. Ao considerarmos o uso dos resultados, a proposição é feita a partir dos conceitos básicos da avaliação emancipatória, que consiste em “um processo de descrição, análise e crítica de uma dada realidade, visando transformá-la.” (SAUL, 2010, p. 65).

Saul (2010) descreve como conceitos básicos da proposta: i) a emancipação, que prevê a consciência crítica da situação e a proposição de alternativas de solução para a mesma; ii) a decisão democrática, que implica em um envolvimento responsável e compartilhado dos elementos que participam de um programa, na tomada de decisão a respeito da proposta avaliativa e rumos do programa; iii) a transformação, gerada coletivamente, visando às alterações substanciais com base nas críticas do mesmo; iv) a crítica educativa, que propõe uma análise valorativa do programa educacional na perspectiva dos avaliadores que atuam no programa.

Desta forma, compreende-se que, há espaço para uma avaliação na perspectiva emancipatória, na medida em que houver um compromisso com uma educação transformadora e que a crítica da realidade não seja cerceada, reservando-se espaço para a criação coletiva de soluções. É à luz destas considerações que se pretendem analisar os dados levantados e indicar possíveis caminhos para superação de possíveis dificuldades encontradas.

Entendemos que o contributo deste estudo está na centralidade da temática como objeto de estudo e interesse acadêmico, confirmado diante do aumento significativo de pesquisas sobre avaliação educacional nos últimos anos (BAUER, 2012), bem como por tratar-se de um tema de interesse social que tem se retratado no espaço ocupado pelos meios de comunicação e, principalmente, pela originalidade do campo de estudo, por projetar foco sobre um sistema de avaliação externa desenvolvido no âmbito de uma rede privada de ensino, de educação básica.

Temos consciência, ao estudar qualquer aspecto do fenômeno educativo – neste caso, a avaliação –, de que há uma carga de subjetividade inerente à complexidade de uma prática social humana, à transitoriedade e ao dinamismo de um processo histórico, à visão única atribuída pelo pesquisador ao objeto da pesquisa, também carregada de impressões e imprecisões próprias do processo investigativo, que por sua vez, emerge de escolhas. Contudo, o rigor científico que se espera de uma pesquisa, aliado a uma sólida fundamentação embasada em autores referenciados e à abertura constante à reflexão e análise das considerações apresentadas são as estacas balizadoras deste exercício tão profícuo de produzir conhecimento.

Como apresentado ao longo do levantamento bibliográfico disposto no primeiro capítulo, não foi encontrada nenhuma pesquisa com processo meta-avaliativo de avaliação externa em larga escala a proposta no presente estudo. Este achado corrobora com as ideias de Gatti (2013) que apresenta a necessidade de pesquisadores em educação se voltarem para estudos e discussões sobre a avaliação em larga escala sob a perspectiva da meta-avaliação.

Além desta introdução, a estrutura textual da tese compreende seis capítulos e as considerações finais. A estrutura disposta objetiva responder às questões de pesquisa, e atender aos objetivos propostos. Como elementos pós-textuais, além das referências bibliográficas, há um apêndice que apresenta os instrumentos utilizados para coleta de dados.

O primeiro capítulo apresenta os caminhos da pesquisa, a revisão da literatura, a abordagem adotada, os instrumentos, os participantes, o *locus* do estudo e as indicações para a análise dos dados. Nele, também são apresentadas as principais noções dos procedimentos e técnicas escolhidos para interpretação dos dados.

O segundo capítulo traz os fundamentos, conceitos, modelos e tipos de meta-avaliação, assim como o seu uso no contexto da pesquisa. Entende-se que esta temática ainda necessita de espaço na produção de conhecimento na área de avaliação e, portanto, um capítulo foi destinado para um aprofundamento das ideias acerca da meta-avaliação.

O terceiro capítulo discorre sobre a trajetória das políticas públicas de avaliação da educação básica brasileira e seu desenvolvimento até esta segunda década do século XXI. Com dados de documentos normativos e a contribuição de pesquisadores que descrevem esta trajetória, o capítulo busca atender ao primeiro objetivo específico proposto.

O quarto capítulo tem como foco a trajetória e as características do programa de avaliação em larga escala da rede objeto da pesquisa. Nele, além da caracterização da rede e do programa, traçamos um panorama do sistema privado de ensino no Brasil<sup>8</sup>.

---

<sup>8</sup> A pedido da mantenedora da rede privada investigada, a identidade da instituição será preservada.

O quinto capítulo apresenta as concepções de avaliação e qualidade educacional expressas no documento norteador do programa de avaliação em larga escala e as concepções coexistentes no contexto educacional brasileiro. Para tanto, se vale do aporte teórico de autores referenciados para fundamentar tais concepções.

O sexto capítulo apresenta o resultado das análises dos ganhos e das dificuldades do programa de avaliação pesquisado, bem como os possíveis caminhos para a superação das dificuldades identificadas. Ao final, uma síntese dessas indicações é formulada, baseada nos dados coletados pelos instrumentos de pesquisa e análise documental.

## CAPÍTULO 1 - CAMINHOS DA PESQUISA

Este capítulo apresenta os caminhos percorridos para o desenvolvimento da pesquisa: a revisão da literatura; a abordagem adotada; as técnicas e os instrumentos utilizados; a descrição do *locus* e dos participantes; e a meta-avaliação como o procedimento da investigação.

As escolhas realizadas no percurso e desenvolvimento da pesquisa foram pautadas por alguns posicionamentos que deram sentido à atitude e à atividade investigativa da pesquisadora. O primeiro baseia-se numa afirmação de Garcia (2011), ao descrever sua experiência como pesquisadora em educação: “hoje sabemos que a dúvida, a incerteza, a insegurança, a consciência de nosso *ainda não saber* é que nos convida a investigar e, investigando, podemos aprender algo que não sabíamos.” (p. 20, itálico da autora). São as dúvidas, e não as certezas, que nos conduzem à investigação. Ou, como a autora afirma, “quem tem certezas não tem motivos para pesquisar.” (GARCIA, 2011, p. 22).

O segundo está pautado na proposição de Stake (1983, p.6): “acredito que a maioria das pesquisas educacionais deva estar a serviço da educação e, claramente, proporcione melhor compreensão de seus problemas práticos.”. Assim, buscando um sólido arcabouço teórico para todas as análises realizadas, o foco é desenvolver uma pesquisa que promova contribuições positivas para a vida das escolas, para educadores interessados pela temática, que poderão se valer das reflexões produzidas por esta investigação, sem que seu destino seja apenas os pares da academia, ou bibliotecas físicas e virtuais. Como afirma Demo (2011, p. 14, grifos do autor):

De si, as ciências sociais tratam da *práxis histórica*, do seu presente, passado e futuro. Teorizar sobre ela é fundamental, mas seria prática inaudita permanecer apenas na teoria. Pesquisar somente para saber já seria proposta alienante, porque desencarna a pesquisa da sua face inserida na realidade histórica, reduzindo-a ao esforço de sistematização de ideias e de especulação dedutiva. Saber está marcadamente ligado a *interesses sociais*, definidos aqui como contraposições dialéticas. Até mesmo acumular saber para cultivar a ignorância é possível e não raro sintomático.

Essa percepção de que o saber está ligado a interesses sociais e que o conhecimento produzido deve oferecer aos sujeitos dos grupos pesquisados condições formadoras e emancipadoras de transformação ancora-se nas ideias de Souza Santos (2018) acerca de que, na ciência pós-moderna, o conhecimento científico deve ser capaz de ensinar a viver e traduzir-se em um saber prático que enriqueça nossa relação com o mundo.

O terceiro posicionamento surge da consciência da complexidade da educação como uma prática social humana, um processo histórico e, portanto, inconcluso e em transformação,

que “emerge da dialética entre homem, mundo, história e circunstâncias.” (GHEDIN; FRANCO, 2011, p. 40). Por essa perspectiva, não poderá ser estudada valendo-se de procedimentos metodológicos que congelam alguns momentos dessa prática social para serem analisados. Deve-se valer de métodos que sejam capazes de “dar conta de apreendê-la em sua natureza dialética, captando não apenas as objetivações de uma prática real concreta, mas também a potencialidade latente de seu processo de transformação.” (p. 40). Esse processo em transformação lhe confere uma subjetividade tal, que não apenas modifica as representações dos envolvidos, mas também interfere na interpretação e ressignificação do fenômeno vivido. Deste modo, provoca alterações também naquele que se debruça sobre os fenômenos educativos como objeto de estudo, sendo fundamental que o método permita a captação dos significados que os indivíduos vão construindo em processo.

### **1.1 Revisão da Literatura**

Considerando as proposições presentemente descritas, a pesquisa iniciou-se com a busca de trabalhos pertinentes ao tema a fim de construir uma visão e uma síntese das produções disponíveis, ampliando assim a compreensão do campo de estudo.

A elaboração de uma investigação em um programa *stricto sensu* requer do pesquisador uma ampla revisão da literatura capaz de situar a temática escolhida e seu objeto de estudo no panorama da produção do conhecimento da área, mapear as questões que já foram levantadas e quais as possíveis lacunas que o estudo poderá preencher, ou qual a contribuição original que oferecerá. Este exercício permite que o problema seja mais bem delimitado e contextualizado em relação ao campo de conhecimento a que pertence. Com esse propósito, nos meses de abril de 2018 e maio de 2019, foram realizadas as pesquisas nas seguintes bases de dados: Biblioteca Digital Brasileira de Teses e Dissertações (BDTD); Catálogo de Teses da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES); e da *Scientific Electronic Library Online* (SciELO), para um levantamento da produção científica e de estudos publicados tendo como referência o descritor “meta-avaliação”, em todos os campos de busca avançada. Como filtro de busca levou-se em conta a necessidade da vinculação com a área da educação e a temática de avaliação de sistemas de ensino. Não houve um recorte temporal, sendo analisados todos os trabalhos indicados pelas bases, nas datas da pesquisa.

Na busca na base de dados da Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), foram encontrados 21 documentos vinculados à área da educação, os quais estão descritos no Quadro 1.

**Quadro 1 - Dissertações e Teses na BDTD tendo como referência a Meta-avaliação**

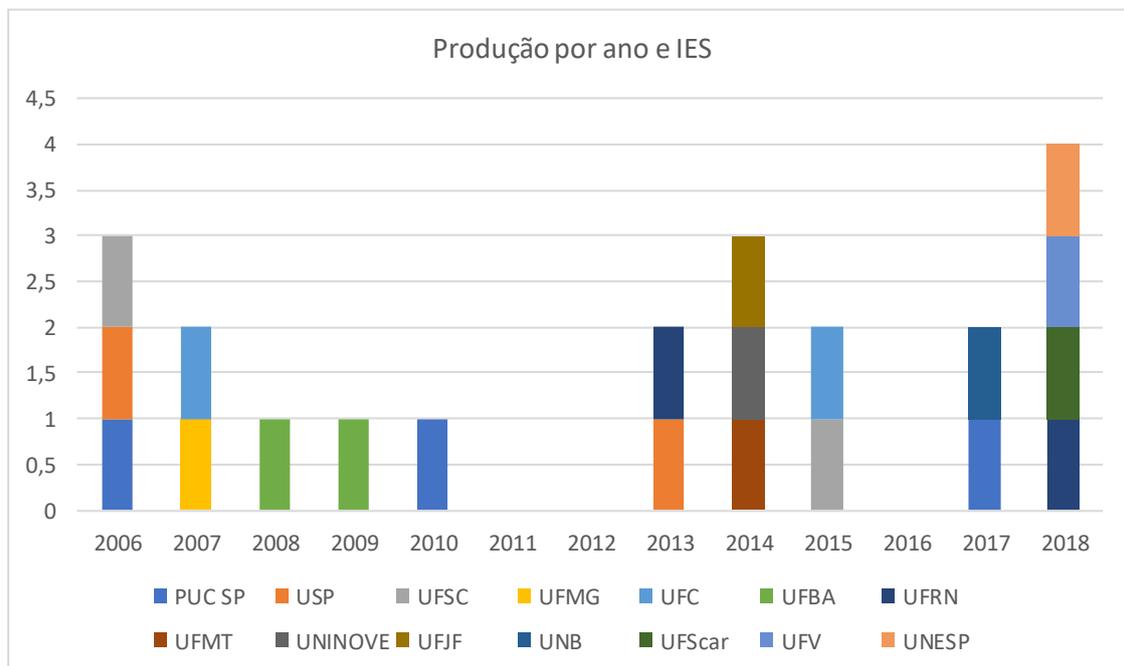
Ano	Título	Autor	Tipo	IES
2006	Um estudo meta-avaliativo de uma experiência de autoavaliação em uma instituição de ensino superior	Gimenes, Nelson Antonio Simão	Dissertação	PUC SP
2006	Uma proposta para caracterizar a atuação do professor na sala de aula	Scarinci, Anne Louise	Dissertação	USP
2006	Modelo de meta-avaliação de processos de avaliação da qualidade de cursos de graduação	Davok, Delsi Fries	Tese	UFSC
2007	Escrevendo a avaliação: a escrita de diários como exercício avaliativo	Luís, Suzana Maria Barrios	Tese	UFMG
2007	Autoavaliação e planejamento participativo no âmbito da Universidade Federal do Ceará: em busca de saberes e da mudança de cultura institucional	Rodrigues, Maria do Socorro de Sousa	Tese	UFC
2008	Meta-avaliação da aprendizagem na perspectiva da negociação comunicativa nos cursos de pós-graduação lato sensu a distância na Bahia	Garcia, Rosineide Pereira Mubarack	Tese	UFBA
2009	As contribuições das políticas de avaliação educacional em larga escala: o caso da avaliação da aprendizagem na Bahia	Dantas, Lys Maria Vinhaes	Tese	UFBA
2010	Meta-avaliação de projetos em educação com o uso das TIC	Machado, Marco Antonio de Jesus	Tese	PUC-SP
2013	Uma avaliação política da política de avaliação da educação superior brasileira	Lopes, Pedro Isaac Ximenes	Dissertação	UFRN
2013	ENEM: limites e possibilidades do Exame Nacional do Ensino Médio enquanto indicador de qualidade escolar	Travitzki, Rodrigo	Tese	USP
2014	Implicações dos projetos de extensão universitária para a formação do professor de Educação Física	Araújo, Carolina Moraes de	Dissertação	UFTM
2014	Meta-avaliação do sistema de avaliação institucional da Escola Técnica Estadual Parque da Juventude do Centro Paula Souza	Perez, Marinês Oliveira	Dissertação	UNINO VE
2014	A inclusão de ciências naturais nos sistemas de avaliação da educação: o caso do Avalia BH	Silva, Fernanda Gomes da	Dissertação	UFJF
2015	Meta-avaliação: uma década do processo de avaliação institucional do SINAES	Pinto, Rodrigo Serpa	Tese	UFSC
2015	Fazendo o ENADE: concepções e usos da avaliação por coordenadores de cursos de graduação na Universidade Federal do Ceará	Silva, Pablo Rodrigo da	Dissertação	UFC
2017	Autoavaliação no SINAES: prática vigente e perspectivas para uma agenda futura	Fernandes, Ivanildo Ramos	Dissertação	UNB
2017	Um estudo meta-avaliativo da avaliação da docência	Bordini, Débora Valim Cirino	Dissertação	PUC SP
2018	Meta-avaliação das políticas de cotas: um	Souza, Andreliza	Tese	UFSCar

	estudo dos processos nas universidades estaduais paranaenses	Cristina		
2018	Avaliação Institucional da UFRN: viabilidade e utilização	Pinheiro, Any Kadidja de Melo Tavares	Tese	UFRN
2018	O SIMAVE e sua relação com o trabalho docente: percepções de diferentes atores do cotidiano escolar	Dutra, Phillippe Rocha	Dissertação	UFV
2018	Avaliação Institucional interna da Universidade Estadual de Goiás: um estudo meta-avaliativo	Goulart, Joana Correa	Tese	UNESP

Pesquisa realizada no dia 19.05.2019, às 13h e 46 min.

Pelos dados levantados, entre os 21 documentos é possível identificar 10 dissertações e 11 teses, de 2006 a 2018, sendo que o maior número de trabalhos em um ano corresponde a quatro, em 2018. Desses 21 trabalhos, 14 são estudos voltados para o ensino superior e sete são voltados para a educação básica. Das 14 IES do Quadro 1, apenas duas são de natureza privada: a PUC de São Paulo e a Uninove, também em São Paulo. O Gráfico 1 apresenta o número de produções por ano e por instituição.

**Gráfico 1 - Distribuição de trabalhos por ano e por IES**



Elaboração: autora

Há o predomínio de apenas uma publicação por IES. As que publicaram mais que um estudo no período foram: USP, UFSC, UFC, UFBA, UFRN com duas produções e, apenas PUC-SP, com três publicações no período. Porém, nenhum deles teve como tema central a

meta-avaliação de avaliações externas ou de sistemas privados de ensino. Das sete dissertações e teses que são relacionadas à educação básica, três não apresentam um procedimento meta-avaliativo como caracterizado neste estudo. Scarinci (2006) e Luís (2007) apresentam a meta-avaliação como a percepção que o professor tem de seu trabalho ou como a avaliação de sua prática pedagógica. Dutra (2018), nas conclusões da dissertação, aponta a urgente necessidade de realização de meta-avaliação que promova reflexão sobre a avaliação em larga escala. Sendo assim, esses autores não se aproximam da meta-avaliação na perspectiva de uma avaliação de avaliação como apontado pelos teóricos da temática. Os quatro demais trabalhos que se aproximam deste estudo por realizarem uma meta-avaliação de uma avaliação vinculada à educação básica são os que seguem.

Lys Maria Vinhaes Dantas (2009), em tese apresentada no Programa de Pós-graduação em Educação da Universidade Federal da Bahia, investigou as políticas de avaliação educacional implementadas em larga escala nas escolas públicas do estado da Bahia, no período de 1999 a 2004. Tendo como referência o *checklist* do *Joint Committee on Standards for Educational Evaluation* (JCSEE) elaborado por Stufflebeam (1999), a pesquisadora adaptou os critérios do Uso e Utilidade para realizar uma meta-avaliação das avaliações de desempenho em Língua Portuguesa e Matemática aplicadas aos alunos do 4º ano das escolas públicas. Entre os achados da pesquisa, está o uso real das avaliações como compartilhamento de uma visão sobre a realidade do ensino e uma proposta de discussão a ser feita não seja voltada sobre o uso ou não uso das avaliações, mas, se os usos feitos contribuem, ou não, para a melhoria da qualidade da educação.

Rodrigo Travitzki (2013), em tese apresentada à Faculdade de Educação da Universidade de São Paulo, estudou o ENEM numa abordagem quantitativa, porém valendo-se de reflexões filosóficas sobre as críticas ao exame e dos possíveis efeitos benéficos do *ranking* entre escolas. As conclusões apontam para os limites do *ranking* como indicador da qualidade escolar, pois embora o ENEM possa avaliar o mérito dos alunos, ele é pouco informativo sobre o mérito das escolas. Ainda que o pesquisador tenha colocado nas palavras-chave o descritor meta-avaliação, este tema não foi abordado na tese.

Marinês Oliveira Perez (2014), em dissertação apresentada ao Programa de Mestrado em Gestão e Práticas Educacionais da Universidade Nove de Julho, estudou o sistema de avaliação institucional de duas escolas técnicas do estado de São Paulo, o qual foi por ela denominado “mecanismo de avaliação das próprias avaliações” (p.8). Sua pesquisa foi realizada por meio de análise documental e questionário semiestruturado de questões abertas

aplicado aos membros da Comissão de Avaliação, retratando o sistema em foco como uma meta-avaliação da avaliação institucional. Foram indicados pontos frágeis e pontos fortes no processo de avaliação institucional e apresentadas sugestões para o aperfeiçoamento do sistema de avaliação.

Fernanda Gomes da Silva (2014), em dissertação apresentada no Programa de Mestrado Profissional em Gestão e Avaliação da Educação da Faculdade de Educação da Universidade Federal de Juiz de Fora, estudou a implementação de testes de Ciências da Natureza no programa de avaliação das escolas públicas municipais de Belo Horizonte – o Avalia BH –, investigando as motivações e os efeitos na gestão escolar, no desempenho dos estudantes e na prática dos professores. Nessa pesquisa, a meta-avaliação foi apresentada como uma sugestão para que os gestores da rede elaborassem um Plano de Ação Educacional que pudesse ampliar os dados da investigação e sistematizasse os principais pontos a serem modificados, criando estratégias que tornassem a política pública compensatória ao ensino de Ciências. Portanto, não foi realizada uma meta-avaliação do sistema de avaliação, mas ela foi apontada como sugestão de estratégia aos gestores municipais.

Como demonstrado até aqui, a única pesquisa que estudou um sistema de avaliação em larga escala de educação básica foi a realizada por Dantas (2009), na rede pública do estado da Bahia. A pesquisadora realizou uma meta-avaliação apoiada em análise documental, para levantar a percepção da Utilidade e Uso dos instrumentos de avaliação de Língua Portuguesa e Matemática do 4º ano do Ensino Fundamental.

No Banco de Teses da Capes, os achados foram semelhantes aos da Biblioteca Digital Brasileira de Teses e Dissertações (BDTD). Foram encontrados 20 documentos, sendo que, das 12 universidades com produção no período de 1989 a 2016, nove são públicas e três são privadas. Dos 20 trabalhos selecionados, 11 são relacionados ao ensino superior e nove são estudos voltados à educação básica. Segue o Quadro 2 demonstrativo:

**Quadro 2 - Dissertações e Teses do Banco da Capes tendo como referência da Meta-avaliação**

Ano	Título	Autor	Tipo	IES
1989	Análise da Avaliação do Ciclo Básico de Alfabetização das Escolas Estaduais de 1º grau de Minas Gerais	Alencar, Rita Maria de Mello	Dissertação	UFRJ
1997	Avaliação Educacional e o avaliador	Vianna, Heraldo Marelim	Tese	PUC-SP
1998	Avaliação - o reflexo da formação acadêmica do psicólogo organizacional	Rodrigues, Rosangela Rocio Jarros	Dissertação	UNESP

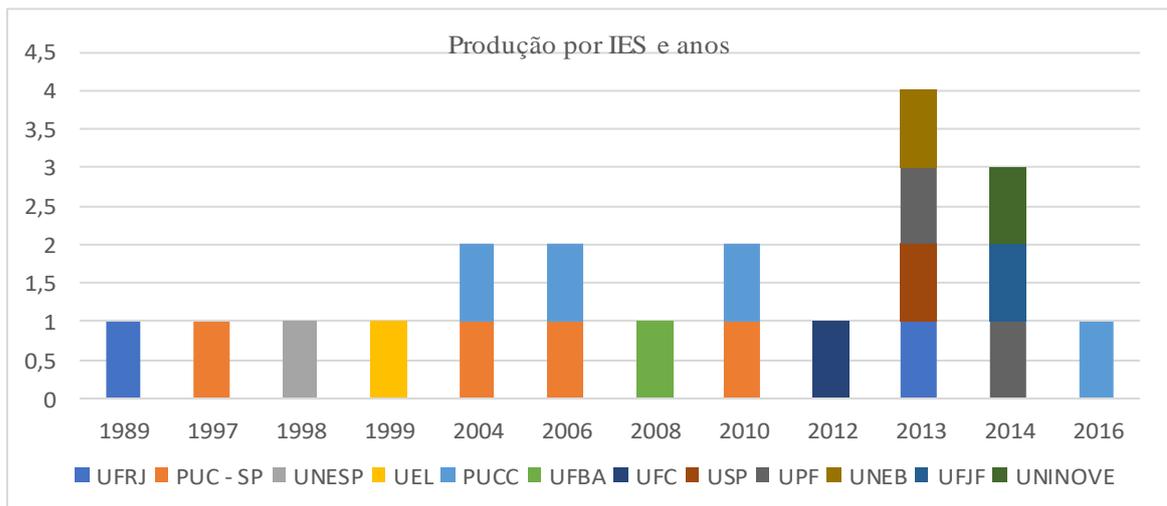
1999	Avaliação em Larga Escala no Paraná: um estudo sobre as provas e a recuperação no Espaço Escolar	Prust, Kátya Ap. de Carvalho	Dissertação	UEL
2004	Meta-Avaliação de docentes no ensino superior	Rios, Mônica Piccione Gomes	Tese	PUC-SP
2004	De Faculdade a Centro Universitário: avaliação institucional como instrumento para uma construção coletiva	Dardengo, Vilma Maia	Dissertação	PUC - Campinas
2006	Um estudo meta-avaliativo de uma experiência de autoavaliação em uma IES	Gimenes, Nelson Antonio Simão	Dissertação	PUC-SP
2006	Avaliação Institucional: um estudo comparativo entre duas universidades	Lobo, Clever Eduardo Zuins	Dissertação	PUC - Campinas
2008	Meta-avaliação da aprendizagem na Perspectiva da negociação comunicativa nos cursos de Pós-graduação Lato-sensu a Distância na Bahia	Garcia, Rosineide Pereira Mubarack	Tese	UFBA
2010	A meta-Avaliação e a extensão universitária: um estudo de caso	Arroyo, Daniela Munerato Piccolo	Dissertação	PUC - Campinas
2010	Meta-avaliação de projetos em educação com uso das TIC	Machado, Marco Antonio de Jesus	Tese	PUC-SP
2012	Docência Universitária: Avaliação Diagnóstica como instrumento transformador da Prática Docente	Pitombeira, Suzete Suzana Rocha	Tese	UFC
2013	ENEM: Limites e Possibilidades do Exame Nacional do Ensino Médio enquanto indicador de qualidade escolar	Oliveira, Rodrigo Travitzki Teixeira de	Tese	USP
2013	Meta-avaliação: caminhos para o aperfeiçoamento de práticas avaliativas no Sistema de Ensino Naval	Aguiar, Natalia Morais Correa Borges de	Tese	UFRJ
2013	Avaliação da Contribuição do Programa Mais Educação para as escolas municipais de Rio Verde - GO: dos números do IDEB à percepção qualitativa dos sujeitos	Moreira, Sebastiana Aparecida	Dissertação	UPF
2013	Avaliação Institucional na Educação Básica: limitações e possibilidades	Oliveira, Cleide Pereira	Dissertação	UNEB
2014	A inclusão de Ciências naturais nos sistemas de Avaliação da Educação: o caso Avalia BH	Silva, Fernanda Gomes da	Dissertação	UFJF
2014	Ensino Fundamental de nove anos: para além do ajuste legal	Fortunato, Raquel Paula	Dissertação	UPF
2014	Meta-avaliação do Sistema de Avaliação Institucional da Escola Técnica Estadual Parque da Juventude do Centro Paula Souza	Perez, Marines Oliveira	Dissertação	UNINOVE
2016	Os efeitos do SINAES no Curso de Turismo na PUC-Campinas: percepção dos gestores, professores e egressos	Pontes, Marina Piason Breglio	Dissertação	PUC - Campinas

Pesquisa realizada em 05.04.2018, às 18h e 20min

Pode-se notar que, das 20 pesquisas encontradas no Banco de Teses da CAPES, 11 foram desenvolvidas em universidades públicas e nove em universidades privadas, sendo que oito dessas pesquisas foram efetuadas nas Pontifícias Universidades Católicas de São Paulo e Campinas. Destaca-se que, neste levantamento, o primeiro trabalho disponível sobre a

temática data de 1989, período em que as primeiras iniciativas de avaliação em larga escala eram construídas no Brasil. A distribuição é equilibrada ao longo dos anos, havendo um aumento no ano de 2013, com quatro trabalhos publicados, conforme ilustra o Gráfico 2.

**Gráfico 2 - Distribuição de Trabalhos por ano e IES no Banco de Teses da Capes**



Elaboração da autora

No Gráfico 2 é demonstrado que houve um lapso de seis anos entre a publicação do primeiro e do segundo trabalho e mais algumas lacunas entre 1999 e 2004. Não houve publicações sobre a temática nos anos 2005, 2007, 2009, 2011, 2015 e, de 2017 a 2019, também não há registros de publicações sobre o tema pesquisado nessa base de dados. Vale destacar que a PUC-Campinas e a PUC-São Paulo são as universidades com maior número de trabalhos publicados (quatro, em cada uma), seguidas da UFRJ e UFP (2) e as demais instituições, UNESP, UEL, UFBA, UFC, USP, UNEB e UNINOVE, cada uma com uma produção, no período.

Dos nove trabalhos relacionados à educação básica, três também estão na relação de pesquisas da BDTD, apresentados acima (TRAVITZKI, 2013; SILVA, 2014; PEREZ, 2014). Dos seis restantes, os três primeiros não estão disponíveis na base de dados por serem publicações anteriores à Plataforma Sucupira<sup>9</sup>, que gerencia o banco de teses e dissertações. Os três restantes têm as características descritas a seguir.

Sebastiana Aparecida Moreira (2013), em dissertação apresentada ao Programa de Pós-Graduação da Faculdade de Educação da Universidade de Passo Fundo, estudou a

<sup>9</sup> Ferramenta *online* do Sistema Nacional de Pós-Graduação da Capes que coleta, disponibiliza e analisa informações sobre este nível de ensino.

contribuição do Programa Mais Educação para as escolas municipais de Rio Verde, GO, comparando os resultados do Ideb entre as escolas que aderiram e as que não aderiram ao programa pela percepção dos sujeitos. A pesquisadora concluiu que o programa contribuiu para a melhoria do aprendizado dos alunos e para a retirada de crianças, adolescentes e jovens vulneráveis da situação de risco. Apesar de aparecer na busca de estudos meta-avaliativos, esse estudo não foi caracterizado como tal, pois não houve avaliação de uma avaliação, e sim de um programa que promoveu atividades em período integral para os alunos de algumas escolas do município.

Cleide Pereira Oliveira (2013) apresentou sua dissertação ao Programa de Pós-graduação Gestão e Tecnologias Aplicadas à Educação (GESTEC), da Universidade do Estado da Bahia, na qual analisou a experiência de avaliação institucional de uma escola de educação básica do sistema municipal de ensino da cidade de Salvador, com o objetivo de avaliar os limites e as possibilidades da prática da autoavaliação pelas escolas em seu cotidiano. Segundo a pesquisadora, trata-se de uma meta-avaliação por ter sido realizada uma avaliação da avaliação institucional, e os resultados obtidos serviram para estimular e instruir outras unidades escolares a adotarem a prática da avaliação institucional a fim de que se torne uma realidade e uma marca da cultura da escola.

Raquel Paula Fortunato (2014), em dissertação apresentada ao Programa de Pós-Graduação da Faculdade de Educação da Universidade de Passo Fundo, RS, analisou em que medida a implementação da Lei nº 11.274/ 2006, que instituiu a obrigatoriedade do ensino fundamental de nove anos, fundou uma nova organização pedagógica, com vistas ao avanço da democratização e à universalização do ensino. A pesquisadora analisou documentos e entrevistou gestores de duas escolas municipais, duas estaduais e uma escola privada do município de Passo Fundo, e considerou que as escolas conseguiram atender parcialmente à construção de uma nova lógica escolar para atendimento da legislação e que houve avanços em medidas para democratizar o ensino e garantir a permanência dos alunos na escola. Por ser um estudo que analisou uma política pública e seus efeitos no sistema educacional, não havendo qualquer menção à avaliação educacional, não consideramos que essa dissertação seja compatível com o objeto de pesquisa em pauta e o descritor da busca no catálogo de teses da Capes.

Em levantamento junto à *Scientific Eletronic Library Online* (SciELO) do armazenamento da produção científica em formato eletrônico nos periódicos brasileiros,

foram detectados dez artigos que se encaixam no descritor definido na pesquisa. No Quadro 3 consta a síntese dos achados.

**Quadro 3 - Artigos disponibilizados na Scientific Eletronic Library Online (SciELO)**

Ano	Artigo/ Periódico	Autor
2004	O uso da avaliação e a avaliação de seu uso (acerca da avaliação da qualidade do ensino). Paidéia (Ribeirão Preto), Vol. 14, n. 29, pp. 255-264	Silva, José Aparecido da
2005	A categoria precisão na avaliação e na meta-avaliação: aspectos práticos e teóricos em um novo enfoque. Ensaio: Avaliação e Políticas Públicas em Educação. Vol. 19, n. 73, pp. 941-964	Letichevsky, Ana Caroline; Velasco, Marley Maria B. R.; Souza, Reinaldo Castro
2007	Um sistema <i>fuzzy</i> de suporte à decisão para meta-avaliação uma nova abordagem e um estudo de caso desenvolvidos no Brasil. Ensaio: Avaliação e Políticas Públicas em Educação, Vol. 15, n. 56, pp. 447-462	Letichevsky, Ana Carolina; Velasco, Marley Maria B. R.; Tanscheit, Ricardo
2010	Meta-avaliação de uma extensão universitária: estudo de caso. Avaliação: Revista da Avaliação da Educação Superior, Vol. 15, n. 2, pp. 131-157	Arroyo, Daniela Muserato P.; Rocha, Maria Silvia de Moura Librandi da.
2011	Meta-avaliação: das abordagens às possibilidades de aplicação. Ensaio: Avaliação e Políticas Públicas em Educação, Vol. 19, n. 73, pp. 941-964	Elliot, Ligia Gomes
2012	Parâmetros e paradigmas em meta-avaliação: uma revisão explorativa e reflexiva. Ciência e Saúde coletiva, Vol. 17, n. 3, pp 695-705	Furtado, Juarez Pereira; Laperriere, Heléne
2013	Políticas de avaliação do ensino naval. Ensaio: Avaliação e Políticas Públicas em Educação, Vol. 21, n. 80, pp. 533-546	Aguiar, Natália Moraes C.B.; Lins, Maria Judith S. da C.
2016	Meta-avaliação: uma década do processo de avaliação institucional do SINAES. Avaliação: Revista da Avaliação da Educação Superior, Vol. 21, n. 1, pp. 89-108	Pinto, Rodrigo S.; Mello, Simone P. T. de; Melo, Pedro A.
2017	A avaliação da educação superior: um comparativo dos instrumentos de regulação entre Brasil e Portugal. Avaliação: Revista da Avaliação da Educação Superior, Vol. 22, n. 1, pp. 35-54	Felix, Glades Tereza; Bertolin, Julio Godoy; Polidori, Marlis Morosini
2018	Integração de processos avaliativos em uma instituição de ensino superior brasileira. Avaliação: Revista da Avaliação da Educação Superior, Vol. 23, n. 1, pp. 58-74	Castro, Rosangela N. A. de; Costa, Eula M. M. Barcelos; Silveira, Everton W. da; Marcório, Aretuza Alves

Pesquisa realizada no dia 19.05.2019, às 10h 15 min.

Como descrito no Quadro 3, o recorte temporal disponível é de 2004 a 2018, sendo que nos anos 2009, 2014 e 2015 não houve publicação de artigos sobre a temática nos periódicos dessa base. Dos dez artigos publicados, apenas dois são dos mesmos autores (LETICHEVSKY; VELLASCO; SOUZA, 2005; LETICHEVSKY; VELLASCO; TANSCHHEIT, 2007). Os demais publicaram somente um artigo sobre a temática pesquisada. Os artigos selecionados se caracterizam por abordarem predominantemente a meta-avaliação de procedimentos avaliativos no Ensino Superior. Nenhum estudo atende à especificidade da avaliação na educação básica ou de sistema privado de ensino.

Com este levantamento bibliográfico, pode-se verificar que a temática escolhida para esta pesquisa está inserida em um campo que oferece amplo espaço para crescimento em investigações. A pesquisa que mais se aproximou do objeto de estudos em questão foi a realizada por Dantas (2009), que avaliou as contribuições de uma avaliação em larga escala em escolas do sistema de ensino público do estado da Bahia. Portanto, tal revisão levou-nos a considerar dois aspectos: procedimentos meta-avaliativos no campo da educação não são recursos utilizados com frequência e não há nenhuma pesquisa com meta-avaliação de sistemas de avaliação em larga escala em redes privadas de ensino. Entende-se que, além de verificar as produções do campo de estudos, foi possível constatar uma lacuna que esta pesquisa poderá contribuir para preencher.

## 1.2 Aspectos Metodológicos

A pesquisa está pautada na abordagem<sup>10</sup> qualitativa, que busca descrever, compreender e explicar fenômenos sociais. Minayo e Sanches (1993) a descrevem como a abordagem que trabalha com valores, crenças, representações, hábitos, atitudes e opiniões. Os autores consideram que:

É no campo da subjetividade e do simbolismo que se afirma a abordagem qualitativa. A compreensão das relações e atividades humanas com os significados que as animam é radicalmente diferente do agrupamento dos fenômenos sob conceitos e/ou categorias genéricas dadas pelas observações e experimentações e pela descoberta de leis que ordenariam o social. A abordagem qualitativa realiza uma aproximação fundamental e de intimidade entre sujeito e objeto, uma vez que ambos são da mesma natureza: ela se envolve com empatia aos motivos, às intenções, aos projetos dos atores, a partir dos quais as ações, as estruturas e as relações tornam-se significativas. (MINAYO; SANCHES, 1993, p. 244)

Nessa abordagem, o pesquisador necessita de uma atitude atenta e aberta, livre de predisposições e preconceitos, com uma genuína humildade, capaz de considerar que todas as pessoas envolvidas na pesquisa são sujeitos que elaboram conhecimentos e produzem práticas adequadas para interferirem nos problemas que identificam (CHIZZOTTI, 2010).

Para VÍCTORA, Knauth e Hassen (2000), a abordagem qualitativa se assenta sobre alguns pressupostos que orientam não apenas os aspectos metodológicos da pesquisa, mas a própria

---

<sup>10</sup> Alguns autores utilizam as expressões “pesquisa qualitativa” ou “método qualitativo” (FLICK, 2009; GUERRA, 2014; CHIZZOTTI, 2010). Utilizamos a expressão “abordagem qualitativa” (MINAYO, 1993; 2009, 2010) por entendermos que denota a lógica de uma perspectiva que vai além dos aspectos técnicos ou metodológicos, envolvendo questões epistemológicas e axiológicas e que atende a inúmeras correntes distintas de pesquisas.

concepção do problema de pesquisa, os pressupostos teóricos e os resultados da pesquisa. São eles: i) o mundo real só existe, de fato, na medida em que nós tomamos parte dele e ele faz sentido para nós; ii) a sociedade é constituída de microprocessos, ou seja, a realidade social não é um todo unitário, mas uma multiplicidade de processos sociais que atuam simultaneamente e em temporalidades diferenciadas, compondo uma totalidade; iii) as sociedades se movimentam a partir das forças da ação individual e grupal, havendo preponderância da ação grupal sobre a individual, e esta não é a soma da ação dos indivíduos. Assim, diante de tais pressupostos, deve-se ter um permanente cuidado, por exemplo, em reconhecer os microprocessos que ocorrem num contexto – em nosso estudo, o educacional – objetivá-los para investigá-los, sem realizar um recorte que os isole ou desvincule, ainda que temporalmente, da realidade social na qual estão inseridos. Para tanto, o pesquisador necessita ser capaz de alcançar uma compreensão global dos fenômenos, o que requer que “mantenha um contato estreito e direto com a situação onde os fenômenos ocorrem naturalmente” (LÜDKE; ANDRÉ, 1986, p. 12).

Guba e Lincoln (2011), ao descreverem as metodologias de investigação numa abordagem construtivista, consideram que os pesquisadores que se inserem na moldura de aprendizes e, portanto, encaram a possibilidade de não saberem o que eles ainda não sabem, pretendendo empregar o instrumento humano para coletar informações, devem usar os métodos que “estiverem mais à mão de um determinado ser humano. Esses métodos são, evidentemente, qualitativos” (p. 196). Os autores argumentam que:

Os seres humanos coletam informações de uma maneira mais adequada e cômoda empregando diretamente seus sentidos: conversando com as pessoas, observando o que elas fazem, lendo documentos a seu respeito, analisando os indícios discretos que elas deixam para trás, respondendo às suas dicas não verbais, entre outros meios. É por este motivo que os métodos qualitativos são preferidos, e *não* porque se configurem como um alicerce para a definição do paradigma construtivista. (GUBA; LINCOLN, 2011, p. 196).

Outra característica dessa abordagem, destacada por esses autores, é que a pesquisa prescinde da formulação de uma hipótese (condição esta, indispensável na abordagem quantitativa pautada no paradigma convencional<sup>11</sup>), ela se estabelece sobre pressupostos que dão origem a um problema ou pergunta norteadora (GUBA; LINCOLN, 2011).

Consideração semelhante é apresentada por Chizzotti (2010), ao caracterizar a pesquisa na abordagem qualitativa. O autor considera que o problema de pesquisa não pode ficar atrelado a uma hipótese previamente aventada. Este decorre de um processo indutivo que vai se definindo e se delimitando na exploração dos contextos em que se realiza a pesquisa.

---

<sup>11</sup> O paradigma convencional será descrito no capítulo 4.

Numa reflexão sobre o processo de análise de estudos com base empírica, buscando facilitar a compreensão dos que desejam se familiarizar com a abordagem qualitativa, Minayo (2012) indica a necessidade de se conhecerem os termos estruturantes das pesquisas qualitativas. Estes são compostos por um conjunto de substantivos e três verbos. Os substantivos são: experiência, vivência, senso comum e ação. Os verbos são: compreender, interpretar e dialetizar. Baseando-se nos estudos do filósofo alemão Martin Heidegger, a autora apresenta o sentido de cada termo:

- Experiência diz respeito ao que o ser humano apreende no lugar que ocupa no mundo e nas ações que realiza. O sentido da experiência é a compreensão de si mesmo e seu significado no mundo da vida.
- Vivência é o produto da reflexão pessoal sobre a experiência. Embora pessoal, a vivência tem um ingrediente do coletivo onde ela ocorre e é única de cada indivíduo, pois depende de sua personalidade, biografia e historicidade.
- Senso comum pode ser definido como o conjunto de conhecimentos provenientes das experiências e vivências que orientam o ser humano nas diferentes situações da vida. É constituído das opiniões, valores, crenças, modos de pensar, agir, sentir e relacionar. Expressa-se pela linguagem, atitudes, condutas e é a base do entendimento humano. A autora o denomina “o chão dos estudos qualitativos.” (MINAYO, 2012, p. 622).
- Ação (humana e social) é definida como o exercício dos indivíduos, dos grupos e das instituições para construir suas vidas e artefatos culturais, vinculado à noção de liberdade para agir e transformar o mundo.
- Compreender é o principal verbo da abordagem qualitativa. Corresponde ao exercício da capacidade de colocar-se no lugar do outro, levando em conta a singularidade do indivíduo “porque sua subjetividade é uma manifestação do viver total” (p. 623). Toda compreensão é parcial e inacabada, tanto do pesquisado como do pesquisador.
- Interpretar, portanto, é o ato que sucede à compreensão. É a apropriação do que se compreende, é elaborar as possibilidades projetadas pelo que foi compreendido.
- Dialetizar é exercitar o entendimento das contradições, pois o ser que compreende o faz tanto na ação como na linguagem, e ambas têm características conflituosas e contraditórias. Não se deve fechar o campo da ideia, e sim discuti-

la como um ato de novos alcances. A noção de que somos limitados no que compreendemos e no que interpretamos deve nos conduzir a uma atitude de busca constante de entendimento das contradições, das desigualdades, dos efeitos das relações de poder e de produção, à luz dos fundamentos teóricos que nos orientam nas discussões e indagações que ocorrem durante todo o percurso investigativo.

As premissas da abordagem qualitativa orientam a escolha das técnicas e instrumentos para a produção do material empírico que privilegiem que os indivíduos pesquisados possam expressar suas experiências, vivências e reflexões sobre o objeto da investigação.

### **1.2.1 Técnicas e Instrumentos de produção do material empírico**

A abordagem qualitativa permite um amplo conjunto de técnicas e instrumentos que sejam capazes de “capturar a perspectiva dos participantes” (LÜDKE; ANDRÉ, 1986, p. 12). Na verdade, a complexidade inerente às pesquisas educacionais e a outras áreas das ciências sociais tem propiciado a constante busca por aperfeiçoar técnicas que se adaptem à natureza de determinado campo de estudos. Flick (2009) menciona a existência de manuais específicos de pesquisa qualitativa para as áreas de Psicologia, outros para a área de Saúde e até mesmo de Gestão, procurando a adaptação de técnicas às perspectivas das pesquisas sociais. Para esta pesquisa, o material empírico será produzido por coleta e análise documental, entrevistas semiestruturadas e questionários *online*.

A análise documental, ainda que, segundo Víctora, Knauth e Hassen (2000), seja pouco utilizada na abordagem qualitativa, consiste em uma valiosa fonte de informações quando aliada a outras técnicas, a fim de complementar ou evidenciar novos fatos. Chizzotti (2010) apresenta a análise de documentos como uma importante técnica que pode agregar um volume qualitativo de dados originais e relevantes. Guba e Lincoln (2011) também destacam a vantagem do uso de documentos como uma fonte repleta de informações sobre o contexto que não deve ser ignorada, quaisquer que sejam as demais técnicas escolhidas. Nesta pesquisa, os documentos coletados e analisados são: o Projeto Pedagógico da rede; o Regulamento do Programa de Avaliação em Larga Escala da Rede (PAER); comunicados informativos e cartas encaminhados aos gestores da rede, aos gestores das unidades escolares, aos pais dos alunos; Boletins de Resultados das unidades escolares e de alunos. Houve também extensa pesquisa na legislação brasileira pertinente ao tema.

Foram realizadas entrevistas semiestruturadas que se constituem como uma das principais técnicas na abordagem de pesquisa qualitativa, por favorecer, por meio da conversação, o levantamento de informações e obtenção de dados em um contato face a face entre o entrevistador e o entrevistado. Por meio de uma interação pessoal, o entrevistador é capaz de avaliar atitudes, gestos, reações, observando o que é dito e como é dito. A entrevista semiestruturada permite correções, esclarecimentos e adaptações que a tornam muito eficaz na obtenção de informações desejadas. Nela, as experiências dos indivíduos e sua percepção do real são as informações por excelência, que se buscam captar. Tais dados só são obtidos pela interação com as pessoas, os quais não se encontram em fontes documentais e que são relevantes e significativos (MARCONI; LAKATOS, 2017).

Lüdke e André (1986) afirmam que “a entrevista ganha vida ao se iniciar o diálogo entre o entrevistador e o entrevistado” (p. 34). É na interação social que os dados subjetivos podem ser obtidos com a contribuição das pessoas. Trata-se de um privilégio, porém, acompanhado de uma responsabilidade acentuada para o pesquisador, que precisa desenvolver um clima de confiança, de diálogo aberto e envolvimento que, na visão de Minayo (2009), são condições essenciais: “A inter-relação que contempla o afetivo, o existencial, o contexto do dia a dia, as experiências e as linguagens do senso comum no ato da entrevista, é condição *sine qua non* do êxito da pesquisa qualitativa” (p. 68, grifos da autora).

Para Minayo (2009), as entrevistas são “*conversas com finalidade* que se caracterizam pela sua forma de organização” (p. 64, grifos da autora). As semiestruturadas são as que combinam perguntas fechadas e abertas, na qual o entrevistado tem a liberdade de discorrer sobre o tema em questão sem se prender rigidamente à indagação formulada.

Flick (2009) sinalizou a possibilidade de as entrevistas serem realizadas remotamente, valendo-se de meios de comunicação como telefone, por recursos da internet, quando os envolvidos se encontram distantes. Considerando que o objeto da presente investigação está inserido no contexto de uma rede de ensino distribuída em todo o território nacional, com distâncias geográficas que inviabilizam o deslocamento da pesquisadora, a entrevista por recurso *online* foi, assim, utilizada para permitir que os participantes fossem ouvidos do lugar onde estivessem.

Levando-se também em conta a extensão territorial, optamos por um questionário *online* com questões fechadas e questões abertas, que tem como vantagem atingir um maior número de participantes em menor tempo. Apesar de alguns autores considerarem o questionário uma técnica da abordagem quantitativa (VÍCTORA; KNAUTH; HASSEN,

2000; CHIZZOTTI, 2010), outros a consideram como uma útil técnica da abordagem qualitativa, quando esse instrumento é capaz de apreender opinião, percepção e preferências dos pesquisados (CHAER *et al.*, 2011; MINAYO, 2010; RIBEIRO; GESSINGER, 2018). Para tanto, requer que as questões sejam bem elaboradas e pré-testadas, para que atendam às condições de uma abordagem qualitativa. Ribeiro e Gessinger (2018), citando Minayo (2005), apresentam as seguintes condições para o questionário constituir-se um importante instrumento de coleta de dados:

a confiabilidade dos questionários e formulários reside na quantidade de respostas idênticas em diferentes intervalos de tempo e espaço, independentemente do pesquisador que aplica o questionário ou formulário. O pesquisador deve ser conhecedor do assunto que vai abordar no questionário, para poder organizá-lo em temas, e deve elaborar cuidadosamente as questões, atentando para o grau de importância, o que vai favorecer a obtenção de informações relevantes e válidas. Os temas devem estar em alinhamento com os objetivos gerais e específicos, e as questões devem ser codificadas para facilitar sua tabulação. Por fim, para que os participantes saibam o que se deseja com a aplicação do questionário ou formulário, deve conter instruções definidas, claras e notas explicativas. (RIBEIRO; GESSINGER, 2018, p. 104).

Segundo Marconi e Lakatos (2017), a realização da pré-testagem favorece a verificação da fidedignidade, da validade e da operatividade. Os autores afirmam que os questionários alcançam, em média, 25% de devolução. Sendo assim, não são indicados para amostragens qualitativas que trabalhem com um número consideravelmente pequeno de participantes.

Podemos concluir, portanto, que haverá uma triangulação de técnicas internas à própria abordagem, segundo Minayo (2011), estratégia que consiste em olhar o mesmo objeto sob diversos ângulos, comparando o resultado de duas ou mais técnicas de coletas de dados e de duas ou mais fontes de informação. Neste caso, a análise documental, a entrevista e o questionário.

### **1.3 O locus da Pesquisa**

A pesquisa tem como *locus* uma rede privada de ensino que está presente em todas as unidades da federação e atende a todos os níveis de ensino: da educação básica ao ensino superior e pós-graduação *stricto sensu*. Este estudo tem como foco a educação básica da rede, que conta com cerca de 350 unidades de educacionais.

As unidades educacionais estão localizadas em todas as capitais dos estados e em municípios de grande, média ou baixa densidade populacional, indistintamente. Há unidades

em regiões centrais e em regiões periféricas das cidades. Portanto, a diversidade é uma das características marcantes da rede, pois além da extensão geográfica – com escolas nas grandes cidades, bem como em municípios ribeirinhos da Amazônia –, também abriga a diversidade populacional das diferentes regiões em que atua.

Com mais de 120 anos de atuação no Brasil, a rede desenvolve seu próprio programa de avaliação em larga escala há oito anos. A referida rede é caracterizada com os detalhes possíveis, no capítulo 4 da tese, em função da preservação da identidade, conforme algures mencionado.

#### **1.4 Os participantes da pesquisa**

O desenho de uma pesquisa qualitativa requer, ao serem determinados o problema de pesquisa, sua delimitação e contextualização, que o pesquisador esteja habilitado a identificar, também, quais participantes serão capazes de contribuir com suas experiências, com as percepções que possuem do problema, com seus saberes e práticas, para que efetivamente a pesquisa retrate o mais fidedignamente o contexto na perspectiva desses sujeitos (LÜDKE; ANDRÉ, 1986).

A escolha dos participantes está relacionada aos objetivos da pesquisa. Tratando-se de uma meta-avaliação, na qual se pretende desvelar os possíveis ganhos e dificuldades de um programa de avaliação em larga escala, a seleção se deu em função daqueles que estão em posições estratégicas de acesso a dados, tomada de decisão e que podem fazer uso do programa de avaliação como possível propulsor da qualidade educacional pretendida. Portanto, foram elencados participantes em funções de gestão em três níveis da estrutura administrativa da rede, num total de 109 gestores, assim caracterizados:

- PLG: Líder gestor para toda a rede na América do Sul e idealizador do programa.
- PGR: Gestores das oito regiões administrativas do Brasil, assim identificados: PGR1, PGR2, PGR3, PGR4, PGR5, PGR6, PGR7, PGR8.
- PGSR: Gestores das aproximadamente 50 sub-regiões do Brasil.

Para os participantes PLG e PGR de 1 a 8, o instrumento utilizado foi a entrevista semiestruturada. Para os participantes PGSR das sub-regiões administrativas, o instrumento utilizado foi o questionário misto *on-line*.

Foram fatores de exclusão do participante da pesquisa duas condições: 1) encontrar-se na função, na ocasião da entrevista ou questionário, há menos de dois anos; 2) não haver participado de nenhuma edição do programa de avaliação em larga escala, na função exercida. Devido a estes fatores, a configuração dos participantes passou a ser de 5 entrevistados: PLG; PGR 1; PGR 2; PGR 3; PGR4. Apenas 4 Gestores Regionais atendiam ao critério de dois anos na função. Porém, ainda assim, obteve-se uma representatividade das regiões geográficas, assim configuradas: Norte, Nordeste, Sul e Sudeste.

Os questionários enviados por e-mail para as cerca de 50 sub-regiões distribuídas em todo o território brasileiro foram endereçados para 105 gestores, e o total de respondentes foi de 47 participantes.

### **1.5 Análise dos dados**

A pesquisa na abordagem qualitativa requer que o tratamento de todo o material coletado na investigação leve em conta as premissas dessa abordagem, tendo em vista, principalmente, a dimensão descritiva e a dimensão interpretativa, que são fundamentais na hora de trabalharmos o material coletado e apresentar os achados mais significativos. É o momento de “confronto entre o quadro de referência do pesquisador e o material empírico recolhido.” (GUERRA, 2014, p. 62).

Sobre essa fase da pesquisa, VÍCTORA, KNAUTH e HASSEN descrevem:

A análise é desenvolvida através da *discussão* que os temas e os dados suscitam. A análise inclui as referências bibliográficas, o modelo teórico, juízos de valor e deve propor conclusões. No momento da discussão, o pesquisador adota e descarta teorias existentes com base na argumentação que seus achados lhe facultam. [...] Os resultados vão sendo apresentados e, concomitantemente, discutidos. (VÍCTORA; KNAUTH; HASSEN, 2000, p. 75, grifo dos autores)

Para os autores, esse momento da pesquisa é o que mais se identifica com a ciência propriamente dita, quando “a explicação da realidade é a tentativa de encontrar os princípios subjacentes aos achados no decorrer da pesquisa.” (VÍCTORA; KNAUTH; HASSEN, 2000, p. 76). O caráter interpretativo e explicativo da pesquisa qualitativa encontra seu amplo sentido no processo de análise dos dados, que requer do pesquisador não somente a habilidade de identificar fenômenos, mas fôlego para entendê-los e explicá-los (VÍCTORA; KNAUTH; HASSEN, 2000).

Romeu Gomes (2009) descreve o processo de análise e interpretação dos dados da pesquisa, destacando que:

[...] quando falamos de análise e interpretação de informações geradas no campo da pesquisa qualitativa, estamos falando de um momento em que o pesquisador procura finalizar seu trabalho, ancorando-se em todo o material coletado e articulando esse material aos propósitos da pesquisa e à sua fundamentação teórica. Nesse sentido, estamos nos referindo a uma etapa final no processo de investigação. No entanto, não podemos desconsiderar dois aspectos importantes. O primeiro deles diz respeito à ideia de que, tanto a análise quanto a interpretação ocorrem ao longo de todo o processo. Já o segundo se refere ao fato de que, em pesquisa qualitativa, às vezes, ao chegarmos à fase final, descobrimos que necessitamos retornar às partes das fases anteriores. (GOMES, R., 2009, p. 81).

O autor argumenta que são utilizados três tratamentos dos dados qualitativos no processo final: a descrição, a análise e a interpretação. Na descrição, as opiniões dos informantes são apresentadas da maneira mais fiel possível. Na análise, decompõem-se os dados, indo além do descrito, buscando relação entre as partes. Na interpretação, busca-se sentido para as falas ou ações para se chegar a uma compreensão ou explicação além do descrito e analisado.

Nesta pesquisa, a análise e interpretação do material empírico considerou as categorias estabelecidas e levantadas, a partir das respostas dos participantes, em constante diálogo com os autores referenciados, valendo-se da análise de conteúdo, em um esforço de triangulação dos dados. Flick (2009) chamou de triangulação, a articulação de mais de um método ou técnica de pesquisa, com a finalidade de ampliar o conhecimento sobre o objeto de estudo, na busca de superar as limitações impostas pela utilização de um único método. O autor considera que a triangulação metodológica contribui para a ampliação, enriquecimento e complementariedade da produção do conhecimento.

A análise de conteúdo é um procedimento ou técnica amplamente utilizada na pesquisa de abordagem qualitativa. Surgiu no início do século XX, num cenário em que o positivismo predominava, influenciado pela corrente behaviorista da psicologia. Portanto, originalmente, foi concebida numa perspectiva quantitativa, visando a uma descrição objetiva e sistemática do conteúdo expresso na comunicação. Já na década de 1970, foi possível perceber avanços no sentido de a técnica ser adaptada para a abordagem qualitativa, ampliando as possibilidades de uso, com a abertura de um vasto campo de aplicação em diversas áreas do conhecimento.

Bardin (1979), em clássica obra sobre o assunto, assim designou a análise de conteúdo:

Um conjunto de técnicas de análise das comunicações, visando obter, por procedimentos, sistemáticos e objetivos de descrição do conteúdo das mensagens, indicadores (quantitativos ou não) que permitam a inferência de conhecimentos relativos às condições de produção/ recepção (variáveis inferidas) dessas mensagens. (BARDIN, 1979, p. 42).

Como vemos na descrição da autora, a análise de conteúdo comporta um conjunto de técnicas que se expande, se adapta e se assenta, inclusive em grande diversidade de posturas teóricas. Como afirma Guerra (2014), diferentes paradigmas podem valer-se dessa técnica, assim como “o tratamento do conteúdo varia, pois, consideravelmente, de pesquisa para pesquisa, e de investigador para investigador.” (GUERRA, 2014, p. 63).

Tratando-se de uma pesquisa cujo material empírico foi coletado a partir de entrevistas semiestruturadas e questionários, propõe-se uma análise temática-categorial, apoiada na estruturação de Bardin (1979), por meio da qual o pesquisador se debruça para além das palavras e sua significação, “buscando-se outras realidades *através* da mensagem” (BARDIN, 1979, p.44, *itálico* do autor).

Como recurso de apoio para análise dos dados qualitativos será utilizado o *software* *WebQDA*, cuja licença foi adquirida pela pesquisadora. Operando de forma semelhante a outros programas disponíveis para este fim, este se diferencia por permitir um trabalho colaborativo de vários pesquisadores simultaneamente. Desenvolvido por um grupo de professores e técnicos da Universidade de Aveiro, é especialmente útil para investigadores que trabalhem com análise de dados qualitativos, favorecendo a categorização, a codificação, a filtragem e o questionamento dos dados que emergem da investigação.

O esforço da análise dos dados requer que o aporte teórico dos autores referenciados seja o balizador para que o pesquisador consiga ir além do levantamento de dados – que, muitas vezes, se constitui numa reunião exaustiva de informações –, para um frutífero cotejamento de ideias, mediante o qual a interpretação dos dados não seja a culminância do estudo, mas o indicativo de que o problema levantado pela pesquisa obteve resposta ou respostas.

## CAPÍTULO 2 - FUNDAMENTOS DA META-AVALIAÇÃO

Este capítulo apresenta os fundamentos da meta-avaliação, os principais modelos utilizados por meta-avaliadores na área educacional e a abordagem empregada na presente investigação.

Estudos e pesquisas recentes apontam para o número crescente de sistemas de avaliação em larga escala nos sistemas públicos de estados e municípios, constituindo-se em instrumento de políticas públicas cuja finalidade anunciada aponta para o alcance e a garantia da qualidade educacional. Sabe-se que o uso de tais avaliações pode estar assentado em interesses, concepções, fundamentos teóricos e valores variados.

Sobre o uso de avaliações no contexto educacional, Domingos Fernandes considera que os sistemas educacionais estão organizados com base em culturas de avaliação diferentes: uns adesevolvem para classificar, certificar, selecionar, com ênfase nos resultados dos alunos; e outros, com o propósito primordial, entre outros fins, de promover a aprendizagem, induzir a práticas inovadoras de ensino, favorecer a gestão e a tomada de decisões em todos os níveis (DOMINGOS, D. 2009). Porém, o autor afirma que há uma insatisfação quase crônica, em muitos países do mundo, com a qualidade do serviço educacional que é prestado pelos sistemas educacionais, em geral, e que governos lançam reformas que, por vezes, ficam aquém do desejável, e sente-se falta da concretização de medidas estruturantes subordinadas a uma agenda educacional forte e que melhore substancialmente a educação. Afirma que:

“No domínio da avaliação das aprendizagens dos alunos os esforços têm-se centrado muito mais nas avaliações externas do que nas avaliações internas, na expectativa de obter informação confiável acerca do que os alunos sabem e de contribuir para melhorar a qualidade do ensino e das escolas” (FERNANDES, D., 2009, p. 111).

Na visão do autor, avaliação externa é uma medida que visa a melhoria da qualidade do ensino, mas trata-se de matéria polêmica para pesquisadores e educadores que consideram duvidoso que, por si mesmas, as avaliações possam cumprir esse papel. Contudo, reconhece que os países continuam investindo em “exames nacionais” e em “estudos internacionais de avaliação das aprendizagens” e que “os exames podem ter certamente um conjunto de vantagens e, talvez em parte, por essa razão, são aplicados em quase todos os países.” (p. 137). O fato de haver uma adesão crescente de sistemas de avaliação externa como um fato percebido não somente no Brasil, mas até mesmo em vários países, como sinalizado pelo autor, favorece ao questionamento: estas “vantagens” são traduzidas em melhoria da qualidade do ensino? Qualidade pautada em quais pressupostos?

Teóricos da área da avaliação sinalizam para os riscos e problemas decorrentes de uma avaliação mal realizada, mal utilizada, ou mesmo mal-intencionada (WORTHEN; SANDERS; FITZPATRICK, 2004; GUBA; LINCOLN, 2011; FERNANDES, D. 2009), sem deixar de considerar a possibilidade de descompasso conceitual entre os envolvidos, a inadequação entre os objetivos e finalidades propostos e aqueles esperados, as falhas de comunicação ou conflitos ideológicos entre os grupos de interesse ou interessados (*stakeholders*) e os avaliadores. Essas possibilidades sempre existiram, desde que se realizam avaliações de natureza formal. Contudo, não se pode deixar de considerar a importância da avaliação de sistemas, programas e práticas, quando devidamente realizados.

É neste contexto de aumento da importância atribuída às avaliações de sistemas de ensino, de dúvidas e críticas levantadas sobre o uso, a função, a eficácia e o alcance dos objetivos propostos que as justificam, que estudos meta-avaliativos podem constituir-se em um promissor recurso para ajudar a julgar a pertinência de avaliações, seus instrumentos, processos e uso dos resultados. Grego (1997) salienta que é no exercício da crítica e da reflexão sobre os processos de avaliação que os estudos na área da meta-avaliação vêm oferecendo um auspicioso caminho. Pois, como afirmou Stufflebeam (2011), “*Good evaluation requires that evaluation efforts themselves be evaluated.*” (p. 99).

A meta-avaliação tem sido considerada a avaliação da avaliação. Segundo Cook e Gruder (1978 *apud* FURTADO; LAPERRIÈRE, 2012), a expressão “avaliações da avaliação” foi criada por Orata, nos anos 1940, refletindo, porém, o sentido estrito da meta-avaliação: a preocupação em qualificar os processos avaliativos.

O conceito de meta-avaliação como a avaliação de valor e mérito de uma avaliação foi introduzido por Michael Scriven, nos Estados Unidos, em 1969. Ampliando a concepção, no ano de 2009, quando completou 40 anos de sua primeira publicação que conceituava o termo, Scriven (2009) publicou um artigo intitulado *Meta-evaluation Revisited*, no qual procurou responder a três questões principais: (i) exatamente o que é; (ii) como podemos justificá-la; e (iii) quando e como deve ser usada. Em sua conclusão, o autor afirma que: “*It is perhaps fair to say that metaevaluation is the conscience of evaluation, just as evaluation is the conscience of society. Long may both continue to improve!*” (SCRIVEN, 2009, p. viii). Consideramos que esta síntese do autor consegue retratar a importância que este processo tem adquirido, não apenas no contexto em que foi escrito, no qual já há uma cultura avaliativa, mas em outros contextos e realidades. Vale ressaltar que, segundo Worthen, Sanders e Fitzpatrick (2004), a

meta-avaliação surgiu no contexto da avaliação educacional, havendo se expandido para áreas nas quais as avaliações de programas e processos tornaram-se essenciais.

Stufflebeam (2001) procurou ampliar a definição de meta-avaliação elaborada por Scriven (1969), ao conceituá-la como:

*[...] the process of delineating, obtaining, and applying descriptive information and judgmental information—about the utility, feasibility, propriety, and accuracy of an evaluation and its systematic nature, competent conduct, integrity/honesty, respectfulness, and social responsibility—to guide the evaluation and/or report its strengths and weaknesses. (STUFFLEBEAM, 2001, p. 185)*

O autor também enfatizou o aspecto de ajuizamento de mérito e valor em suas abordagens ao tema, sendo que a ideia de mérito está ligada ao atendimento a critérios e padrões instrumentais, e a concepção de valor, ligada ao atendimento das necessidades dos *stakeholders* (STUFFLEBEAM; SCHINKFIELD, 1985).

Díaz (2001) considera que a meta-avaliação não é uma teoria alternativa à avaliação. Ela não só transcorre pelos atos e dados avaliados como transcende a própria avaliação. Porém, ela se estabelece a partir dos mesmos referenciais adotados para a avaliação. Em outras palavras, as mesmas bases epistemológicas que orientam as concepções de avaliações, orientam as concepções acerca da meta-avaliação.

Em uma ampla análise da literatura sobre meta-avaliação, Díaz (2001) faz uma síntese dos conceitos de meta-avaliação em duas abordagens distintas: a primeira, que coincide com a definição de Scriven de 1969, faz referência a um conteúdo prático, operacional, e de aplicação imediata do conceito. Nessa perspectiva, a meta-avaliação é uma fase, ou uma etapa, usualmente terminal, do trabalho ou projeto de avaliação. É definida como a atividade ou conjunto de atividades que consistem na elaboração de listas de comprovação e na aplicação dessas listas e, com base nelas, construir um juízo de valor sobre a avaliação como resultado. Essas listas incluem critérios e escalas de valores para medir variáveis, cuja inclusão responde a esses critérios. A segunda aceção é mais compreensiva e geral. Nesse caso, por meta-avaliação se entende dar sentido e explicar o significado do conjunto de atividades humanas que se reconhecem como avaliação. Mas, além de uma clarificação terminológica, o propósito da meta-avaliação é a explicação da avaliação como um processo social complexo. Essa abordagem se refere ao objeto da meta-avaliação – a avaliação –, lidando não só com uma rede de ações sociais complexas, como também com as suas referências e explicações do caráter axiológico, ideológico e epistemológico, fundamentalmente. Portanto, podemos considerar que os mesmos paradigmas que fundamentam a avaliação, também fundamentam as ideias sobre meta-avaliação.

Semelhante análise é feita por Cotera e Matamoros (2011), que, ao apresentarem uma resenha do desenvolvimento do conceito de meta-avaliação, fazem uma distinção de duas vertentes de autores que tratam da temática. A primeira está centrada nos estudos de Scriven (1967, 1991) e de Stufflebeam (1981, 1987), que, juntamente com outros estudiosos, entendem que a meta-avaliação é uma investigação sistemática, cujo objetivo é emitir um juízo acerca da qualidade e dos méritos de uma avaliação, tornando-se assim, no objeto de estudo do meta-avaliador. Para isso, consideram necessário o desenvolvimento de uma metodologia que tenha um conjunto de critérios apropriados. Embora estejam de acordo no tocante a que as avaliações devam chegar a conclusões que sejam validas interna e externamente, autênticas, com critérios e *standards* técnicos, entendem que estes não são suficientes para julgar os achados ou dados gerados pelas avaliações. Sendo assim, a meta-avaliação necessita de critérios específicos, tanto para o aperfeiçoamento e benefício da avaliação como para validar a própria avaliação.

Já em consulta a outros autores, Cotera e Matamoros (2011) constatam que a meta-avaliação, no âmbito dos programas educativos, pode ser entendida como *“la investigacion que trata de valorar la calidad de la propia metodología de evaluación empleada y, por tanto, el valor de las consecuencias de la aplicación de tal metodología.”* (GENTO PALACIOS; CASTILLO ARREDONDO, 1995 *apud* COTERA; MATAMOROS, 2011 p. 252). Nessa perspectiva, a meta-avaliação é considerada como a evolução da avaliação, com a qual se trata de averiguar até que ponto a avaliação realizada alcançou e ampliou seus objetivos, até que ponto cumpriu a missão a que fora designada e até que ponto os instrumentos utilizados são adequados para obter as informações desejadas (RUIZ RUIZ, 1996, *apud* COTERA; MATAMOROS, 2011).

Para esses autores, a meta-avaliação não é um processo essencialmente técnico, mas tem uma natureza ética e política. É um processo de aprendizagem que gera outros processos, não menos importantes, já que pode ser aplicada em qualquer momento da avaliação, desde seu planejamento até uma avaliação já completada, com os resultados e a opinião dos avaliadores (COTERA; MATAMOROS, 2011).

O *Sistema Centroamericano de Evaluación y Armonización de la Educación Superior* (SICEVAES), responsável por promover a cultura da qualidade, pela acreditação e avaliação das instituições superiores da América Central, em fevereiro de 2005, emitiu um documento no qual orientava a respeito da importância da meta-avaliação em todos os processos avaliativos institucionais e da análise do impacto que os processos de autoavaliação e

avaliações externas estavam produzindo nas universidades da América Central. Nesse documento, o sistema apresenta como objetivo da meta-avaliação “*asegurar la calidad de los servicios evaluativos, evitar enfrentarse a las prácticas ilegales y a los servicios que no son de interés público, señalar el camino para el perfeccionamiento de la profesión y promover una mayor comprensión de la empresa evaluativa* (SICEVAES, 2005, p. 6). Consideram-se, assim, que os diferentes modelos possíveis de meta-avaliação refletem os distintos marcos referenciais existentes sobre os processos avaliativos.

## 2.1 Tipos, modelos e padrões de Meta-avaliação

Como descrito, a meta-avaliação é uma forma de avaliação. Portanto, segue as mesmas premissas da avaliação. Analisando os elementos que caracterizam os modelos ou padrões adotados, podem-se identificar com mais precisão os pressupostos que a orientam. Nesta subseção, apresentamos as classificações de diferentes autores, levando em consideração distintos elementos da meta-avaliação que indicam seus marcos referenciais.

### 2.1.1 Tipos de Meta-avaliação

Bustelo (2003, *apud* CONTERA; MATAMOROS, 2011) classifica a meta-avaliação em quatro tipos básicos, levando em conta seu papel, seu conteúdo, o momento e o agente que a realiza:

- 1) **Considerando seu papel:** a meta-avaliação pode ser formativa ou somativa. É formativa, quando se dá ao longo do processo, com o propósito de melhorar o desenho e o processo avaliativo. É somativa, quando se dá ao final, com a finalidade de recapitular o processo completo de avaliação.
- 2) **Considerando o conteúdo ou a fase:** pode ser do desenho, do processo e do resultado. A meta-avaliação do desenho implica na análise e valoração do planejamento: os propósitos, os objetivos, como foi concebido, a contextualização, a coerência interna, as estratégias de implementação etc. Também leva em conta os custos, as expectativas, a utilidade e o calendário, se é realístico. A análise do processo considera as dificuldades na implementação da avaliação, os ajustes necessários, a efetividade das equipes de trabalho e ajustes de calendário. E, finalmente, a meta-avaliação dos resultados leva em conta as informações obtidas com os resultados e seus efeitos. Tanto as informações obtidas por meio dos dados como as evidências produzidas por meio de juízos, recomendações,

interpretações disponíveis nos relatórios são analisadas para ajudar no entendimento da utilidade da avaliação.

- 3) **Considerando o momento da avaliação:** Antes da implementação do processo (ex-ante) ou durante o processo ou imediatamente após o término (ex-post-facto).
- 4) **Considerando o agente que realiza:** Agente interno ou externo. Na primeira, os mesmo que realizaram a avaliação desenvolvem a meta-avaliação. Logo, são os responsáveis pelo padrão de qualidade de todo o processo. A meta-avaliação por agente externo é desenvolvida por um avaliador que não esteve envolvido no processo de avaliação. Acredita-se que esse tipo dá mais credibilidade ao processo e reconhecimento externo.

### 2.1.2 Abordagens somativas e formativas

Elliot (2011), em um estudo sobre distintas abordagens de meta-avaliação de diversos autores, classificou-as como somativas e formativas. Como abordagens somativas, a autora apresentou os seguintes modelos: os Padrões de Avaliação do *Joint Committee on Standards for Educational Evaluation*; a Lista Chave de Verificação da Avaliação de Scriven (2007); os Critérios Fundamentais, a Abordagem de segunda opinião e a Abordagem híbrida de Davidson (2005). Como abordagem formativa, a autora também apresenta duas sugestões de Davidson (2005). Descrevemos, a seguir, os modelos, segundo a classificação de Elliot (2011).

#### 2.1.2.1 Padrões de Avaliação do Joint Committee on Standards for Educational Evaluation

Durante a década de 1960, os principais teóricos norte-americanos da avaliação<sup>12</sup> começaram a discutir o que seria uma boa avaliação ou uma avaliação ruim. A princípio, produziram listas de diretrizes para avaliação e as trocaram informalmente entre si. A seguir, vários avaliadores passaram a publicar as propostas de critérios de meta-avaliação ou relatórios de avaliação. Worthen, Sanders e Fitzpatrick (2004) afirmam que tais publicações foram, em geral, bem recebidas, e vários autores dos critérios propostos para a meta-avaliação tentaram torná-los úteis para os consumidores da avaliação, pois pensavam que talvez os clientes de estudos avaliatórios tivessem mais condições de julgar a qualidade e pertinência

---

<sup>12</sup> O capítulo 4 detalha esse período, sendo que os principais teóricos norte-americanos citados são: Lee Cronbach, Michel Scriven, Daniel Stufflebeam e Robert Stake (CALDERÓN; BORGES, 2013).

das avaliações. Contudo, os autores afirmam que o conjunto de diferentes critérios propostos se revelou desconcertante para ambos: clientes e avaliadores.

O cenário de ausência de consenso entre os avaliadores, diante de inúmeros conjuntos de critérios propostos, trouxe como consequência um ambicioso projeto iniciado em 1975: a criação de um comitê misto que desenvolveria um amplo conjunto de diretrizes sobre as quais houvesse concordância geral em termos da qualidade da avaliação. Assim, foi criado o *Joint Committee on Standards for Educational Evaluation*, formado por representantes de 15 instituições profissionais norte-americanas e canadenses, além de inúmeros avaliadores profissionais, cujo trabalho teve como resultado a publicação, em 1981, do *Standards for evaluations of educational programs, projects, and materials*. Como podemos observar no título dado aos padrões estabelecidos, foram concebidos, inicialmente, para a avaliação em educação. Posteriormente, em 1988, um conjunto de diretrizes foi publicado para a área de recursos humanos, mas em 1994, uma segunda edição foi publicada com uma ampliação dos padrões e escopos que atendessem à avaliação de programas em várias áreas. Recebeu o título de *The program evaluation standards*. Para essa revisão, também houve a mobilização de representantes de 15 organizações e associações norte-americanas, além de representantes de outros países que utilizavam os padrões produzidos pelo *Joint Committee* – como é conhecido. Em 2004, foi iniciado um processo de revisão e validação dos Padrões para publicação de uma terceira edição. Semelhante aos anos anteriores, inúmeros participantes e usuários contribuíram com sugestões que, após diversas apresentações e revisões, culminaram com a apresentação da última versão, em 2011. Os padrões de avaliação do *Joint Committee* estão agrupados em cinco grandes atributos: utilidade, viabilidade, adequação, precisão e responsabilização<sup>13</sup>. Assim, os padrões de avaliação, num total de 30, são divididos entre os cinco atributos, conforme consta no Quadro 4.

**Quadro 4 - Lista de Verificação dos Padrões de Avaliação do Joint Committee**

ATRIBUTOS	PADRÕES DE AVALIAÇÃO
UTILIDADE	U1 – Credibilidade do avaliador
	U2 – Atenção às partes interessadas
	U3 – Objetivos negociados
	U4 – Valores explícitos

<sup>13</sup> Os atributos recebem esta denominação segundo a tradução realizada por Stark (1999 *apud* Pinto, 2015). Outros autores traduzem os atributos como “utilidade, viabilidade, conveniência, precisão” (Furtado; Laperrière 2012); “utilidade, exequibilidade, adequação, precisão e responsabilização” (Elliot, 2011); “utilidade, viabilidade, ética e precisão (Penna Firme; Letichevsky, 2010); “utilidade, viabilidade, propriedade e exatidão” (Ristoff, 2000).

	U5 – Informações relevantes
	U6 – Processos e produtos significativos
	U7 – Comunicação e relatórios apropriados e no prazo
	U8 – Preocupação com as consequências e influência
VIABILIDADE	V1 – Gerenciamento do projeto
	V2 – Procedimentos práticos
	V3 – Viabilidade contextual
	V4 – Uso de recursos
ADEQUAÇÃO	A1 – Orientação responsiva e inclusiva
	A2 – Acordos formais
	A3 – Direitos humanos e respeito
	A4 – Clareza e justiça
	A5 – Transparência e divulgação
	A6 – Conflitos de interesse
	A7 – Responsabilidade fiscal
PRECISÃO	P1 – Conclusões e decisões justificadas
	P2 – Informações válidas
	P3 – Informações confiáveis
	P4 – Descrições explícitas do programa e do contexto
	P5 – Gestão da informação
	P6 – Projetos e análises adequados
	P7 – Raciocínio de avaliação explícito
	P8 – Comunicação e relatórios
RESPONSABILIZAÇÃO	R1 – Documentação da avaliação
	R2 – Meta-avaliação interna
	R3 – Meta-avaliação externa

Fonte: *Evaluation Checklist Projet*; elaboração da autora.

Os padrões de avaliação do *Joint Committee* são conhecidos e adotados mundialmente. Principalmente nos Estados Unidos, onde há inúmeras organizações de avaliação, acreditação e *survey*, e com tantos avaliadores profissionais, há amplo uso da matriz desenvolvida pelo comitê. Para Whorten, Sanders e Fitzpatrick (2004), “as Diretrizes do Joint Committee continuam sendo o melhor quadro de referência que existe.” (p. 600).

No Brasil, vários pesquisadores em suas teses de doutorado ou dissertações de mestrado também utilizam essas diretrizes, na maioria das vezes, parcialmente. Como apresentou Elliot (2011), de seis exemplos de estudos meta-avaliativos produzidos por pesquisadores brasileiros em distintas áreas do conhecimento (BERLINER, 2010; FIGUEIRÓ; THULER; DIAS, 2008; GADELHA, 2010; HEDLER; TORRES, 2009; PEREIRA, 2010; SEROUR, 2010), cinco utilizaram os padrões de avaliação do *Joint Committee*, e um pesquisador utilizou o modelo desenvolvido por Davok (2006) em sua tese de doutorado.

### 2.1.2.2 Lista Chave de Verificação da Avaliação de Scriven (*Key Evaluation Checklist*)

Scriven (2015) considera que sua Lista de Verificação é a mais antiga e mais detalhada lista produzida para ser utilizada por avaliadores profissionais, nas mais diversas áreas, para avaliação de programas, serviços, produtos, práticas e em meta-avaliações. Para o autor, cada vez mais, todas as áreas de atividades e conhecimento necessitam de processos avaliativos e são amplamente beneficiadas pela *Key Evaluation Checklist (KEC)*, justificando essa afirmação pelas frequentes atualizações às quais sua lista foi submetida. A última versão disponível em seu site [www.michelscriven.info](http://www.michelscriven.info) data de novembro de 2015.

A lista é um guia para avaliadores e meta-avaliadores. É composta por quatro componentes principais que representam níveis para se realizar a meta-avaliação. O primeiro componente denominado *Preliminares*, segundo o autor, é parte essencial da avaliação, momento em que se busca a identificação da avaliação como se fosse um posfácio dos resultados obtidos. O segundo componente é chamado de *Fundamentos*, quando são determinados o contexto e a natureza do objeto da avaliação, e a investigação realmente se inicia. Também nesse nível, os valores definidos em relação ao padrão adotado na avaliação são descritos. O terceiro componente é denominado *Subavaliações*, quando o processo é propriamente detalhado ou, como descreve o autor, a fase de “apuração dos fatos” é iniciada. Nela, se consideram os resultados ou impactos da avaliação, os custos, a comparação com outros programas e a capacidade de utilizar componentes da avaliação em outras situações. O último nível, das *Conclusões*, é o componente no qual uma síntese é realizada na forma de relatório com uma apresentação concisa dos resultados, apontando possíveis fraquezas e forças, o que foi feito e o que poderia ter sido feito. E, para o autor, é nesse nível que o avaliador deve realizar uma meta-avaliação, lançando um olhar crítico para a avaliação realizada. O Quadro 5 traz uma representação da lista de Scriven.

**Quadro 5 - Lista Chave de Verificação de Avaliação de Scriven – KEC**

<b>Componentes</b>	<b>Pontos de Verificação</b>
PRELIMINARES	A1 - Sumário executivo
	A2 – Esclarecimentos
	A3 – Desing e métodos
FUNDAMENTOS	B1 – Antecedentes e contextos
	B2 – Descrições e definições
	B3 – Consumidores (impactos)
	B4 – Recursos
	B5 – Valores

SUBAVALIAÇÕES	C1 – Processos
	C2 – Resultados (efeitos)
	C3 – Custos
	C4 – Comparações
	C5 – Generalizações
CONCLUSÕES	D1 – Síntese
	D2 – Recomendações, explicações e previsões
	D3 – Responsabilidade e justificativa
	D4 – Relatório e suporte
	D5 – Meta-avaliação

Fonte: Scriven (2015)      Elaboração da autora

Para Scriven (2009), ao realizar uma meta-avaliação, todos os avaliadores ou meta-avaliadores devem zelar arduamente para considerar os aspectos essenciais dessa prática, a fim de alcançar sempre padrões mais elevados de julgamento, fazendo o melhor possível para alcançar o objetivo de garantir confiabilidade a todo o corpo de conhecimento desenvolvido na área de avaliação.

### 2.1.2.3 Critérios fundamentais, Abordagem de Segunda Opinião e Abordagem Híbrida ou mista de Davidson

Elliot (2011) e Pinto (2015) apresentam as três abordagens que E. Jane Davidson descreveu em seu livro *Evaluation methodology basics*, publicado em 2005. Davidson (2005) dedicou um capítulo à meta-avaliação. A primeira abordagem desenvolvida pela autora é denominada *Critérios Fundamentais*.

Os critérios fundamentais de Davidson (2005, *apud* ELLIOT, 2011), que orientam uma meta-avaliação buscando seu mérito e qualidade intrínseca, são: Validade, Utilidade, Conduta, Credibilidade e Custos. Numa publicação simplificada de seu método de avaliação pela ação traduzido para o espanhol, denominada *Principios Básicos de la Evaluación para la Acción*, Davidson (2013) procurou apresentar de forma simples, porém não simplista, cada critério e sua importância para a adesão e envolvimento das partes interessadas (*stakeholders*). Destacamos, no Quadro 6, as principais ideias envolvidas em cada critério.

#### Quadro 6 - Critérios Fundamentais de Davidson para a meta-avaliação

Critérios	Considerações
Validade	Consiste na combinação adequada de conhecimentos especializados dos atores envolvidos na avaliação, ajudando a assegurar que todas as vozes sejam ouvidas.
Justiça	Trata-se da oportunidade de participação aos membros da comunidade, aos reais e potenciais

	beneficiários dos programas e aqueles a quem o programa impacta.
Credibilidade	Em que medida merecem crédito os resultados e suas fontes, levando em conta o contexto, a independência e imparcialidade, o trato com os possíveis conflitos de interesse e a <i>expertise</i> dos meta-avaliadores.
Utilização	A participação dos interessados influencia na utilização, não só dos resultados, mas do conhecimento, do <i>know-how</i> e da capacidade de avaliação desenvolvida no próprio processo. A meta-avaliação deve verificar em que medida os resultados da avaliação estão sendo utilizados.
Custos	A consideração dos custos, não apenas monetários, mas de tempo investido, pois a relação custo-benefício deve ser equilibrada, ponderando se os resultados valem o que custaram.

Fonte: Davidson, 2013; elaboração da autora.

Finalmente, Elliot (2011) faz uma apresentação das duas abordagens que Davidson (2005) apresentou em seu capítulo sobre meta-avaliação, denominadas de segunda opinião e híbrida. A primeira é aquela na qual o interessado solicita que uma nova equipe realize outra avaliação do mesmo objeto, valendo-se dos mesmos caminhos metodológicos da primeira avaliação. No final da segunda avaliação, os relatórios são comparados para que se verifique se as conclusões são semelhantes ou se há discrepâncias. Caso haja discrepâncias, o meta-avaliador, o avaliador e as equipes precisam trabalhá-las para que se possa chegar a um retrato o mais fiel possível.

A abordagem híbrida é assim descrita por Elliot:

Nessa abordagem, híbrida ou eclética, que conjuga dois procedimentos, o meta-avaliador faz o julgamento da avaliação como ela é, da maneira como foi realizada e apresentada, acrescentando uma verificação cruzada de aspectos que podem ser questionados quando a avaliação for divulgada. Isso ocorre quase sempre quando os *stakeholders* não concordam com alguns dos resultados encontrados. É uma maneira sistematizada de verificar a qualidade da avaliação e de clarificar pontos de discordância nos resultados ou em sua interpretação. (ELLIOT, 2011, p. 956).

A autora ainda tece considerações sobre a abordagem formativa que visa à melhoria da avaliação em processo, enquanto está sendo desenvolvida, apresentando duas sugestões de Davidson (2005) para a realização desse tipo de meta-avaliação. A primeira seria a equipe de avaliação contratar um consultor especialista em medidas ou conteúdo para apreciar e comentar o plano de avaliação. Um painel pode ser utilizado com a participação de avaliadores especialistas e membros da organização na qual ocorre a avaliação. A participação dos membros da organização aumenta as chances de aceitação posterior dos resultados da avaliação. A segunda possibilidade é contar com observadores independentes como meta-avaliadores. Atuam ao longo do processo,

com a avaliação em andamento, apontando e cobrindo pontos que julgam deficientes para serem corrigidos antes mesmo dos resultados finais.

### **2.1.3 Modelo de critérios paralelos e critérios de autenticidade de Guba e Lincoln**

Em contraposição aos padrões e critérios desenvolvidos por comitês e associações de avaliadores, estão os critérios estabelecidos por Guba e Lincoln (2011). Ao escreverem sobre a qualidade de uma avaliação, os autores consideram que os critérios utilizados para a realização da avaliação da adequação de uma avaliação estão enraizados em determinados pressupostos. À semelhança dos paradigmas que regem os modelos de avaliações, não se pode apreciar uma avaliação de natureza construtivista por meio de critérios e padrões positivistas. Desse modo, os autores tecem críticas a padrões desenvolvidos por comitês e associações que estabelecem *standards* convencionais, considerando-os inaplicáveis a avaliações realizadas num paradigma construtivista. Dentre os vários modelos de padrões e critérios desenvolvidos com a finalidade de avaliar a qualidade de uma avaliação, consideram que os padrões elaborados pelo *Joint Committee (JCSEE)*, já descrito, não são adequados, mas podem ser aceitos, conquanto não sejam em grande medida eficazes para a avaliação da qualidade de uma avaliação de quarta geração<sup>14</sup>. Portanto, em contraposição aos principais padrões para a prática de avaliações de avaliações, criaram uma relação de critérios para julgar a qualidade e adequação de uma avaliação na perspectiva do paradigma construtivista.

Em uma abordagem positivista, os principais critérios seriam: validade interna e externa, confiabilidade e objetividade. Para os autores, esses aspectos refletem uma exigência positivista de neutralidade, de isenção de tendenciosidade, isenção de valores ou preconceitos que caracterizam a estrutura ontológica e epistemológica do positivismo. Por isso, em contraposição e de forma paralela, numa perspectiva construtivista, recorrem a uma elaboração de critérios de fidedignidade que julgam significativos nessa abordagem. São eles: credibilidade, transferabilidade, dependabilidade e confirmabilidade. Já como critérios de autenticidade, os autores consideram: a imparcialidade, a autenticidade ontológica, a autenticidade educativa, a autenticidade catalítica e a autenticidade tácita. O Quadro 7 traz uma breve descrição de cada critério.

#### **Quadro 7 - Critérios Paralelos e de Autenticidade de Guba e Lincoln**

---

<sup>14</sup> Avaliação de quarta geração é a designação do modelo proposto pelos autores para uma avaliação numa abordagem construtivista, num paradigma que se contrapõe ao positivismo, mais bem detalhado no capítulo 4.

<b>CRITÉRIOS PARALELOS</b>	<b>Credibilidade:</b> a validade é resultado da correspondência entre as realidades construídas pelos respondentes ou grupos de interesse e a reconstrução destas, pelo avaliador.
	<b>Transferabilidade:</b> o processo empírico que confirma o grau de semelhança entre os contextos de origem e de destino e a aplicabilidade a cada situação.
	<b>Dependabilidade:</b> a estabilidade dos dados ao longo do tempo, permitindo mudanças ao longo do processo.
	<b>Confirmabilidade:</b> os dados, as interpretações e os resultados das investigações estão enraizados em pessoas, e não no método, nem nas interpretações do avaliador, apenas.
<b>CRITÉRIOS DE AUTENTICIDADE</b>	<b>Imparcialidade:</b> diz respeito a diferentes estruturas de valores subjacentes ao processo de avaliação.
	<b>Autenticidade ontológica:</b> aprimoramento da experiência consciente e amadurecimento dos indivíduos ou grupos.
	<b>Autenticidade educativa:</b> o grau em que a compreensão e a apreciação dos respondentes em relação às construções de pessoas externas são ampliadas.
	<b>Autenticidade catalítica:</b> o grau em que os processos de avaliação estimulam e facilitam determinadas ações.
	<b>Autenticidade Tácita:</b> diz respeito ao grau de autonomia e poder de ação dos interessados e participantes.

Elaboração: autora.

Desse modo, os autores consideram que não é possível realizar uma meta-avaliação numa abordagem construtivista, utilizando-se critérios positivistas, de modo que a visão do papel do avaliador, da relação entre os interessados e o avaliador, bem como de todo o processo, deve ser pautada numa perspectiva coerente com o paradigma que definiu as bases originais da avaliação.

## 2.2 O meta-avaliador

Desde a sistematização de critérios e padrões para a avaliação da qualidade das avaliações, formulados por iniciativa de comitês e associações de avaliadores da América do Norte, na década de 1970, ao uso sistemático por países da Europa e da América Central, como descrito anteriormente, a meta-avaliação adquiriu um caráter profissional, de domínio técnico, que requer um amplo conhecimento por parte de quem a realiza.

Whorten, Sanders e Fitzpatrick (2004), ao descreverem o papel do meta-avaliador e seu perfil, citaram as ideias de Brinkerhoff *et al.* (1983, *apud* WHORTEN; SANDERS; FITZPATRICK, 2004), que dizem que “[Os meta-avaliadores] não só devem ser competentes o bastante para *fazer* a avaliação original como também têm de ser capazes de dizer se esta é boa ou ruim e de convencer os outros de que eles sabem a diferença” (p. 607, grifo dos autores). Para os autores, o desafio seria de um superavaliador, com competências especiais para realizar avaliações de avaliações e de avaliadores e, preferencialmente, um segundo avaliador com qualificações extraordinárias. Contudo, concluem que é difícil encontrar tal profissional e, como em “muitas outras atividades, simplesmente fazemos o melhor possível” (WHORTEN; SANDERS; FITZPATRICK, 2004, p. 607).

Sendo assim, indicam três possibilidades de meta-avaliadores: (i) O avaliador original realizando a meta-avaliação; para os autores, essa não é a melhor opção, mas, em caso de necessidade, é melhor o próprio avaliador aferir seu trabalho do que não haver uma reavaliação. (ii) O cliente, o interessado ou o consumidor, como meta-avaliador. Certamente, essa abordagem dependerá muito da competência técnica do cliente para julgar a partir de critérios previamente estabelecidos. Também há riscos nessa estratégia. (iii) Avaliadores competentes, sendo que, em geral, os meta-avaliadores externos têm mais credibilidade que os internos, e uma equipe tem mais condições de contribuir do que um único avaliador.

Compreendemos que o perfil do meta-avaliador requer as mesmas competências que se exigem de um avaliador. Além de competências técnicas e a imparcialidade em seus julgamentos, citamos resumidamente os princípios orientadores que a *American Evaluation Association* (AEA) coloca como expectativa de modo de vida profissional e que devem permear a vida cotidiana dos que se dedicam à avaliação como carreira: (i) Investigação sistemática: os avaliadores devem fazer investigações sistemáticas sobre seu objeto de estudo; (ii) Competência: preparar-se para o desempenho da sua atividade, procurando manter e aumentar suas competências e desenvolvimento profissional; (iii) Integridade/ honestidade: em todo o processo, o avaliador deve preservar esses princípios que garantirão a credibilidade de sua atuação; (iv) Respeito pelas pessoas: respeitar a dignidade, a segurança, a autoestima

dos participantes, dos interessados e dos clientes, presando a ética e preservando o sigilo, quando assim for necessário; (v) Responsabilidade pelo bem-estar geral e público: levar em conta as diversidades de interesses e valores relacionados não apenas aos clientes ou interessados, mas os interesses da sociedade como um todo (WHORTEN; SANDERS; FITZPATRICK, 2004).

Ainda que se considerem todos os princípios e sejam tomados todos os cuidados para a realização de uma meta-avaliação fidedigna, sempre há riscos de erros ou problemas. Porém, como bem afirmou Scriven (2009, p. v), em artigo no qual revisita o tema após 40 anos da primeira publicação: “*A partial meta-evaluation is better than none*”. O autor considera que, no mundo real, é muito difícil a realização de uma meta-avaliação completa e isenta. Porém, a qualidade do avaliador pode atenuar essas dificuldades.

Em uma meta-avaliação desenvolvida por meio de uma abordagem formativa e colaborativa, os meta-avaliadores reconhecem que também aprendem com os avaliados. Portanto, é um processo conjunto, instrutivo, consensual. O meta-avaliador continua sendo um técnico, porém é um analista de dados humanos. Além das competências técnicas, desenvolve competências inter-relacionais, que Guba e Lincoln (2011) consideram como as principais: “*paciência, humildade, transparência, adaptabilidade e senso de humor.*” (p. 292).

Nessa perspectiva, o meta-avaliador reconhece que a avaliação é um processo sociopolítico, e, sendo assim, os aspectos sociais, culturais, políticos e contextuais são partes essenciais e tão importantes quanto às questões técnicas. Considera a avaliação como um processo contínuo, sujeito a reconstruções, adaptações, levantando mais perguntas do que respostas. Ele assume uma função de colaborador, aprendiz, modelador e agente de mudança. A competência técnica é redefinida, ampliada, e a abertura para sempre aprender potencializa essa competência (GUBA; LINCOLN, 2011).

Para efeito desta pesquisa, nos situamos na condição de pesquisadora que tem como foco um estudo meta-avaliativo, cuja abordagem descrevemos a seguir.

### **2.3 A meta-avaliação nesta pesquisa**

Como é perceptível, a literatura e os modelos sobre meta-avaliação concentram-se em autores de vários países, principalmente dos Estados Unidos, onde as experiências com avaliações profissionais, agências de acreditação e avaliação, e a tradição em organização de padrões, critérios e métodos de avaliação de programas, processos, instituições e serviços já atravessa décadas e tem sido adaptada em vários países, tais como a Holanda, que possui um

sistema de meta-avaliação gerido pelas universidades, por meio de uma associação privada e independente (AMARAL, 1994 *apud* GREGO, 1997); no Reino Unido, como uma atividade institucionalizada pelo “*Higher Education Quality Council*”(GREGO, 1997); nos países da América Central, que criaram o *Sistema Centroamericano de la Calidad de la Educación Superior* (SICEVAES), que orienta a realização de meta-avaliações nas instituições de ensino superior da região, atestando-se, portanto, que esse procedimento está presente em várias atividades e vários países.

No Brasil, os procedimentos meta-avaliativos foram intensificados nos últimos anos com o crescimento dos sistemas de avaliação desenvolvidos por meio das políticas públicas de educação e demandas por avaliações de programas em diversas áreas, principalmente na área da saúde. Contudo, como constatado no levantamento bibliográfico, a literatura sobre o tema na área educacional ainda é escassa. Davok (2006) considerou que há uma lacuna tanto na disseminação de relatórios de meta-avaliação quanto de referências teórico-metodológicas de como conduzi-las. Podemos destacar que há autores brasileiros que têm atuado com meta-avaliações e produzido literatura sobre suas experiências, principalmente, ligadas ao ensino superior (GREGO, 1997; RISTOFF, 2000; PENNA FIRME; LETICHEVSKY, 2002; GIMENES, 2007), assim como há publicações de experiências meta-avaliativas na área da saúde (HARTZ, 1997; FURTADO, 2006; FURTADO; LAPERRIÈRE, 2012; HARTZ; FELISBERTO; SILVA, 2008).

Gatti (2013) considera que:

Os pesquisadores em educação precisam voltar-se mais ao estudo e à discussão das avaliações externas em larga escala, quer quanto aos aspectos relativos às políticas e ações educacionais, quer quanto aos aspectos teóricos e aos procedimentos implementados, à luz de uma perspectiva também pedagógica. (GATTI, 2013, p. 65).

Para a autora, os estudos meta-avaliativos podem contribuir para análises consistentes quanto às políticas avaliativas adotadas no país. Considera que a perspectiva da meta-avaliação permite:

(1) Manter nossa capacidade ‘crítica’ em estado de atenção, evitando subjugar-se por opiniões ‘prontas’ e justificativas pseudo-revolucionárias, (2) buscar certo domínio sobre os métodos empregados (o que permite uma visão consciente quanto à avaliação e seus resultados: limites, problemas, contribuições); (3) trabalhar fatos e dados da avaliação educacional com uma visão prospectiva de aprimoramento, alavanca para desenvolvimento de domínios de conhecimentos, saberes, formas de agir; (4) traçar novos caminhos de socialização de saberes, lembrando o direito de todos ao conhecimento, à apropriação de saberes necessários a uma melhor existência e aos cuidados de si e do outro; (5) fugir ao domínio de estereótipos,

reconhecer as diversidades e acreditar que é possível melhorar, mudar uma condição, respeitar situações. (GATTI, 2013, p. 65).

Diante do rol de contribuições da meta-avaliação listadas pela autora, entendemos que ainda há um amplo espaço para avançarmos no sentido de transformarmos a meta-avaliação em um procedimento necessário, diante de tantas avaliações que têm sido desenvolvidas no âmbito educacional nas últimas décadas. Como demonstrado no levantamento bibliográfico, há, ainda, algumas práticas de avaliação que são consideradas equivocadamente como meta-avaliações e quase nada tem sido produzido sobre meta-avaliação de avaliações externas ou, ainda, institucionais, na educação básica.

A meta-avaliação, nesta pesquisa, vem atender ao objetivo geral: identificar os ganhos e dificuldades de um Programa de Avaliação em Larga Escala desenvolvido em uma rede privada de ensino de educação básica. Semelhantemente, contribuirá alcançar outros objetivos específicos. Tais como: analisar os ganhos e dificuldades do Programa de Avaliação em Larga Escala; e apontar possíveis caminhos para a superação das dificuldades identificadas.

Alcançar esse objetivo constitui-se o propósito de uma meta-avaliação na visão de Patton (1997 *apud* FURTADO; LAPERRIÈRE, 2012, p. 697), que considera que “[...] a meta-avaliação proveria informações sobre as limitações e as potencialidades da avaliação realizada, aumentando sua credibilidade e permitindo aos grupos de interesse julgarem e contextualizarem os resultados obtidos”.

Portanto, a escolha por uma meta-avaliação numa abordagem formativa, sem a utilização de *standards*, sem controle de processos, sem a busca de eficiência ou eficácia, mas sim da compreensão do contexto onde dá-se o processo avaliativo em análise, a busca de dar sentido e explicar o significado do conjunto de atividades que se reconhecem como avaliação em larga escala, a percepção da avaliação educacional como um processo social complexo, tendo como referência o caráter axiológico, ideológico e epistemológico, que rege essas práticas, situa a meta-avaliação aqui pretendida numa perspectiva colaborativa, tendo como referência as ideias de avaliação pautadas no paradigma emergente.

## **CAPÍTULO 3 - TRAJETÓRIA DAS AVALIAÇÕES EM LARGA ESCALA DA EDUCAÇÃO BÁSICA NO BRASIL**

Este capítulo descreve a trajetória da avaliação em larga escala da educação básica brasileira, especificamente do Sistema de Avaliação da Educação Básica (Saeb), cujo início do desenvolvimento data do final dos anos 1980, e sua evolução até os dias atuais. Tendo em vista que o Saeb se tornou referência para a criação de inúmeros sistemas de avaliação em todo o território brasileiro e, especificamente, foi modelo para a criação do programa de avaliação que é o objeto desta pesquisa, a descrição da sua trajetória oferece a oportunidade de contextualizarmos o cenário que propiciou o surgimento, o desenvolvimento e a manutenção de tal iniciativa numa rede privada de ensino.

O recorte histórico se inicia no período da penúltima década do século XX, enquanto a extensão da análise se limita à avaliação da educação básica em nível nacional. Contudo, a percepção das condições que favoreceram o desenvolvimento do modelo de avaliação em larga escala, adotado pelo Brasil, faz-se necessária para que a trajetória seja delineada em um contexto que permita uma compreensão dos inúmeros fatores que influenciaram a criação e o desenvolvimento do sistema de avaliação, tal qual o temos neste primeiro quarto do século XXI. Um sistema que não nasceu pronto, mas foi construído por etapas que marcam tanto a evolução da própria teoria da avaliação educacional quanto a institucionalização das organizações necessárias para a sua realização (BROOKE; ALVES; OLIVEIRA, 2015). Aparentemente, uma trajetória simples, por ser recente. Porém, não se torna simples quando buscamos compreender as motivações políticas, a complexidade de uma avaliação educacional sistêmica num país com as características territoriais do Brasil, as mudanças de governo nesse período e a lenta formação de um campo teórico próprio.

### **3.1 Iniciativas embrionárias do Sistema de Avaliação da Educação Básica**

Dependendo da perspectiva pela qual se analisa o fenômeno, distintos fatores podem ser considerados como propulsores ou geradores do processo de construção do sistema de avaliação em larga escala brasileiro. Ao escrever sobre a trajetória do Saeb, Pestana (2016) considerou que:

Embora já faça parte do senso comum, o entendimento de que o agir do homem na transformação cotidiana do mundo se dá em razão de múltiplas determinações e é resultado de uma combinação ativa (talvez única), de momentos históricos, de situações econômicas, sociais, políticas, culturais e psicológicas, de grupos sociais e de indivíduos revela a complexidade associada à análise de qualquer experiência

social. Nesse sentido, qualquer reflexão sobre o processo de criação e desenvolvimento do Sistema de Avaliação da Educação Básica (Saeb) não esgota a possibilidade de análise e, por isso, é sempre necessária. (PESTANA, 2016, p. 72).

É compreensível que uma análise seja marcada por nuances próprias do momento em que é realizada, do olhar dos atores que a realizam e da própria natureza multiforme do tema em questão, no caso, a avaliação em larga escala. E, talvez, seja essa a riqueza da possibilidade de se debruçar sobre um único tema e dele extrair perspectivas diversas, ainda que utilizando-se as mesmas fontes ou os mesmos métodos de análise, porém, tendo a complexidade característica da educação, como prática social que se transforma e se modifica, que garante sua polissemia.

Em uma análise histórica das experiências brasileiras com avaliação educacional de porte mais extenso, Gatti (2002) estabelece a década de 1960 como um ponto de partida, por ter sido, nesses anos, que surgiu a preocupação específica com processos avaliativos escolares baseados em critérios mais claramente enunciados e com técnicas mais objetivas. Os testes objetivos passaram a ser utilizados em larga escala pelos educadores, e os vestibulares unificados, adotados pelas universidades em 1969, como um exemplo dessa modalidade de instrumental de avaliação.

Horta Neto (2018) considera que é impossível analisar o crescimento da avaliação educacional no Brasil sem levar em conta a *expertise* no desenvolvimento de testes cognitivos entre os anos 1960 e 1980. Destaca a importância do papel que a Fundação Carlos Chagas viria a ter no Sistema de Avaliação da Educação Básica, a partir das experiências desenvolvidas por esse órgão nos grandes concursos vestibulares do estado de São Paulo, na década de 1960. Do mesmo modo, Gatti (2009) destaca a iniciativa da Fundação Getúlio Vargas, no Rio de Janeiro, que, em 1966, criou o Centro de Estudos de Testes e Pesquisas Psicológicas que elaborou testes para as últimas séries do ensino médio nas áreas de Língua Portuguesa, Matemática, Ciências da Natureza e Ciências Humanas. A pesquisa também incluía um questionário sobre características socioeconômicas dos alunos e suas aspirações. Essa, portanto, pode ser considerada a primeira iniciativa ampla, no Brasil, para verificação da aquisição de conhecimento associada a diferentes variáveis. Para Horta Neto (2018) e Gatti (2009), esses seriam os embriões para as bases do conhecimento necessário para a construção dos testes cognitivos que fariam parte das avaliações nacionais.

Todavia, as iniciativas e experiências embrionárias que ocorreram na década de 1960 podem ser contextualizadas em uma perspectiva mais ampla do cenário mundial, que implicariam as ideias e a prática educacional brasileiras.

Nesse período, no contexto global, a guerra fria e a polarização entre os Estados Unidos e a União Soviética, após a Segunda Guerra Mundial, são apresentadas por Vianna (1995) como um dos fatores desencadeadores da expansão do campo da avaliação:

A avaliação, após o trauma provocado pela constatação da deficiência tecnológica associada à carência educacional no mundo ocidental, principalmente nos Estados Unidos, com o lançamento do Sputnik, no dia 4 de outubro de 1957, tornou-se impositiva. Há todo um esforço no sentido de recuperar o tempo educacional que fora perdido, sendo criados novos currículos, e a avaliação passou a ter papel de relevância no desenvolvimento de novas estratégias de ensino. (VIANNA, 1995, p.7).

Saul (2010) aponta que, na década de 1960, houve uma forte influência norte-americana no cenário intelectual e educacional brasileiro, mas com uma defasagem de cerca de duas décadas. Algumas experiências realizadas nesse período se basearam em referências teóricas e práticas desenvolvidas nos Estados Unidos. Horta Neto (2007) estabelece como marco de referência, na ocasião, a relação entre avaliação e qualidade, com destaque ao grande levantamento educacional em larga escala realizado nos Estados Unidos, que deu origem ao *Relatório Coleman*. Esse relatório baseou-se em uma pesquisa do governo americano envolvendo 645 mil alunos em cinco diferentes níveis de ensino, com o objetivo de verificar, entre diversas escolas, qual era a variação dos conhecimentos adquiridos pelos alunos. Com ampla divulgação em vários países, inclusive no Brasil, os resultados deram margem a debates e estudos acadêmicos sobre a qualidade educacional.

Fernandes e Gremaud (2009) também entendem que a referida década se constituiu um marco da avaliação. Atribuem ao *Relatório Coleman* uma importância significativa na mudança de foco das avaliações em larga escala. Após sua ampla divulgação, inúmeros pesquisadores norte-americanos debruçaram-se sobre os dados produzidos, tanto para refutá-los como para aprofundarem ainda mais a compreensão sobre os impactos e o papel das escolas no contexto socioeconômico, e assim, a ênfase no processo avaliativo ganhou novo impulso.

Do mesmo modo, Rothen (2018) afirma que, na literatura sobre avaliação, é recorrente a visão de que as ideias vigentes nas políticas internacionais de avaliação da educação foram geradas, principalmente, nos Estados Unidos e na Inglaterra. Há, portanto, um consenso da influência das referências teóricas na área de avaliação, advindas do eixo anglo-saxão no início da segunda metade do século XX e que, por inúmeros fatores de natureza política,

econômica, social e cultural, aprofundaram-se nas últimas décadas do século passado, a partir das diretrizes de órgãos internacionais, como o Banco Mundial (BM), a Organização das Nações Unidas para a Educação, a Ciência e a Cultura (Unesco), a Organização para a Cooperação e Desenvolvimento Econômico (OCDE), que influenciaram o estabelecimento de políticas de avaliação educacional de forma não linear, porém marcante, nas distintas realidades de países latino-americanos. A autora Dirce Freitas menciona como fatores externos, além das recomendações das agências da Organização das Nações Unidas (ONU) relativas à melhoria da qualidade do ensino, que a própria divulgação de experiências de avaliação de vários países e pesquisas focadas no efeito escola e na eficácia escolar <sup>15</sup> também impulsionaram a organização de um aparelho avaliativo (FREITAS, D., 2013).

Riscal e Luiz (2016), ao fazerem um breve histórico das avaliações externas e de larga escala em educação, apresentam o contexto internacional, desde o final da Segunda Guerra Mundial, como um cenário propício para o surgimento de iniciativas de cooperação entre países e a criação de organismos supranacionais, cujo ideário comum aponta enfaticamente para a educação como um meio de melhoria da situação econômica dos indivíduos, por meio da capacitação para o mercado de trabalho, bem como de um fator de crescimento econômico e redução da pobreza nos países. Os autores consideram a OCDE e seus objetivos, o Banco Mundial e o FMI, como organismos conexos que atuaram no financiamento e monitoramento dos países membros e convidados, influenciando diretamente nas políticas do campo educacional, na segunda metade do século passado.

Coelho (2008) destaca a atuação de organismos internacionais, como a Unesco e o Banco Mundial, grandes impulsionadores das reformas educacionais nas últimas décadas do século XX que, segundo a autora, lançaram bases para a formulação e aprovação da Lei de Diretrizes e Bases da Educação Nacional (BRASIL, 1996), assim como para a consolidação do sistema de avaliação da educação básica no país. De igual modo, Dourado (2020) considera inegável a influência dos organismos multilaterais na agenda brasileira, por meio de parâmetros internacionais, e cita a Unesco, o Banco Mundial e a OCDE como indutores de políticas no campo educacional, com destaque para a avaliação estandardizada e as proposições sobre qualidade da educação.

---

<sup>15</sup> Escolas eficazes é o nome dado ao conjunto de estudos e pesquisas desenvolvidas inicialmente nos Estados Unidos, em reação às conclusões do relatório Coleman (1966), cujo propósito era coletar informações sobre escolas que faziam a diferença e produziam resultados positivos em alunos em contexto de vulnerabilidade ou nível socioeconômico desfavorecido. Brooke e Soares (2008, p. 10) definiram o efeito escola como sendo “o quanto um dado estabelecimento escolar, pelas suas políticas e práticas internas, acrescenta ao aprendizado do aluno”.

É nesse contexto que Bruns e Luque (2014), analistas do Banco Mundial, ao estudarem o desempenho educacional na América Latina e no Caribe, sinalizaram que, nas últimas décadas do século XX, houve uma forte expansão da cobertura educacional na região, com o aumento da oferta de vagas, oportunizando o acesso da maioria da população em idade de escolarização ao sistema de ensino público. A maioria dos países alcançou a conclusão universal do ensino fundamental e o aumento das taxas de escolaridade no ensino médio. Contudo, a forte expansão não foi acompanhada, necessariamente, pelo acesso a um ensino de qualidade e, no Brasil, conforme afirmou Maluf (1996), os indicadores de altas taxas de evasão e retenção revelaram-se entre os “piores da América Latina” (MALUF, 1996, p.11).

Houve, portanto, fatores internos e externos que favoreceram o crescente interesse pela temática da avaliação no Brasil, criando condições para as primeiras iniciativas que propiciaram o desenvolvimento de um sistema de avaliação nos anos subsequentes. Duas experiências avaliativas significativas marcaram a década de 1970. Horta Neto (2007) considera como a primeira experiência de avaliação na educação, uma iniciativa da Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (CAPES), órgão ligado ao Ministério de Educação e Cultura (MEC), iniciada em 1976, promovendo a avaliação de programas de pós-graduação existentes no país. A segunda experiência se deu por iniciativa do Programa de Estudos Conjuntos de Integração Econômica Latino-americana (Eciel). Gatti assim o descreve:

um estudo, no Brasil e outros países da América Latina, sobre os determinantes dos níveis de escolaridade e do rendimento escolar obtidos por alunos com diferentes características pessoais e socioeconômicas. Aos alunos foi aplicado questionário para levantamento de dados sobre situação socioeconômica, atitudes com relação ao processo escolar e aspirações, e, um exame de compreensão de leitura e de ciências. (GATTI, 2009, p. 9).

No mesmo período, foi realizado um estudo para a elaboração de um instrumento de medida para verificar a situação dos alunos nas primeiras séries do ensino fundamental abrangendo Leitura, Escrita e Matemática. O instrumento foi testado com amostras de todas as regiões geográficas do Brasil, em um projeto que derivou da experiência com outros estudos durante a década de 1970, no então estado da Guanabara (hoje, cidade do Rio de Janeiro), além de outros pequenos estudos feitos em convênio com o Centro Brasileiro de Pesquisas Educacionais / Inep/ MEC. Sobre esse período, Gatti considera que:

Não houve, no entanto, nos anos imediatamente subsequentes outras iniciativas com a intenção de avaliações mais abrangentes, embora alguns ensaios localizados tenham sido desenvolvidos. Não se observa, nesse período, preocupação de administrações públicas com uma avaliação sistemática do rendimento escolar dos alunos das redes de ensino. (GATTI, 2009, p. 9).

Uma exceção a ser considerada de uma municipalidade que se preocupou com a questão da avaliação de sua rede de ensino foi a cidade de São Paulo que, em 1980, avaliou todas as crianças do terceiro estágio da educação infantil e da 1ª, 3ª, 5ª e 7ª séries do ensino fundamental e dos primeiros anos do ensino médio, com provas de Língua Portuguesa, Matemática e Ciências. O objetivo era verificar o nível de escolaridade dos alunos da rede municipal e utilizar tais dados para planejamento das atividades escolares e extraescolares. Porém, com a mudança administrativa do município, em 1982, os dados obtidos na avaliação não foram utilizados nem foi dada continuidade ao estudo.

Planejado no ano de 1977, mas implementado nos anos iniciais da década de 1980, um projeto de avaliação de grande porte que pode ser considerado um marco na história dos estudos de avaliação de políticas e programas foi o Programa de Expansão e Melhoria do Ensino no Meio Rural do Nordeste Brasileiro, o Projeto Edurural. Com o objetivo de expandir o acesso às quatro primeiras séries do ensino, diminuir as taxas de repetência e evasão e melhorar o rendimento dos alunos de escolas rurais, o projeto foi desenvolvido mediante uma parceria do Ministério da Educação e das secretarias de Educação locais, com financiamento próprio desse ministério e do Banco Mundial, este último com algumas exigências quanto à execução do programa, entre elas, que houvesse uma avaliação externa que fornecesse subsídios para as definições do desenvolvimento da iniciativa. Os estudos e projetos de avaliação ficaram a cargo da Fundação Cearense de Pesquisa, que ficou responsável pelos aspectos institucionais, e da Fundação Carlos Chagas, que teve como incumbência a avaliação do rendimento escolar (GATTI; VIANNA; DAVIS, 1991).

O Edurural oportunizou um estudo transversal com coletas de dados nos anos de 1981, 1983 e 1985, sendo que 60 municípios de três estados do Nordeste – Piauí, Pernambuco e Ceará –, seiscentas escolas e cerca de seis mil alunos foram envolvidos na avaliação do rendimento escolar. Porém, além do rendimento, avaliaram-se as formas de gerenciamento do projeto, do gerenciamento local, os professores, as famílias, as Organizações Municipais de Ensino, levando em conta as características das escolas, com estudos de casos e análise qualitativa dos dados levantados. Segundo Gatti (2009), seminários foram realizados sobre os resultados durante e depois do processo, e a metodologia de coleta e análise dos dados representou um exemplo de futuros estudos que se poderiam realizar, visando “propugnar uma escola mais condizente com as necessidades das populações menos favorecidas socialmente.” (GATTI, 2009, p. 10).

Essa experiência envolvendo pesquisadores da Universidade Federal do Ceará e da Fundação Carlos Chagas contribuiu para o desenvolvimento de *expertise* para um movimento de avaliações de sistemas de ensino que ocorreu em todo o país, no final dos anos 1980. Vianna assim a descreve:

Quase no final da década de 80, foi iniciado um amplo programa de avaliação do rendimento de alunos das escolas de 1º Grau da Rede Pública em todo o País, por iniciativa do **Instituto Nacional de Estudos e Pesquisas Educacionais – INEP**. Inicialmente, em 1987, essa avaliação (Vianna e Gatti, 1988; Vianna, 1989; Vianna, 1989; Gatti, Vianna, Davis, 1991; Gatti, 1993) objetivou identificar, na diversidade do quadro brasileiro, pontos curriculares críticos; verificar o desempenho em aspectos cognitivos básicos de alunos de 1ª, 3ª, 5ª e 7ª séries; e subsidiar os professores para uma recuperação de seus alunos em aspectos básicos do currículo escolar. Essa avaliação, ao final, abrangeu uma amostra de 27.455 alunos de 238 escolas em 69 cidades de vários Estados da Federação existentes à época, inclusive no então Território do Amapá. O projeto teve seus desdobramentos, com novas avaliações, a pedido da Secretaria de Estado de Educação do Paraná, envolvendo, inicialmente, alunos de 2ª e 4ª séries de escolas de sete cidades (Vianna e Gatti, 1988) e depois, uma amostra de alunos de escolas oficiais de 22 outras cidades (Vianna, 1991), com os mesmos objetivos dos trabalhos propostos pelo INEP. (VIANNA, 1995, p.19, grifos do autor).

Essas iniciativas propiciaram a discussão de vários problemas que ocorriam nos sistemas educacionais, apontados por pesquisadores da área, tais como a repetência e a evasão escolar que indicavam um fracasso na escola básica. Nesse contexto, o MEC promoveu uma oficina de trabalho com educadores para discutir a questão. Colocava-se a possibilidade de uma nova estrutura política e a necessidade de uma mudança de atuação do Ministério da Educação, “levando-o a ter um papel orientador e um papel avaliador como referência para políticas e avanços na área da educação básica.” (GATTI, 2009, p.11).

No cenário político, no início da década de 1980, o Brasil entrava em um processo de redemocratização após duas décadas de regime militar, quando o então vice-presidente José Sarney assumiu a presidência do país, após a morte de Tancredo Neves. Inúmeras reformas políticas, sociais e econômicas foram instituídas, sendo, uma das principais, o estabelecimento de uma assembleia nacional constituinte. A mudança esperada pelos educadores que discutiam o fracasso da escola básica foi alocada para a expectativa de que a Constituição Brasileira oportunizasse uma nova configuração e atuação do Ministério da Educação, favorecendo as mudanças necessárias.

Promulgada em 5 de outubro de 1988, a Constituição da República Federativa do Brasil salvaguardou a educação como direito fundamental, em seu capítulo II (Dos direitos sociais), e estabeleceu, em seu artigo 206, que: “O ensino será ministrado com base nos seguintes princípios: [...] VII - garantia de padrão de qualidade” (BRASIL, 1988).

A expressão “garantia de padrão de qualidade”, que surge como princípio constitucional, ainda necessitaria ser esclarecida e resolvida por legislação infraconstitucional. Gomes (2016), que trabalhou como assessor legislativo do Senado Federal de 1987 a 1998, portanto abrangendo o período da Assembleia Nacional Constituinte de 1988 até a promulgação da Lei de Diretrizes e Bases (LDB), em 1996, narra em um depoimento documentado suas memórias desse período, como testemunha e coadjuvante de inúmeros debates internos e externos da casa legislativa, em que se descortinava uma “aurora de redemocratização” e a percepção de responsabilidades maiores ao Poder Legislativo, para restabelecer o Estado Democrático de Direito. Nesse depoimento, descreve que a preocupação com a avaliação externa, como uma forma de controle diante da descentralização da responsabilidade do Estado prevista nas reformas, já se fazia presente nas discussões dos parlamentares das comissões e subcomissões da Assembleia Constituinte. Contudo, tratava-se de um tema cercado por fortes resistências, especialmente, na educação superior.

Gomes (2016) detalha, ainda, que o grande foco de discórdia na Subcomissão da Educação, Cultura e Esporte estava na defesa dos setores público e privado, com lutas históricas sobre a transferência de recursos públicos. Para resolver esse impasse, o relator Artur da Távola chegou a um entendimento com representantes do setor privado, que se transformou no Artigo 209 da Constituição Federal, cujo foco voltou-se a resultados e não a processos, quando estabelece: “O ensino é livre à iniciativa privada, atendidas as seguintes condições: [...] II – autorização e avaliação da qualidade pelo Poder Público” (BRASIL, 1988). Dessa forma, apesar de a avaliação ter sido tema amplamente discutido diante da nova ordem política que se estabelecia, figurando no texto da Carta Magna, ela afeta apenas ao setor privado, tendo como sujeito avaliador (ou regulamentador) o Estado.

Castro (2016) aponta como a principal mudança institucional derivada da Constituição de 1988, que repercutiu diretamente na organização do sistema educacional, a alteração do pacto federativo, concedendo autonomia aos municípios para organizarem seus próprios sistemas de ensino, dando origem a mais de 5 mil sistemas educacionais, tornando muito mais complexa a coordenação de uma política nacional de educação básica. Essa estrutura descentralizada necessitaria de outras formas de acompanhamento e controle por parte do Estado.

Ainda seriam esperados oito anos para que a Lei de Diretrizes e Bases da Educação Nacional (LDB) esclarecesse a organização da educação no país, as incumbências de cada esfera governamental, a composição dos sistemas de ensino (federal, estaduais e municipais),

além de pronunciar-se sobre outros tantos temas, entre eles, sobre avaliação. O Projeto de autoria do parlamentar Darcy Ribeiro ocupou-se da avaliação nos diferentes níveis escolares, com dispositivos e normas que indicavam uma preocupação com o processo de aprendizagem. Mas, em consonância com a Constituição, que garantia padrão de qualidade como princípio de ensino, o Projeto pronunciou-se sobre o processo nacional de avaliação e ficou garantido, no Artigo 9º da Lei de Diretrizes e Bases da Educação Nacional (Lei 9394/ 1996), que a “União incumbir-se-á de: [...] VI – assegurar processo nacional de avaliação do rendimento escolar do ensino fundamental, médio e superior, em colaboração com os sistemas de ensino, objetivando a definição de prioridades e a melhoria da qualidade de ensino” (BRASIL, 1996). No entanto, anos antes da promulgação da LDB, as iniciativas de avaliações em escala nacional da educação básica já estavam em andamento e em processo de consolidação.

### **3.2 Os oito primeiros ciclos do Saeb – 1990 a 2003**

Vianna (1995) narra que, após a experiência de avaliação nacional realizada em 1987, no ano de 1988, por meio da Secretaria de Ensino Fundamental, o Ministério da Educação solicitou ao grupo de avaliadores que atuara por iniciativa do Inep, que uma sistemática de avaliação fosse implementada no país. Assim, foi apresentada a proposta de criação do Sistema Nacional de Avaliação das Escolas Públicas de 1º Grau (Saep). Dirce Freitas (2013) sinaliza que essa “proposta previa organização, articulação, complementaridade e direção dos estudos de avaliação até então realizados, e pretendia organizar nesse sentido as iniciativas futuras.” (FREITAS, D., 2013). Em função da falta de recursos, o Sistema só foi implementado em 1990, e foi então realizada a primeira aferição nacional (MALUF, 1996). Vale notar que o Saep surgiu como um programa do Ministério da Educação, sem que houvesse uma normativa legal que o criasse.

O então coordenador de avaliação do Inep, Orlando Pilati (2015), descreveu que, nessa primeira aferição de 1990, foi utilizado o instrumental elaborado pela Fundação Carlos Chagas na pesquisa *Avaliação do Rendimento de Alunos de Escolas de 1º grau da Rede Pública*. Para a elaboração dos instrumentos, professores da rede oficial colaboraram e procuraram definir os elementos principais e indispensáveis numa sequência curricular, já que não havia um currículo supostamente ideal. Assim, assegurar-se-ia que os testes aferissem as propostas curriculares das Unidades Federadas.

Para a aplicação dessa primeira aferição, Horta Neto (2018) descreve que foi organizado um encontro, com a duração de duas semanas, com servidores de cada uma das secretarias estaduais de Educação, com o propósito de capacitação e discussão dos instrumentos. A participação dessas secretarias era fundamental para que o sistema, de fato, fosse implementado. Na primeira aplicação, apenas os estados de Mato Grosso do Sul e Piauí não participaram, devido a uma greve de professores. Mas nos anos seguintes, ambos os estados realizaram a aplicação por conta própria. Pestana (2016) afirma que, nessa fase inicial, com o novo contorno do pacto federativo, a efetiva contribuição das administrações estaduais de Educação, com a participação em termos técnicos, operacionais e financeiros, fez-se indispensável. A autora afirma que:

Nessa interação, que previa divisões de responsabilidades, foi possível aprofundar o debate sobre o significado, os meios utilizados e as consequências do uso da avaliação como política de melhoria da qualidade da educação, além de aspectos técnicos do processo de avaliação. Como resultado, essa dinâmica auxiliou na formação de uma primeira leva de técnicos e especialistas aptos a operar o sistema que se desenhava. (PESTANA, 2016, p. 75).

Em um documento elaborado pela Coordenação de Avaliação Educacional no Inep (1994) intitulado *Sistema Nacional de Avaliação da Educação Básica: Objetivos, Diretrizes, Produtos e Resultados*, que traz informações a respeito dos primórdios do sistema de avaliação e um histórico das duas primeiras aferições (1990/ 1993), é-nos apresentado o marco da sistemática de avaliação estabelecido em sua proposta original:

- desenvolver e/ ou aprofundar as capacidades avaliativas das unidades gestoras do sistema educacional (o próprio Ministério, as Secretarias Estaduais e Órgãos Municipais de Educação), através da constituição progressiva de quadros especializados quanto aos diferentes aspectos da avaliação enquanto instrumento seja de planejamento ou monitorização, seja provocador de mudanças qualitativas e;
- regionalizar a operacionalização do processo avaliativo, de forma a possibilitar sua gestão diretamente pelas instâncias locais (equipes técnicas das Secretarias, Instituições de Pesquisa, Universidades etc.), criando nexos e estímulos para o desenvolvimento local de uma infraestrutura de pesquisa e avaliação educacional;
- propor uma estratégia metodológica em condições de articular e relacionar os resultados dos diversos trabalhos de pesquisa e avaliação já realizados ou em vias de implementação;
- conhecer e construir parâmetros do rendimento dos estudantes quanto aos aspectos cognitivos das propostas curriculares;
- identificar, nas propostas curriculares, os pontos que apresentam maiores deficiências quanto à formação dos professores ou que devem ser discutidos quanto ao momento mais adequado em que devem ser tratados tendo em vista o desenvolvimento intelectual das crianças. (INEP, 2004)

O mesmo documento também apresenta os três eixos sobre os quais o Saeb foi originalmente desenvolvido:

- (a) Democratização da Gestão - Até que ponto a gestão educacional está se tornando mais eficiente e democrática?

(b) Valorização do Magistério - Tem havido mudanças nas condições de trabalho e na competência pedagógica do professor?

(c) **Qualidade de Ensino** - 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? (INEP, 1994, grifos do original).

Para responder a essas questões, elas foram decompostas em dimensões e indicadores mais concretos. Com a efetiva distribuição de responsabilidades entre os entes federativos, o sistema proposto era construído sobre esses eixos de política educacional, sendo que o rendimento do aluno era um entre vários indicadores do eixo “Universalização com qualidade”, conforme se pode observar no Quadro 8.

**Quadro 8 - Eixos, dimensões e indicadores da avaliação da educação básica**

Eixo	Dimensão	Indicador
Universalização com qualidade	Escolarização	Matrícula
		Taxa de escolarização real
	Retenção	Taxas de evasão, aprovação e repetência
	Produtividade	Taxa de sobrevivência
	Qualidade	Rendimento do aluno Competência do professor Custo-aluno direto Condições do ensino
Valorização do professor	Competência	Conteúdos e projeto pedagógico
	Condições de Trabalho	Níveis salariais
		Taxas de professores concursados Vigência de Estatuto do Magistério
Gestão Educacional	Integração estado/município	Diagnóstico de gestão
		Equalização custo aluno
	Descentralização	Diagnóstico de Gestão
		Grau de autonomia da escola
		Melhoria da gestão escolar
	Desburocratização	Diagnóstico de gestão
	Eficiência	Relação custo docente/não docente
Diagnóstico de gestão		
Democratização	Racionalização de recursos	
		Diagnóstico de gestão
		Melhoria da gestão escolar

Fonte: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (1992, p.3).

Na proposta original do Saep (1988), ficou demonstrado que a avaliação iria muito além do rendimento do aluno, incidindo também sobre a gestão escolar, a situação e a competência dos professores e o custo-aluno direto. Dirce Freitas (2013) afirma que:

Com esta proposta, a ação avaliativa nacional visava estimular o desenvolvimento de condições e capacidades subnacionais para acionar, articulada e inter-

relacionadamente, a pesquisa e a avaliação na política e gestão do ensino fundamental. Seria tarefa federativa cuja organicidade, promovida pelo governo federal, preservaria as autonomias dos entes federativos. (FREITAS, D., 2013, p. 82).

Pode-se aquilatar o desafio que se desenhava em um sistema de avaliação que previa 23 indicadores distribuídos em 11 dimensões, sendo que a equipe da Fundação Carlos Chagas desenvolveu apenas o instrumento que mediria o rendimento do aluno. Como afirma Horta Neto (2018), numa época em que a internet engatinhava no Brasil, os recursos computacionais do ministério e das secretarias estaduais eram praticamente inexistentes, e os dados estatísticos coletados demoravam anos para serem concluídos. Além desses fatores, havia a falta de recursos que fez com que o plano original da avaliação acontecer a cada dois anos não pudesse ser cumprido em 1992.

Na primeira aplicação, em 1990, os alunos participantes foram selecionados aleatoriamente a partir das listas de frequência das turmas da amostra probabilística de alunos das turmas de 1ª, 3ª, 5ª e 7ª séries do ensino fundamental, por meio de sorteio. As provas abrangiam conteúdos selecionados como comuns aos currículos das secretarias de Educação dos estados, nas áreas de Português, Matemática e Ciências. Na ocasião, os alunos do 1º e 3º ano realizaram as provas semiobjetivas de Português e Matemática, e os de 5º e 7º ano, provas objetivas de Português, Matemática, Ciências e Redação. As provas do 1º e 3º ano possuíam 30 questões, cada uma, e cada aluno selecionado só realizava o teste para uma disciplina. Para as turmas de 5º e 7º ano, as provas também continham 30 questões cada, e o aluno que fizesse a prova de Português também faria a redação. As correções seguiam o modelo clássico e as provas semiobjetivas (Português e Matemática da 1ª e 3ª séries), assim como as provas de redação da 5ª e 7ª séries, foram corrigidas por equipes de professores e especialistas das secretarias estaduais de Educação, seguindo instruções contidas em grades de correção por série e disciplina, especialmente preparadas pela equipe central do projeto. As provas objetivas (Português, Matemática e Ciências da 5ª e 7ª séries) foram corrigidas via computador, para o que foi elaborado, dentro do programa Saeb, um módulo específico de correção (INEP, 1994). A Tabela 1 sintetiza os dados da primeira aplicação.

**Tabela 1- Primeira aplicação do Saeb -1990**

Abrangência Escolas Públicas Amostrais	Série/EF	Disciplinas	Instrumento	Nº de questões
	1ª	Português	Prova semi-objetiva	30
		Matemática	Prova semi-objetiva	30
		Português	Prova semi-objetiva	30

	3°	Matemática	Prova semi-objetiva	30
	5°	Português e redação	Objetiva	30
		Matemática	Objetiva	30
		Ciências	Objetiva	30
	7°	Português e Redação	Objetiva	30
		Matemática	Objetiva	30
		Ciências	Objetiva	30

Elaborada pela autora.

Os dados para a caracterização da gestão escolar foram levantados por meio de questionário que envolveu 50 itens fechados, que abordavam inúmeros aspectos ligados à dimensão pedagógica da função administrativa do ensino e ao relacionamento interno e externo da unidade escolar. Os aspectos verificados foram:

- dos serviços educacionais oferecidos pela escola;
- das formas de gestão e participação na escola;
- dos níveis de autonomia da escola;
- da situação da infraestrutura física (prédios, equipamentos etc.);
- do aproveitamento do “tempo -pedagógico”;
- das estratégias de planejamento, execução e avaliação dos planos da escola;
- das formas de articulação da escola com o seu meio;
- dos processos de ensino-aprendizagem preconizados;
- das consequências pedagógicas (rendimento do aluno, taxa de aprovação, evasão, repetência etc.), relacionadas a diferentes estilos de gestão. (INEP, 1994).

Os dados foram organizados com a finalidade de delinear a figura do diretor, a organização e funcionamento da escola, assim como as práticas da gestão nas escolas que atendessem ao critério de possuir mais de uma sala de aula em funcionamento. O estudo sobre a atuação docente também contou com a aplicação de questionário com 53 itens fechados, apresentado junto com os demais instrumentos da escola, com a finalidade de investigar a situação funcional, o papel e as competências docentes, assim como de captar alguns aspectos das representações dos professores do ensino público a respeito de sua prática docente. Para tanto, os elaboradores do instrumento levaram em conta o quadro de referência analítico de desenvolvimento histórico do perfil docente no Brasil (tradicional, escolanovista, tecnicista, reprodutivista, crítico). Além de levantarem aspectos do ideário pedagógico do professor brasileiro, foram coletados dados básicos de sua formação, experiência, condições de trabalho e carreira e sua participação na gestão escolar. As principais linhas abordadas no questionário para os docentes foram:

- níveis de formação e capacitação docente;
- condições de trabalho e experiência profissional;
- percepção do “papel” do professor;
- formas de utilização do “tempo pedagógico”;
- materiais e estratégias de ensino utilizados;

- principais problemas no ensino;
- níveis de participação na gestão escolar;
- resultados pedagógicos (níveis de aprendizagem, evasão, repetência, etc.) de diferentes estratégias e perfis docentes. (INEP, 1994).

O estudo dos dados levantados nos questionários “tentou uma primeira aproximação do perfil qualitativo da prática docente dos professores: quem são, como atuam e qual a eficácia do seu trabalho”. (INEP, 1994). Esta iniciativa, contudo, restringiu-se à primeira aplicação.

Em 1991, os dados levantados na primeira aferição foram processados e, visando à expansão da avaliação para o segundo grau, o sistema passou a ser denominado Sistema Nacional de Avaliação da Educação Básica (Saeb) (INEP, 1994). A segunda aplicação ocorreu em 1993 com o financiamento dos estados, pois o MEC não possuía orçamento para a operacionalização da avaliação.

Com os resultados obtidos na primeira aplicação, as equipes das secretarias estaduais solicitaram algumas reformulações nos questionários de docentes e gestores, bem como nos testes de rendimento dos alunos. Assim, na segunda aplicação, os questionários dos gestores e dos docentes foram simplificados, abordando os mesmos aspectos da primeira aferição e privilegiando aqueles relacionados a perfil e à prática profissional. Já os testes de rendimento passaram por um processo bem mais profundo de modificações. Com a ajuda de equipes técnicas da secretaria de Educação de vários estados e professores de universidades, foram montados grupos de trabalho para atuarem nos vários estágios do processo, desde a escolha dos conteúdos essenciais das áreas, a elaboração de um banco de questões, até a versão final de um instrumento que pudesse ser utilizado em todo o território nacional. Como resultado, os testes de Português e Matemática de 1ª e 3ª séries passaram a ter 20 questões semiobjetivas cada um, e os testes de 5ª e 7ª séries mantiveram as 30 questões para cada área: Português, Matemática e Ciências (INEP, 1994).

Numa análise dos documentos oficiais relacionados especificamente aos instrumentos contextuais do Saeb, Bonamino (2016, p. 117) observa que os questionários dos dois primeiros ciclos “enfaticavam as condições intraescolares intervenientes no desempenho do aluno” sem levar em conta fatores sociodemográficos e culturais dos estudantes, privilegiando as questões ligadas ao perfil e às práticas de docentes e gestores. Esses instrumentos passaram por várias alterações ao longo dos ciclos de avaliação, na medida em que se sentiu a necessidade da adoção de uma abordagem mais sociológica e a preocupação de explicitar as

referências teóricas, conceituais e metodológicas que orientaram o processo de sua elaboração.

Dirce Freitas, em uma análise das características e pressupostos da avaliação da educação básica no Brasil (FREITAS, D., 2013), afirma que na proposta do Saep (1988) os objetivos declarados indicavam um realce à democratização, descentralização e autonomia da política para o ensino fundamental. A ação avaliativa visava estimular o desenvolvimento de condições e capacidades subnacionais “para acionar, articulada e inter-relacionadamente, a pesquisa e a avaliação na política e gestão do ensino fundamental” (p. 82). O governo federal preservaria a autonomia dos entes federativos no modelo proposto originalmente. Contudo, segundo a autora, esses objetivos não se concretizaram:

Com dois ciclos de levantamentos efetuados (1990 e 1993), o SAEB trazia objetivos baseados no pressuposto de que o governo federal deveria robustecer sua capacidade de controle dos sistemas educacionais, por meio da avaliação externa centralizada e focada no desempenho cognitivo dos estudantes, privilegiando os resultados dos sistemas educacionais. (FREITAS, D., 2013, p. 82).

Em 27 de dezembro de 1994, o ministro da Educação Murilo de Avellar Hingel, nos últimos dias do governo do presidente Itamar Franco, cria oficialmente o Sistema Nacional da Educação Básica por meio da edição da Portaria nº 1.795, que, além de apresentar os objetivos do sistema de avaliação, estabeleceu a formação de um conselho diretor de composição diversificada, com a ampla participação dos entes federativos e de entidades representativas. A principal competência desse conselho seria elaborar o Plano Nacional de Avaliação da Educação Básica. Horta Neto (2018) afirma que essa portaria garantia o caráter descentralizado e participativo que norteou o desenho inicial do Saeb. Contudo, Dirce Freitas analisou que os objetivos declarados na portaria deixaram claro que “a descentralização de decisões e ações e o estímulo às autonomias subnacionais não seria a tônica da avaliação nacional da educação básica. Nem tampouco a qualidade do ensino seria considerada de forma a abranger condições, processos e resultados.” (FREITAS, D., 2013, p. 82).

Os objetivos gerais estabelecidos no Artigo 1º da Portaria MEC nº 1.795, de 1994 são:

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 aferição 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 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)

Castro (2016) ressalta que, na construção da agenda de reformas educacionais, nesse período, também foi importante o debate promovido no Brasil após a I Conferência Mundial de Educação para Todos, ocorrida, em 1990, na cidade de Jomtiem, na Tailândia. O Brasil foi um dos signatários da Declaração Mundial sobre Educação para Todos, que reconhecia a educação de qualidade como direito fundamental de todos e o seu papel para a construção de uma sociedade mais justa. O compromisso brasileiro foi firmado em 1994 com a realização da primeira Conferência Nacional de Educação para Todos, quando definiram-se metas para serem incorporadas no Plano Decenal de Educação para Todos (1994-2004). Rabelo (2013) descreve que esse plano consistia, basicamente, em “um conjunto de diretrizes políticas voltado para a recuperação da escola fundamental no país.”. Portanto, é nesse contexto que o Saeb se estabelece, tendo como um de seus objetivos, conhecer o sistema educacional brasileiro em profundidade (RABELO, 2013, p. 3).

No ano de 1995, o Saeb passa por uma significativa reestruturação de natureza metodológica. Da Teoria Clássica de avaliação, adotada nas duas primeiras aplicações, incorporaram-se as técnicas estatísticas da Teoria de Resposta ao Item<sup>16</sup> (TRI), “exatamente pela possibilidade que essa teoria trazia de comparação dos desempenhos dos estudantes ao longo dos anos”, afirma Rabelo (2013, p. 3). Outras mudanças ou processos foram introduzidos a partir desse ano. Castro (2009) e Dirce Freitas (FREITAS, D., 2013) identificam como avanços do período, a construção de Escalas Comuns de Proficiência, com a definição de métricas interpretadas em termos do que os alunos conhecem, compreendem e são capazes de fazer. Essas escalas interpretam os resultados do desempenho a partir de uma ampla cobertura dos conteúdos, competências e habilidades em diferentes graus de complexidade e nas diferentes séries avaliadas. Os itens das provas passaram a ser elaborados com base numa Matriz de Referência Curricular, que foi construída a partir de ampla consulta nacional a partir dos currículos estaduais, livros didáticos adotados nas séries avaliadas e conteúdos trabalhados efetivamente nas escolas brasileiras.

Na edição do Saeb de 1995, todas as unidades da federação participaram; a avaliação incluiu em sua amostra o ensino médio, e a rede privada de ensino passou a participar de

---

<sup>16</sup> Para Klein (2009, p. 127), a TRI “é um conjunto de modelos matemáticos onde a probabilidade de respostas a um item é modelada como função da proficiência (habilidade) do aluno (variável latente, não observável) e de parâmetros que expressam certas propriedades dos itens.”

forma voluntária. Pestana (1998) e Castro (2016) apontam que houve, nesse ano, uma redefinição das séries que seriam avaliadas, adotando-se as séries finais de ciclos avaliativos organizados em três: as quatro primeiras séries do ensino fundamental (1ª a 4ª); as outras quatro séries (5ª a 8ª) e o ensino médio. Essa informação, contudo, difere dos dados disponíveis no portal do Inep, que apresenta ainda em 1995, a avaliação das turmas de 1ª, 3ª, 5ª e 7ª séries, como nas aplicações anteriores e apenas para as escolas públicas.

Porém, as modificações produzidas com a incorporação da TRI e a escala de proficiência para Língua Portuguesa e Matemática, na visão de Brooke, Alves e Oliveira (2015), fizeram com que o Saeb “começasse a assumir características atuais de um sistema de monitoramento do Ensino Fundamental e Médio, capaz de oferecer informações seguras sobre avanços nos repertórios de aprendizagem das crianças e jovens brasileiros.” (BROOKE; ALVES; OLIVEIRA, 2015, p. 85).

Horta Neto (2018) avalia, contudo, que as alterações promovidas em 1995, com o aprimoramento dos instrumentos utilizados e a criação de Matrizes de Referência, que resultaram num “novo desenho para o Saeb” (p. 44), apequenou a avaliação da educação brasileira, quando comparada com o projeto original, pois concentrou-se apenas nas dimensões dos testes cognitivos e questionários contextuais – que coletavam dados sobre os alunos, como a caracterização socioeconômica, cultural e de hábitos de estudos; dos professores, sobre perfil e práticas pedagógicas; e dos diretores, sobre perfil e práticas de gestão. Semelhantemente, Bonamino e Franco (1999) constataram que, nesse terceiro ciclo de avaliação, houve um abandono de uma perspectiva mais participativa de professores e técnicos dos sistemas de ensino – que ocorrera nos dois ciclos anteriores –, tendo concorrido para um modelo de terceirização da elaboração, aplicação e análise dos resultados por instituições privadas.

Vale ressaltar que, em 1995, assume a presidência do Brasil, Fernando Henrique Cardoso e, desse período, Horta Neto (2007) destaca a elaboração do Plano Diretor da Reforma do Aparelho do Estado:

Neste Plano estão indicadas algumas diretrizes que influenciaram determinantemente o processo educacional brasileiro. Dentre elas, podemos citar a descentralização, como uma nova visão para a gestão das instituições públicas, balizadas pelos princípios de eficiência, autonomia e participação social; esta última considerada como fator de “melhoria da qualidade e eficiência na prestação de serviços pelo setor público”. (Brasil/ Mare, 1995a, p. 5) (HORTA NETO, 2007, p. 3).

Esse Plano preconizava a relevância da qualidade dos serviços prestados pelo Estado e a elaboração de sistemática de avaliação de cada órgão da administração pública. Assim, como bem caracteriza Dias Sobrinho, “a avaliação como estratégia de governo se fortalece sobretudo em virtude da adesão dos governos brasileiros ao neoliberalismo<sup>17</sup>”. Na análise desse autor, “desde 1995, o Ministério de Educação passou a dedicar grande importância à avaliação como controle e instrumentalização de suas políticas reformistas” (DIAS SOBRINHO, 2003, p. 75).

Ainda sobre esse período e sobre essa política instaurada, Dias Sobrinho e Ristoff (2003, p. 38) afirmam:

Esta conjuntura pode ser bem definida pela expressão ‘Estado Avaliador’, que caracteriza o estado forte, no controle do campo social, e liberal relativamente à economia. Segundo esta lógica, o estado confere maior liberdade à gestão dos meios e processos, como fator de eficácia e produtividade, podendo assim justificar a diminuição dos financiamentos e criar as facilidades para expansão da privatização e da mentalidade empresarial em educação. Por outro lado, exerce um rígido controle sobre os fins e produtos, através de mecanismos que chama de avaliação, para consolidar os modelos desejados e orientar o mercado.

Destaca-se que o modelo de avaliação adotado como política pública pelo governo do presidente Fernando Henrique Cardoso não se constituiu como fenômeno isolado. Afonso (2000), em uma análise sociológica da avaliação, insere o “Estado Avaliador” numa conjuntura mundial de competitividade internacional, diante da necessidade de manter e criar padrões de inovação científica e tecnológica, de concorrência, de livre mercado e globalização. Assim, uma cultura gerencialista se instala no setor público, acompanhada de mecanismos de controle e responsabilização, tanto em países de governos neoliberais como em países de governos socialistas<sup>18</sup>.

Nesse contexto, em 1996, foi promulgada a Lei de Diretrizes e Bases – Lei 9.394 –, acerca da qual, Horta Neto (2007) fez a consideração de que, “pela primeira vez a legislação refere-se à coleta de informações e à avaliação com o objetivo de atuar sobre a qualidade do ensino”. A União, em colaboração com os sistemas de ensino, deve assegurar um processo nacional de avaliação de rendimento escolar com o objetivo de definir prioridades para melhorar a qualidade do ensino. Pestana (2016) acrescenta que, além da avaliação tornar-se uma atribuição da União, “passa também a ser uma política do MEC, que começa a

---

<sup>17</sup> Dias Sobrinho e Ristoff (2003, p.38) explicam o conceito de neoliberalismo como o modelo de governo que flexibiliza os meios de produção para obter mais ganhos de eficiência, respeitando o princípio da democracia liberal, favorecendo a liberdade de escolha visando a maior eficiência, diversificação e diferenciação, favorecendo a privatização.

<sup>18</sup> Afonso (2000) cita o exemplo da França como país de governo socialista que adotou a avaliação das políticas públicas com impulso importante no período.

diversificar e ampliar o sistema nacional de avaliação, com a criação do Exame Nacional de Cursos – Provão – em 1996 e do Enem em 1998” (PESTANA, 2016, p. 79).

A quarta aferição do Saeb, realizada em 1997, pode-se dizer que manteve as características processuais presentes no ciclo de 1995. Nesse ano, porém, por meio do Decreto nº 2.146 de 1997, o ministro Paulo Renato de Souza concedeu ao Inep uma nova estrutura e novas responsabilidades, conferindo ao Instituto, entre outras, a atribuição de “planejar, orientar e coordenar o desenvolvimento de sistemas e projetos de avaliação educacional, visando o estabelecimento de indicadores de desempenho das atividades de ensino no país.” (BRASIL, 1997). Para Horta Neto (2007), há uma mudança perceptível do objetivo central do Inep, deixando de concentrar-se nas atividades de pesquisa, priorizando, a partir de então, as atividades de avaliação da educação. Sobre esse período, Pestana (2016, p 78) afirma que:

Quanto ao arranjo institucional, verificou-se o aprofundamento da centralização das atribuições relativas ao sistema de avaliação no Inep e a gradativa diminuição da participação de estados e municípios na tomada de decisão ou discussão de aspectos técnicos, operacionais ou políticos, sendo as equipes estaduais e municipais apenas eventualmente mobilizadas. As instâncias políticas do Conselho Nacional de Secretários de Educação (Consed) e da União Nacional dos Dirigentes Municipais de Educação (Undime) tendem a se manifestar ou ser convocadas somente nos momentos de divulgação dos resultados.

Sobre esse quarto ciclo, Bonamino e Franco (1999) observaram que o processo foi delegado a um único órgão, em todas as fases, e assim descrevem:

Do ponto de vista do processo, o ciclo de 1997 do SAEB se caracteriza por ter concentrado numa única fundação grande parte do processo de formulação dos itens e o conjunto de etapas concernentes à aplicação das provas e à análise dos resultados. Foi essa fundação quem se incumbiu da elaboração dos itens Língua Portuguesa, Matemática e Ciências, que passaram a compor o banco de itens do MEC, responsabilizando-se também pela sua pré-testagem em dez capitais brasileiras. Entretanto, as provas foram preparadas pelo MEC, a partir do conjunto de itens pré-testados que faziam parte do banco. Com as provas já prontas, a Fundação CESGRANRIO se encarregou da preparação dos cadernos de testes e de sua aplicação a uma amostra de mais de 200 mil alunos definida pelo MEC, da preparação dos coordenadores estaduais e da análise dos resultados obtidos com a aplicação dos testes. (BONAMINO, FRANCO, 1999, p.119, 120)

Após a análise das quatro primeiras aplicações do Saeb, esses autores concluíram que houve uma mudança significativa dos objetivos do Sistema de Avaliação que preconizava, originalmente, a criação de uma “cultura de avaliação no âmbito dos órgãos gestores das redes de ensino”, migrando para a “prioridade em monitorar as políticas e a qualidade da educação”. Associam essa mudança de objetivo à interferência direta dos técnicos do Banco Mundial no financiamento e na indução da política avaliativa (BONAMINO; FRANCO, 1999, p. 126).

Ainda que a metodologia adotada no ciclo de 1995 buscasse a garantia da comparabilidade entre os vários ciclos de aplicação, algumas mudanças foram efetuadas no público-alvo de 1997 que sugeririam cautela sobre uma análise comparativa mais apurada dos resultados obtidos pelos alunos nos testes cognitivos. Nessa edição, conforme descrito no relatório dos resultados do Saeb publicado pelo Inep/ MEC (1997), “foram excluídos do universo os alunos de turmas multisseriadas e das escolas federais, os estudantes das escolas rurais da Região Norte e também os alunos da 3ª série do ensino médio matriculados em cursos profissionalizantes.”. Houve, portanto, alguma alteração das características da população avaliada. Também nesse ciclo, as Matrizes de Referência foram elaboradas com uma ampla consulta nacional sobre os conteúdos praticados nas escolas brasileiras de ensino fundamental e médio, incorporando a análise de professores, pesquisadores e especialistas sobre a produção científica em cada área que se tornou objeto de conhecimento escolar (BRASIL, 1997).

Sobre esse ciclo, Horta Neto (2007, p. 10) afirma que,

O ciclo de 1997 do SAEB pode ser considerado como o passo decisivo para a institucionalização da avaliação da educação básica no Brasil. A partir deste ciclo, poucas foram as modificações adotadas aos ciclos seguintes. Como inovação, ele trouxe a construção de uma escala de proficiência única para cada disciplina avaliada, permitindo assim a comparação entre os resultados obtidos pelo ciclo de 1995 do SAEB e abrindo a possibilidade de se fazer o mesmo com os próximos ciclos.

O ciclo de 1999 teve como marco normativo a Portaria MEC nº 839, publicada pelo ministro Paulo Renato de Souza, criando o Sistema Nacional de Avaliação da Educação Básica (Saeb), com o “objetivo geral de monitorar a qualidade do ensino, a equidade e a eficiência do sistema de educação básica”; revogando a Portaria 1.795/94 e dando nova estrutura ao modelo de avaliação, conforme descrevem Santos, Horta Neto e Junqueira (2017, p. 14):

Em 1999, a Portaria MEC nº 839 (Brasil. MEC, 1999) deu nova estrutura ao Saeb, instituindo, em lugar de um modelo de governança colaborativa, um Comitê Consultivo do Sistema de Avaliação da Educação Básica (ComSaeb), com finalidade de dar apoio ao Inep no exercício de suas competências relacionadas ao desenvolvimento do sistema. Essa alteração ratificou algo já praticado anteriormente, pois os instrumentos, a data de aplicação e a forma de divulgação de resultados já eram decididos pelo Inep. A proposta original de colaboração com os outros entes da Federação foi substituída pelo apoio logístico das secretarias estaduais à aplicação dos testes.

Dessa forma, o modelo idealizado originalmente no primeiro ciclo, com a participação efetiva dos três níveis dos entes federativos, foi substituído pelo modelo que predomina até os

dias atuais, com a coordenação, planejamento e operacionalização centralizados no Inep que, desde 1997, fora completamente reformulado, ampliado e transformado em autarquia capaz de fazer frente aos novos desafios. Castro (2016) ressalta que, nesse período, foi iniciada a construção do Banco Nacional de Itens baseado na Matriz de Referência que atenderia ao próprio Saeb e, posteriormente, viria apoiar os estados e municípios na criação dos sistemas de avaliação subnacionais.

A edição desse ciclo trouxe como alteração para os testes cognitivos a inclusão da disciplina Geografia. Além da área de Língua Portuguesa e Matemática, presente em todos os ciclos, algumas disciplinas foram acrescentadas e, mais adiante, retiradas ao longo dos demais ciclos de aplicação. A partir de 2001, Língua Portuguesa e Matemática passaram a ser as únicas áreas avaliadas até a edição de 2013, que incluiu Ciências, em caráter experimental, até nova alteração ocorrida em 2019.

Quanto aos instrumentos contextuais, Bonamino (2016) analisou que, a partir do ciclo de 1999, foi reconhecida a necessidade do enriquecimento com medidas socioeconômicas, tais como indicadores de renda dos pais e bens e serviços disponíveis nas residências dos alunos, propiciando análises do desempenho cognitivo relacionado com fatores escolares e extraescolares. O Inep passou a publicar relatórios detalhados com o perfil das escolas, “utilizando informações coletadas nos questionários do diretor, do professor e do estado de conservação dos estabelecimentos de ensino.” (BONAMINO, 2016, p. 118).

Contudo, é no ciclo de 2001 que o Saeb define claramente o objetivo de adotar, além de medidas cognitivas, medidas contextuais que aperfeiçoem a coleta de dados sobre os alunos por meio de variáveis relacionadas com a caracterização sociodemográfica; o capital cultural; o capital social; a motivação e a autoestima; as práticas de estudo e a trajetória escolar. Também foi incluído um novo item para mensurar a ocupação dos pais ou responsáveis (INEP, 2002, p. 8). Da mesma forma, essa perspectiva sociológica foi estendida ao questionário dos professores e dos diretores. Segundo o relatório do Inep, de 2002, , com essa configuração dos instrumentos contextuais, pretendia-se:

[...] obter medidas que permitam captar, além do que os alunos aprendem, como os resultados educacionais se distribuem tendo em vista: i) a origem social dos alunos; ii) como evoluem as condições de qualidade da escolarização; iii) como as condições de qualidade da escolarização se distribuem em função da origem social dos alunos; e iv) quais fatores escolares promovem eficácia e equidade na educação brasileira. (BRASIL, 2002, p. 8).

Esse esforço em aproximar-se dos pressupostos teóricos da sociologia da educação estendeu-se até o ciclo de 2003, tendo sido incorporado nos questionários contextuais

aspectos relacionados à discriminação racial, aos alunos beneficiados pelo Programa Bolsa Escola<sup>19</sup> e ao perfil de estudantes trabalhadores, procurando estabelecer relações desses aspectos com os resultados dos testes cognitivos. Nesse período, o Inep divulgou os resultados do Saeb 2001 com uma nova forma de apresentar os resultados da proficiência dos alunos, utilizando uma classificação em estágios: muito crítico, crítico, intermediário, adequado e avançado (BRASIL, 2003).

Dessa forma, o ciclo de 2003 manteve as características do ciclo de 2001. Bonamino (2016) afirma que até essa edição, o Saeb adotou uma abordagem sociológica do desempenho escolar que se refletiu nos questionários contextuais do aluno, do professor e do diretor, demonstrados nos relatórios publicados pelo Inep. Porém, para a autora, houve um “abandono paulatino da perspectiva sociológica que tinha pautado as experiências iniciais de avaliação em larga escala” (p. 120) nos ciclos posteriores, e os questionários não mantiveram essas características, sendo a ênfase passou a ser conferida aos resultados cognitivos da avaliação.

Enfim, o Saeb foi gestado e teve sua primeira edição no governo do presidente Fernando Collor de Melo, nos anos 1990/1991. Sua segunda edição deu-se no governo de Itamar Franco, tendo sido oficialmente criado no apagar das luzes dessa administração. Consolidou-se como política educacional no governo de Fernando Henrique Cardoso, que institucionalizou a avaliação da educação básica e fortaleceu a estrutura organizacional que garantiu a continuidade do modelo após seu mandato.

Em 2003, com a eleição do presidente Luiz Inácio Lula da Silva, criou-se uma expectativa de mudanças significativas nas políticas educacionais. Castro (2016) cita os seminários organizados pelo Inep, em abril de 2003, sob o título “Avaliar para quê?”, como iniciativas de desmontar os sistemas de avaliação, dentro das quais se questionava a existência e discutia-se a extinção do Prova, Saeb e Enem. A autora afirma que, apesar dessas tentativas, os sistemas sobreviveram e foram ampliados.

Semelhantemente, Pestana (2016), comentando essa transição de perfil de governo e o seminário organizado pelo Inep, cita uma declaração do então presidente do Inep, Otaviano Helene, de que se iria

debater os processos de avaliação e adequá-los à situação de um novo governo que tem um compromisso social com a área da educação, diferente do governo anterior, que tinha compromisso com um perfil liberalizante ou mercantilista de educação. (SEMINÁRIO 2003, *apud* PESTANA, 2016, p. 79).

---

<sup>19</sup> Programa Nacional de Renda Mínima vinculada à Educação – “Bolsa Escola”. Criado no Governo de Fernando Henrique Cardoso por meio da Lei nº 10.219, de 11 de abril de 2001.

Brooke, Alves e Oliveira (2015), sobre a permanência e consolidação de uma política de avaliação educacional que transcendeu a dois modelos de governos de matrizes distintas, concluíram:

A retomada da expansão e consolidação dos sistemas estaduais ao longo dos últimos dez anos reflete um ambiente novo em que a avaliação se naturalizou, se tornando elemento indispensável da administração educacional. A continuidade dada ao SAEB e ao Enem pelo governo federal durante o primeiro mandato do presidente Lula e a posterior criação da Prova Brasil, do IDEB e da ANA foram fundamentais para a criação desse ambiente suprapartidário que permitiu a despolitização da avaliação e sua gradual incorporação por todos os governos, independente do seu lugar no espectro político. Ou seja, a adoção da avaliação por um governo popular atenuou sua identidade como política educacional neoliberal e amainou as resistências. (BROOKE, ALVES, OLIVEIRA, 2015, p. 574).

Esses autores ainda consideram que “difícilmente, a avaliação teria avançado com tamanha força, patrocinada por governos de tão distintas colorações, se não oferecesse informações valiosas para a gestão de todos os sistemas educacionais.” (p. 574). Contudo, havendo se afastado da concepção original, quando criado, de regionalizar a operacionalização do processo avaliativo e a articulação progressiva entre os níveis de gestão, como pode-se afirmar que “oferece informações valiosas para a gestão”? Quais informações valiosas seriam estas? Os autores reconhecem que “é difícil encontrar evidência pela conexão entre a avaliação e a melhoria dos processos educacionais, fora as declarações dos órgãos governamentais responsáveis” (p. 574). Estas evidências necessitam ser encontradas por meio de estudos e pesquisas que possibilitem identificar as razões da continuidade e transcendência das avaliações externas.

Com os objetivos sendo alterados ao longo do tempo e com mudanças significativas nos anos seguintes, o Saeb concluiu seu oitavo ciclo de aplicação e mais de uma década de existência com algumas características preservadas até então, conforme demonstra a Tabela 2.

**Tabela 2 - Evolução dos ciclos do Saeb – 1990 a 2003**

PARTICIPANTES	ANO DE REALIZAÇÃO						
	1990	1993	1995	1997	1999	2001	2003
ESCOLAS	4.790	3.111	2.839	1.933	6.890	6.935	6.270
DIRETORES	3.546	3.111	2.839	1.933	6.890	6.820	6.000
DOCENTES	17.814	8.756	4.967	18.077	53.815	21.754	17.000
ALUNOS	108.875	179.527	90.499	167.196	279.764	287.719	275.081
SÉRIES AVALIADAS	1ª 3ª 5ª7ª	1ª3ª5ª7ª	4ª8ª2º3º	4ª8ª3º	4ª8ª3º	4ª8ª3º	4ª 8ª 3º

<b>DISCIPLINAS</b>	Port. Mat. Cie. Red.	Port. Mat. Cie.	Port Mat.	Port. Mat. Cie*.	Port. Mat. Geo. Cie*. Hist.	Port. Mat.	Port. Mat
--------------------	----------------------------	--------------------	--------------	------------------------	--------------------------------------	---------------	--------------

Fonte: Elaboração da autora com base em dados dos relatórios Brasil, Inep, 1994, 2002, 2004.

\*No 3º ano do ensino médio, Ciências corresponde a Biologia, Física e Química.

Bonamino e Souza (2012) consideram que, desde a sua criação, o Saeb, com suas características originais, configurou-se como uma avaliação com desenho apropriado para diagnosticar e monitorar a qualidade da educação básica, sem atribuição de consequências diretas para as escolas e para o currículo. Portanto, pode ser classificado como avaliação de primeira geração, quando não há uma política de responsabilização ou atribuição de consequências a agentes escolares. A forma de aplicação dos testes por base amostral, a divulgação dos resultados de forma agregada, sem possibilidade de medir o desempenho individual de alunos ou das escolas, são elementos que não permitiram a introdução de políticas de responsabilização.

### 3.3 O Saeb de 2005 à segunda década do século XXI

Durante os dois mandatos do presidente Luiz Inácio Lula da Silva – 2003 a 2010 –, a avaliação educacional se consolidou como política pública e, juntamente com outras pautas que caracterizaram a agenda educacional desse governo, avançou por meio de um robusto sistema de avaliação nos vários níveis de ensino. Pestana (2016) afirma que a utilização dos testes se ampliou nesse período, bem como a obrigatoriedade do ensino, a implantação da educação integral e o atendimento das demandas das minorias. As três últimas, como conquistas importantes e pautas históricas da esquerda brasileira. Para a autora, as mudanças ocorridas no sistema de avaliação durante esse período, permitindo as políticas de responsabilização, foram contraditórias ao ideário de um governo de esquerda.

No ano de 2005, a Portaria MEC nº 931 normatizou um novo modelo para o Saeb que, suprimida da sigla a palavra “nacional”, passa a se chamar Sistema de Avaliação da Educação Básica. 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). A portaria, nos seus artigos 1º e 2º, descreve da seguinte forma as características e objetivos das duas avaliações:

A ANEB manterá os objetivos, características e procedimentos da avaliação da educação básica efetuada até agora pelo SAEB realizado por meio de amostra da população, quais sejam:

- a) a ANEB tem como objetivo principal avaliar a qualidade, equidade e a eficiência da educação brasileira;
- b) caracteriza-se por ser uma avaliação por amostragem, de larga escala, externa aos sistemas de ensino público e particular, de periodicidade bianual;
- c) utiliza procedimentos metodológicos formais e científicos para coletar e sistematizar dados e produzir informações sobre o desempenho dos alunos do Ensino Fundamental e Médio, assim como as condições intra e extraescolares que incidem sobre o processo de ensino e aprendizagem;
- d) as informações produzidas pela ANEB fornecerão subsídios para a formulação de políticas públicas educacionais, com vistas à melhoria da qualidade da educação, e buscarão comparabilidade entre anos e séries escolares, permitindo, assim, a construção de séries históricas;

[...] A Avaliação Nacional do Rendimento Escolar – ANRESC tem 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 as políticas estabelecidas pelas diretrizes da educação nacional;
- d) oportunizar informações sistemáticas sobre as unidades escolares. Tais informações serão úteis para a escolha dos gestores da rede à qual pertencem. (BRASIL, 2005).

O Inep regulamentou a aplicação de ambas as avaliações por meio de duas portarias que definiram as regras para o ciclo de 2005. A Avaliação Nacional da Educação Básica (Aneb) manteve as características e objetivos do Saeb aplicado até 2003, garantindo a comparabilidade com as edições anteriores, e a Anresc – também chamada de Prova Brasil – passou a avaliar de forma censitária, nos primeiros ciclos, alunos do 5º e 9º anos do ensino fundamental apenas de escolas públicas localizadas em zona urbana com, pelo menos, 30 alunos matriculados em cada série ou ano. Nessa primeira aplicação, apenas os conhecimentos de Língua Portuguesa foram avaliados.

Para a edição de 2007, a Aneb não teve alteração, mantendo as características da edição de 2005. Já a Anresc foi regulamentada pela Portaria Inep nº 47, de 3 de maio de 2007, que, pela primeira vez, estabeleceu objetivos específicos para a Prova Brasil:

- I – aplicar instrumentos (provas e questionários) em escolas da rede pública de ensino localizadas na zona urbana, que possuam pelo menos 20 alunos matriculados em cada uma das turmas de 4ª e 8ª série 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 regular;
- II – aplicar testes de Matemática e de Língua Portuguesa com foco em resolução de problemas e leitura respectivamente, definidos nas Matrizes de Referência do Sistema de Avaliação da Educação Básica;

III – fornecer informações sobre as unidades escolares, que sejam úteis aos gestores da rede à qual pertencam as escolas avaliadas. (BRASIL, 2007)

Um aspecto importante destacado por Fernandes (2016) é que, mesmo sendo voluntária a participação das secretarias de Educação na Prova Brasil, desde a sua primeira edição, a adesão desses órgãos em nível estadual foi completa, e a das redes municipais ficou próxima da totalidade.

No ano de 2007, foi publicado o Decreto nº 6.094, de 24 de abril, que dispôs sobre o Plano de Metas Compromisso Todos pela Educação, no qual foi estabelecido um regime de colaboração entre os entes federativos visando à mobilização social pela melhoria da qualidade da educação básica, mediante programas e ações de assistência técnica e financeira. Entre outras providências, o Decreto criou o Índice de Desenvolvimento da Educação Básica (Ideb), explicitado no Artigo 3º, como segue:

A qualidade da educação básica será aferida, objetivamente, com base no IDEB, calculado e divulgado periodicamente pelo INEP, a partir dos dados sobre rendimento escolar, combinados com o desempenho dos alunos, constantes do censo escolar e do Sistema de Avaliação da Educação Básica – SAEB, composto pela Avaliação Nacional da Educação Básica – ANEB e a Avaliação Nacional do Rendimento Escolar (Prova Brasil). (BRASIL, 2007).

Reunindo dois conceitos – o fluxo escolar (taxa de aprovação, reprovação e abandono) e o desempenho dos alunos nos testes –, esse indicador é calculado com base nos dados fornecidos pelo Censo Escolar, pela Prova Brasil e pelo Saeb, possibilitando dispor de um valor para o país, para cada estado, cada município e cada escola.

Para Dirce Freitas (2013), o Ideb explicitou nova ênfase dada pelo governo federal à qualidade do ensino “cuja melhoria foi projetada em termos de metas bianuais e de patamar (mínimo) a ser alcançado até 2021 – a média nacional 6,0” (FREITAS, D., 2013, p.75). Na opinião da autora, operando apenas com os indicadores fluxo escolar e desempenho cognitivo, “o Ideb acaba contribuindo para que a melhoria da qualidade do ensino seja vista e enfrentada de forma estreita.”. Nessa linha, conclui que:

O IDEB possibilitou a projeção e o monitoramento de metas de progresso nos resultados do ensino fundamental e médio, tornando-se uma das principais ferramentas da política educacional brasileira na esteira da racionalidade econômica e gerencial que se impôs com as transformações sociais das últimas décadas. (FREITAS, D., 2013, p. 75).

Bonamino e Souza (2012) consideram que o princípio básico que orientou a formulação desse indicador é o de que a qualidade da educação envolve que o aluno aprenda e

passa de ano. Sendo que os resultados da Prova Brasil passaram a ser amplamente divulgados, e o Ideb passou a ser o principal indicador adotado para traçar metas educacionais, tanto pelo governo federal como pelas redes estaduais, municipais e pelas próprias escolas.

Para Fernandes e Gremaud (2009), ao combinar as notas da Prova Brasil/Saeb com as taxas de aprovação, o indicador coíbe tanto a reprovação indiscriminada como a prática de aprovar alunos que nada aprenderam. Justificam que o Ideb surgiu para contornar os riscos de agravar os problemas de fluxo que um sistema de *accountability*<sup>20</sup> tradicional poderia gerar em um país como o Brasil, no qual o modelo de descentralização da gestão da educação básica pode favorecer alguma distorção. Os autores argumentam que a principal justificativa para a criação do Ideb foi o não agravamento dos problemas de repetência e evasão diante de um sistema de avaliação externa que fosse focado apenas no desempenho dos estudantes. Apontam que o indicador também é utilizado para o estabelecimento de metas para redes e escolas e, assim, pode-se obter um maior comprometimento das secretarias de Educação de estados e municípios “para que, até 2021, o Brasil atinja o estágio educacional atual dos países desenvolvidos<sup>21</sup>” (FERNANDES; GREMAUD, 2009, p. 18).

Ainda sobre a fixação de metas pelas escolas, sobre o modelo de divulgação dos resultados que favorece a mobilização social e a ênfase do sistema de *accountability*, os autores concluem que,

O Brasil possui um sistema educacional descentralizado, com mais de cinco mil redes de ensino com autonomia para gerir suas escolas. A vantagem da descentralização é que ela tem maior potencial para gerar inovações. Mas para que todos se beneficiem é preciso que os casos de sucesso sejam identificados e difundidos. Por outro lado, a falta de resultados comparáveis entre redes pode levá-los ao isolamento. Um sistema nacional de avaliação possibilita que: i) inovações de sucesso sejam compartilhadas por todos; ii) o MEC identifique as redes em pior situação, orientando ações de assistência técnica e financeira; e iii) a opinião pública local tenha elementos para cobrar seus representantes por melhorias na educação. Enfim, o plano de metas Compromisso Todos pela Educação é um sistema de *accountability* compatível com o federalismo existente no país. (FERNANDES; GREMAUD, 2009, p. 18).

É importante ressaltar que a divulgação dos resultados do Ideb tem tido repercussão na mídia e, para Fernandes (2016) de modo geral a opinião pública tem demonstrado boa

<sup>20</sup> Este termo não tem palavra equivalente em português. Na educação, o *accountability* remete à integração entre avaliação, prestação de contas e responsabilização (AFONSO, 2012). Schneider & Nardi (2014), numa análise das concepções predominantes, propostas por vários autores, sintetizam que a despeito das alterações que o conceito de *accountability* vem sofrendo, a concepção matricial remete seu significado ao direito de alguém exigir prestação de contas e à obrigação ou dever, de quem é chamado a prestar contas.

<sup>21</sup> Fernandes (2016) descreveu que, para chegarem a uma referência de meta para 2021, após escutarem diversos especialistas, resolveu-se adotar um padrão externo a ser alcançado como média: o desempenho educacional observado para os países da Organização para a Cooperação e Desenvolvimento Econômico (OCDE). O cálculo levou em conta uma análise estatística de comparabilidade entre as notas do Brasil no Programa Internacional de Avaliação de Estudantes (PISA) e as notas da Prova Brasil de 2005. Dessa forma, foi possível projetar uma trajetória do Ideb ao longo do tempo, estabelecendo um desempenho desejável ao final do período.

aceitação ao sistema de avaliação. Contudo, inúmeros autores consagrados no campo da avaliação educacional fazem críticas ao modelo avaliativo brasileiro, especificamente à Prova Brasil e ao Ideb (AFONSO, 2018; BAUER, 2010; BONAMINO; SOUZA, 2012; COELHO, 2008; DOURADO; OLIVEIRA, 2009; DOURADO *et al.*, 2007; FERNANDES, 2018; HORTA NETO, 2018; MINHOTO, 2016; NARDI *et al.*, 2014.; SOUZA, 2014; TREVISOL *et al.*, 2015), sendo que as principais incluem aspectos metodológicos mal formulados, a dificuldade de tomar um conceito único de qualidade educacional em um país grande e diverso como o Brasil, pressões excessivas sobre professores e gestores, falta de entendimento e utilização dos resultados da avaliação nas escolas e sistemas educacionais estaduais e municipais, entre outras.

Num estudo sobre o Ideb e a construção de um modelo de *accountability* na educação brasileira, Schneider e Nardi (2014), após uma análise detalhada do índice: criação e normatização, propósitos anunciados, metas fixadas, mobilização da sociedade com a divulgação dos resultados, visibilidade dos dados produzidos e as consequências (positivas ou negativas); consideram que:

[...] as evidências apontadas são suficientemente indicativas de que o IDEB constitui uma fase inicial de construção de um modelo de *accountability* em educação no Brasil. No entanto, pouco se pode dizer das possibilidades de avanço deste modelo para a construção de um sistema mais amplo, que incorpore preocupações efetivas com as dimensões éticas, de justiça, democracia e participação crítica e ativa, haja vista sua aproximação com protótipos pautados em perspectivas neotayloristas, nomeadamente as que informam o campo educacional. (SCHNEIDER; NARDI, 2014, p. 22.).

Para esses autores, a política em que o Ideb se inscreve não contempla as concepções de *accountability* que possam ser articuladas com uma política educacional de qualidade. Ainda que a avaliação preceda uma etapa de responsabilização, o índice coloca em causa apenas o desempenho dos estudantes e das escolas, sem levar em conta outros fatores que intervêm nas condições de qualidade. Concluem que:

Um sistema de prestação de contas e de responsabilização, como o que o IDEB procura integrar, não pode prescindir de uma avaliação e de uma autoavaliação referenciadas, congruentes com o postulado de uma educação democrática na qual as políticas educacionais contemporâneas estão inscritas. (SCHNEIDER; NARDI, 2014, p. 23).

Sobre as críticas, Reynaldo Fernandes (2016), que era o presidente do Inep quando o índice foi criado, considera que,

De modo geral, não se diferenciam muito das relatadas na literatura internacional sobre *accountability* na educação: 1) os programas são incompletos, pois não consideram todos os resultados importantes das escolas; 2) são injustos, ao

responsabilizar os professores por aspectos sobre os quais eles não possuem total controle; e 3) podem gerar distorções como o estreitamento curricular e a exclusão de alunos com maiores dificuldades de aprendizado. (FERNANDES, 2016, p. 106).

Apesar das críticas, autor considera que a Prova Brasil e o Ideb são uma importante contribuição ao sistema de avaliação brasileiro – que considera um dos melhores do mundo – e que “alguns aprimoramentos seriam bem-vindos” (p.106). Esta possibilidade de aprimoramento sinalizada, precisa ser considerada para qualquer política pública voltada para a educação, tendo em vista o dinamismo característico da área.

Para Dirce Freitas (2013), com a avaliação censitária no ensino fundamental, foi “ampliada a possibilidade de controles sociais acionados pela publicidade do governo federal, envolvendo organizações sociais parceiras, agências de formação de opinião pública” (FREITAS, D., 2013), sendo que, além do controle exercido pelo governo, possibilitaram-se controles sociais mediante a divulgação dos resultados das aferições com maior abrangência. Por serem divulgados os resultados das escolas e tornados acessíveis para consulta pública, disponíveis aos pais e à sociedade em geral, Bonamino e Sousa (2012) consideram que a Prova Brasil representa um modelo de avaliação de segunda geração. Nessa classificação, a responsabilização é apenas branda, já que as equipes escolares podem receber alguma pressão dos pais ou da comunidade, favorecendo a mobilização para melhoria da qualidade da educação. Não há, contudo, consequências mais expressivas das avaliações sobre os professores ou diretores na prestação de contas do resultado do trabalho, nem o envolvimento de sanções ou premiações pelos resultados divulgados.

Em 2009, a décima edição do Saeb foi normatizada pela Portaria Inep nº 87, de 7 de maio, que estabeleceu objetivos gerais para o sistema:

- 1) oferecer subsídios à formulação, à reformulação e ao monitoramento de políticas públicas e programas de intervenção ajustados às necessidades diagnosticadas;
- 2) identificar problemas e as diferenças regionais do ensino;
- 3) produzir informações sobre os fatores do contexto socioeconômico, cultural e escolar que influenciam o desempenho dos alunos;
- 4) proporcionar aos agentes educacionais e à sociedade uma visão dos resultados dos processos de ensino e aprendizagem e das condições em que são desenvolvidos;
- 5) desenvolver competência técnica e científica na área de avaliação educacional, ativando o intercâmbio entre instituições educacionais de ensino e pesquisa. (BRASIL, 2009)

O critério quantitativo de alunos para a aplicação da Prova Brasil foi alterado novamente, de pelo menos 20 alunos matriculados em cada turma da escola, para 20 alunos

matriculados nas escolas em cada série ou ano escolares. No mais, a aplicação da avaliação repetiu as regras da edição anterior.

Com relação à Aneb, os objetivos específicos publicados em portarias anteriores foram mantidos e parte da amostra de alunos foi extraída da população da Prova Brasil. Também os objetivos específicos da Anresc/Prova Brasil, divulgados em sua criação, foram mantidos, aos quais foram acrescentados os seguintes:

- 1) produzir informações sobre o desempenho dos alunos, assim como sobre as condições intra e extraescolares que incidem sobre o processo de ensino e aprendizagem, no âmbito das redes de ensino e unidades escolares;
- 2) fornecer dados para o cálculo do Ideb. (BRASIL, 2009)

As características do teste cognitivo foram mantidas nas edições de 2009 e 2011. Quanto aos questionários contextuais, em análise dos documentos, Bonamino (2016) sinaliza algumas alterações nessas duas edições. Na versão de 2009, a autora revela que alguns blocos de questões foram excluídos dos questionários de diretores, como os itens sobre recursos didáticos disponíveis na escola, enquanto outros itens relacionados à infraestrutura foram incluídos em 2011. Também nessa edição foi excluído um item sobre o conhecimento do diretor acerca dos resultados da escola, o qual foi rerepresentado na edição de 2013.

Em relação ao questionário contextual dos professores, também na versão de 2011 houve algumas alterações, quando foi incluído um bloco de questões sobre os seus hábitos culturais. As alterações mais profundas nesse documento ocorreram na edição de 2013, quando o questionário passou a detalhar aspectos sobre a formação continuada de professores, aproveitamento do tempo em sala de aula, relacionamento com diretor e trabalho de equipe. A autora da análise afirma que “no contexto da criação da Prova Brasil, a relativa estabilidade dos questionários contextuais propícia à comparabilidade dos dados ao longo do tempo, é acompanhada pelo tratamento tópico e descontínuo dispensado à revisão e atualização.” (BONAMINO, 2016, p. 115). Abstrai-se, assim, que a amplitude de dados gerados pelos instrumentos contextuais ao longo das edições, principalmente após 2005, não favoreceu um interesse real pelas medidas sociais, escolares e pedagógicas que esses dados apontam. Há ênfase contínua e excessiva sobre as medidas cognitivas, ao passo que não há interesse nem reflexão necessários por parte dos diretores e professores diante dos dados gerados por esses instrumentos, conclui Bonamino (2016).

É nesse período, após a repercussão da divulgação dos resultados do Ideb de 2009, que se inicia um movimento de alguns gestores da rede privada de ensino, objeto desta pesquisa, buscando um modelo de avaliação externa em larga escala que fosse capaz de apresentar

dados sobre a qualidade educacional da rede. A primeira divulgação do Ideb após a pactuação do compromisso do Plano de Metas Compromisso Todos pela Educação demonstra uma evolução positiva na primeira fase do ensino fundamental, um expressivo aumento da proficiência dos alunos das séries iniciais no Saeb. A expectativa de que a avaliação externa em larga escala, e o índice que ela compunha, contribuíssem para melhora da qualidade da educação no Brasil, fez com que, em 2011, fosse reunida uma comissão de educadores representando várias regiões do país para discutirem uma proposta de criação e implementação do programa. Essa iniciativa culmina em outubro de 2012, com a aplicação das primeiras avaliações, inicialmente apenas para alunos do 5º ano do ensino fundamental, como detalhado no capítulo quatro.

Em 2013, com a publicação da Portaria nº 482, em 7 de junho, foi revogada a Portaria MEC nº 931, de 21 de março de 2005. Um novo formato é desenhado para o Saeb, que passou a ser composto por três avaliações. A Avaliação Nacional da Alfabetização (ANA) passou a integrar o sistema que, então, foi composto por: a Aneb, a Anresc/Prova Brasil e a ANA. As duas primeiras avaliações continuaram mantendo suas características, procedimentos e objetivos até então realizados. À ANA, com o objetivo de avaliar a qualidade, a equidade e a eficiência do Ciclo de Alfabetização da rede pública e produzir informações sistemáticas sobre as unidades escolares, foram atribuídas as seguintes características:

- I - ser uma avaliação censitária, de larga escala, externa aos sistemas de ensino público, aplicada anualmente no Ciclo de Alfabetização;
- II - a utilização de procedimentos metodológicos formais e científicos para coletar e sistematizar dados e produzir índices sobre o nível de alfabetização e letramento dos alunos do Ciclo de Alfabetização do ensino fundamental, conforme disposto no art. 30 da Resolução CEB/ CNE nº 7, de 14 de dezembro de 2010, da Câmara de Educação Básica do Conselho Nacional de Educação, e sobre as condições intraescolares que incidem sobre o processo de ensino e aprendizagem;
- III - 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;
- IV - concorrer para a melhoria da qualidade do ensino, redução das desigualdades e democratização da gestão do ensino público nos estabelecimentos oficiais, em consonância com as metas e políticas estabelecidas pelas diretrizes da educação nacional; e
- V - oportunizar informações sistemáticas sobre as unidades escolares. (BRASIL, MEC, 2013)

A Avaliação Nacional da Alfabetização (ANA) foi desenvolvida pelo Inep no contexto do Pacto Nacional pela Alfabetização na Idade Certa (PNAIC), instituído pela Portaria MEC nº 867, de 4 de julho de 2012. Por esse instrumento, o Ministério da Educação e as secretarias de Educação estaduais, municipais e o Distrito Federal reafirmaram e ampliaram o compromisso de alfabetizar as crianças até os oito anos de idade, ao final do 3º ano do ensino fundamental.

A mesma portaria previa uma avaliação externa para aferição da aprendizagem em Língua Portuguesa e Matemática, como forma de apoiar os sistemas públicos no compromisso assumido.

No dia 21 de junho de 2013, foi publicada pelo Inep a Portaria nº 304, que regulamentava as aplicações do Saeb no referido ano. A Prova Brasil e a ANA foram restritas às escolas públicas, e as áreas de Ciências foram, “em caráter experimental para efeito de matrizes e escalas”, incorporadas aos cadernos do 9º ano. Essa portaria ainda transferiu às secretarias estaduais e municipais a responsabilidade de informar ao Inep quais escolas indígenas que, devido às características de seus projetos pedagógicos, optassem por não participar da Prova Brasil.

No primeiro ano, uma aplicação-piloto da ANA foi realizada de forma censitária em todas as escolas públicas com turmas de 3º ano do ensino fundamental, localizadas em zonas rurais ou urbanas, e de forma amostral, em escolas com turmas multisseriadas. Já, no ano seguinte, 2014, participaram do ciclo todas as escolas públicas urbanas e rurais com, no mínimo, dez estudantes matriculados no 3º ano do ensino fundamental de nove anos. Cada teste de Matemática continha 20 itens objetivos, e os de Língua Portuguesa continham 17 itens de Leitura e 3 itens de produção escrita, totalizando 20 itens (INEP, 2015).

Em 25 de junho de 2014, foi sancionada<sup>22</sup> a Lei 13.005 que aprova o Plano Nacional de Educação (PNE) para o decênio 2014 – 2024. Na referida lei, além de um artigo dedicado ao Sistema Nacional de Avaliação da Educação Básica, há a confirmação do Ideb como um dos indicadores produzidos pelo sistema de avaliação, cuja responsabilidade de cálculo e divulgação compete ao Inep. Entre as metas estabelecidas para o período de vigência do Plano, uma, especificamente, está relacionada ao fomento da qualidade da educação, propondo o aumento progressivo das médias do Ideb para o país ao longo dos anos. Um elenco significativo de estratégias é projetado com vistas ao alcance da meta de:

Fomentar a qualidade da educação básica em todas as etapas e modalidades, com a melhoria do fluxo escolar e da aprendizagem, de modo a atingir as seguintes médias nacionais para o IDEB

IDEB	2015	2017	2019	2021
ANOS INICIAIS DO ENSINO FUNDAMENTAL	5,2	5,5	5,7	6,0
ANOS FINAIS DO				

<sup>22</sup> Lei sancionada pela presidente Dilma Rousseff, que governou o Brasil no período de 01 de janeiro de 2011 a 31 de agosto de 2016.

ENSINO FUNDAMENTAL	4,7	5,0	5,2	5,5
ENSINO MÉDIO	4,3	4,7	5,0	5,2

Fonte: BRASIL (2014).

Um sistema de acompanhamento das metas também é detalhado no Plano e, periodicamente, relatórios têm sido produzidos pelo Inep, com essa finalidade<sup>23</sup>. Consolidaram-se, nessa lei, os indicadores do rendimento e do fluxo escolar como representantes da qualidade da escola, embora figure, em seu texto, a previsão de indicadores de avaliação institucional que levem em conta as características do perfil do alunado e do corpo dos profissionais da educação; as relações entre corpo docente, técnico e discente; a infraestrutura e os recursos pedagógicos; os processos de gestão, entre outros aspectos que possam ser considerados relevantes.

Diante da normatização advinda do PNE, com vistas a buscar um novo modelo que valorizasse as potencialidades das informações produzidas, discutisse o que seria a qualidade da educação, e minimizasse o uso limitador dos índices, Horta Neto, Junqueira e Oliveira (2016, p. 28) afirmaram que:

Além de consagrar o Sistema Nacional de Avaliação da Educação Básica (Sinaeb) no corpo da Lei e de constituí-lo, conforme Art. 11, em “fonte de informação para a avaliação da qualidade da educação básica e para a orientação de políticas públicas desse nível de ensino”, o atual PNE introduz uma modificação de extrema relevância para o debate educativo atual no que diz respeito às informações produzidas pelas avaliações em larga escala: alçou os indicadores institucionais ao mesmo patamar dos indicadores de rendimento. Essa guinada não é pouca e nos estimula a pensar um novo uso e um maior aproveitamento de tais informações.

Assim, nos anos de 2015 e 2016, um grupo de pesquisadores da Diretoria de Avaliação da Educação Básica do Inep e especialistas de várias instituições realizaram uma série de debates para a construção de uma proposta que reunisse um elenco de indicadores capazes de fornecer aos gestores públicos e à sociedade civil, informações mais diversificadas e que contemplassem várias dimensões do processo educativo.

Partindo do princípio de que o processo educacional é multidimensional e que seria necessário construir um conjunto amplo de indicadores que abarcasse um espectro maior de dimensões, indo além do desempenho dos alunos nos testes cognitivos, os pesquisadores sugeriam que:

<sup>23</sup> Relatórios de monitoramento do PNE disponíveis em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/gestao-do-conhecimento-e-estudos-educacionais/estudos-educacionais/relatorios-de-monitoramento-do-pne>.

[...] é preferível disponibilizar um sistema de indicadores que, além de abarcar as principais dimensões educacionais, diferencie as informações referentes ao contexto, aos recursos, aos processos e aos resultados, pois, mesmo apresentando limitações, ele é capaz de oferecer um retrato mais amplo e coerente da realidade enfocada, ao articular os aspectos mensurados e, assim, potencializar sua capacidade interpretativa. (HORTA NETO; JUNQUEIRA; OLIVEIRA, 2016, p. 30).

Das metas definidas no PNE, emanou a proposta sintetizada no Quadro 9.

### Quadro 9 - Diretrizes e dimensões do Sistema Nacional de Avaliação da Educação Básica

Diretriz	Dimensão
Universalização do atendimento escolar	Acesso e Permanência
	Trajetória
	Infraestrutura
	Aprendizagens
Melhoria da Qualidade do Aprendizado	Práticas Pedagógicas
	Ambiente Educativo
	Formação para o Trabalho e Cidadania
Valorização dos profissionais da educação	Formação Inicial e Continuada
	Carreira e Remuneração
	Satisfação Profissional
Gestão democrática	Financiamento
	Planejamento e Gestão
	Participação
	Inclusão e Equidade
Superação das desigualdades educacionais	Direitos Humanos, Diversidade e Diferença
	Contexto Socioeconômico e Espacial
	Intersetorialidade e Sustentabilidade

Fonte: Santos, Horta Neto e Diniz (2017).

Com base nessa proposta, o Ministério da Educação editou a Portaria nº 369, de 5 de maio de 2016, que instituiu o Sistema Nacional de Avaliação da Educação Básica (Sinaeb), com o objetivo de:

[...] assegurar o processo nacional de avaliação da educação básica em todas as etapas e modalidades, considerando suas múltiplas dimensões, na perspectiva de garantir a universalização do atendimento escolar, por meio de uma educação de qualidade e democrática, a valorização dos profissionais da educação e a superação das desigualdades educacionais. (BRASIL, 2016).

Essa portaria traz detalhadamente várias definições a respeito do Sistema e, entre as principais, podemos destacar: i) a produção de indicadores de avaliação institucional; ii) a

criação de indicadores de qualidade das condições de oferta como uma referência para a definição do custo aluno; iii) a agregação do cálculo do Índice de Diferença do Desempenho esperado e verificado – o IDD; iv) criação de um Comitê de Governança do Sinaeb, composto por representantes de várias entidades; v) a definição de uma Avaliação Nacional da Educação Infantil, além das avaliações já regulamentadas como: Avaliação Nacional de Alfabetização; Avaliação da Alfabetização – Provinha Brasil; Avaliação Nacional da Educação Básica e Avaliação Nacional do Rendimento Escolar – Prova Brasil.

Horta Neto (2018) afirma que, apesar de a portaria incluir o Índice de Diferença do Desempenho esperado e verificado, que não fazia parte da proposta original, o restante dela foi fiel ao estudo produzido. O autor também destaca a previsão do comitê de governança como uma garantia da transparência ao processo avaliativo e cuidado com as relações entre os diferentes entes federativos, com representação dos estados, municípios, da sociedade civil e da comunidade científica.

Nesse contexto, o cenário político do Brasil convivia com um processo de *impeachment* da presidente Dilma Rousseff, cuja tramitação no Congresso iniciara em dezembro de 2015. Com a admissibilidade do processo no Senado Federal, o vice-presidente assumiu interinamente a partir de 12 de maio de 2016. Antes mesmo do encerramento do processo de *impeachment*, em 31 de agosto de 2016, o Ministério da Educação, por meio da Portaria nº 981, de 28 de agosto, revogou a Portaria nº 369, alegando que, com “as revisões da Base Nacional Comum Curricular – BNCC ainda [...] em curso”, o Sinaeb deveria pautar-se nas recomendações oriundas desse processo, e seriam mantidas as avaliações já realizadas pelo Inep, permanecendo a estrutura anterior até o ano de 2017.

No ano de 2016, foi realizada a Avaliação Nacional da Alfabetização (ANA), cujos resultados foram divulgados no final do ano de 2017 por meio de relatório publicado pelo Inep, atendendo, assim, à periodicidade bianual estabelecida por ocasião da sua regulamentação.

A edição do Saeb de 2017 foi regulamentada pela Portaria MEC nº 564, de 20 de abril, e trouxe como alteração, em relação à edição de 2015, a aplicação da avaliação de forma censitária para o último ano do ensino médio, com a possibilidade de adesão das escolas privadas, mediante assinatura de termo de adesão e pagamento de um valor correspondente ao número de alunos participantes.

Para Horta Neto (2018), a publicação dessa portaria contradiz a justificativa apresentada pelo MEC para “barrar” o Sinaeb, ao ampliar a aplicação do Saeb para todas as

escolas de ensino médio, inclusive as da rede privada que aderissem e pagassem pelos testes. Considera que, “neste caso específico, amplia-se a avaliação, mesmo que a BNCC para o ensino médio esteja ainda longe de ser finalizada, em contradição com o argumento utilizado para revogar a portaria que instituiu o Sinaeb.” (HORTA NETO, 2018, p. 51).

No ano de 2018, a Avaliação Nacional da Alfabetização não foi realizada com o argumento de que os instrumentos deveriam ser ajustados à Base Nacional Comum Curricular, que fora homologada pelo ministro da Educação em 20 de dezembro de 2017 e instituída pela Resolução CNE/CP nº 2, de 22 de dezembro de 2017.

Por meio do Decreto Presidencial nº 9.432, publicado no dia 29 de junho de 2018, Michel Temer regulamentou a Política Nacional de Avaliação e Exames da Educação Básica. Estabeleceu, o decreto, que a política de avaliação será orientada pela Base Nacional Comum Curricular, observando o que está disposto na Lei de Diretrizes e Bases da Educação, e terá como objetivos:

- I – diagnosticar as condições de oferta da educação básica;
- II – verificar a qualidade da educação básica;
- III – oferecer subsídios para o monitoramento e aprimoramento de políticas educacionais;
- IV – aferir as competências e habilidades dos estudantes;
- V – fomentar a inclusão educacional de jovens e adultos; e
- VI – promover a progressão do sistema de ensino. (BRASIL, MEC, 2018)

O Saeb é considerado então, “um conjunto de instrumentos que permite a produção e a disseminação de evidências, estatísticas, avaliações e estudos a respeito da qualidade das etapas que compõem a educação básica: educação infantil, ensino fundamental e ensino médio” (BRASIL, MEC, 2018).

Interessante destacar que o decreto vem regulamentar uma política de avaliação cuja base estrutural e operacional constituiu-se ao longo de, pelo menos, 28 anos. O Saeb, cuja sigla representa um “sistema” de avaliação, é considerado então como “conjunto de instrumentos”. Sobre isso, alguns meses antes, Horta Neto (2018) já escrevia que:

Nos últimos tempos, a avaliação se tornou sinônimo de política educacional. E isso é um equívoco tremendo. A avaliação é um instrumento importantíssimo para aprimorar os processos educacionais e políticas públicas da área, mas é apenas um instrumento. Uma política educacional é algo muito mais amplo que isso, envolve uma série de programas, projetos, intersetorialidade e deve apontar de forma inequívoca aonde se quer chegar, e não apenas definir metas de aprendizado como vem acontecendo desde o primeiro governo de Luiz Inácio Lula da Silva. (HORTA NETO, 2018, p. 51).

A sinalização para a necessidade de um espectro muito mais amplo que – apenas – denominar um conjunto de instrumentos de avaliação de política nacional parecia antever a publicação do decreto presidencial. A busca por uma proposta mais abrangente, que envolva múltiplas dimensões do processo educacional, que seja fruto da discussão ampla com os diferentes níveis de governo, sistemas e atores educacionais, ainda não se concretizou pela edição do Decreto.

No apagar das luzes do governo de Michel Temer, no dia 26 de dezembro de 2018, o Inep publica a Portaria nº 1.100, que regulamentaria a realização do Saeb em 2019. Porém, com a posse do presidente eleito, Jair Messias Bolsonaro, entre mudanças de ministros e presidentes do Inep, a realização do Saeb de 2019 foi regulamentada pela Portaria Inep nº 366, de 2 de maio, que também o reestrutura significativamente. A partir de então, deixam de existir as siglas ANA, Anresc e Aneb, e todas as avaliações da educação básica passam a ser identificadas como Saeb. Porém, a maior reformulação ocorreu em maio de 2020, quando foi editada pelo então ministro da Educação, Abraham Weintraub, a Portaria nº 458, alterando substancialmente o sistema nos seguintes pontos: i) A prova será realizada anual e censitariamente, em escolas públicas e privadas; ii) Serão aplicadas provas e questionários nos alunos de todos os anos ou séries do ensino básico; iii) Os resultados poderão ser utilizados como mecanismo de acesso à educação superior. Os objetivos foram ressignificados e passaram a ser os seguintes:

I - construir uma cultura avaliativa, ao oferecer à sociedade, de forma transparente, informações sobre o processo de ensino-aprendizagem em cada escola, comparáveis em nível nacional, anualmente e com resultados em tempo hábil, para permitir intervenções pedagógicas de professores e demais integrantes da comunidade escolar;

II - produzir indicadores educacionais para o Brasil, suas regiões, unidades da federação, municípios e instituições escolares, tendo em vista a manutenção da comparabilidade dos dados, permitindo, assim, o incremento de séries históricas;

III - avaliar a qualidade, a equidade e a eficiência da educação praticada no País em seus diversos níveis governamentais;

IV - 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

V - desenvolver competência técnica e científica na área de avaliação educacional, por meio de intercâmbio entre instituições de ensino e pesquisa, bem como de servidores do Inep, docentes e gestores da educação de todos os entes envolvidos. (BRASIL, MEC, 2020)

Ao longo dessas três décadas de avaliação da educação básica no Brasil, descritas neste capítulo, aliado a um movimento global de avaliação de estudantes, que pode ser caracterizado pelo surgimento, nesse período, do *Programme for International Student Assessment* (Pisa), os resultados dessas avaliações têm ocupado significativo espaço de

divulgação nos meios de comunicação. Com isso, a opinião pública e o interesse da sociedade em geral têm se voltado para os indicadores e há um crescente movimento na direção de avaliações próprias nos vários níveis de governo – estaduais e municipais. Principalmente após a criação do Ideb, estudiosos (HORTA NETO *et al.*, 2016; BAUER *et al.*, 2015) têm relatado o crescimento do número de estados e municípios que realizam avaliações em seus sistemas. Muitos destes, com uma forte política de responsabilização que caracteriza a terceira geração da avaliação no Brasil, conforme descrita em Bonamino e Souza (2012).

Há um crescente número de pesquisadores debruçando-se sobre a temática, como sinaliza Bauer (2012), apresentando o aumento significativo de pesquisas sobre avaliação de sistemas no Brasil, num levantamento que a autora realizou no Banco de Teses e Dissertações da Capes. Principalmente após a criação da Prova Brasil e do Ideb, o número de publicações aumentou significativamente. Nesse panorama, muitas contribuições, críticas e proposições têm sido apontadas para o futuro do sistema de avaliação. A complexidade do próprio ato de avaliar, associado à realidade de um país de dimensão continental, a necessidade de articulação entre os diferentes níveis de governo, o uso adequado dos dados coletados – que vão muito além do resultado dos testes –, a necessidade de discussão ampla sobre as dimensões que abrangem a qualidade educacional; esses e muitos outros aspectos ainda se configuram desafios para os anos futuros.

Ainda que complexo, que envolva recursos financeiros e humanos; de alguma forma o interesse por este modelo de avaliação tem se ampliado e tem sido assimilado por sistemas de ensino, instituições e redes privadas. A aderência a sistemas de avaliação externa e em larga escala não tem sido volátil ou passageira, como pode-se concluir pela análise da trajetória até aqui descrita. É possível inferir também, que esta estabilidade e adesão sejam indicativos de que as potencialidades deste modelo de avaliação são reconhecidas e está alcançando seus propósitos.

É nesse cenário que surge uma iniciativa de uma rede privada de ensino em desenvolver um programa de avaliação externa em sua rede, presente em quase todo o território brasileiro, cuja descrição será objeto do próximo capítulo.

## **CAPÍTULO 4 – PROGRAMA DE AVALIAÇÃO EM LARGA ESCALA DA REDE PRIVADA DE ENSINO: TRAJETÓRIA E CARACTERÍSTICAS**

Este capítulo apresenta a trajetória do Programa de Avaliação em Larga Escala da rede de ensino objeto desta pesquisa. Inicia-se com uma breve descrição da história do sistema de ensino privado no Brasil, seguida do arcabouço normativo delimitado pela Constituição Federal de 1988, e pela LDB 9394 de 1996, que garantem o direito do ensino à iniciativa privada e, então, o programa de avaliação externa é descrito.

Para a caracterização da rede de ensino e do programa de avaliação externa nela desenvolvido, faz-se uma descrição do sistema privado de ensino no cenário nacional. Esta contextualização busca favorecer a compreensão da singularidade do objeto do estudo, observando-se que a maioria dos estudos sobre avaliação externa concentra-se em sistemas públicos de ensino ou em instituições privadas de ensino superior.

Conforme apontado no capítulo anterior, houve um aumento significativo da avaliação em larga escala nos sistemas públicos de ensino em estados e municípios brasileiros, principalmente a partir da criação da Prova Brasil e do Ideb. Pesquisadores publicaram levantamentos desse crescimento nos últimos anos (HORTA NETO *et al.*, 2016; BAUER *et al.*, 2015), demonstrando que mesmo em cenários de mudanças e distintos modelos ideológicos de governo, tanto na esfera nacional como nas esferas subnacionais, sistemas de avaliação têm se estabelecido e consolidado ao longo do tempo. Foi no contexto desse crescimento que a rede privada de ensino, a que se refere o presente estudo, desenvolveu seu Programa de Avaliação em Larga Escala – aqui denominado PAER –, cuja primeira edição deu-se em 2012 e que será caracterizado neste capítulo.

### **4.1 Uma breve história do ensino privado no Brasil**

A história da educação brasileira mescla-se com a história da iniciativa privada na educação – principalmente de natureza confessional –, se considerarmos que o primeiro estabelecimento de ensino no país foi fundado pela ordem dos Franciscanos, na Bahia, em 1533. No contexto da colonização brasileira, conforme registrou Alves, “não é de se estranhar que a primazia do ensino tenha recaído sobre a escola confessional.” (ALVES, 2009, p.71).

Saviani (2008), por sua vez, afirma que a primeira instituição escolar brasileira foi fundada por jesuítas em 1549. Seja esta data, esta ordem religiosa, ou as citadas por Alves (2009), há consenso entre os historiadores acerca de que, até 1759, o ensino no Brasil ficou a

cargo das ordens religiosas. Alves (2009) aponta que diversos autores citam que haveria cerca de 20 estabelecimentos de ensino dos jesuítas. A expulsão da Companhia de Jesus de Portugal e do Brasil, por ordem do Marquês de Pombal, encerra uma fase da educação privada brasileira, mas não a sua atuação no país. Outras ordens religiosas, como os Dominicanos, Beneditinos, Oratorianos, Carmelitas e os próprios Franciscanos desenvolviam atividades de ensino no Brasil, com estruturas próprias, eficazes e autônomas, e não interromperam suas atividades. A presença de outras instituições particulares assegurou a continuidade do ensino privado.

Desse período, Tobias (1986) descreve que,

A reforma de Pombal, porém, expulsando os jesuítas, confiscando-lhes, de modo abrupto e sem planejamento, as escolas, os bens e a educação, esfacelou, da noite para o dia, a educação brasileira. Nas aldeias, os padres da Companhia de Jesus tinham escolas de ler e escrever e, na época da expulsão, possuíam 17 colégios e seminários, 36 missões e 25 residências. (TOBIAS, 1986, p. 96)

Alves (2009) reitera que a intenção de Pombal era, por meio dos agentes educativos pagos e controlados pelo Estado, formar os quadros administrativos e políticos a serviço do próprio Estado. Saviani (2008), fazendo uma periodização da história da educação brasileira, sobre esse período afirma que a reforma pombalina, representada pelas “Aulas Régias”, foi a primeira tentativa de instaurar uma “escola pública estatal inspirada nas ideias iluministas, segundo a estratégia do despotismo esclarecido” (p.2). Nos anos seguintes, houve tentativas descontínuas e intermitentes de organizar a educação como responsabilidade do poder público representado pelo governo imperial até a década de 1890.

Sem a presença dos jesuítas, Tobias (1986) descreve que os brasileiros recorreram a professores leigos, improvisados de outras profissões para a organização do ensino a ser oferecido pelo Estado. Muitas vezes, os padres e capelães de engenho eram indicados para esse mister. Muitos desses professores leigos eram frutos da educação do jesuíta. Sendo assim, o autor assinala que a Companhia de Jesus foi expulsa do Brasil, mas não da educação brasileira. Porém, como destacou Saviani (2010), as ideias inspiradas no laicismo originadas no iluminismo passaram a ter espaço para circularem no âmbito educacional.

Saviani (2010), ao investigar sobre a história das ideias pedagógicas no Brasil, afirma que no período de 1759 a 1932, coexistiram as vertentes religiosa e leiga da pedagogia tradicional. A primeira fase desse período, na visão de Alves (2009), foi marcada por um retrocesso pedagógico, quando os ensinos primário e secundário viveram uma situação de quase abandono. Com a vinda da família real e a instalação do governo português no Brasil,

em 1808, a ênfase foi dada ao ensino superior e alguns ofícios técnicos. Dessa forma, até 1824, quando o Brasil promulgou sua primeira Constituição, no nível primário e secundário, tanto as escolas confessionais quanto as escolas administradas pelo poder público passaram por um período de marasmo e esquecimento. Inspirada no modelo da Constituição francesa, essa nova Constituição instituiu um Sistema Nacional de Educação, propondo a criação de escolas primárias em todas as cidades, colégios e liceus nas capitais, bem como o ensino superior nas grandes cidades. Porém, inúmeros fatores, capitaneados pelo nascimento do aparelho burocrático do Estado, após a independência de Portugal, reduziram os recursos para a organização da educação pública, mesmo no nível primário, sendo mantida a prioridade no ensino superior, em função da necessidade da criação de novos quadros na administração pública. Alves (2009) cita a criação das Faculdades de Direito de São Paulo (1825) e de Recife (1827) como exemplos dessa prioridade e afirma que a iniciativa privada, a partir de então, iniciou um processo de expansão (ALVES, 2009, p. 73).

Se, até então, predominava o ensino privado ligado à confissão e ordens católicas, em 1834 um Ato Adicional à Constituição influenciou significativamente a educação no Brasil. A nova legislação descentralizou a direção da educação cuja consequência foi a desarticulação entre o ensino primário e o secundário. Às províncias coube a responsabilidade sobre o ensino primário e, ao ensino superior, sob a administração do governo central, coube a responsabilidade de organizar o ensino secundário como preparatório para as faculdades. Ao final do ciclo da mineração, com o deslocamento das atividades rurais para as atividades comerciais nas cidades, surgiu a burguesia, como classe emergente, disputando espaço com a aristocracia do período colonial. A necessidade de garantir escolaridade para seus filhos fez com que a pequena burguesia reivindicasse o direito de frequentar a mesma escola que a aristocracia. Dessa forma, Alves (2009) considera que surge um novo modelo de ensino privado, aquele que é financiado diretamente pelos pais.

A ausência de recursos das “Províncias” para organizar seu próprio ensino, público e gratuito, especialmente em nível secundário, abriu espaço para que a iniciativa privada assumisse tal tarefa. Neste contexto, ela pode, pouco a pouco, conforme as circunstâncias em cada Província, ampliar seu espaço. Consolida-se o ensino privado financiado diretamente pelos pais dos alunos, nos moldes que subsiste até hoje. (ALVES, 2009, p. 74).

Ainda, sob a diligente atuação das elites brasileiras, do clero e de intelectuais, em 1854, o imperador aprovou a Lei de Liberdade do Ensino, consolidando a livre iniciativa na educação por meio de escolas privadas autônomas em relação ao Estado, garantindo uma nova expansão ao segmento. No mesmo período, foi criada a Inspeção Geral da Instrução,

“encarregada de controlar o ensino primário e secundário, público e privado; a fixação de normas e mecanismos para regulamentar a atuação privada no ensino.” (ALVES, 2009, p. 74).

O ensino privado continuou a expandir-se com uma proliferação significativa de escolas católicas mantidas por congregações religiosas da Europa. Concomitantemente, os ideais liberais, anticlericais e republicanos expandiam-se na sociedade brasileira, de modo que, em 1879, foi decretada a reforma do ensino articulada por Leôncio de Carvalho que, entre outras medidas, ampliou a liberdade de ensino, garantindo a manifestação de outras tendências pedagógicas e o livre credo religioso dos alunos. Saviani (2010) ainda ressalta, entre outras inovações dessa reforma, a subvenção do ensino particular pelo Estado e a possibilidade de contratação de professores particulares.

Desse período pré-republicano, Saviani (2010) destaca um caso paradigmático de escolas privadas criadas por Abílio César Borges: o Barão de Macahubas (1824-1891). Médico de formação, o barão pouco atuou na área da saúde, tendo se dedicado por pouco tempo como diretor-geral dos estudos da província da Bahia, afastando-se da função por divergências políticas. Contudo, desse breve período de cerca de um ano no ensino público, desenvolveu ideias pedagógicas que pode colocar em prática nas três instituições privadas por ele criadas: o Ginásio Baiano em Salvador, o Colégio Abílio do Rio de Janeiro e o Colégio Abílio de Barbacena. Suas linhas pedagógicas foram influenciadas por sua formação religiosa católica e pelas ideias do liberalismo incorporadas em virtude de suas viagens à Europa, de onde trouxe compêndios, materiais pedagógicos e professores. Suas ideias caracterizaram-se, principalmente, por defender uma disciplina branda; não aplicar castigos aos alunos; novos processos de alfabetização; métodos renovados de estudo das línguas vivas; valorização dos professores; distribuição gratuita de livros didáticos, sendo que a maioria destes, produzida por ele próprio. Seus colégios também se notabilizaram pelas personalidades que neles estudaram, como Castro Alves e Rui Barbosa. O Colégio do Rio de Janeiro ainda funcionou até 1911, dirigido por seus filhos.

Saviani (2010) destaca que o modelo exitoso de escola privada produzido por Abílio César Borges acendeu a discussão sobre a liberdade de ensino. Assim descreve:

De fato, quando diretor-geral da Instrução Pública, Abílio César Borges preconizou o exercício, por parte do poder público, de um controle e fiscalização rigorosa das escolas privadas, posicionando-se, portanto, contra a chamada “liberdade de ensino”. Depois, já na iniciativa privada, ele propôs, em 1873 ao Conselho Diretor da Instrução do Município da Corte que fosse decretado o ensino inteiramente livre, à vista da ineficiência da inspeção oficial e considerando que o Estado não se deve intrometer em algo que é dever natural e impreterível das famílias. [...] A partir de 1870, liberais e conservadores irmanaram-se na defesa da liberdade de ensino. E a

pujança da iniciativa privada da qual são exemplos os colégios fundados por Abílio César Borges, assim como a crença de que com liberdade plena haveria ainda maior expansão, fortaleceram a referida tese. (SAVIANI, 2010, p. 158).

Alves (2009) afirma que, “com o advento da República instaura-se um período dos mais importantes para o ensino privado” (p. 74). O regime federativo mantinha sob a responsabilidade das antigas províncias, agora estados, a responsabilidade da instrução. Os ideais liberais e as novas exigências produtivas pressionaram a abolição e a Proclamação da República, movimentos que afetaram fortemente a instrução popular.

Após a Proclamação da República, o modelo federativo de descentralização, previsto em 1834 por meio do Ato Constitucional, definitivamente é adotado. O movimento positivista republicano fez-se sentir na reforma de ensino instituída em 1890 por Benjamin Constant, por meio do Decreto nº 981. Embora limitada ao Distrito Federal e com abrangência para os níveis primário e secundário, a reforma era referência para a organização do ensino nos estados. As propostas básicas eram “a laicização, a gratuidade e a preponderância das disciplinas científicas, em ruptura ao modelo humanista clássico; numa palavra, uma escola estatal acessível a todos. A reforma foi um fracasso” (ALVES, 2009, p. 75). Os próprios positivistas, dos quais Benjamin Constant era um dos líderes, criticaram amplamente a reforma, afirma Saviani (2010). Como uma das principais causas do fracasso, Alves (2009) indica “a falta de apoio das classes dirigentes e dos intelectuais que não queriam prejudicar a formação de seus filhos no ensino privado.” (p. 75).

Schulz (2003), num estudo sobre a atividade educativa das confissões protestantes no Brasil, data das últimas décadas do século XIX o estabelecimento de escolas primárias e secundárias das diferentes confissões que migraram para o Brasil. Como fatores que contribuíram para a inserção do sistema de ensino protestante de forma definitiva no Brasil, o autor listou:

- a) o acordo que foi firmado entre Portugal e Inglaterra, país capitalista e protestante, passou a exercer influência e condicionamentos;
- b) as mudanças em nível da estrutura de produção escravocrata para a mão de obra livre, estimulando a imigração (se mera coincidência ou não, deu-se a expansão missionária pelas agências das missões norte-americanas e a expansão da influência, por alguns, chamado imperialismo americano, sobre a América Latina);
- c) as mudanças em nível ideológico, pela assimilação e infiltração das ideias modernas, pelos jovens que iam estudar na Europa e retornavam com uma nova concepção de mundo;
- d) o interesse, principalmente da ideologia liberal, em quebrar o monopólio religioso e, para isso, era preciso separar Igreja e Estado;
- e) a crise da Igreja Católica e a sua decadente influência, principalmente nos estados onde a “modernização” teve maior espaço, como São Paulo. (SCHULZ, 2003, p. 37).

Assim, nesse período, surgiram e se expandiram inúmeras instituições privadas de confissão protestante que ainda mantêm redes de ensino em vários pontos do território brasileiro, tais como: presbiterianos, metodistas, batistas, adventistas e luteranos (SCHULZ, 2003).

Após a Proclamação da República, uma nova constituição é promulgada em 1891. Nela, os princípios de federalização do Estado e descentralização do poder, característicos do modelo republicano, foram estabelecidos, e o ensino obrigatório, laico e gratuito foi garantido. Contudo, a garantia constitucional esbarrava em inúmeros obstáculos estruturais, tanto de recursos materiais, como humanos, nas esferas descentralizadas de poder: estados e municípios que se estruturavam na nova estrutura de governo.

Nos anos seguintes, até o início do século XX, várias reformas de ensino empreendidas pelo Estado não conseguiram modificar de maneira significativa o sistema escolar. Para Alves (2009), “a despeito de seus méritos, estas reformas sempre ignoraram os conflitos sociais, econômicos, políticos e ideológicos.”. Na falta de recursos humanos e materiais por parte do Estado, “o ensino privado vai rapidamente ocupar os espaços vazios, sobretudo no ensino secundário” (p. 75). Desse período, vale, contudo, destacar uma ampla reforma iniciada no estado de São Paulo, que em 1892 implantou a organização das escolas primárias em grupos escolares. Saviani (2005) descreve que, além da organização administrativa e pedagógica, compreendendo a definição de conteúdos curriculares e métodos de ensino, um elemento importante foi a iniciativa de construir prédios públicos vistosos “que rivalizavam com a igreja, a câmara municipal e as mansões mais importantes tanto da capital como das principais cidades do interior.” (p. 30). Esse modelo originário de São Paulo disseminou-se pelos demais estados. Contudo, Saviani (2005) afirma que essa forma de organização da escola primária aprofundou a seleção e formação de elites. “A questão da educação das massas populares ainda não se colocava.” (p. 31). Essa questão só virá à tona na reforma paulista de 1920, década caracterizada por aquilo que Nagle (1974) denominou entusiasmo pela educação.

Marçal (1993) caracteriza a década de 1920 pelo declínio da oligarquia, pela crise do modelo agrário-comercial-exportador e o impulso à industrialização e o modelo nacional-desenvolvimentista. A classe burguesa fortalecida passa a reivindicar mudanças. Além da queda da oligarquia e ascensão da burguesia industrial, as revoluções, a Semana de Arte Moderna, movimentos sociais e políticos, como a organização do Partido Comunista, são o

pano de fundo para o surgimento, no Brasil, do movimento da Escola Nova, cujas ideias principais estavam centralizadas na defesa do ensino leigo, universal, gratuito e obrigatório, com ênfase na importância do Estado na educação e, por esta, a reconstrução nacional. O pensamento escolanovista apelava para o humanismo científico-tecnológico como solução para os problemas do país. A ideia de que o homem e a máquina podem conviver bem, buscando o progresso a serviço do homem.

O liberal pragmatismo da Escola Nova influenciou inúmeras reformas educacionais em vários estados da federação, de 1920 a 1929. Contudo, nesse período, recrudescer uma tensão entre as ideias dos pioneiros da Escola Nova e o pensamento conservador católico que, na visão de Saviani (2010), mobilizou a Igreja Católica para uma “estratégia de resistência ativa”, que “foi acionada de forma mais efetiva” (p.180). Sobre o episódio, ele relata:

Com essa força organizativa, os católicos constituíram-se no principal núcleo de ideias pedagógicas a resistir ao avanço das ideias novas, disputando, palmo a palmo com os renovadores, herdeiros das ideias liberais laicas, a hegemonia do campo educacional no Brasil a partir dos anos de 1930. (SAVIANI, 2010, p. 181).

Essa disputa por protagonismo garantiu a restituição, à Igreja Católica, de sua influência política no Estado Novo, que fora enfraquecida com a Proclamação da República. Em 1930, foi criado o Ministério da Educação e Saúde Pública e, em 1931, seu titular Francisco Campos publicou inúmeras reformas. Dentre elas, decretou a volta do ensino religioso à escola pública como uma das conquistas dessa investida estratégica de resistência.

A situação política confusa no início da década de 1930, após a Revolução que depôs o presidente Washington Luiz e estabeleceu o governo provisório de Getúlio Vargas, a insatisfação da sociedade em geral, o descontentamento de setores políticos e a inabilidade do governo de tomar medidas imediatas no campo educacional levam à mobilização de um grupo de educadores comprometidos com as reformas da década de 1920 e com os ideais escolanovistas, os quais produziram o *Manifesto dos Pioneiros da Educação Nova*. Saviani (2010) considera que, como o próprio nome apresenta, o Manifesto constituiu-se um instrumento político. O objetivo anunciado no documento – a reconstrução educacional do Brasil – abarcava não somente os aspectos filosóficos do movimento, mas a própria organização e administração do sistema educacional, comprometido com o princípio de uma escola comum e única, e com a garantia, pelo Estado, da laicidade, gratuidade, obrigatoriedade e coeducação em todos os níveis de ensino.

Ao *Manifesto dos Pioneiros da Educação Nova*, também houve uma reação de intelectuais ligados à Igreja Católica. Marçal (1993) detalha um período subsequente de

conflito de ideias, que foi marcado por inúmeros congressos, seminários, conferências que defendiam o pensamento e propostas dos pioneiros, por um lado, e o pensamento e propostas dos católicos, por outro lado. Afirma, Marçal, que Getúlio Vargas aproveitou muito bem essas divergências na área da educação, quando as reivindicações educacionais se confundiam com os interesses políticos, para solidificar-se no poder, procurando atender tanto a um como a outro grupo.

Esse duplo apoio pode ser avaliado na Constituição de 1934, que garantia o ensino religioso facultativo nas escolas públicas, ajuda financeira às escolas privadas e confessionais, a liberdade de organização de um sindicato para a educação católica e a origem do movimento sindical patronal no ensino privado (ALVES, 2010). Inicia-se, então, um acirrado período de disputa entre o ensino público e o ensino privado.

Souto (2012), num estudo sobre a relação entre o ensino privado e o ensino público na história da educação brasileira, com respeito a esse momento, considera que:

Em meio a essa conjuntura, a luta que ora se trava em defesa da escola pública é um verdadeiro divisor de águas na história pedagógica da nação, propiciando o choque inevitável entre duas mentalidades existentes no país, a saber, católicos e liberais. A luta que se abriu em nosso país entre os partidários da escola pública e os da escola particular é, no fundo, a mesma que se travou em vários países, entre a escola religiosa (ou o ensino confessional), de um lado, e a escola leiga (ou o ensino leigo), de outro. (SOUTO, 2012, p. 38).

No ano de 1937, sob o regime do Estado Novo<sup>24</sup>, uma nova Constituição é outorgada, dedicando menos espaço à educação, mas garantindo que “a arte, a ciência e o ensino são livres à iniciativa individual e a de associações ou pessoas coletivas públicas e particulares” e constitui “dever do Estado contribuir, direta e indiretamente, para o estímulo e desenvolvimento de umas e de outras.” (BRASIL, 1937).

Os anos seguintes são marcados por uma luta entre os que defendiam uma escola pública de qualidade e laica em oposição aos que defendiam o ensino privado. Alves (2009) considera a época “de grandes conflitos entre *privatistas e publicistas*. Tal conflito colocava em xeque a livre iniciativa na educação.” (p. 75).

Em 1945, um golpe de Estado depôs Getúlio Vargas, e o Brasil passa por um período de abertura democrática que, até então, não experimentara. Uma nova Constituição é promulgada em 1946, que manteve todos os direitos adquiridos anteriormente pelo ensino privado.

---

<sup>24</sup> Regime político instituído em 1937, inspirado na ditadura de Antônio de Oliveira Salazar, em Portugal. Durou até 1945, quando Getúlio Vargas foi deposto pelas forças armadas.

Brzezinski (2010), ao recuperar a história dos embates de projetos antagônicos de sociedade e de educação, entre os defensores do ensino público e os defensores do ensino privado, considera o longo período de treze anos de tramitação da primeira Lei de Diretrizes e Bases da Educação, no Congresso Nacional, como consequência desse antagonismo e do acirramento do debate e das divergências entre os educadores liberais e os intelectuais católicos.

Saviani (2010) também narra detalhadamente esse momento da história em que o conflito escola particular *versus* escola pública “empolgou a opinião pública” (p. 292) e os meios de comunicação da época – os principais jornais em circulação –, e as revistas, que cada segmento editava, encarregavam-se de municiar a opinião pública com as ideias de ambas as posições.

A Lei 4.024 de 21 de dezembro de 1961 foi promulgada, garantindo a liberdade para a iniciativa particular. Alves (2009) argumenta que a primeira LDB promoveu uma renovação dos estabelecimentos de ensino do setor privado e deu “às escolas maior autonomia na sua organização administrativa, disciplinar e didática, desafiando o ensino privado a se reorganizar, questionando o modelo educacional vigente” (p. 75). Já Saviani (2010), considerando a posição dos liberais diante da Lei de Diretrizes e Bases, comenta que “a vitória só não foi completa em razão das concessões feitas à iniciativa privada” (p. 307), deixando de levar em conta um importante aspecto defendido pelos pioneiros da educação nova em sua defesa do ensino público a todos.

Diante de diferentes reivindicações das diversas camadas da sociedade, ocorreu o golpe de estado, e os militares assumiram o poder no ano de 1964. Há um período inicial de recuperação econômica seguido de um acentuado desenvolvimento do setor industrial. Saviani (2010) salienta que, com o regime militar, a ruptura deu-se no nível político e não no âmbito socioeconômico. Na educação houve, portanto, continuidade dos princípios estabelecidos pela Lei nº 4.024, de 1961, pois atendiam aos interesses do regime, e as reformas se deram apenas nas bases organizacionais, com a promulgação da Lei 5.692, de 1971, “tendo em vista ajustar a educação aos reclamos postos pelo modelo econômico do capitalismo de mercado.” (SAVIANI, 2010, p. 364). A principal preocupação da nova Lei de Diretrizes e Bases era com a profissionalização, buscando a terminalidade profissional no nível médio.

É nesse período da década de 1970, que Alves (2009) situa o crescimento do ensino privado, com a seguinte consideração:

Em função do crescimento demográfico da classe média, a partir dos anos 1950 e 1960, e da incapacidade da escola confessional em acompanhar o crescimento da demanda, os estabelecimentos escolares mantidos por educadores e empresários da educação se veem em vertiginoso crescimento a partir dos anos 1970. O desenvolvimento econômico do país, a precária qualidade do ensino público e a crescente qualidade do serviço educacional prestado pela escola particular propiciam o crescimento do setor privado, tendo à frente empresários e dirigentes extremamente competentes e organizados na gestão educacional. É o *boom* do ensino privado, em um primeiro momento na educação básica, e, a partir dos anos 1990, também no ensino superior. (ALVES, 2009, p. 75).

Vale destacar que a ideia de qualidade educacional neste período está associada principalmente à questão da oferta de vagas e acesso, como está descrito no capítulo seis deste estudo. Portanto, a ideia de precariedade da qualidade do ensino público precisa ser considerada e discutida à luz deste contexto histórico.

Gomes (2016), ao descrever o cenário nacional no início da década de 1980, pré-constituente, relembra a elevada inflação, o aviltamento do câmbio da moeda nacional, a campanha de Diretas Já e a dura recessão econômica que empobrecia o país, obrigando os governos a efetuarem cortes orçamentários, especialmente no setor social, sendo a educação um dos alvos dessa medida. É nesse contexto, que uma Assembleia Nacional Constituinte é organizada com o desafio hercúleo de ser aberta a ouvir sugestões de legisladores e não partir de um anteprojeto, como a de 1946. Para compreender áreas tão específicas como economia, educação, saúde, orçamento, cultura etc., além da necessidade de buscar apoio em um corpo técnico-científico das áreas, a estrutura de subcomissões temáticas que apresentavam sugestões às comissões temáticas, que por sua vez, apresentavam à comissão de sistematização e, finalmente ao plenário, garantia um processo mais democrático, porém, longo. O autor, que atuou como assessor legislativo do Senado Federal durante esse período afirma:

Na educação, a base da Subcomissão da Educação, Cultura e Esporte, que, em correspondência com as lutas históricas, logo se dividiu em defensores dos setores público e privado, como um duradouro pomo de discórdia, interessante como bandeira política e bem menos proveitoso do ponto de vista financeiro, como a Assessoria evidenciou aos senhores constituintes, pelo seu volume histórico. (GOMES, 2016, p. 58).

O autor usou a expressão “falha tectônica básica” (p. 59) para se referir ao conflito público *versus* privado que se desencadeou durante todo o processo da constituinte e estendeu-se ao longo do tempo de elaboração da nova Lei de Diretrizes e Bases da Educação que, por oito anos, ainda seria gestada, após a promulgação da Carta Magna, em 1988.

Brzezinski (2010) descreve esse momento:

No período de tramitação do projeto da LDB atual (1989-1996), o conflito público *versus* privado, a disputa entre dois projetos de sociedade e de educação voltaram ao cenário político-ideológico, porém marcaram presença novos atores. Um fenômeno contribuiu para isto: as escolas básicas confessionais católicas e algumas escolas públicas foram paulatinamente retirando-se do cenário educacional, para dar lugar às escolas privadas laicas em decorrência das profundas transformações da sociedade brasileira e das relações de produção que passaram a favorecer a versão perversa do capitalismo globalizado, neoliberal, cujo maior regulador é o mercado. Muitos aspectos então impulsionaram uma nova força hegemônica no campo educacional, com nítida tendência de conceber as instituições educacionais tanto da Educação Básica quanto da Educação Superior como empresas de ensino, e estas mobilizaram um grupo de atores sociais configurado pelos proprietários de escolas particulares e seus lobistas que se pautaram exclusivamente pela lógica do lucro. (BRZEZINSKI, 2010, p. 190).

Nesse cenário de disputa ideológica retratado pelos autores, o ordenamento jurídico brasileiro que estabelece a estrutura normativa da educação foi constituído e mantém-se desde o final do século XX. Sobre essa tensão na relação público x privado, Cury (2018) considera que:

A liberdade de ensino é considerada tanto um direito civil do indivíduo e da família quanto um princípio da educação. A relação entre educação pública e educação privada não é neutra, já que se trata da presença dentro do campo da educação, um campo de disputas. Nesse sentido, a inserção de princípios no ordenamento legal torna a relação coberta de tensão. (CURY, 2018, p. 871).

Esses aspectos históricos nos oportunizam a percepção de que ainda caminhamos em um campo de tensão, quando nos debruçamos sobre um objeto de estudo vinculado a um sistema privado de ensino. Porém, com a chancela normativa que o garante e que consideraremos a seguir.

#### **4. 2 O ordenamento jurídico constitucional e na LDB para o ensino privado**

Como descrito, a presença das escolas particulares na educação nacional foi constante, mas variável nos distintos períodos históricos. Além do que, sempre tiveram sua figura reconhecida em todas as constituições brasileiras, que garantiram a “liberdade de ensino”; liberdade esta, que foi limitada pela obediência à legislação educacional que impunha certo controle, dependendo da própria conjuntura histórica.

Quanto ao ensino regular e privado, a Constituição, neste primeiro quarto de século, estabelece nos artigos 206 e 209 que:

Art. 206 – O ensino será ministrado com base nos seguintes princípios:

[...] III – Pluralismo de ideias e de concepções pedagógicas e coexistência de instituições públicas e privadas de ensino;

[...]

Art. 209 – O ensino é livre à iniciativa privada, atendidas as seguintes condições:

I – cumprimento das normas gerais da educação nacional;

II – autorização e avaliação de qualidade pelo poder público. (BRASIL, 1988)

Cury (2009, p.3) considera que “a atual formulação legal se diferencia um pouco de legislações passadas em que a educação escolar privada era uma ‘concessão’ do Estado. Ao relacionar a condição de coexistência institucional prevista no inciso III do Art. 206, com a condição estabelecida no inciso II do Art. 209, o autor afirma que a categoria concessão foi deslocada para a categoria autorização e avaliação de qualidade. Dessa forma, “há pressupostos para a inserção da iniciativa privada na educação escolar regular e condições para seu estabelecimento. Submetendo-se aos processos de autorização e avaliação [...]”. (CURY, 2009, p. 3).

A LDB 9394 de 1996 adicionará outras condições para a existência e estabelecimento do ensino pela iniciativa privada, conforme descrito nos artigos abaixo:

Art. 7 – O ensino é livre à iniciativa privada, atendida as seguintes condições:

I – cumprimento das normas gerais da educação nacional e do respectivo sistema de ensino;

II – autorização de funcionamento e avaliação de qualidade pelo Poder Público;

III – capacidade de autofinanciamento, ressalvado o previsto no Art. 213 da Constituição Federal;

[...]

Art. 19 – As instituições de ensino dos diferentes níveis classificam-se nas seguintes categorias administrativas:

[...]

II – privadas, assim atendidas as mantidas e administradas por pessoas físicas ou jurídicas de direito privado;

III – comunitárias, na forma da lei.

§ 1º - As instituições de ensino a que se referem os incisos II e III do caput deste artigo podem qualificar-se como confessionais, atendidas a orientação confessional e a ideologia específicas. (BRASIL, 1996).

Pode-se identificar, no ordenamento legal, que entre as escolas privadas existem as que seguem os mecanismos do sistema contratual de mercado e constituem-se com fins lucrativos, e as que, embora tenham mecanismos privados de oferta, não são lucrativas. Estas, diferenciam-se entre si como filantrópicas, confessionais e comunitárias.

Ao discorrer sobre a diferença entre concessão e autorização que distingue a atual relação do Estado com as instituições privadas de ensino, Cury (2009) considera que a educação escolar é um serviço público de Estado, que tem a titularidade plena e o exercício pleno da função de ensinar. “Já, a iniciativa privada possui uma titularidade compartilhada mediante condições, entre as quais a autorização e a avaliação da qualidade.” (CURY, 2009, p. 6). O autor, discorrendo sobre essa coexistência, afirma:

O ensino privado é, pois, um serviço público (impróprio) autorizado e cuja consequência maior – aliás, posta em lei – é o respeito às normas gerais da educação e a avaliação de qualidade, de acordo com o artigo 209 da Constituição e o artigo 7 da LDB. Desta forma, o poder público credencia os executores deste bem, autorizando-os e exercendo adequada fiscalização sobre eles. (CURY, 2009, p. 6).

Ainda, para o autor:

É nesse sentido que a avaliação talvez seja a mais profunda ruptura estabelecida na atual organização nacional. Ela entra na Constituição e de certo modo substitui o conceito de concessão. A avaliação ainda não se constitui como uma cultura, o que exigiria processos mais permanentes e consolidados. Contudo, pode-se dizer que ela já foi acolhida como princípio e tem demonstrado seu peso na busca de maior legitimidade por parte das instituições. (CURY, 2009, p. 13).

Diante da atual formulação legal, a coexistência do público e do privado nas instituições de ensino ainda se caracteriza por um campo de tensão. Como ressalta Brzezinski (2010, p. 203), “não foram abrandadas as contradições e as disputas entre o ensino público e o ensino privado em nosso país”. Porém, o padrão de referência exigido pela Constituição e reafirmado pela LDB, de qualidade educacional, é para ambos os modelos de entidade: pública ou privada. Avaliar a qualidade é, cada vez mais, uma condição, não só para a autorização, mas essencialmente para a continuidade e permanência de um sistema de ensino: público ou privado.

Em face do exposto, faz-se necessário pontuar que o uso das palavras sistemas e redes de ensino não é uma opção isenta de possíveis equívocos ou livre de suscitar diferentes acepções. Saviani, desde sua tese de doutoramento, em 1978, já apontava o uso equivocado da palavra sistemas para a área educacional no Brasil e, após a promulgação da LDB, analisando detalhadamente o novo ordenamento jurídico, considerou

O termo “sistema”, em relação à educação, é empregado com acepções diversas, o que lhe confere um caráter de certo modo equívoco. No entanto, partindo da educação como fenômeno fundamental, é possível superar essa aparência e captar o seu verdadeiro sentido. Com efeito, a educação aparece como uma realidade irreduzível nas sociedades humanas. Como assistemática, ela é indiferenciada, ou seja, não se distinguem ensino, escola, graus, ramos, padrões, métodos etc. Quando o homem sente a necessidade de intervir nesse fenômeno e erigi-lo em sistema, então ele explicita sua concepção de educação enunciando os valores que a orientam e as finalidades que preconiza, sobre cuja base se definem os critérios de ordenação dos elementos que integram o processo educativo. E surgem as distinções: ensino (como transmissão de conhecimentos e habilidades), escolas (como locais especialmente preparados para as atividades educativas), articulação vertical e horizontal (graus e ramos) etc. (SAVIANI, 1999, p. 120).

Tomando como premissa o conceito de sistema como a “unidade de vários elementos *intencionalmente* de modo a formar um conjunto coerente e operante” (SAVIANI, 1999, p. 120) (grifo do autor), Saviani considera que um sistema educacional envolve um conjunto de atividades que se cumprem, tendo em vista uma finalidade e, para tanto, precisam ser organizadas mediante normas próprias que lhe confirmam um grau de autonomia e que sejam

comuns a todos os seus integrantes. Por essa razão, o autor conclui que só podemos considerar sistema educacional, aquele que tem legitimidade para legislar, ou seja, a esfera pública. A esfera privada não deve, na concepção de Saviani (1999), ser chamada de sistema e sim de rede, pois a ela, não compete legislar. O uso da expressão “sistema privado” só deveria ocorrer por analogia. Porém, o abuso de analogias da palavra gera imprecisões. Sugere, ainda, que os legisladores deveriam ter um esclarecimento preciso do conceito de sistemas para evitar dificuldades na interpretação e clareza dos textos legais.

Em virtude dessas possíveis imprecisões, a Câmara de Educação Básica do Conselho Nacional de Educação homologou, no ano 2000, o Parecer nº 30, estabelecendo que sistemas de ensino

são o conjunto de campos de competências e atribuições, voltadas para o desenvolvimento da educação escolar que se materializam em instituições, órgãos executivos e normativos, recursos e meios articulados pelo poder público competente, abertos ao regime de colaboração e respeitadas as normas gerais vigentes. (CNE, 2000)

Ao analisar essa temática, Macedo-Silva, Silva e Santos (2017) consideram que muitos confundem o uso das palavras rede e sistema de ensino, como se fossem a mesma coisa, sendo que podem soar parecidos, mas não são iguais. Ambos “são compostos por órgãos que devem funcionar por um mesmo objetivo.” Mas a rede é parte do sistema. Afirmam então, que “rede de ensino é um sistema sustentado por um mesmo tipo de fonte financeira, que pode ter algum regimento complementar, além daquela prevista em Lei.” (MACEDO-SILVA; SILVA; SANTOS, 2017, p. 4).

A rede privada, portanto, está contida no sistema público de ensino, sendo composta por instituições “mantidas e administradas por pessoas físicas ou jurídicas de direito privado” (BRASIL, 1996).

Do mesmo modo, Saviani (1999) considerou que a competência de legislar em matéria de educação, desenvolver normas próprias que organizem o respectivo sistema de ensino é que o caracteriza como tal. Do contrário, o fato de ter um único mantenedor e financiador, indiferentemente público ou privado, mas sem a competência de criar as próprias normas de organização e funcionamento, caracterizaria uma rede pública ou privada, e não um sistema.

Em face do exposto, neste estudo, utilizamos a nomenclatura rede para um conjunto de instituições de ensino sustentadas pela mesma mantenedora, com objetivo comum e regimento comum. Desta forma, distinguimos rede de estabelecimentos privados de ensino, cuja pessoa física ou jurídica que o mantém administra uma única instituição. Estes fazem

parte da rede privada de determinado sistema educacional, mas não se constituem em si, uma rede de ensino.

### 4.3 Cenário do ensino privado no final da segunda década do século XXI

O cenário descrito a seguir leva em conta, basicamente, aspectos quantitativos da presença do ensino privado na educação básica brasileira, na segunda década do século XXI. Além dos dados que indicam a quantidade de estabelecimentos privados e matrículas, estudos recentes apontam o crescimento do segmento conforme descritos a seguir.

O Censo Escolar, pesquisa estatística que oferece um diagnóstico sobre a educação básica e que é publicada anualmente pelo Inep, apresentou em 2019 os dados que seguem (Tabela 3) sobre estabelecimentos e matrículas do ensino privado, referentes ao ano de 2018.

**Tabela 3 - Educação Básica no Brasil – Estabelecimentos e Matrículas**

<b>Unidade federativa</b>	<b>Estabelecimentos Total</b>	<b>Estabelecimentos Privados</b>	<b>%</b>	<b>Alunos Total</b>	<b>Alunos Escolas Privadas</b>	<b>%</b>
Acre	1.566	42	2,6%	264.593	12.486	4,7%
Alagoas	2.985	593	19,8	865.501	150.526	17,3%
Amapá	841	80	9,5%	220.269	21.310	9,6%
Amazonas	5.318	307	5,7%	1.165.535	87.921	7,5%
Bahia	16.682	2.835	16,9%	3.485.631	574.868	16,4%
Ceará	7.519	1.603	21,3%	2.161.816	418.273	19,3%
D. Federal	1.237	577	46,6%	657.869	195.716	29,7%
Esp. Santo	3.084	352	11,4%	881.826	107.358	12,1%
Goiás	4.629	1.134	24,4%	1.447.842	278.532	19,2%
Maranhão	11.748	1.067	9%	1.993.909	237.326	11,9%
Mato Grosso	2.708	430	15,8%	880.844	110.979	12,5%

M. G. do Sul	1.737	419	24,1%	680.108	91.860	13,5%
Minas Gerais	16.058	3.987	24,8%	4.364.668	700.911	16%
Pará	10.757	976	9%	2.294.276	235.365	10,2%
Paraíba	4.868	932	19,1%	969.806	195.188	20,1%
Paraná	9.511	2.180	22,9%	2.572.007	455.405	17,7%
Pernambuco	8.502	2.346	27,5%	2.232.556	524.547	23,4%
Piauí	4.459	462	10,3%	895.308	112.476	12,5%
Rio de Janeiro	11.325	4.749	41,9%	3.573.417	1.111.732	31%
R.G. do Norte	3.485	674	19,3%	820.485	171.324	20,8%
R.G. do Sul	9.946	2.672	26,8%	2.294.325	440.229	19,1%
Rondônia	1.241	158	12,7%	416.212	43.503	10,4%
Roraima	849	49	5,7%	166.147	15.778	9,4%
Sta. Catarina	6.238	1.086	17,4%	1.610.086	273.024	16,9%
São Paulo	29.611	11.128	37,5%	10.018.115	2.403.782	23,9%
Sergipe	2.098	422	20,1%	543.464	120.406	22,1%
Tocantins	1.608	174	10,8%	397.631	43.960	11%
Brasil	180.610	41.434	22,9%	47.874.246	9.134.785	19%

Fonte: Sinopse Censo Escolar 2019/ INEP; elaboração da autora.

É, portanto, de se considerar que cerca de  $\frac{1}{4}$  dos estabelecimentos de ensino brasileiros e  $\frac{1}{5}$  das matrículas estão na rede de ensino privada. Os dados representam as escolas localizadas em zona urbana e rural e todos os segmentos que compreendem a educação básica: educação infantil, ensino fundamental e ensino médio. Portanto, as escolas privadas correspondem a cerca de 20% dos estabelecimentos e do número de alunos, de todo o Brasil. Da Tabela 3, vale destacar que, percentualmente, o Distrito Federal, os estados do Rio de

Janeiro e de São Paulo possuem maior índice de estabelecimentos e, conseqüentemente, de matrículas na rede privada, estando os três bem acima da média nacional.

Se levarmos em conta os números absolutos, temos os Estados de São Paulo, Rio de Janeiro e Bahia com o maior número de estabelecimentos, e São Paulo, Rio de Janeiro e Minas Gerais com o maior número de matrículas na rede privada. Os estados com o menor percentual de estabelecimentos privados e de matrículas, em relação ao total de escolas, são Acre, Roraima e Amazonas.

As unidades federativas com maior percentual de escolas privadas e matrículas são também as que apresentam os maiores índices de rendimento mensal domiciliar per capita da população, de acordo com dados disponibilizados pelo Instituto Brasileiro de Geografia e Estatística (IBGE), referentes ao ano de 2019, conforme consta no Quadro 10.

**Quadro 10 - Melhor rendimento mensal domiciliar per capita (2019)**

<b>Posição</b>	<b>Unidade Federativa</b>	<b>Valor</b>
1°	Distrito Federal	R\$ 2.686,00
2°	São Paulo	R\$ 1.946,00
3°	Rio de Janeiro	R\$ 1.882,00

Fonte: IBGE, 2019; elaboração da autora.

Outrossim, as unidades federativas com o menor percentual de estabelecimentos privados e matrículas (Acre, Roraima e Amazonas) não são as que apresentam menor rendimento mensal domiciliar, conforme os dados do IBGE, descritos no Quadro 11:

**Quadro 11 - Pior rendimento mensal domiciliar per capita (2019)**

<b>Posição</b>	<b>Unidade Federativa</b>	<b>Valor</b>
25°	Pará	R\$ 807,00
26°	Alagoas	R\$ 731,00
27°	Maranhão	R\$ 636,00

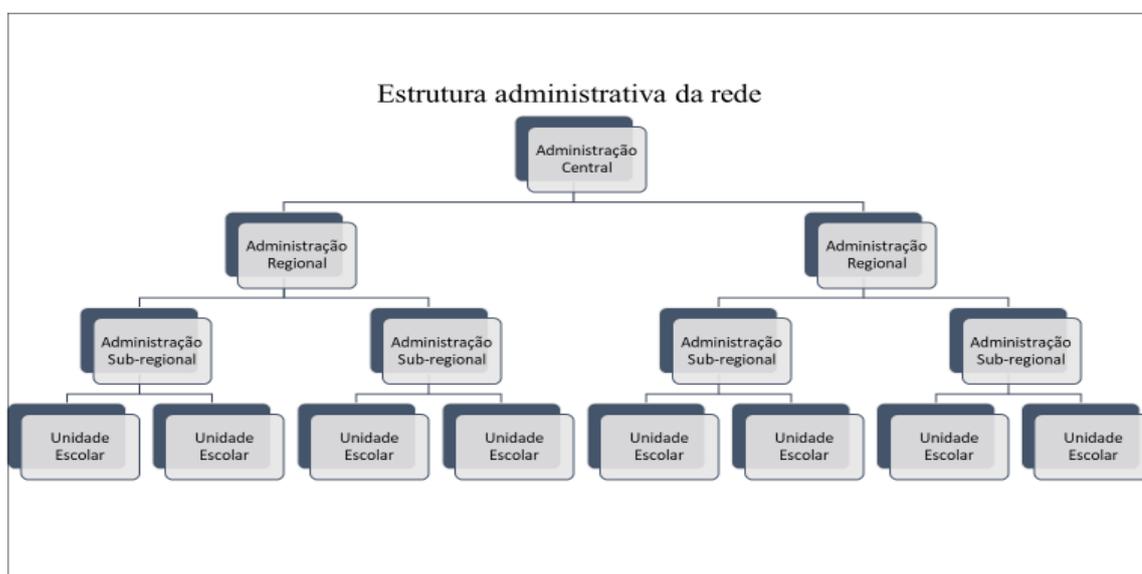
Fonte: IBGE, 2019; elaboração da autora

Porém, há estudos recentes que estabeleceram correlação entre o aumento significativo de escolas privadas de educação básica e o aumento da renda das famílias brasileiras nas primeiras duas décadas do século XX.

#### 4. 4 O Programa de Avaliação Externa da Rede (PAER)<sup>25</sup>

A rede de ensino objeto deste estudo está presente em todas as unidades da federação, embora com uma concentração maior de escolas nas regiões sul e sudeste. No ano de 2019, contava com cerca de 350 unidades de educação básica e 230.000 alunos. Esses números representam, aproximadamente, 1,5% de todas as escolas privadas do Brasil e 2,4% do total de alunos matriculados no ensino privado no país. Essa instituição atua na educação básica, desde a última década do século XIX, e possui também unidades dedicadas ao ensino superior, desde a primeira metade do século XX. É de natureza filantrópica e confessional com uma estrutura administrativa que mantém uma administração central para toda a América do Sul, que inclui o Brasil e mais oito países. Esses países são subdivididos em macrorregiões administrativas e estas, por sua vez, são divididas em sub-regiões. O Brasil está dividido em oito regiões administrativas que se subdividem em cerca de 40 sub-regiões, as quais lideram diretamente as atividades das unidades escolares (PROJETO PEDAGÓGICO, 2019). Esta estrutura administrativa está representada no organograma ilustrado na Figura 1.

**Figura 1 - Organograma da estrutura administrativa da rede**



Fonte: DO, 2017; elaboração da autora.

<sup>25</sup> Os dados sobre o PAER foram obtidos em documentos disponibilizados pela Administração do Programa e que foram enviados às unidades escolares, administrações regionais, alunos e pais desde 2012, com uso devidamente autorizado à pesquisadora e identificados no texto.

Na administração central, o gestor da rede para todo o Brasil é também o diretor do PAER. É assessorado por uma equipe de três pessoas que inclui: um pedagogo, um estatístico e um pesquisador da área de ciências sociais. Também há uma equipe que gerencia a tecnologia da informação e mantém uma plataforma que disponibiliza os dados para toda a rede.

Em cada sede de administração regional há um gestor e uma coordenadora pedagógica que implementam as ações ligadas ao PAER em sua região. Da mesma forma, em cada sub-região administrativa há um gestor e uma coordenadora pedagógica que respondem pelas ações ligadas ao Programa em contato direto com as unidades escolares. Com este modelo de estrutura administrativa, a rede atua com uma capilaridade que garante um padrão de procedimentos e processos na operacionalização do Programa, expressos em um documento organizador elaborado em 2017 (DO).

No ano de 2011, a administração central da rede convidou um grupo de educadores de algumas regiões do Brasil para trabalhar nas bases do Programa de Avaliação Externa da Rede, visando a sua implantação em 2012, com testes cognitivos de Língua Portuguesa e Matemática para o 5º ano do ensino fundamental. Em 2013, foi estendido para o 9º ano, e em 2014, o ciclo foi concluído com o 2º ano do ensino médio (DO, 2017). Com aplicação anual, em 2019, o Programa chegou a sua 8ª edição.

Trata-se, portanto, de um programa de aplicação anual, censitária, com aplicação de provas de Língua Portuguesa e Matemática nos estudantes de 5º e 9º anos do Ensino Fundamental e provas nas quatro áreas do conhecimento para o 2º ano do Ensino Médio. As provas podem ser impressas ou *on-line*, cabendo à unidade escolar a escolha. Acompanha um questionário contextual para alunos e professores das áreas e anos avaliados. Ao longo das edições, demonstra uma estabilidade de aplicação das provas, tendo sido implementado todos os anos, ininterruptamente. O documento organizador apresenta o PAER, resumidamente, assim:

O [...] é um projeto do [...], criado e implementado a partir do ano de 2012, com o aval e apoio de todas as sedes administrativas a nível do país. Tem como propósito desenvolver um sistema de avaliação externa de escolas e colégios da rede [...] de educação, em todo o território da divisão, procurando produzir informação que permita assegurar a qualidade da educação oferecida. Para isto, uma equipe estável de especialistas membros do programa aplica anualmente diferentes tipos de avaliações a alunos, professores e outros participantes do processo educativo. Os resultados são analisados pela mesma equipe responsável, e as avaliações oportunamente compartilhadas com os gestores da [...] em todos os níveis. (DO, 2017, p. 3).

Como justificativa para a criação do programa, o documento descreve:

Um dos objetivos da educação [...] consiste em garantir uma educação de qualidade a todos os seus estudantes. Implícito na ideia de qualidade da educação está o

conceito de equidade, como princípio que promove justiça, imparcialidade e igualdade de oportunidades para todos. Neste sentido, mais do que alcançar um objetivo em um dado momento, a qualidade educacional deve ser vista como uma meta constante [...]. Os programas de avaliação externa têm sido criados justamente para contribuir na verificação do crescimento permanente antes mencionado. Os governos de muitos países e organismos internacionais têm implementado sistemas e mecanismos de avaliação com o propósito de recolher evidências do progresso obtido no caminho da qualidade e equidade da educação. No caso do Brasil, e as diferentes iniciativas de avaliação coordenadas pelo Instituto Nacional de Estudos e Pesquisas Educacionais “Anísio Teixeira” (INEP); ou o Programa Internacional de Avaliação de Estudantes (PISA, em sua sigla em inglês), coordenado pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE), como dois exemplos. (DO, 2017, p. 4)

“Retratar a realidade da qualidade da educação”, conforme os princípios e filosofia de educação norteadores da rede, “através de informações que incentivarão a adoção de medidas que promovam o crescimento”, é descrito como o objetivo geral do PAER (DO, 2017, p. 5). Os objetivos específicos derivados do objetivo geral são:

- 1) Promover a elaboração e adoção de modelos de qualidade da educação, de acordo com a filosofia de educação [...], a missão da organização, e o contexto sociocultural de cada região.
- 2) Planificar, desenvolver e aplicar as avaliações do programa nas instituições educacionais de nível fundamental e médio da rede.
- 3) Estudar e propor melhorias nas avaliações e processos desenvolvidos pelo programa.
- 4) Divulgar o programa entre os diferentes membros da comunidade educacional, usando diferentes estratégias de comunicação.
- 5) Prover informação para a tomada de decisões na gestão estratégica da educação básica em nível nacional.
- 6) Criar condições para proteger a integridade e uso adequado dos dados produzidos pelo programa.
- 7) Sugerir estratégias que permitam estender a adoção de boas práticas detectadas no sistema de educação, e propor medidas que permitam superar as deficiências reveladas pelo programa de avaliação.
- 8) Difundir e comunicar os resultados das avaliações através de documentos científicos em revistas, seminários, congressos, conferências e *workshops*. (DO, 2017, p. 5).

Destacamos, portanto, que promover a qualidade da educação é o compromisso reafirmado no primeiro objetivo específico. Não há, contudo, nenhum esclarecimento no documento sobre a concepção de qualidade educacional da rede, nem de qual crescimento se trata. Esta questão será retomada na análise dos dados no capítulo seis.

Por ocasião da apresentação dos resultados da primeira edição do PAER, a equipe de coordenação do Programa enviou uma carta às administrações regionais, sub-regionais e às unidades escolares, no início do ano de 2013, com as explicações da sistemática da apresentação dos resultados e as bases dos cálculos das médias e da representação em níveis de proficiência dos alunos e, do desempenho das unidades escolares. Nessa carta, a equipe informou que “a elaboração e os resultados da **Prova do PAER** seguem o mesmo rigor científico da **Prova Brasil**” (Carta, 2013) [grifo do autor]. Nos dois primeiros anos de

implantação do Programa, a rede estabeleceu uma parceria com o Centro de Seleção e Promoção de Eventos (CESPE) da Universidade de Brasília, em seu segmento de avaliação educacional. Toda a infraestrutura de aplicação das provas, o banco de questões, o *know-how* disponível referente ao trabalho desenvolvido pelo Centro nas diversas avaliações do Inep, foram fornecidos para a equipe gestora da rede no processo de implantação do PAER.

Portanto, o Programa teve toda a estrutura de aplicação de provas, correção e disposição dos resultados, baseados nos princípios da Prova Brasil, utilizando a mesma descrição dos níveis da escala de desempenho de Matemática e Língua Portuguesa do Saeb, com as habilidades indicadas para cada nível. Utiliza-se também da TRI para obtenção da proficiência dos alunos, e a forma de divulgação dos resultados é semelhante à da Prova Brasil, com as mesmas escalas de proficiência.

Desde a terceira edição, um banco de itens está sendo montado com uma equipe de educadores da rede, representando as várias regiões do Brasil. De acordo com o documento de apresentação do PAER, os itens são pré-testados e calibrados em diferentes graus de dificuldade e possuem bom parâmetro de qualidade. Juntamente com o caderno de provas e a folha de respostas, cada unidade recebe: a) Carta aos pais, informando o objetivo da avaliação e o uso restrito dos resultados; b) Termo de compromisso para o recebimento do material e das provas que deve ser assinado pelo coordenador pedagógico da unidade; c) Termo de Compromisso para o aplicador das provas que não deve ser o professor dos alunos; d) Manual de atribuições do aplicador da prova; e) Formulário de controle da aplicação das provas. Desde 2016, as provas estão também sendo aplicadas *on-line* para as escolas que desejarem participar por esse meio. É possível optar com antecedência, considerando se a unidade possui equipamentos e internet estável para que todos os alunos realizem a prova *on-line* nas mesmas condições de dia, horário e duração da prova impressa.

Em todas as edições do PAER, a coordenação central tem enviado uma carta explicativa aos pais com o intuito de apresentar os objetivos do Programa e solicitar o apoio da família para participação do aluno. Na carta aos pais está descrito que o objetivo do Programa “é contribuir para o contínuo aperfeiçoamento da qualidade acadêmica da Educação Básica da Rede [...]” e que o PAER “possibilita retratar a realidade do desempenho acadêmico de cada aluno, de cada escola, nos municípios, nos estados e no Brasil como um todo.” (CARTA, 2012). Além de apresentar o objetivo, a carta informa que o resultado da avaliação deve ser utilizado para o planejamento pedagógico da unidade escolar e da rede, deixando clara a não interferência dessa avaliação na avaliação escolar do aluno, “isso porque a prova

PAAEB visa avaliar e contribuir para aprimorar a qualidade da Educação da rede” (CARTA, 2012).

A Tabela 4 descreve a edição de 2019, como a quantidade de questões por prova e os anos em que são aplicadas.

**Tabela 4 - Provas do PAER – edição de 2019**

Avaliação Censitária Anuais	Ano	Disciplinas	Instrumento	Nº de questões
	5º EF	Língua Portuguesa	Prova objetiva	22
		Matemática	Prova objetiva	25
	9º EF	Língua Portuguesa	Prova objetiva	30
		Matemática	Prova objetiva	23
	2º EM	Linguagem e seus códigos	Prova objetiva	23
		Matemática	Prova objetiva	23
		Ciências da Natureza	Prova objetiva	23
		Ciências Humanas	Prova objetiva	23

Fonte: Dados disponibilizados pela equipe gestora; elaboração da autora.

Como mencionado, prova é acompanhada de um questionário contextual para todos os alunos participantes, contendo 32 questões para alunos do 5º e do 9º ano do ensino fundamental e o mesmo número para alunos do 2º ano do ensino médio. Também é enviado um questionário para os professores que atuam no 5º ano, os professores de Língua Portuguesa e Matemática do 9º ano e para os professores das áreas do conhecimento avaliadas no 2º ano do ensino médio. O questionário dos professores é composto de, aproximadamente, 30 questões.

O número médio de participantes a cada edição é de cerca de trinta mil alunos, conforme pormenoriza o Quadro 12.

**Quadro 12 - Número de participantes nas edições do PAER**

Ano	PAAEB			
	PAAEB TOTAL	5 ANO	9 ANO	2EM
2012	11907	11907		
2013	25252	12981	12271	
2014	33091	13584	12870	6637
2015	33559	14438	11766	7355
2016	33379	14198	12158	7023
2017	33790	14988	12318	6484
2018	35685	15828	12736	7121
2019	37933	17283	13565	7085
2020	35136	16176	12308	6652
<b>TOTAL</b>	<b>279732</b>			

Fonte: Equipe do PAER 2020

De acordo com o documento que estruturou o PAER (DO, 2017), as provas são aplicadas no mesmo dia e horário para todas as unidades educacionais do país. Quando finalizadas, inicia-se um processo de devolução dos materiais utilizados durante a prova pelo correio para a coordenação do Programa, e as folhas de respostas dos alunos que realizaram as provas na modalidade escrita são digitalizadas na própria unidade e enviadas através do sistema do programa para o imediato processamento. Após o processamento de todas as respostas, a equipe de coordenação do PAER realiza as análises psicométricas e estatísticas correspondentes. Uma vez completadas essas análises, os resultados passam a ser compartilhados com os gestores de várias maneiras, conforme o nível de gestão: administração central; administração regional; administração sub-regional; gestores da unidade escolar e alunos. As formas de divulgação dos resultados estão representadas no Quadro 13:

**Quadro 13 - Formas de divulgação dos resultados do PAER**

<b>Formas de Divulgação</b>	<b>Finalidade</b>	<b>Destinatário</b>
Sistema do próprio PAER	Informar aos gestores de cada nível administrativo os resultados por edição, por ano, por disciplina, por tema, por habilidade, de cada unidade e por aluno, tanto da prova como do questionário contextual. Permite gerar gráficos e tabelas.	<ul style="list-style-type: none"> <li>• Administração Central;</li> <li>• Administração Regional;</li> <li>• Administração Sub-regional;</li> <li>• Unidade Escolar</li> </ul>
Painéis de Controle ( <i>Dashboards</i> )	Permitir a visualização de grande quantidade de dados em forma gráfica e útil para apreciar a evolução de certos indicadores, particularmente os das avaliações cognitivas.	<ul style="list-style-type: none"> <li>• Administração Central;</li> <li>• Administração Regional;</li> <li>• Administração Sub-regional;</li> <li>• Unidade Escolar</li> </ul>
Informes escritos pela equipe do PAER	O detalhamento das análises dos resultados com comentários que ressaltam aspectos a serem destacados ou que requerem atenção, principalmente para os construtos que compõem o questionário contextual.	<ul style="list-style-type: none"> <li>• Administração Central;</li> <li>• Administração Regional</li> </ul>
Assessoria	Assessorar as administrações regionais no uso	<ul style="list-style-type: none"> <li>• Administração</li> </ul>

presencial da equipe do PAER	de todos os recursos disponibilizados para análise dos resultados como ferramenta de gestão.	Regional
Carta impressa para alunos	Descrever o desempenho do aluno por nível, por disciplina, em cada habilidade requerida nas provas.	<ul style="list-style-type: none"> <li>• Alunos</li> </ul>

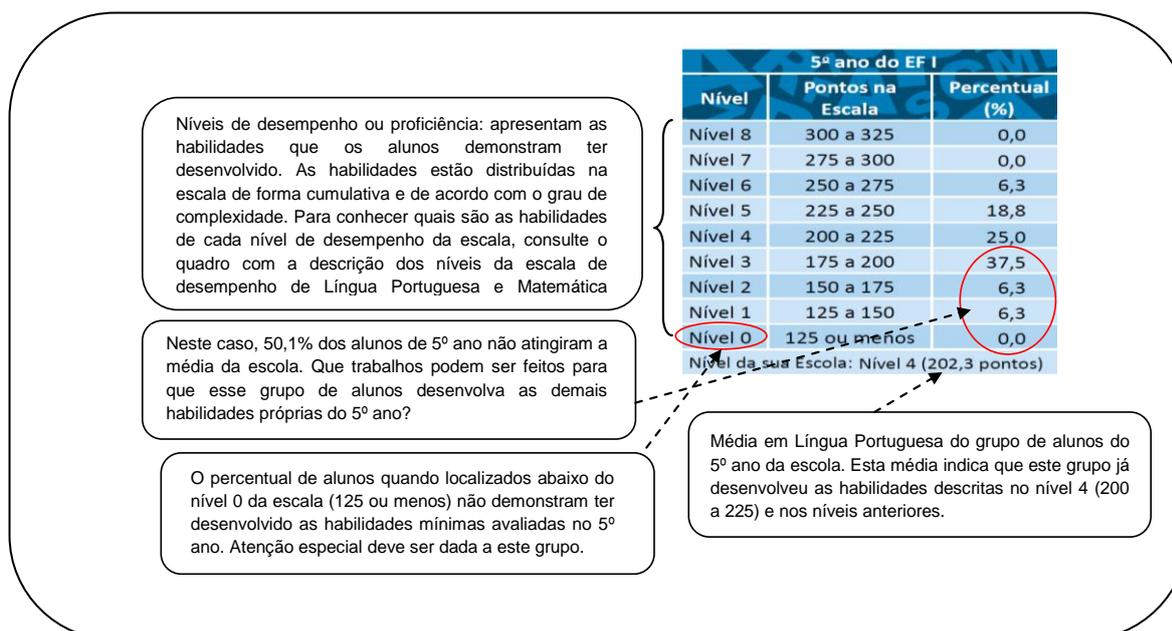
Fonte: DO (2017); elaboração da autora.

Conforme registrado no Quadro 13, há várias formas de os resultados serem compartilhados e acessados. A principal delas é a plataforma do Programa, à qual apenas os usuários registrados em cada nível administrativo podem ter acesso. A unidade escolar, via plataforma, acessa e imprime as cartas destinadas aos alunos, com o resultado individual.

Vale destacar que a visualização do resultado é restrita ao nível hierárquico da esfera administrativa, ou seja, apenas a administração central visualiza todos os resultados. Cada sede de administração regional visualiza apenas os resultados das estruturas a ela subordinadas (conforme demonstrado na Figura 1). Não é possível que o gestor visualize as informações das demais estruturas semelhantes a dele num sentido horizontal do organograma. Não há, nos documentos, uma justificativa para esta regra que será objeto de análise no capítulo seis deste estudo.

Como o PAER foi estruturado com as mesmas bases da Prova Brasil, a forma de apresentação dos resultados utiliza a descrição em uma escala construída a partir da média de proficiência ou média de desempenho, que evidencia as habilidades desenvolvidas pelos alunos. Para que os gestores pudessem compreender a forma de interpretação dos níveis de desempenho representados nas escalas, os resultados iniciais foram acompanhados por um documento denominado “Descrição dos níveis da escala de desempenho” produzido pelo Inep para o Saeb, utilizados para auxiliar na leitura do desempenho apresentado pelos alunos, pois utiliza como referência os níveis expressos em números que vão de 0 a 500 com um intervalo de 25 a 25 pontos para cada nível, tanto em Matemática como em Língua Portuguesa. O valor numérico de cada nível varia de acordo com a escala de proficiência de cada etapa escolar, de forma crescente. Por essa escala, portanto, a média é 250, e o desvio padrão, 50. Em carta explicativa enviada aos gestores das regiões, sub-regiões e unidades escolares para favorecer a interpretação dos resultados, foi apresentada a Figura 2.

**Figura 2 - Exemplo de distribuição percentual dos alunos na escala de proficiência**



Fonte: Carta de resultado do PAER aos gestores de sub-regiões (2013, p. 3).

O boletim de resultados apresenta a proficiência média dos alunos de uma unidade ou região administrativa em comparação com o desempenho médio dos demais níveis administrativos e a quantidade de participantes. A seguir, temos a Figura 3 extraída de um boletim de desempenho de uma unidade escolar, no ano de 2013, com as médias de Língua Portuguesa e Matemática, e o número de participantes em cada nível administrativo superior à unidade escolar, até o Brasil.

**Figura 3 - Comparativo de desempenho dos alunos e participantes de uma unidade escolar e os demais níveis administrativos da rede**

5º ano do EF I		Quantos participaram		9º ano do EF II	
<b>Estudantes participantes</b>				<b>Estudantes participantes</b>	
12.807		Escolas	Brasil	12.129	
4.424		Escolas	União	4.346	
416		Escolas	Associação/	475	
43		Da sua Escola		62	
Língua Portuguesa	Matemática	<b>Proficiências Médias</b>		Língua Portuguesa	Matemática
220,5	219,0	Escolas	do Brasil	268,6	280,3
221,8	224,3	Escola	União	270,5	282,9
224,9	230,5	Escolas	Associação/	280,9	300,2
233,0	231,4	Da sua Escola		286,2	305,6

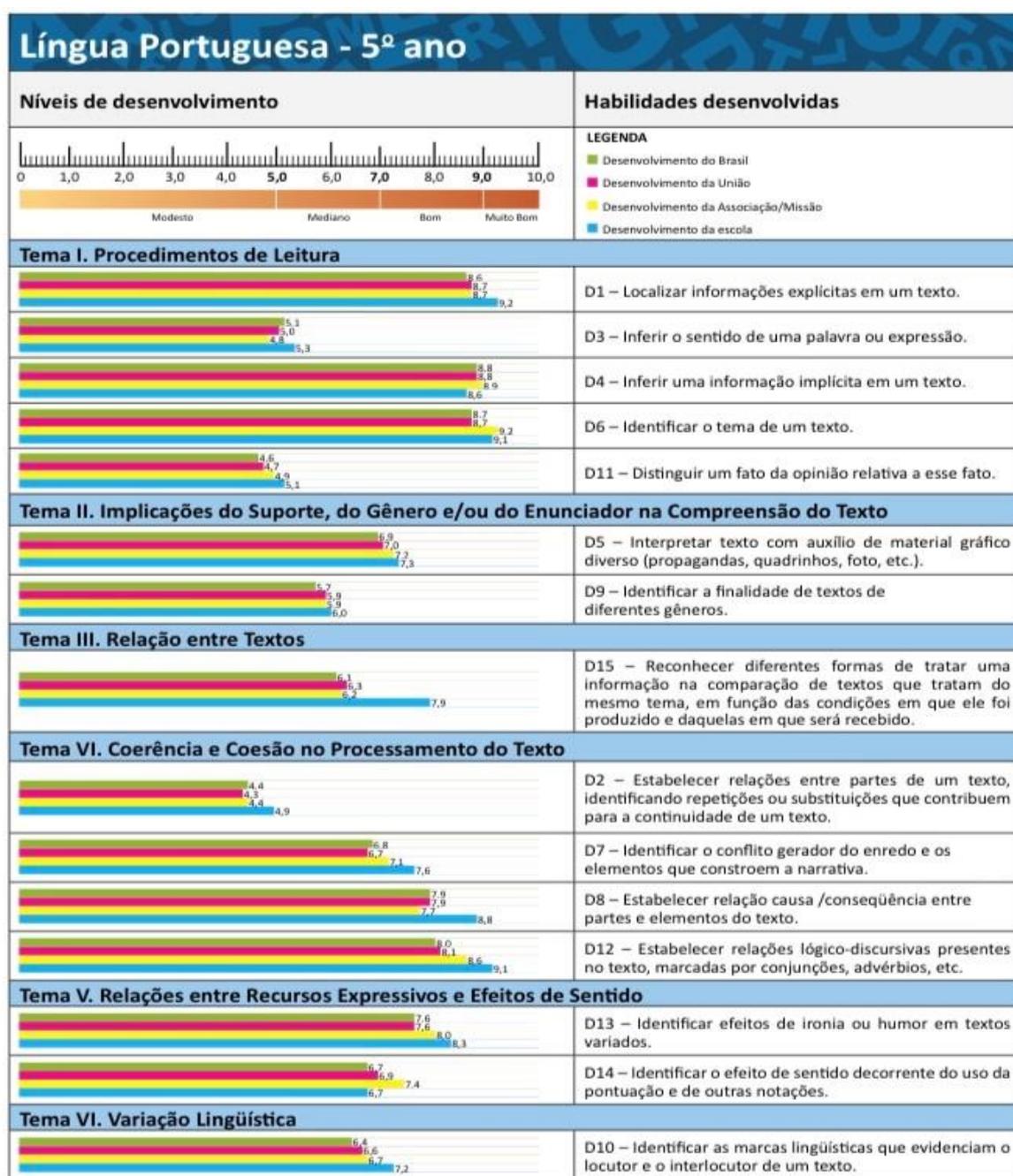
Fonte: Boletim de desempenho da unidade escolar (2013, p.1); editado para ocultar nome da rede

A Figura 3 apresenta um quadro que dispõe informações sobre o número de participantes, no primeiro quadrante, e sobre a média alcançada, no quadrante abaixo. O lado laranja corresponde ao 5º ano, e o lado verde, ao 9º ano. Este quadro também possibilita a comparação das médias nos diferentes níveis administrativos descritos no centro, na parte branca.

Até o ano de 2017, os resultados eram processados e disponibilizados para os participantes, em todos os níveis, no primeiro semestre do ano posterior à aplicação. A partir de 2018, os resultados são divulgados no mesmo ano da aplicação. Esse compartilhamento, otimizado pela possibilidade de geração de dados, gráficos e comparabilidade, no ano em exercício, permite aos gestores a utilização das informações para intervenções e para o planejamento do ano subsequente.

Acompanhando o boletim de resultados, é enviado aos diferentes níveis administrativos da rede um quadro do desempenho dos alunos em cada habilidade, traçando um comparativo com a média de desempenho dos demais níveis. Ou seja, o gestor da unidade escolar consegue visualizar o desempenho de seus alunos nas diferentes habilidades, em relação à sub-região administrativa, à região e ao Brasil. Na Figura 4, há um exemplo de relatório comparativo do desempenho de alunos do 5º ano de uma unidade escolar em relação aos descritores de Língua Portuguesa, possibilitando ao gestor comparar com os resultados médios dos demais níveis administrativos.

Figura 4 - Comparativo dos níveis de desenvolvimento da Unidade Escolar



Fonte: Boletim de Desempenho da Unidade Escolar (2013, p. 3).

As figuras 2, 3 e 4 tem caráter informativo para detalhar o processo de comunicação com gestores, pais e alunos, apresentado no quadro 13. Elas apresentam as ideias e conceitos da equipe gestora do PAER que expressam alguns pontos a serem questionados. Como exemplo, mencionamos a adjetivação dos níveis de desenvolvimento dos alunos na figura 4. Quais os critérios utilizados para a caracterização de “modesto”; “mediano”; “bom” e “muito

bom”)? Não há explicações nos documentos enviados às unidades escolares sobre estes conceitos, mas é fato que refletem uma iniciativa classificatória. Reafirma-se que estes documentos foram utilizados nas cinco primeiras edições do programa e que, desde 2016/2017, utilizam uma plataforma cujos dados são apresentados por meio de gráficos com outra configuração, inacessível para consulta da pesquisadora.

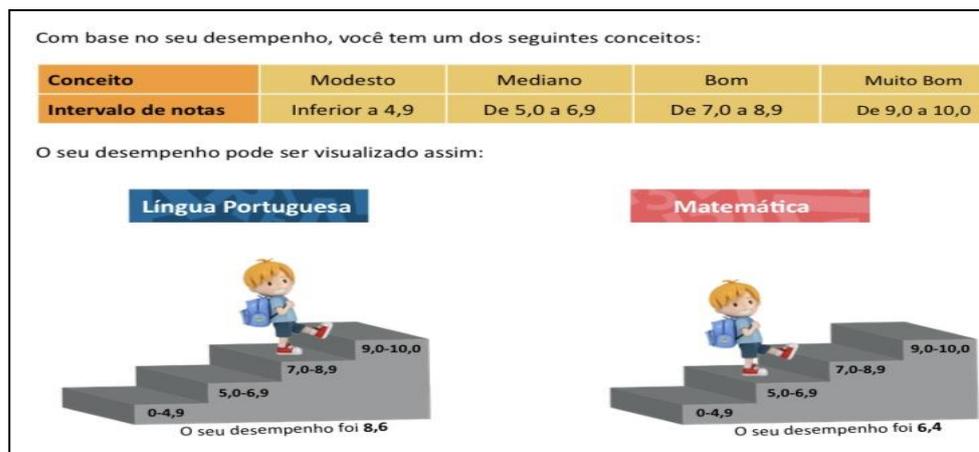
Dessa forma, podemos considerar que, ao receber os resultados de uma edição, é oportunizado ao gestor construir um juízo de valor por critério e normativo, segundo a classificação descrita por Horta Neto (2010), sobre a qual ele faz a seguinte consideração:

Os juízos de valor podem ser expressos basicamente de três formas. Quando a comparação é realizada a partir de referências definidas previamente, diz-se que o juízo de valor é referido a critério. Assim, podem ser estabelecidos distintos níveis de desempenho, indicando níveis de aprendizagem que se esperam para cada aluno, possibilitando construir uma escala que contenha diferentes níveis de aprendizagem de uma determinada área, cada um deles apontando o que se espera do aluno em diferentes fases da sua vida escolar. Quando a comparação é entre indivíduos ou instituições que fazem parte de um mesmo grupo, o juízo de valor está baseado na comparação entre as unidades de análise utilizadas (no caso, indivíduos ou instituições), e diz-se que o juízo de valor é normativo ou que está referido à norma, numa referência à curva normal da estatística. A terceira forma de expressar um juízo de valor é conhecida como a do crescimento da aprendizagem, em que se procura analisar qual a variação do indivíduo em relação a um ponto de partida ou linha de base anterior. (HORTA NETO, 2010, p. 91).

Pelos relatórios disponibilizados é possível realizar uma comparação do indivíduo com os critérios estabelecidos pelos descritores de cada nível de desenvolvimento, assim como uma comparação entre indivíduos e as outras unidades, outras regiões e o Brasil. Portanto, podemos considerar que é possível tecer dois tipos de juízos com o relatório de uma edição da avaliação: o juízo de valor referido a critério e o juízo de valor normativo. Caso o gestor acompanhe a evolução dos resultados em um intervalo de tempo em que possa ter uma referência de ponto de partida, também esses relatórios anuais permitem um juízo de crescimento do conjunto de aprendizagens envolvidas nas habilidades esperadas. Desse modo, o propósito de análise e uso dos relatórios como instrumento de gestão pode ser alcançado.

Para os alunos de 5º e 9º ano do ensino fundamental, o resultado também é acompanhado de uma carta explicativa com um infográfico que mostra o conceito de níveis e os descritores dos tópicos e as habilidades correspondentes, informando o desempenho individual. Diferente dos resultados das unidades, os resultados dos alunos são representados em 4 níveis distribuídos em intervalos de notas que vão de 0 a 10,0, classificados em quatro conceitos: modesto, mediano, bom e muito bom. A Figura 5 foi extraída de uma carta de resultados de um aluno do 5º ano (2013).

**Figura 5 -Resultado disponibilizado para alunos do 5º ano (2013)**



Fonte: Carta de resultado do PAER enviada a aluno (2013, p. 1).

Os alunos também recebem um gráfico descritivo do desempenho individual em cada tópico ou descritor e suas respectivas habilidades nas áreas de Língua Portuguesa e Matemática. Verifica-se que nesta representação, a adjetivação já comentada anteriormente, na descrição da figura 4, é repetida, sem explicações no documento, qual o critério utilizado para esta classificação.

Para o ensino médio, os resultados, tanto para as unidades como para os alunos, são demonstrados em uma escala de 0 a 1000 pontos, em cada um dos quatro eixos cognitivos das áreas: Linguagens, Códigos e suas Tecnologias; Ciências da Natureza e suas Tecnologias; Ciências Humanas e suas Tecnologias e Matemática e suas Tecnologias. Nesse nível de ensino, o boletim vem acompanhado de uma descrição dos níveis de desenvolvimento por competências e habilidades. Para os alunos, esse demonstrativo é individual. Para as unidades e regiões administrativas, a descrição é comparada com as médias alcançadas pelas demais instâncias, conforme apresentada na Figura 4. A Figura 6 traz os resultados obtidos por um aluno do 2º ano do ensino médio, no ano de 2014, quando foi realizada a primeira edição do PAER para esse nível de ensino.

**Figura 6 - Desempenho apresentado no boletim de aluno de ensino médio**

Considerando uma escala de 0 a 1000 você tem o seguinte resultado:

Eixo Cognitivo	Proficiência Média
Ciências Humanas e suas Tecnologias	466,1
Ciências da Natureza e suas Tecnologias	418,6
Linguagens, Códigos e suas Tecnologias	523,6
Matemática e suas Tecnologias	397,9

Fonte: Boletim de desempenho aluno do ensino médio (2014, p. 1).

A plataforma desenvolvida e utilizada desde 2016 favoreceu as análises estatísticas dos dados e das séries históricas das unidades escolares e das regiões e sub-regiões administrativas. A correlação entre dados do desempenho cognitivo dos alunos com os dados dos questionários contextuais, tanto dos alunos quanto dos professores, permite que os gestores possam estabelecer a correlação, por exemplo, entre o desempenho dos alunos em Língua Portuguesa e a frequência com a qual o docente dessa disciplina solicita tarefas de casa. Até 2015, essas análises eram realizadas apenas pela equipe gestora do PAER, na administração central. Com esse recurso, podendo realizar análises estatísticas e comparativas, os responsáveis pelo Programa esperam “apoiar os gestores na elaboração de estratégias de melhora contínua” (DO, 2017, p. 3).

Um dos exercícios propostos por parte da equipe que organizou o PAER foi a comparação das médias de proficiência dos alunos da rede, por unidade federativa, com os mesmos resultados do Saeb nestas unidades, já que, nos primeiros dois anos, foram utilizados o mesmo banco de itens da Prova Brasil e critérios de aplicação das provas cognitivas. No ano em que os primeiros resultados foram divulgados (fevereiro de 2013), a equipe gestora do programa enviou os resultados do Saeb 2011, por unidade federativa, para que fossem analisados pelos gestores regionais a título de possível comparação. Entende-se que esta sugestão de comparação poderia fortalecer a confiança no programa que estava surgindo.

Este exercício de comparação proposto com os dados de 2011 do Saeb não tinha série correspondente no PAER pois a primeira aplicação do foi em 2012, ano que não havia aplicação da Prova Brasil. Assim, a título de comparação entre os resultados, apresentamos na Tabela 5 a média de proficiência por unidade federativa dos alunos do 5º ano do ensino fundamental em Língua Portuguesa e Matemática, de ambas as avaliações, sendo que no ano de 2013, não houve resultado do PAER no estado do Piauí. As cinco melhores médias de cada coluna estão destacadas, favorecendo a visualização e comparação.

**Tabela 5 - Média de Proficiência em Língua Portuguesa e Matemática de alunos do 5º ano PAER e SAEB por unidade federativa no ano de 2013**

UF	PAER MAT	SAEB MAT	PAER LP	SAEB LP
AC	188,95	211,20	203,52	198,83
AL	204,33	186,40	208,97	173,83

AP	183,95	182,04	192,69	171,16
AM	194,57	202,32	207,70	190,51
BA	206,72	194,54	214,84	180,07
CE	217,34	207,38	216,28	196,01
DF	197,33	229,66	201,79	213,67
ES	234,13	218,56	233,36	202,59
GO	206,83	221,57	213,76	206,10
MA	210,97	178,90	215,87	169,41
MT	211,25	207,34	215,01	193,02
MS	217,07	217,76	219,34	201,71
MG	223,62	231,13	223,34	213,48
PA	194,78	183,07	210,45	173,46
PB	190,36	195,85	230,26	181,62
PR	231,13	232,67	227,29	212,68
PE	205,13	200,41	210,15	186,04
PI	-	196,38	-	182,50
RJ	218,96	217,77	217,44	202,72
RN	223,63	194,30	217,30	181,86
RS	229,52	22,89	229,95	206,39
RO	204,74	213,10	209,10	196,30
RR	203,29	204,44	214,58	188,06
SC	225,94	230,83	225,56	213,67

SP	223,64	230,85	221,35	212,39
SE	198,78	196,20	212,22	180,48
TO	212,15	205,21	222,15	189,90
BRASIL	209,96	211,21	215,93	195,91

Fonte: INEP (2015) PAER (2014); elaboração da autora.

Esta tabela permite comparar os resultados entre ambas as avaliações externas e fazer algumas análises. Dos estados que apresentaram os melhores desempenhos, um destaque para o Paraná e Santa Catarina, que estão entre os melhores resultados nas duas avaliações em larga escala, seguidos do estado de São Paulo. No PAER, dos cinco melhores resultados em Matemática, três são da região sul e dois da região sudeste. Já no Saeb, entre os cinco melhores resultados em Matemática, dois são da região sul, dois da região sudeste e do Distrito Federal, na região centro-oeste. Em Língua Portuguesa, no PAER, permanecem três estados da região sul entre as melhores médias, um estado da região nordeste e um estado da região sudeste. Para o Saeb, o mesmo cenário de Matemática se repete em Língua Portuguesa, com três estados da região sul, um estado da região sudeste e o Distrito Federal na região centro-oeste apresentando as melhores médias. Na média geral de proficiência dos alunos do 5º ano, em Matemática, o resultado nacional do Saeb é 1,25 ponto maior que do PAER. Já em Língua Portuguesa, o desempenho dos alunos no PAER é 20,02 pontos acima do Saeb, na média nacional.

Com esta breve análise da Tabela 5, apenas levando em conta os estados com os melhores desempenhos nas duas avaliações, percebe-se uma tendência da região sul do Brasil em apresentar as melhores médias de proficiência dos alunos do 5º ano. Por outro lado, revela também uma similaridade entre as duas avaliações em larga escala, pelo tipo de instrumentos utilizados, o rigor técnico empregado, cuja comparação sugerida pelos gestores, confirmaria a confiabilidade da avaliação em larga escala e abriria um leque de outras oportunidades de análises de dados e micro dados gerados amplamente no decurso das edições de ambas as avaliações para aqueles que se debruçassem sobre estes dados. Após oito edições do PAER, seis delas completas, nos níveis fundamental e médio da educação básica, é possível inferir que há uma série histórica de dados e de indicadores que propiciam estudos meta-avaliativos significativos e que muitas outras pesquisas poder-se-iam desenvolver com o mesmo objeto de estudos ora descrito.

## CAPÍTULO 5 – CONCEPÇÕES DE AVALIAÇÃO E QUALIDADE EDUCACIONAL

Neste capítulo, apresentamos um relato da evolução das concepções de avaliação e qualidade educacional no transcorrer do século XX, que influenciaram a formação de um campo teórico estruturado nas últimas décadas do século passado e que se desenvolve como um crescente campo de estudos ao longo das primeiras décadas do presente século. O foco do capítulo será, portanto, a apresentação dos paradigmas<sup>26</sup> que coexistem e estão subjacentes às práticas avaliativas, quer seja na dimensão da sala de aula, quer seja na dimensão institucional, ou nas políticas públicas de avaliação de sistemas que estão em vigência no contexto educacional brasileiro. Salvaguardando a polissemia do termo qualidade e considerando que avaliação e qualidade estão intrinsecamente relacionadas, ao considerarmos o ato de avaliar como um ato de investigar a qualidade de uma realidade (LUCKESI, 2018), o capítulo reserva lugar para a discussão da qualidade educacional, à luz de autores referenciados.

Avaliação é uma palavra que transita no senso comum e está presente no cotidiano das pessoas envolvendo as ideias de julgamento, decisão, escolha, apreciação, opção, entre outras. O ato de avaliar está intimamente relacionado com a capacidade de raciocinar, tomar decisão, resolver problemas, ponderar – capacidades estas, que o ser humano adquire ao longo de seu processo de desenvolvimento como indivíduo. Esta prática também se desenvolveu ao longo da história da humanidade, quando a necessidade de examinar, atribuir valor a algo, tomar decisões individuais ou coletivas, advindas da vida em sociedade, estruturou um percurso evolutivo do termo, cujo significado não é novo.

Souza Santos (2018) esclarece que a ciência pós-moderna tenta dialogar com outras formas de conhecimento e, para o autor, a mais importante de todas as formas é o conhecimento do senso comum, por meio do qual “no cotidiano orientamos nossas ações e damos sentido à nossa vida.” (SOUZA SANTOS, 2018, p. 95). Invertendo a lógica do racionalismo moderno, quando o grande salto era elevar o conhecimento do senso comum para o conhecimento científico, o mais importante na lógica da ciência pós-moderna é o salto inverso, no qual o conhecimento científico produzido se torne sabedoria para a vida, ele “só se

---

<sup>26</sup> O conceito tem por base as ideias de Patton (*apud* GUBA e LINCOLN, 2011), segundo as quais “Paradigma é uma visão de mundo, uma perspectiva geral, uma forma de decompor a complexidade do mundo real.” Assim como o conceito de Thomas Kuhn, ao considerar paradigma como “as realizações científicas universalmente reconhecidas que, durante algum tempo, fornecem problemas e soluções para uma comunidade de praticantes desta ciência.” (THOMAS KUHN, 1998, p. 13).

realiza enquanto tal na medida em que se converte em senso comum.” (SOUZA SANTOS, 2018, p. 98).

Tendo essa referência, encurtar a distância entre a percepção do senso comum sobre a avaliação e as ideias produzidas pelos estudiosos da área, tornando-as acessíveis e claras, reconhecemos que é uma jornada desafiadora, mas que se configura um dos compromissos de quem se propõe a produzir conhecimento. Num contexto em que a avaliação “na” e “da” educação tornou-se tema de interesse público, em que a sociedade em geral passou a receber informações sobre resultados e desempenho de unidades escolares a sistemas de ensino, a compreensão das ideias estruturantes que orientam as políticas públicas voltadas à avaliação da qualidade do ensino deveriam constituir-se campo de estudos de interesse não somente restrito à academia, mas, como bem sinaliza Vianna (2003), um conhecimento disponível para todos os interessados, incluindo as famílias e a população em geral, conhecimento que se transforme em senso comum.

De procedimentos individuais para coletivos, desde os primórdios da história humana, o campo conceitual da avaliação foi se constituindo e se desenvolvendo a partir da evolução dos fenômenos sociais e do modelo de ciência que prevaleceu num dado momento, para explicar tais fenômenos. No entanto, foi na educação que a avaliação encontrou um espaço privilegiado e, ao mesmo tempo, carregado de complexidade, assim como é complexo o próprio fenômeno educativo. Diferentemente da facilidade de compreensão do termo em seu uso cotidiano, assentado no senso comum, este adquire a polissemia e a complexidade que envolve a própria educação como prática social humana, que emerge de um processo histórico e, se histórico, é inconcluso e em transformação. Esta especificidade que caracteriza o campo da educação requer um cuidado com a amplitude das concepções teóricas que buscam explicá-la, a fim de que não sofram reduções, nem fragmentações que a descaracterizem – risco constante que corremos nas ciências sociais. De igual forma, consideramos que esse cuidado deve ser aplicado no trato com as ideias acerca da avaliação educacional, quando sua abrangência tem se ampliado e diversificado, exigindo maior clareza do que, de fato, se pretende estudar.

### **5.1 Primórdios da avaliação educacional**

O ato de avaliar é constitutivo de todo ser humano e este é um dos modos de conhecermos a realidade, “sob a ótica da qualidade de tudo que existe, seja ela natural ou

cultural.” (LUCKESI, 2018, p. 22). Sendo um ato que praticamos com o propósito de conhecermos a realidade, o praticamos a todo instante, quer de maneira informal, por meio do senso comum, em condutas que, por vezes, supomos que não passariam por uma avaliação<sup>27</sup>, quer através da reflexão sistemática e crítica, como um ato consciente e formal. Para o autor, não existe ato humano que não seja precedido de um ato avaliativo, seja de modo intencional ou consciente, ou de modo habitual e corriqueiro.

De igual modo, Worthen, Sanders e Fitzpatrick (2004) consideram a avaliação uma forma básica de comportamento humano, podendo ser praticada de modo informal ou formal. As avaliações informais prescindem de procedimentos sistemáticos e evidências coletadas formalmente, mas não ocorrem no vácuo. “A experiência, o instinto, a generalização, e o raciocínio podem, todos eles, influenciar o resultado das avaliações informais” (WORTHEN; SANDERS; FITZPATRICK, 2004, p.37). Já a avaliação formal envolve esforços de sistematização, de busca de informações acuradas e estruturação de critérios para sua prática. É esta avaliação formal e sistemática que, praticada ao longo da história, com finalidades, propósitos e usos distintos, transformou-se e disseminou-se nas mais distintas áreas da vida humana, que lança as bases para o estudo do campo da avaliação educacional.

Registros de procedimentos avaliativos formais podem ser encontrados em períodos históricos remotos, conforme relatados na literatura acadêmica, tais como a verificação de aptidões morais de candidatos a funções públicas na Grécia ou em exames de seleção para serviços públicos na China, remontados a 2.000 anos a. C. (DIAS SOBRINHO, 2003; WORTHEN; SANDERS; FITZPATRICK, 2004). Ambos os exemplos já carregavam um grau de formalidade e interesse público, como processo que se vinculava à dimensão social.

Ao longo da história, diante das transformações e da complexidade da vida em sociedade, a avaliação foi demarcando espaço em várias áreas e atividades. Quer seja na administração pública, quer seja na esfera produtiva, ou nas relações de trabalho, entre outras dimensões sociais, os procedimentos avaliativos foram adotados com vistas, predominantemente, a processos de seleção, de classificação ou mesmo para a racionalização da gestão por meio de informações úteis que tais procedimentos podiam oferecer. Contudo, Dias Sobrinho (2003, p. 15) considera que foi “na educação que a avaliação encontrou seu lugar privilegiado”, tanto como prática política e pedagógica como campo de estudo.

---

<sup>27</sup> Para Luckesi, até mesmo as condutas e atos autônomos do cotidiano são precedidos por algum tipo de avaliação do sistema nervoso autônomo que viabiliza a resposta. Esse mecanismo autônomo de tomada de decisões, que trabalha informações fornecidas pelas fontes sensoriais e da memória, Sawrey e Telford (1972) também atribuem a uma configuração única de perceptos, conceitos, ideais e sistemas de valores de cada indivíduo, que constitui um repertório de respostas prontas ou automáticas.

Em virtude da natureza do objeto de estudo desta pesquisa, concentraremos a análise nas práticas avaliativas consideradas precursoras da avaliação educacional propriamente dita, cujas bases possuem interfaces com a avaliação formal de várias áreas e campos de estudos, mas que adquiriu sua identidade nas primeiras décadas do século XX. Ainda dentro do campo da avaliação educacional, consideraremos, principalmente, as iniciativas que originaram as avaliações de programas ou sistemas de ensino, reconhecendo que os primeiros estudos na área se concentravam na avaliação da aprendizagem e, somente na última metade de último século, é que a avaliação em larga escala se constitui um campo de estudo específico com esforços de investigação e construção teórica.

Worthen, Sanders e Fitzpatrick (2004) remontam a Sócrates o uso de avaliações verbais com seus discípulos, como evidência da avaliação na educação, na antiguidade. Contudo, é na idade moderna, nos séculos XVI e XVII, que encontramos o modelo de escola tal qual temos hoje, fruto do crescimento populacional, da nova forma de organização do Estado e, principalmente, devido “às demandas do purismo religioso por parte da Igreja Católica, em sua disputa com o Protestantismo emergente”, cujas tradições religiosas investiam em angariar fiéis por meio da educação, organizando escolas (LUCKESI, 2018, p. 97). A atuação dos católicos, principalmente representada pelos padres jesuítas, e as ideias protestantes representadas pela atuação do bispo protestante John Amós Comênio, formaram a base para o modelo de escola que se constituiu no mundo ocidental.

Na obra publicada em 1599 pelos padres jesuítas, a *Ratio atque Institutio Studiorum Societatis Jesus*, conhecida como *Ratium Studiorum*<sup>28</sup>, havia a prescrição do modelo de ação pedagógica a ser seguido em todas as escolas administradas pela Companhia de Jesus. Nesse documento, há um tópico intitulado *Dos exames escritos e orais*. Nele, além da prescrição de exames escolares gerais, há a indicação do registro do acompanhamento da aprendizagem do estudante em uma Pauta do Professor ou caderneta de registro. Uma banca examinadora teria acesso a esse documento e, após os exames finais, decidiria pela promoção, ou não, de cada estudante. As regras para a aplicação dos exames trazem orientações que, na visão de Luckesi (2018), estão sendo cumpridas até hoje no cotidiano de muitas escolas.

Semelhantemente, a obra *Didática Magna* de Comênio, publicada originalmente na língua tcheca, em 1632, e publicada em latim, em 1654, prescrevia os exames escolares e defendia provas regulares no final de cada aula, de cada dia, de cada semana, de cada mês, de

---

<sup>28</sup> Obra normativa que organizava a educação desenvolvida pela Ordem dos padres jesuítas. Disponível na internet <https://obrascaticas.com/editorarealeza/download/o-metodo-pedagogico-dos-jesuistas-parte-ii-ratio-studiorum/>

cada semestre, para estimular o estudante a aprender. Luckesi (2018) descreve que Comênios propôs ao poder público que nomeasse um Escolarca, personagem responsável por elaborar e aplicar uma prova em todos os estudantes e em todas as escolas existentes, ao final do ano letivo, com o propósito de saber se a aprendizagem dos estudantes estava satisfatória. Segundo o autor, “esta medida destinava-se a avaliar o desempenho da instituição escolar e do seu sistema, desde que o desempenho do estudante individual seria detectado pela série de provas sucessivas durante o ano letivo” (p. 105). Seria, essa proposição, um embrião de uma avaliação de sistemas de ensino? De acordo com Luckesi (2018), em plena metade do século XVII, Comênios “já tinha um olhar a respeito da responsabilidade do sistema de ensino na produção de resultados desejados através do exercício da ação pedagógica.” (LUCKESI, 2018, p. 105).

Assim, predominaram nos séculos XVI e XVII as proposições da pedagogia jesuítica e comeniana, com atos avaliativos denominados exames escolares, com a correlação entre o uso dos exames e a eficiência do processo de ensino e da aprendizagem. Porém, em ambas as proposições, havia um genuíno interesse de que todos os alunos aprendessem. A principal função das práticas avaliativas era diagnóstica e para conduzir a aprendizagem. O uso seletivo dos exames era uma consequência e não uma finalidade.

Profundas mudanças ocorreram no tecido social ao final do século XVIII. O marco da Revolução Francesa, movimento revolucionário da burguesia, desencadeia uma mudança no sistema político de um país, mas esse movimento se estende como ondas, com profundas mudanças nos âmbitos social, cultural e, principalmente, econômico do mundo ocidental. A sociedade se organiza em torno da economia do capital e a educação, como prática social, também é impactada por essa mudança estrutural de modelo social. O segmento social que detinha a posse dos meios de produção, munido de uma lógica política a serviço de seus interesses, excluiu a maioria da população dos bens econômicos, sociais ou culturais, definindo a educação como um fator de exclusão social, nessa lógica de poder social e político controlado pelo capital.

Nos séculos XVIII e XIX, o modelo excludente da sociedade do capital foi se instalando, e a educação consolidou o uso seletivo dos resultados da avaliação da aprendizagem, conforme assinala Luckesi:

O uso seletivo dos resultados da avaliação da aprendizagem na escola é, pois, compatível com as práticas pedagógicas que foram ganhando forma e se sedimentando ao longo da modernidade e que atingiram e atingem a todos nós. Esse uso dos resultados da avaliação da aprendizagem expressa, no seio da prática pedagógica escolar, o modelo da sociedade do capital que emergiu com a

modernidade e, junto com outros fatores, realiza sua mediação. (LUCKESI, 2018, p. 111).

Worthen, Sanders e Fitzpatrick (2004) descrevem que, durante o século XIX, na Grã-Bretanha, a insatisfação com os programas educacionais e sociais gerou movimentos de reforma, dentre as quais, foram nomeadas comissões governamentais para avaliarem as instituições. “Isso levou aos sistemas – existentes até hoje – de inspeções gerais nas escolas da Inglaterra e da Irlanda” (p. 62). Vianna (1995) associa essas transformações sociais ao impacto da Revolução Industrial e, nesse período, houve na Inglaterra uma intensa busca pela eficiência nacional, sendo a educação considerada elemento essencial para o alcance dessa eficiência. Assim como ocorreu na Inglaterra, também nos Estados Unidos algumas iniciativas surgiram com o intuito de discutir programas de natureza educacional. A partir da iniciativa de Horace Mann, em 1845, com a coleta e uso de dados das escolas do estado norte-americano de Massachusetts, o processo avaliativo adquire uma natureza formal. O Comitê Educacional de Boston fez uso “em 1845 e 1846 de provas impressas sobre vários assuntos (o primeiro exemplo de avaliação em larga escala do desempenho estudantil que serviu de base para comparação entre escolas).” (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 62). Para os autores, “esses dois processos ocorridos em Massachusetts foram as primeiras tentativas de estimar objetivamente o desempenho dos alunos com o intuito de avaliar a qualidade de um sistema escolar de grande porte.” (p. 62).

Outras iniciativas de avaliação de vários sistemas educacionais ocorreram no final do século XIX, nos Estados Unidos. Vale ressaltar que nesse período o paradigma positivista, próprio das ciências físico-naturais, dominava todas as iniciativas de pesquisa, assim como as iniciativas de avaliação formal, nas diversas áreas (DIAS SOBRINHO, 2003).

A Psicologia, como ciência, emerge no final do século XIX, impulsionada pela ideia de avaliação de diferenças individuais e habilidades mentais, sob forte influência da teoria da evolução proposta por Charles Darwin. Alguns estudiosos no continente europeu – em Londres, Francis Galton; na Alemanha, Emil Kraepelin; em Paris, Alfred Binet –, valendo-se da ideia de mensuração, desenvolveram os primeiros testes sistematizados e padronizados sobre evidências de diferenças individuais em variados aspectos de processos mentais e comportamentais ou mesmo de desordens psicológicas (SILVA, 2003).

De todos os testes de habilidades mentais desenvolvidos nos anos finais do século XIX, e que dominaram a maior parte do século XX, os testes desenvolvidos por Alfred Binet, com o propósito de avaliar a inteligência geral, foram os mais amplamente utilizados. A pedido do governo francês, entre os anos de 1895 e 1904, Binet e seu mais direto colaborador

Theodosius Simon, desenvolveram testes de mensuração da inteligência com o propósito de diferenciar os casos de atrasos escolares ligados a deficiências intelectuais dos atrasos associados a condições ambientais desfavoráveis. Esses testes foram identificados como “escolásticos”, em função do uso associado diretamente ao desempenho escolar (SILVA, 2003).

Portanto, nesse período, avaliação e medição tornam-se conceitos intercambiáveis. Como afirma Dias Sobrinho (2003), “avaliação confunde-se com medição e se insere basicamente no campo da psicologia. Daí, ter se tornado amplamente disseminado o conceito de psicometria” (p. 16). Silva (2003) define psicometria como “o domínio da avaliação psicológica [...], a qual constitui na abordagem científica para a mensuração das características psicológicas” (p.148). A testagem psicológica representa um dos campos mais amplos da psicologia aplicada. É na psicometria, portanto, que a avaliação escolar, nesse período histórico, se inspira e se desenvolve.

A respeito da influência desse momento sobre a avaliação e a educação, Dias Sobrinho pontua:

Se ainda no século XIX os testes tiveram enorme importância para a evolução posterior da avaliação, em seus primórdios identificada quase exclusivamente como medidas psicológicas, e da própria psicologia, em grande monta dedicada à psicometria, é só no começo do século XX que a avaliação começou, gradualmente, a se desenvolver como prática aplicada à educação. Acompanhando os avanços do conhecimento na área da psicologia e certamente se beneficiando deles, desenvolveu-se com grande êxito o movimento e a cultura dos testes e medidas. (DIAS SOBRINHO, 2003, p. 17).

Ressalta-se, portanto, que a Psicologia foi gestada e inserida no paradigma positivista das ciências naturais e, conseqüentemente, a avaliação-medição foi desenvolvida no âmbito científico da psicologia. Esta constatação é necessária para que possamos compreender as primeiras construções teóricas sobre avaliação educacional, que encontramos no início do século XX.

## **5.2 A avaliação educacional no início do século XX**

As bases da avaliação educacional foram lançadas no século XIX e, nas primeiras três décadas do século XX, ela se desenvolve apoiada amplamente nos conhecimentos adquiridos na prática de mensurações e testagens oriundas da, então jovem ciência, psicologia.

Dias Sobrinho (2003), utilizando uma periodização desenvolvida pelos autores Stufflebeam e Shinkfield, que descrevem cinco períodos básicos da avaliação, considera essas três primeiras décadas do século XX o primeiro período da avaliação, conhecido como pré-Tyler. Gradualmente, a avaliação se incorpora como prática aplicada à educação, fortemente arraigada à cultura de testes e medidas. A figura central desse período foi o psicólogo norte-americano Edward L. Thorndike, cuja grande contribuição foi o desenvolvimento de uma fundamentação teórica sobre a possibilidade de medir mudanças em seres humanos. Havia um movimento favorável à introdução de provas objetivas na educação, e suas ideias a respeito do assunto foram logo abraçadas quando ele publicou, em 1903, *Educational psychology* e, no ano seguinte, *Mental and social measurement*. A liderança intelectual de Thorndike foi além do uso de testes com fins de classificação. Segundo Dias Sobrinho (2003), ele “foi o precursor do movimento que mais tarde veio a se estabelecer como “gestão científica”, um mecanismo que visava dotar a educação de maior eficiência.” (p. 17). Suas ideias alcançaram todos os níveis da educação em seu país.

Nesse período, avaliação e medida eram sinônimos. Vianna (1995) afirmou que “avaliação e medida por meio de testes confundem-se nessa época – e ainda continuam a se confundir nos dias atuais.” (p. 10). O centro de interesse dos estudiosos do começo do século era o aperfeiçoamento dos instrumentos de medição. A avaliação consistia, basicamente, em testes de mensuração, verificação e quantificação da aprendizagem dos estudantes (DIAS SOBRINHO, 2003).

Worthen, Sanders e Fitzpatrick (2004), das primeiras décadas do século XX, afirmam:

A década de 1920 viu o rápido surgimento de testes baseados nas normas dos programas, criados para avaliar os níveis individuais de desempenho. Em meados de 1930, mais da metade dos Estados Unidos tinha alguma forma de testes estaduais, e os testes padronizados, baseados nas normas de referência dos programas, inclusive testes para avaliar o aprendizado, assim como testes de personalidade e perfis de interesse, transformaram-se num grande negócio. (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 63).

É na década de 1930 que tem início a segunda fase ou segunda geração dos períodos avaliativos, com os estudos desenvolvidos por Ralph Tyler, considerado o maior expoente dessa fase. Tyler foi o criador da expressão “avaliação educacional” e “pai” da avaliação educativa, centrando seus estudos na avaliação baseada nos objetivos educacionais (DIAS SOBRINHO, 2003). Suas ideias exerceram grande influência nos teóricos da avaliação que o sucederam e, ainda que essas ideias fossem direcionadas para a educação, seu pensamento pôde ser aplicado em outros serviços ou setores que desenvolveram avaliação de programas.

Com a constatação, após a Primeira Guerra Mundial, de que os currículos escolares precisavam passar por uma profunda reforma e que o modelo de avaliação focado apenas na mensuração e levantamento de dados não era satisfatório para o propósito de mudanças que escolas secundárias americanas necessitavam, o Departamento de Pesquisa Educacional, da Universidade de Ohio, onde Ralph W. Tyler atuava, desenvolveu o programa *Eight-Year Study*, no qual os alunos de trinta escolas secundárias com flexibilidade de currículos não convencionais foram acompanhados em seu desempenho acadêmico, por oito anos de estudos, sendo quatro no ensino médio e quatro na universidade. O propósito do programa era “demonstrar que os alunos capacitados de acordo com esses currículos não convencionais mesmo assim seriam capazes de ter um bom desempenho na faculdade” (GUBA; LINCOLN, 2011, p. 34). Ao longo dos oito anos, foram desenvolvidos testes para avaliar se “os alunos haviam ou não aprendido o que seus professores *pretendiam que eles aprendessem*. Esses resultados desejados de aprendizagem foram chamados de *objetivos*” (p. 34) (grifos dos autores). A avaliação educacional, para Tyler, pretendia que os professores “aprimorassem seus cursos e os instrumentos de medida [...]” e “pudessem verificar a congruência entre os conteúdos curriculares e as capacidades desenvolvidas” (VIANNA, 1995, p. 10).

Enquanto na primeira fase, a ênfase estava na mensuração, agora, na segunda fase, a finalidade surge da necessidade de aperfeiçoar os currículos que estavam sendo desenvolvidos e confirmar se eles estavam funcionando. Guba e Lincoln (2011) afirmam que “foi assim que nasceu a avaliação de programas” (p.35). Dessa maneira, a ênfase está na análise de pontos fortes e fracos com respeito aos objetivos estabelecidos.

Segundo Saul (2010), a avaliação da aprendizagem, na concepção de Tyler, está alicerçada em seu modelo para elaboração de currículo, que assume um caráter de planejamento e sistematização semelhante ao processo de produção industrial. Suas ideias calcadas em pressupostos positivistas, que apresentavam um modelo de avaliação educacional com uma metodologia sofisticada e uma relação entre as medidas e os resultados desejados da aprendizagem, lançaram as sementes para muitos outros estudos que refinaram e aperfeiçoaram suas ideias nas décadas seguintes. No Brasil, Saul (2010) cita Dalila Sperb, Marina Couto e Lady Tina Traldi, como autoras que seguiram as ideias de Tyler e nele se inspiraram, na área de currículo. Com cerca de uma década de atraso, as ideias chegavam ao país e eram difundidas no meio acadêmico brasileiro.

Worthen, Sanders e Fitzpatrick (2004) afirmam que as ideias de Tyler exerceram uma grande influência sobre os teóricos que o sucederam por serem baseadas numa “argumentação

lógica, cientificamente aceitável, fácil de adotar pelos avaliadores” (p. 130). Assim como o fato de que muitos deles tinham a mesma formação metodológica de mensuração pré-pós-testes de comportamento, que ele tanto enfatizava. “Tyler via a avaliação como o processo de determinação da medida em que os objetivos de um programa são realmente alcançados.” (p. 130). Dias Sobrinho (2003), sobre a centralidade das ideias de Tyler sobre avaliação, assim observa:

O papel essencial da avaliação, segundo Tyler, é averiguar até que ponto os objetivos educacionais traçados estão sendo alcançados pelo currículo e pelas práticas pedagógicas, ou seja, a determinação do grau em que ‘mudanças comportamentais estão ocorrendo’. (Tyler, 1976: 98) (DIAS SOBRINHO, 2003, p. 19).

Constata-se, então, que o paradigma da racionalização científica, comprometido com uma ideologia utilitarista tão peculiar ao sistema de produção industrial, em forte desenvolvimento no começo do século XX, permeou a pedagogia por objetivos e direcionou a avaliação educacional. Pode-se identificar, com certa facilidade, que os pressupostos formulados por Tyler estenderam suas raízes ao longo das décadas seguintes e chegam até as primeiras décadas do século XXI, adaptados e ampliados. Assim, considera Dias Sobrinho:

Tal como a pedagogia por objetivos, que assume um conjunto de preocupações e ideias e vai se atualizando ao longo do tempo, também a avaliação comprometida com ela, ou seja, com a ideologia da eficiência e da racionalidade instrumental, tem recebido ampla acolhida. Ela se manifesta, por exemplo, na modalidade da *accountability*, atualmente muito praticada, que pode ser entendida como a responsabilidade de demonstrar eficiência e produtividade. Por exercer funções reguladoras e classificatórias, é hoje em dia amplamente utilizada por governos e agências contatadas por eles. (DIAS SOBRINHO, 2003, p. 21).

As décadas seguintes ainda foram profundamente marcadas pelas ideias de avaliação educacional desenvolvidas por Tyler. A década de 1940, segundo Worthen, Sanders e Fitzpatrick (2004), viu um período de consolidação dos métodos mais antigos de avaliação, com o esforço de elaboração de provas padronizadas pelos professores e, na década seguinte, esforços para a melhoria da abordagem tyleriana, com a ênfase no ensino, nos educadores, na elaboração de objetivos explícitos e mensuráveis com a oferta de taxonomias de objetivos educacionais, no “domínio cognitivo (Bloom, Engelhart, Furst, Hill e Krathwohl, 1956).” (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 65).

Seguindo a periodização adotada por Dias Sobrinho (2003), o intervalo de 1946 a 1957 foi denominado era da inocência e “se caracterizou por um certo descrédito, não só quanto à avaliação, mas também à própria educação.” (p. 21). Um dos motivos foi o reconhecimento de que os testes psicométricos careciam de mais critérios científicos. Os

resultados passaram a ser questionados pela imprecisão e falta de rigor, bem como a estigmatização dos indivíduos testados.

Contudo, o ano de 1957 foi marcado por um fato que abalou o *establishment* norte-americano. Os soviéticos lançaram ao espaço, com êxito, o *Sputinik I*, e esse evento produziu uma reação imediata e em cadeia, atingindo diretamente a educação. Segundo Guba e Lincoln (2011), esse fato colocou em evidência as possíveis falhas na educação americana e na abordagem de avaliação educacional centrada em objetivos. Vianna (1995) afirma que um trauma foi provocado pela constatação da deficiência tecnológica e carência educacional dos Estados Unidos e, diante disso, a avaliação passou a ter um papel de relevância diante da necessidade de desenvolvimento de novos currículos e novas estratégias de ensino.

Com a promulgação da Lei Educacional de Defesa Nacional, em 1958, Worthen, Sanders e Fitzpatrick (2004) explicam que milhões de dólares foram disponibilizados para a criação de programas novos de desenvolvimento de currículos, principalmente nas áreas de Matemática e Ciências, e inúmeras instituições foram criadas para o desenvolvimento de amplos programas de avaliação patrocinados pelo poder público, na grande maioria. Os estudos resultantes dessas iniciativas revelaram o empobrecimento conceitual e metodológico da área da avaliação.

A maioria dos estudos dependia de conceitos e técnicas importados da pesquisa das ciências sociais e comportamentais que eram muito apropriados para essas áreas, mas pouco adequados para a avaliação de programas escolares. O trabalho teórico relacionado diretamente à avaliação (em contraposição à pesquisa) não existia, e logo ficou evidente que o melhor em termos de reflexão teórica e metodológica da pesquisa social e comportamental não tinha diretrizes a oferecer sobre a forma de realizar muitas das partes constituintes da avaliação. Assim, os cientistas e implementadores da área da educação tiveram de aproveitar tudo o que puderam da pesquisa aplicada nas áreas sociais, comportamentais e educacionais. (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 66).

Os esforços empreendidos no início da década de 1960 não se revelaram satisfatórios, apesar da alocação de recursos, do envolvimento das “melhores cabeças pensantes na área educacional” (VIANNA, 1995, p.11) e da realização de avaliações apoiadas em grande rigor técnico. A abordagem passou a apoiar-se em conceitos e métodos das Ciências Sociais<sup>29</sup>, e muitos instrumentos utilizados revelaram-se inadequados para os propósitos definidos. Surgiu, nesse momento histórico, a figura do avaliador como um profissional especialista para

---

<sup>29</sup> Destaca-se aqui que, embrionariamente, a avaliação educacional apoiou-se nas técnicas psicométricas e estatísticas concebidas nas Ciências Naturais. Já em seu desenvolvimento inicial, apoiou-se nas Ciências Sociais. Porém, ambas ainda ancoradas no racionalismo, com a busca pelo nexo causal e foco na regularidade dos comportamentos, com vistas de previsão de futuros fenômenos. O modelo da racionalidade foi insuficiente, na medida em que as Ciências Sociais se desenvolveram e se afirmaram, sendo capazes de produzir um campo teórico próprio (SOUZA SANTOS, 2018).

atuar nos grandes projetos de reestruturação educacional do país, e as universidades começaram a se dedicar à formação de recursos humanos para os trabalhos de avaliação. Contudo, houve muita improvisação de professores que se tornaram avaliadores sem a formação necessária. Vianna (1995) afirma que “isso produziu avaliações bastante comprometidas, que pouco serviram para avaliar a efetiva eficiência dos sistemas educacionais, seus currículos e programas” (p. 12). Os reflexos desse período estenderam-se pelas décadas seguintes.

### 5.3 A avaliação educacional como campo de estudos

O reconhecimento da insuficiência do modelo de avaliação adotado até então, pautado em objetivos e na descrição de resultados, mobilizou estudiosos a buscarem novos modelos e a estruturação da área, respondendo às demandas por mudanças profundas nas políticas sociais e no sistema educacional norte-americano, julgado ineficaz até então.

Dias Sobrinho (2003) aponta-nos que, em 1963, Cronbach escreveu um artigo documentando os problemas do modelo avaliativo adotado até então, exigiu novas diretrizes e propôs que fosse superada a ideia de avaliação organizada *por* e *para* objetivos por um modelo orientado para a decisão e para o processo. A avaliação do processo e das estruturas específicas, na visão de Cronbach, era mais importante para as decisões de mudança. Suas ideias repercutiram, chamaram a atenção de outros intelectuais acadêmicos e ajudaram a criar uma concepção de avaliação bem mais ampla, especialmente na década seguinte.

Guba e Lincoln (2011) consideram que as ideias defendidas por Cronbach e abraçadas por outros estudiosos que visavam se contrapor às ideias tylerianas deram origem a uma terceira geração de avaliação, cujo procedimento manteria as funções técnicas e descritivas da geração anterior, mas apoiar-se-ia no juízo de valor e na existência de critérios para que o avaliador exercesse o papel de julgador. Nesse cenário, surgem autores que se dedicaram a tornar a avaliação educacional um campo de estudos definido que exigia uma construção teórica própria e com base científica.

Calderón e Borges (2013), descrevendo esse período, relacionaram os principais autores que contribuíram para a construção desse campo teórico nos Estados Unidos e sua forte influência no Brasil:

[...] foram outros quatro autores norte-americanos os que contribuíram decisivamente, na década de 70, para o desenvolvimento teórico da área da avaliação educacional: Lee Cronbach, Michel Scriven, Daniel Stufflebeam e Robert Stake. Os estudos destes autores, somados aos de Malcom Pasllet, David Hamilton,

e Barry Macdonald, exerceram uma forte influência na trajetória brasileira da avaliação. (CALDERÓN; BORGES, 2013, p. 262).

Dias Sobrinho (2003) denomina essa fase como o quarto período, ou realismo, que abrange os anos 1958–1972. Desse momento, além de destacar a contribuição de Cronbach, o autor faz um destaque à grande contribuição de Scriven, ao fazer uma distinção entre funções e objetivos do processo avaliativo. Os objetivos seriam invariáveis, e as funções se refeririam ao emprego das informações para tomadas de decisões no processo. Outra contribuição de Scriven foi a distinção entre avaliação formativa e avaliação somativa. A primeira é aquela realizada ao longo do processo com o objetivo de aprimorar as ações, e a última, a que se realiza no final do processo para verificar resultados.

Na década de 1970 inúmeros modelos de avaliação surgiram, e o campo passou por uma efervescência. Assim a descreve, Vianna (1995, p. 13):

A partir dos anos 70, a avaliação educacional torna-se um campo profissional definido, exigindo especialização aprofundada, com a exclusão de improvisações supostamente técnicas, mas pouco sólidas conceitualmente, num campo em que atuavam diferentes profissionais, inclusive administradores. A avaliação, pelo menos no contexto norte-americano, deixa de ser uma ‘terra de ninguém’. Simultaneamente, surgem importantes revistas especializadas, algumas de alta qualidade técnica, associando avaliação e políticas públicas, avaliação e planejamento, entre outras, permitindo assim, a disseminação de novas ideias, a formulação de teorias e modelos, e, sobretudo, a divulgação de importantes estudos.

Na periodização utilizada por Dias Sobrinho (2003), apoiado em Stufflebeam e Skinkfield, a década de 1970 representa a quinta e última fase, denominada período do profissionalismo. O campo de estudos e vasta produção teórica se consolida nessa década, e o autor afirma que “a avaliação ganha importância e visibilidade para além das salas de aula e das instituições educacionais. Como resultado dos avanços teóricos, surge a meta-avaliação ou avaliação das avaliações.” (p. 23).

Esse período pode ser representado com a frase de Dias Sobrinho (2003): “Devemos ressaltar que a avaliação, nesse momento, desloca o seu centro: dos objetivos para a tomada de decisões.” (p. 24). E ainda destaca o esforço de superação dos esquemas limitantes da mensuração e objetivos predefinidos que deu origem à contribuição essencial, que “foi caracterizar a avaliação como *juízo de valor*” (p. 24) (itálico do autor). Num esforço de expandir o alcance meramente de testagem e descrição, avança na dimensão de avaliar a qualidade do próprio processo. No caso da educação, a efetividade do próprio ensino.

A dimensão valorativa que se instala a partir de então, incorporando o juízo de mérito ou valor, envolve uma subjetividade e complexidade que reclamam a consideração de inúmeros aspectos humanos, sociais, políticos, éticos, uma dimensão em que os consensos

precisavam ser construídos. As contradições vão muito além de questões epistemológicas, elas resultam de concepções de mundo. Assim, construir consensos requer uma nova postura, muito mais democrática e política do que técnica e científica. Sendo assim, num esforço de síntese dos períodos transcritos pelo autor, o Quadro 14 descreve as principais ideias desenvolvidas ao longo das sete primeiras décadas do século XX.

**Quadro 14 - Periodização, segundo Stufflebeam e Skinkfield**

<b>Período</b>	<b>Designação</b>	<b>Foco</b>	<b>Função</b>	<b>Principais Teóricos</b>
1º - Últimos anos do século XIX até a década de 1930	Pré - Tyler	Avaliação-medição centrada em testes.	Verificação, mensuração e quantificação da aprendizagem dos estudantes.	E. L. Thorndike
2º - 1934 a 1946	Avaliação por objetivos	Avaliação de objetivos educacionais.	Determinar em que grau os objetivos traçados foram alcançados.	Ralph Tyler
3º - 1946 a 1957	Era da Inocência	Certo descrédito à educação e avaliação.	Vários testes produzidos no período.	Ralph Tyler
4º - 1958 - 1972	Realismo	Avaliação para tomada de decisão	Emprego das informações para tomada de decisões.	Lee J. Cronbach Michael Scriven
5º - 1973 e anos seguintes	Profissionalismo	Avaliação como um julgamento de valor	Determinar o mérito e valor de um objeto.	Daniel Stufflebeam Robert Stake Michael Scriven

Fonte: Dias Sobrinho (2003); elaboração da autora

Observa-se no Quadro 14 uma evolução e um alargamento do campo da avaliação, adquirindo complexidade na medida em que se apresentaram os desafios conjunturais de cada período. Aspectos econômicos, culturais, sociais e políticos passaram a justificar e pressionar a ampliação do campo que, ao crescer em complexidade, incorporou os conflitos resultantes da mudança de perspectiva, que, antes, apoiava-se nas bases epistemológicas positivistas e naturalistas das ciências duras, para uma perspectiva dinâmica e contraditória das ciências sociais. A cada período, a ênfase ou propósito procurou suprir falhas ou carências do período anterior, a partir de novas demandas que surgiam, na medida em que a avaliação adquiriu centralidade na solução de problemas envolvendo “campos de valores e expectativas a respeito da vida e do futuro.” (DIAS SOBRINHO, 2003, p. 29).

Guba e Lincoln (2011), apresentando uma visão histórica dos cem anos de desenvolvimento das ideias acerca da avaliação, o fazem igualmente em períodos que designam gerações, como já sinalizamos ao longo do texto. Essas gerações são delimitadas pelo contexto histórico, pelo propósito que as pessoas tinham em mente ao conduzirem as avaliações e pelos vários significados a ela atribuídos pelos teóricos e avaliadores. Também sinalizam essa evolução de construção e reconstrução de significados em razão de “inúmeras influências interagentes.” (GUBA, LINCOLN, 2011, p. 28). O Quadro 15 é um esforço de síntese da divisão em gerações ao longo do século XX, proposta pelos autores.

**Quadro 15 - Periodização, segundo Guba e Lincoln**

<b>Período</b>	<b>Designação</b>	<b>Ênfase</b>
1ª Geração: final do século XIX a 1930 ≅	Geração da Mensuração	Uso de testes para investigar qualquer variável que se desejasse investigar e pudesse ser mensurada.
2ª Geração: Década de 1930 a 1957	Geração da Descrição	Abordagem caracterizada pela descrição dos padrões de pontos fortes ou fracos, a partir de objetivos estabelecidos.
3ª Geração: 1958 a década de 1970 ≅	Geração do Julgamento de Valor	O juízo de valor era a parte essencial da avaliação, e o avaliador, um julgador ou juiz.
4ª Geração: Final da década de 1970	Geração da Negociação	Enfoque responsivo e construtivista, com base no qual todos os envolvidos interagem de forma a respeitar a dignidade, integridade e privacidade.

Fonte: GUBA; LINCOLN, 2011; elaboração da autora

Podemos visualizar, tanto no Quadro 14 como no Quadro 15, uma evolução de sentidos e de sofisticação da avaliação ao longo do tempo. A transição de um período ou de uma geração para a outra foi estimulada por deficiências encontradas na geração anterior, quer pelo tipo de abordagem, por procedimentos adotados, quer pela incapacidade de apontar caminhos frente aos resultados obtidos pela avaliação, ou mesmo pela constatação do caráter fundamentalmente humano e centrado em valores, que se configurou nas últimas gerações. Pode-se notar que a percepção dos autores quanto ao foco central de cada período, em ambos os quadros, é muito semelhante, havendo como diferença um acréscimo de um período intermediário (1946 – 1957), denominado era da inocência por Stufflebeam e Skinkfield, e uma diferença no foco do último período de cada quadro.

O fato de surgir um novo período ou uma nova geração não significa que as ideias, as práticas ou mesmo os instrumentos adotados pela geração anterior deixaram de existir ou não coexistam. Sobre isso, Dias Sobrinho (2003) considera que, embora a avaliação educacional

tenha tido uma grande evolução desde o início do século XX, adquirindo novos significados e aperfeiçoando tecnologias e metodologias,

[...] ainda hoje vige uma tendência em certos círculos a confundi-la com medição e a utilização largamente de testes de verificação de aprendizagem, uns de caráter individual e limitado às salas de aula, alguns outros estendendo-se a âmbitos internacionais, que muitas vezes ultrapassam os limites da educação. Com efeito, com todo o alargamento do campo da avaliação, a sofisticação de seus instrumentos, a intensificação de suas atividades e o fortalecimento de sua importância política consolidados ao longo dos tempos, os velhos testes de aprendizagem ainda hoje constituem o núcleo mais comum de importantes programas de avaliação educacional conduzidos por governos nacionais, organismos internacionais e outras agências. Ampliou-se consideravelmente o seu campo de aplicação dos limites mais estreitos da sala de aula, aos cenários nacionais e mesmo transnacionais. (DIAS SOBRINHO, 2003, p. 26).

Seja na sistematização em períodos históricos de Dias Sobrinho (2003), seja na divisão em gerações proposta por Guba e Lincoln (2011), o movimento que se aprofunda durante as décadas de 1960 e 1970, de questionamento da validade do paradigma científico positivista para atender aos processos de avaliação cada vez mais complexos, exerce uma pressão por mudança de modelo que culmina com o surgimento de um novo paradigma, orientado por pressupostos e por uma lógica que se contrapõem aos do modelo anterior. Ambos coexistem e direcionam desde políticas de avaliação educacional até as práticas avaliativas de sala de aula no cotidiano das escolas. Para identificá-los, utilizaremos a designação de Souza Santos (2018): paradigma dominante, pautado nos princípios da ciência moderna e tradicional; e paradigma emergente, que busca romper com a estrutura científica para uma dimensão política, ética e social, conforme descreveremos a seguir.

#### **5.4 Os paradigmas avaliativos**

Antes de apresentarmos os principais pressupostos de cada paradigma que condicionam as ideias acerca da avaliação educacional, entendemos ser necessária a confirmação do sentido do uso do termo paradigma neste estudo.

Worthen, Sanders e Fitzpatrick (2004) consideram que paradigma “é, muito simplesmente, um modelo ou exemplo. Em termos bem genéricos, poderíamos dizer que é uma filosofia ou escola de pensamento.” (p. 113). Porém, os autores reconhecem que os filósofos da ciência dão ao termo um significado especial, usando-o para descrever:

[...] visões de mundo integradas de indivíduos ou grupos que determinam como esses indivíduos ou grupos percebem a verdade e tentam compreendê-la. A visão de mundo (paradigma) resultante abrange uma crença epistemológica e uma crença ontológica que, juntas, governam as percepções e opções feitas na busca da verdade científica. (WORTHEN; SANDERS; FITZPATRICK, 2004, p. 113).

Como afirma Moraes (1996), “toda formulação teórica traz consigo um paradigma do qual decorre todo um sistema de valores que influencia não somente o processo de construção do conhecimento, mas também a maneira de ser, de fazer, e de viver/ conviver.” (p. 3). Dias Sobrinho, citando Patton, descreve o termo como “uma visão de mundo, uma perspectiva geral, um modo de esmiuçar a complexidade do mundo real.” (PATTON, *apud* DIAS SOBRINHO, 2003, p. 85). Para Guba e Lincoln (2011),

Os paradigmas são chamados de visão de mundo não apenas porque estabelecem como premissa um conjunto de crenças de apoio sobre as características do mundo sobre como investigá-las, mas também porque eles têm implicações profundas no modo como construímos nossos assuntos políticos, como julgamos se uma atividade é moral ou ética e como mantemos a justiça nas relações sociais. Eles representam mais do que um pano de fundo filosófico – o cenário – no qual metodologias de investigação ressonantes – o enredo – podem ser construídas e testadas. Eles são, em última análise, as pedras de toque de nossa vida. (GUBA; LINCOLN, 2011, p. 133).

Portanto, a escolha pelo termo paradigma para designar o escopo de ideias que estão subjacentes às políticas e práticas de avaliação educacional deve-se à compreensão de que este termo, utilizado por vários estudiosos da área, carrega em si uma amplitude semântica que traduz melhor as dimensões ontológicas, epistemológicas e axiológicas que sustentam uma construção teórica.

Falando sobre a ideia de dois paradigmas em conflito, Dias Sobrinho (2003) entende que muitas vezes eles não se apresentam em confronto ou oposição simples. Necessariamente, não teríamos a posição “ou um, ou outro” dentre eles, optando por um e negando o outro. O autor considera que “frequentemente se combinam e se entrecruzam”. Muitas vezes, para enfrentar a complexidade da teoria e prática da avaliação, faz-se necessário que se faça uso de uma multiplicidade de recursos metodológicos, sem que um paradigma exclua o outro. Entretanto, afirma que, “comumente são apresentados de forma isolada e esquemática, como se constituíssem campos que se excluíssem. É verdade que apresentam muitas diferenças e têm seus defensores e acusadores.” (DIAS SOBRINHO, 2003, p. 85).

Um paradigma foi organizado por esquemas objetivos, pautado numa ética utilitária, pragmática, na gestão racional e desenvolvimentista natural da ciência e da tecnologia, e por uma visão elementarista e mecanicista de mundo. O outro paradigma, conforme Margarida Fernandes (1998), apresenta uma outra perspectiva da realidade, que se situa no campo das epistemologias subjetivistas, naturalistas, fenomenológicas, que encara as verdades relativas às experiências humanas e as valoriza (FERNANDES, 1998).

Neste estudo, adotamos as nomenclaturas designadas por Souza Santos (2018) que, ao primeiro denomina paradigma dominante e, ao segundo, paradigma emergente. Ambos são descritos a seguir, bem como suas implicações na avaliação educacional.

#### **5.4.1 O paradigma dominante**

A evolução histórica das concepções de avaliação no século XX desvela que o modelo de ciência que prevaleceu no início do século influenciou toda a formulação teórica inicial e a prática da avaliação no contexto educacional. O paradigma que orientou os quatro primeiros períodos ou as três primeiras gerações descritas no tópico anterior foi denominado por Guba e Lincoln (2011) de paradigma científico. Fortemente apoiado na rigidez da racionalidade científica e nos conceitos da ciência moderna, esse paradigma é facilmente identificado pelos princípios que o sustentam.

Domingos Fernandes (FERNANDES, D., 2009), ao analisar os escritos dos autores Guba e Lincoln, refere-se ao paradigma que dominou as três primeiras gerações de avaliação, como paradigma psicométrico, ou velho paradigma. Moraes (1996) o chama de paradigma tradicional. Para descrever as implicações desse paradigma na educação e, mais especificamente, na avaliação educacional, apoiamo-nos nos estudos de vários autores (AFONSO, 2000; DIAS SOBRINHO, 2003; FERNANDES, D., 2009; FRANCO, 1990; GUBA; LINCOLN, 2011; LUCKESI, 2018; MORAES, 1996; SAUL, 2010; WORTHEN; SANDERS; FITZPATRICK, 2004). De maneira sucinta, à luz dos estudos indicados, procuraremos descrever os princípios desse paradigma, retomando suas bases epistemológicas para compreender suas implicações na educação e, conseqüentemente, na avaliação educacional.

O princípio fundamental de que a razão e a sensação são as formas de se obter o conhecimento e que, pela experimentação e observação controlada, pode-se objetivamente produzir o conhecimento científico pautado na regularidade das leis naturais reduz tudo que é observável pelos sentidos, como digno de ser estudado. A dicotomia entre o conhecimento científico e o senso comum, bem como a rígida separação entre natureza e ser humano se constituem fundamentos do paradigma dominante. Nessa perspectiva, a separação entre sujeito/ objeto requer um distanciamento e neutralidade que é condição indispensável para o exercício da ciência.

Souza Santos (2018), sobre algumas ideias centrais do paradigma, descreve:

As ideias que presidem à observação e à experimentação são as ideias claras e simples a partir das quais se pode ascender a um conhecimento mais profundo e rigoroso da natureza. Essas ideias são as ideias matemáticas. [...] Deste lugar central da matemática na ciência moderna derivam duas consequências principais. Em primeiro lugar conhecer significa quantificar. O rigor científico afere-se pelo rigor das medições. As qualidades intrínsecas do objeto são, por assim dizer, desqualificadas e em seu lugar passam a imperar as quantidades em que eventualmente se podem traduzir. O que não é quantificável é cientificamente irrelevante. Em segundo lugar, o método científico assenta na redução da complexidade. O mundo é complicado e a mente humana não o pode compreender completamente. (SOUZA SANTOS, 2018, p. 28).

Portanto, essa fragmentação do objeto, dos fatos, dos fenômenos, é condição inicial para que se observe e se meça com rigor. Desse modo, “conhecer significa dividir e classificar para depois determinar relações sistemáticas entre o que se separou.” (SOUZA SANTOS, 2018, p. 28). O conhecimento causal invoca a formulação de leis que determinam regularidades, as quais permitem a previsão de futuros fenômenos. Assim, o conhecimento científico reduz a complexidade da vida e a complicação do mundo, que a mente humana tem dificuldade de compreender completamente. O conhecimento é, portanto, baseado na formulação de leis e estas garantem ordem e estabilidade no mundo. Daí, o sucesso da ideia de que a racionalidade dos estudos da natureza poderia ser transportada tal e qual para estudos da sociedade. Se é possível descobrir leis naturais, também é possível descobrir leis da sociedade e no comportamento humano, de modo que as ciências sociais nasceram sob essa premissa. Um dos desdobramentos desses princípios epistemológicos é o de que fatos sociais, para serem convertidos em objetos de estudo, devem ser isolados do sujeito que o estuda e precisam ser tratados como coisas, exatamente como o cientista trata os fenômenos naturais. “Isto implica em considerar, de um lado, os fatos sociais desprovidos de historicidade, movimento e contradição e, de outro lado, o cientista social, ‘subjeto’, mas com a tarefa de se esforçar para estudar a realidade, da qual participa, como se não fizesse parte dela.” (FRANCO, 1990, p. 64).

Há uma visão fragmentada do processo de aquisição do conhecimento que é transferida para a visão da natureza, da sociedade, do ser humano e para a própria dinâmica da vida. Evidentemente, quando essa visão de mundo busca compreender a educação, o faz com essa perspectiva também separatista e fragmentada.

As implicações desse paradigma na educação e nas amplas dimensões envolvidas no fenômeno educativo, tais como a noção de aprendizagem, de currículo, dos métodos de ensino, dos materiais didáticos, do papel do professor, entre tantas outras, devem-se ao fato da

educação lidar diretamente com questões epistêmicas, e as concepções pautadas em princípios científicos carregados de objetividade geram segurança para os que desejam reproduzir o modelo e garantir os resultados obtidos ao longo do tempo. Contudo, o paradigma vai muito além das questões epistêmicas e apresenta respostas metodológicas e pragmáticas para os temas educacionais, dentre eles, a avaliação.

A avaliação educacional pautada no paradigma dominante tem, predominantemente, a finalidade de classificar e selecionar os alunos. A ênfase é dada na quantificação de capacidades, comportamentos ou atitudes passíveis de mensuração. A tendência é uma avaliação centrada nos resultados ou nos produtos. A aprendizagem dos alunos é uma realidade que pode ser avaliada de forma objetiva e neutra, sem interferência do avaliador que deve ter uma atitude técnica e isenta. Acredita-se que é possível avaliar sem qualquer interferência de natureza subjetiva por parte dos avaliadores, e estes são imunes ao contexto e às possíveis interações no processo avaliativo. Os instrumentos são, principalmente, os testes objetivos que permitem a quantificação das aprendizagens dos alunos e que possibilitam determinar o que os alunos sabem ou não sabem, já que, nesse paradigma, a aprendizagem dos alunos é uma realidade objetiva, passível de ser mensurada em sua totalidade por meio do processo de avaliação.

Moraes (1997) assim descreve a avaliação, nessa perspectiva paradigmática:

Em termos de avaliação, as provas assumem um papel central determinando o comportamento do aluno, privilegiando a memória e a capacidade de expressar o que foi acumulado. A preocupação do professor é que o aluno adquira conhecimentos específicos, de maneira determinada, e o anseio do jovem é estudar para tirar boas notas, tendo como referencial aquilo que lhe será cobrado. Acrescentando-se a isso as “galerias” nos murais das escolas com prêmios e punições, a competição entre alunos e uma avaliação classificatória, seletiva. O diploma é visto como um princípio organizador importante, símbolo do coroamento e da consagração de todo o ciclo de estudos. O conhecimento possui uma natureza estática, podendo ser representado por um pergaminho pendurado na parede simbolizando o seu ponto de chegada, o “final da linha”. (MORAES, 1997, p. 52).

Os avaliadores, nesse paradigma, estão comprometidos em prestar contas ao gestor ou administrador da instituição que contratou a avaliação, sem envolver qualquer outro grupo de interesse (*stakeholders*), sejam eles os alunos, sejam os pais ou a comunidade. Como a ênfase está nos resultados e na classificação, normalmente, estes são apresentados na forma de *rankings*, o que privilegia a competição e a concorrência entre alunos e entre instituições. Sendo a ênfase nos resultados, estes são investidos de autoridade e não devem ser discutidos nem contestados. Em função disso, os valores, o contexto, as diferenças não são levados em conta ou não entram em confronto com os dados ou com os resultados, pois são a

representação da realidade e, portanto, expressão da verdade. Outra consequência dessa ênfase nos resultados é que as avaliações são predominantemente somativas<sup>30</sup>, ou seja, ocorrem no final da ação ou levam em conta o produto e não o processo.

A rigidez produzida pela excessiva dependência do método científico e pela objetividade associada aos processos de quantificação geram avaliações pouco ou nada contextualizadas, pois a contextualização enfraquece o controle. Assim, as diferenças não são levadas em conta, tendo como consequência a exclusão dos que não atingem os padrões pré-estabelecidos como aceitáveis.

A incapacidade desse modelo paradigmático em atender e responder satisfatoriamente aos problemas da própria ciência, às mudanças sociais que desencadeiam expectativas de mudanças na educação geram uma tensão pela necessidade de um novo paradigma. Como considera Domingos Fernandes (2009, p. 81):

Os paradigmas transformam-se e evoluem através do desenvolvimento, por parte de dada comunidade científica, de novos conceitos, novas concepções, e valores que se destinam a procurar responder a problemas e a questões a que o paradigma dominante já não responde de forma satisfatória. É então que, em sua perspectiva, surge o que designou por *revoluções científicas* que abrem caminho à introdução de novas práticas e de novas formas de abordar as questões e que se traduzem por rupturas mais ou menos dramáticas uma vez que o *novo paradigma*, ou o paradigma emergente, pode questionar e pôr em causa todo o sistema de concepções, de valores e de interesses diversos que sustentam o *velho paradigma*. ( Grifos do autor)

Citando as três primeiras gerações de avaliação definidas por Guba e Lincoln (2011), Domingos Fernandes (2009) considera que a essência do paradigma dominante permeou todas as iniciativas descritas e ainda se faz vigente nos sistemas educacionais que convivem com um período de tensão pela necessidade de renovação e de melhorias do ensino, das aprendizagens e das escolas. A essa tensão que suscita a transição do modelo paradigmático, Souza Santos (2018) denomina de crise do paradigma dominante e assim a descreve:

[...] a crise do paradigma da ciência moderna não constitui um pântano cinzento de cepticismo ou de irracionalismo. É antes o retrato de uma família intelectual numerosa e instável, mas também criativa e fascinante, no momento de se despedir, com alguma dor, dos lugares conceituais, teóricos e epistemológicos, ancestrais e íntimos, mas não mais convincentes e securizantes, uma despedida em busca de uma vida melhor a caminho doutras paragens onde o otimismo seja mais fundado e a racionalidade mais plural e onde finalmente o conhecimento volte a ser uma aventura encantada. (SOUZA SANTOS, 2018, p. 62).

---

<sup>30</sup> A classificação foi feita primeiramente por Scriven, que faz a distinção entre a avaliação formativa – que se faz ao longo do processo; e a avaliação somativa – que é realizada depois de terminado o processo (SCRIVEN *apud* DIAS SOBRINHO, 2003, p. 23). Depois, essa mesma classificação foi ampliada por Bloom, que incluiu a avaliação diagnóstica – a qual ocorre antes de uma ação –, e considera somativa, aquela que está comprometida com a qualidade final da ação (BLOOM, *apud* LUCKESI, 2018, p. 174).

Assim, a crise no paradigma dominante aponta-nos para o perfil do paradigma emergente, cujo surgimento deve-se à necessidade de contraposição e de respostas aos desafios não contemplados e não atendidos pelo paradigma anterior.

#### **5.4.2 O Paradigma Emergente**

As sementes do paradigma emergente foram lançadas no solo da ciência moderna no início do século XX, com as descobertas da teoria quântica na Física, que não somente levantaram questões concernentes a problemas físicos, como aqueles relacionados ao próprio método das ciências naturais e exatas, e com a teoria da relatividade proposta por Einstein, em 1905, revolucionando o pensamento moderno. Essas descobertas esfacelaram os principais conceitos da visão de mundo cartesiana e da mecânica newtoniana relacionados à noção de tempo e espaço absolutos, à objetividade científica, à causalidade e vários outros conceitos até então sustentados pela ciência.

Souza Santos (2018) enumera algumas das principais características do paradigma emergente, tais como: i) deixou de ter sentido e utilidade a distinção dicotômica entre as ciências naturais e ciências sociais; ii) o conhecimento tende a não ser dualista e rompe com distinções e fragmentações disciplinares; iii) o conhecimento é total, como também é local, garantindo pluralidade metodológica; iv) todo conhecimento é autoconhecimento, rompendo com a dicotomia sujeito/ objeto e traduzindo-se em um saber prático; v) todo conhecimento visa constituir-se em senso comum, dialogando com outras formas de conhecimento, gerando uma nova racionalidade.

Guba e Lincoln (2011), ao denominar o paradigma que se contrapõe ao dominante, o chamam de construtivista, e segundo esses autores, este é também chamado de naturalista, hermenêutico ou interpretativo (com algumas ligeiras diferenças de significado entre si). Afirmam que ele “existe há várias centenas de anos, mas não foi amplamente aceito ou compreendido, particularmente nos países de língua inglesa.” (p. 97). Ao fazerem uma comparação entre o paradigma dominante e o seu sucessor, destacam três aspectos das crenças básicas que sustentam um paradigma: o ontológico, o epistemológico e o metodológico. A título de síntese, transcrevemos no Quadro 16 o esquema que os autores fizeram para estabelecer essa comparação.

**Quadro 16 - Comparação entre o sistema de crenças convencional e o construtivista**

<b>Crenças convencionais</b>	<b>Crenças construtivistas</b>
<p><i>Ontologia</i> A <i>ontologia</i> realista sustenta que existe uma única realidade que é independente do interesse de qualquer observador e é regida por leis naturais imutáveis, muitas das quais em forma de causa-efeito. Define-se verdade como um conjunto de postulados isomórfico em relação à realidade.</p>	<p>A <i>ontologia</i> relativista sustenta que existem múltiplas realidades socialmente construídas, não governadas por nenhuma lei natural, seja ela causal ou não. Define-se “verdade” como as construções mais fundamentadas (quantidade e qualidade das informações) e mais esclarecidas (intensidade segundo a qual as informações são compreendidas e utilizadas) sobre as quais exista consenso (embora possa haver inúmeras construções subsistentes que atendam simultaneamente a esse critério).</p>
<p><i>Epistemologia</i> A <i>epistemologia objetivista dualista</i> sustenta que um observador pode (aliás, deve) exteriorizar o fenômeno em estudo, mantendo-se à parte e distante dele (um estado em geral chamado de “dualismo sujeito-objeto”) e rejeitando todo e qualquer julgamento de valor que o influencie.</p>	<p>A <i>epistemologia subjetivista</i> monista sustenta que o investigador e o investigado estão entrelaçados de tal forma que as constatações de uma investigação são a <i>criação literal</i> do processo de investigação. Observe que essa postura efetivamente põe por terra a distinção clássica entre ontologia e epistemologia.</p>
<p><i>Metodologia</i> A <i>metodologia intervencionista</i> despe o contexto de suas influências (variáveis) contaminadoras (que provocam confusão/ confusionais), para que desse modo a investigação convirja para a verdade e explique a natureza tal como na realidade se apresenta e de fato funciona, do que decorre a capacidade de prever e controlar.</p>	<p>A <i>metodologia hermenêutica</i> encerra uma dialética contínua entre interação, análise, crítica, reiteração, reanálise e assim por diante, dando origem a uma construção conjunta/ articulada (entre todos os indagadores e respondentes ou entre as visões ética e êmica) sobre um determinado caso.</p>

Fonte: Guba e Lincoln (2011).

O Quadro 16 nos apresenta diferenças básicas entre os paradigmas, e os autores consideram que essa distinção dificulta qualquer conciliação entre ambos. Contudo, para Domingos Fernandes (2009), a coexistência de ambos no sistema educacional não é difícil de admitir. Como exemplo, o autor cita a ação de um professor em sala de aula, numa perspectiva do paradigma emergente, e uma avaliação externa do sistema de ensino adotando uma perspectiva psicométrica do paradigma dominante. Porém, as implicações educacionais do paradigma emergente desvelam as diferenças entre ambos.

Segundo Moraes (1997), a cosmovisão quântica que referencia o paradigma emergente traz uma visão holística e sistêmica de mundo e uma consciência da unidade da vida e do estado de inter-relação entre todos os fenômenos: físicos, biológicos, psicológicos, sociais, culturais e, conseqüentemente, os educacionais. Portanto, uma primeira implicação dessa nova visão na educação é a promoção de uma compreensão mais abrangente e adequada dos aspectos envolvidos na multidimensionalidade do processo educacional, reconhecendo a interconectividade dos problemas, “substituindo a compartimentação por integração,

desarticulação por articulação, descontinuidade por continuidade, tanto na parte teórica quanto na práxis da educação.” (MORAES, 1997, p. 86).

A visão sistêmica e a percepção de que as relações entre os fenômenos são dinâmicas faz com que a educação seja vista como processo, em que o mundo é concebido em movimento e, portanto, os fenômenos educacionais estão em permanente estado de mudança e transformação. Esse dinamismo requer muita criatividade, flexibilidade e capacidade inovadora. Consequentemente, surge uma outra visão de como o aluno aprende e adquire conhecimento. Não mais, o conhecimento é algo que se transmite ou que provém da sensação e da percepção do sujeito em observação do objeto, mas é uma construção que se dá pela força da relação indissociável entre sujeito e objeto, relação esta que promove uma ação transformadora do sujeito sobre o objeto e do sujeito sobre si mesmo. Portanto, o conhecimento também está em processo.

A visão de integralidade do indivíduo faz com que o aluno seja visto como um ser único, singular, que possui diferentes habilidades para aprender e estilos de aprendizagem distintos. Portanto, o respeito às diferenças e à diversidade caracterizam esse paradigma, bem como uma educação inclusiva em que as diferenças sejam respeitadas, mas que o aluno seja parte do coletivo, capaz de integração e interação com o outro e com a comunidade nos processos coletivos de construção do conhecimento. Esses espaços coletivos vão muito além da sala de aula e da escola. Como afirma Moraes (1996), uma escola expandida, que “sinaliza a importância da superação das barreiras existentes entre a escola e a comunidade, aluno e professor, escola e escola, pais e pais” (p. 16), reconhecendo que o conhecimento trafega em muitos e novos ambientes e que a educação se dá ao longo da vida e vai muito além dos conteúdos curriculares.

Nessa perspectiva, a avaliação educacional adquire características muito distintas daquelas que carrega no paradigma dominante. É reconhecida como um processo complexo, subjetivo, no qual se reconhece a impossibilidade de apreender na totalidade o que, de fato, os alunos sabem ou são capazes de fazer. A avaliação das aprendizagens desenvolve-se em processo, com uma participação ativa dos alunos em interação entre si e com o professor, reconhecendo que ambos interferem no processo e o influenciam por intermédio de seus conhecimentos, experiências e concepções. Assim, o uso de uma diversidade de estratégias e instrumentos de avaliação, aplicados ao longo do processo de ensino, acompanhando o desenvolvimento dos alunos, os processos que utilizam para resolução de problemas e execução de tarefas, privilegiando operações mentais mais sofisticadas, tais como descrição,

análise e interpretação, dão ênfase aos aspectos qualitativos, muito mais que aos quantitativos. Como afirma Domingos Fernandes (2009), as principais características da avaliação nesse paradigma “sugerem que a ênfase está na compreensão dos processos cognitivos dos alunos que, para tanto, são descritos, analisados e interpretados qualitativamente.” (FERNANDES, D. 2009, p. 83). Portanto, as abordagens qualitativa e holística são indispensáveis para que a avaliação seja capaz de trazer à luz a totalidade das situações educativas, sem ser restringida apenas por resultados de exames ou testes.

Outra característica da avaliação educacional na perspectiva do paradigma emergente é o encorajamento à cultura da autoavaliação com critérios consensualmente encontrados, enfatizando a participação, a cooperação e a autonomia dos alunos, num ambiente colaborativo. Portanto, em contraponto à competição, a avaliação deve encorajar a colaboração e cooperação. Esse clima de colaboração e cooperação pode esbarrar em pontos em que não haja consenso entre os envolvidos e que demandam a habilidade de negociação, quando a clareza, a transparência e o maior número possível de informações estejam disponíveis para todos os interessados.

Para Domingos Fernandes (2009), a coexistência dos dois paradigmas nos sistemas educacionais, em função do crescente número de avaliações em larga escala – cujas características são predominantemente do paradigma dominante – deve-se ao fato de que os sistemas educacionais não estão isolados e interagem e influenciam outros sistemas sociais, dos quais também podem depender para resolver os problemas que têm de enfrentar. Essas avaliações serviriam de referência e meios de comparabilidade entre os sistemas, e como explicação sobre essa situação de coexistência, descreve:

Pode haver várias explicações para essa situação. Desde as vantagens e desvantagens de cada uma das formas de avaliação quanto a questões de validade, de confiabilidade, até a evolução na elaboração de itens em exames ou provas de larga escala, ou à preocupação crescente em contextualizar dados. Mas há também uma questão que não devemos ignorar. Os sistemas educacionais e formativos têm de criar procedimentos que resultem em referências que as sociedades e seu funcionamento lhes exigem como é o caso da certificação ou da avaliação externa. Os sistemas educacionais, é importante que se diga, não vivem isolados dos outros sistemas sociais, interagem com eles, influenciam e são influenciados e, por si mesmos, não conseguem resolver todos os problemas que têm de enfrentar. Precisam que as sociedades também contribuam para sua solução. (FERNANDES, D. 2009, p. 84).

O sistema educacional assume, assim, um compromisso com o sistema social e uma relação de interdependência. Essa relação envolve o desafio do oferecimento de uma educação de qualidade, mas uma “qualidade distribuída com equidade, voltada para a

melhoria do processo de aprendizagem e capaz de garantir a equidade nos pontos de chegada visando à igualdade de oportunidade e de tratamentos.” (MOARES, 1996, p. 17).

Os documentos normativos que regulam a criação e a organização dos sistemas de avaliação desenvolvidos para a educação básica, no Brasil, preconizam que as avaliações em larga escala visam à “melhoria da qualidade do ensino” (BRASIL, 1996). Contudo, faz-se necessário analisar os pressupostos que envolvem a concepção de qualidade subjacente aos documentos oficiais que se expressam em políticas educacionais.

### **5.5 Concepções de qualidade educacional**

A melhoria da qualidade da educação é apresentada como o principal propósito do sistema de avaliação da educação básica desenvolvido e implementado no Brasil, nas últimas décadas. A temática da qualidade na educação não é nova, embora, nos últimos anos, tenha adquirido uma centralidade nas discussões e agendas de governos, de movimentos sociais, da mídia e entre pesquisadores do campo da educação. Apesar da relevância e centralidade do tema, está longe de alcançar algum consenso, está carregado de ambiguidade, marcado pela polissemia, “dado se tratar de um produto histórico, construído socialmente, que reflete posicionamentos políticos e ideológicos orientados por diferentes setores e perspectivas diversas.” (NARDI; SCHINEIDER; RIOS; 2014, p. 361).

Esquinsani e Dametto (2018) consideram que, “embora qualidade seja uma noção recorrente em diversos discursos contemporâneos acerca dos processos educacionais, reside sob tal noção, uma constelação de significados.” (p. 297). Afirmam que por trás de um aparente consenso, há concepções dissonantes, senão radicalmente opostas, e que “o convergente desejo de uma educação de qualidade e as múltiplas concepções atreladas a tal objeto trazem à tona a amplitude e a fragilidade dessa noção.” (ESQUINSANI; DAMETTO, 2018, p. 297).

Apesar da multiplicidade de concepções, o atributo “qualidade”, presente nos discursos educacionais, adquiriu legitimidade ao tornar-se princípio constitucional, meta consensual e unanimidade como padrão desejável do direito à educação. A noção de qualidade da educação ou do ensino pressupõe parâmetros comparativos que permitam a distinção entre uma boa ou uma má qualidade. Tratando-se a educação de um fenômeno social, a noção de boa ou má, será determinada pelo sistema de valores predominantes numa determinada sociedade que pode sofrer variações, de acordo com o momento histórico, com a

conjuntura temporal e espacial. Como afirma Azevedo (2011), seu conteúdo “se relaciona com o modo pelo qual se processam as relações sociais, produto dos confrontos e acordos dos grupos e classes que dão concretude ao tecido social em cada realidade.” (AZEVEDO, 2011, p. 422).

Gadotti (2013) afirma que “qualidade na educação implica em saber de que educação estamos falando, já que não existe uma só concepção de educação.” (p. 6). Nesse sentido, os mesmos paradigmas que regem as perspectivas filosóficas determinantes das visões de mundo, de homem, de conhecimento e que dão sentido a tudo que se relaciona ao fenômeno educativo, também dão sentido à subjetividade inerente à ideia de qualidade educacional. Morosini (2009) sustenta que a concepção de qualidade não é clara, mas a autora está certa de que é um construto imbricado no contexto dos paradigmas de interpretação da sociedade e do papel da educação. Nessa perspectiva, a qualidade educacional ora é concebida como uma construção social de inclusão democrática e participativa, ora é concebida como eficiência de processos e eficácia dos resultados de natureza performática. Sobre isso, Azevedo (2011) declara:

É importante lembrar que o conteúdo da qualidade é uma construção histórica e, portanto, variável, de acordo com os projetos de sociedade e, por conseguinte, de educação, em implementação ou em luta em cada conjuntura. Desta perspectiva, quando se está tratando do conceito de qualidade na educação, é necessário ter presente que são muitos os significados a ele atribuídos. Os próprios conteúdos do atributo qualidade, na medida em que resultam da reflexão/ ação humanas, historicamente, vêm se construindo e têm variado de acordo com os interesses de grupos e classes sociais. Por conseguinte, seu significado predominante – mas não unívoco – varia de acordo com os projetos hegemônicos em distintos contextos históricos, sem, no entanto, deixar de expressar uma contaminação de elementos oriundos de outras significações, peculiares às forças sociais em luta. (AZEVEDO, 2011, p. 423).

Sendo uma construção histórica, valemo-nos da análise do desenvolvimento da noção de qualidade da educação que Oliveira e Araújo (2005) traçaram, demarcando três significados distintos construídos e que foram adotados simbólica e concretamente na sociedade brasileira desde a primeira metade do século XX. Para os autores, é na década de 1940 que, no Brasil e no mundo, o tema da qualidade torna-se uma preocupação central no debate educacional. A necessidade de expansão de oportunidade de escolarização era evidente para a sociedade brasileira, e a oferta limitada de vagas foi o condicionador para a percepção de qualidade da escola, cujo acesso era insuficiente para todos e fruto de uma rigorosa seleção. “Portanto, a definição de qualidade estava dada pela possibilidade ou impossibilidade de acesso.” (OLIVEIRA; ARAÚJO; 2005, p. 8). Este foi, portanto, o primeiro indicador de qualidade: a expansão de oportunidades de acesso à escolarização.

No Brasil, o processo de expansão da escolarização obrigatória foi alcançado nas duas últimas décadas do século XX e, com a inclusão de uma grande parcela da população que não tinha acesso à educação, evidenciaram-se tensões, contradições e grandes diferenças, não apenas geradas pelos novos usuários das escolas, mas pelo sucateamento e precarização do trabalho docente, pelos insumos de segunda categoria disponibilizados para a quantidade crescente de prédios escolares, cuja principal preocupação era aumentar o número de vagas. Segundo Oliveira e Araújo (2005), “a escolarização da população gerou obstáculos relativos ao prosseguimento dos estudos [...] visto que não tinham as mesmas experiências culturais dos grupos que tinham acesso à escola anteriormente, e esta não se reestruturou para receber esta nova população.” (p. 9). Surge, assim, o segundo indicador ou condicionante que determinou o conceito de qualidade educacional – a possibilidade de continuidade e permanência dos alunos que progridem, ou não, no sistema de ensino – o fluxo retratado pela taxa de repetência. Se o obstáculo anterior era o acesso, nesse período, o obstáculo tornou-se a permanência com sucesso no sistema escolar. Os autores assim descrevem essa fase:

A partir da comparação entre a entrada e a saída de alunos do sistema de ensino, era medida a qualidade da escola. Se a saída se mostrasse muito pequena em relação à entrada, a escola ou sistema como um todo teria baixa qualidade. Com uma política pouco direcionada de expansão da escolarização mediante a construção de escolas, o Brasil, apesar do aumento expressivo do número de matrículas na etapa obrigatória de escolarização, chegou ao final da década de 1980 com uma taxa expressiva de repetência: de cada 100 crianças que ingressava na 1ª série, 48 eram reprovadas e duas evadiam (Brasil, Ministério da Educação, 1998), o que evidenciava a baixa qualidade da educação oferecida à população brasileira. (OLIVEIRA; ARAUJO, 2005, p. 10).

Inicia-se, então, um período de iniciativas de regularização do fluxo e de adoção de estratégias e programas para reverter essa tendência de reprovação e evasão. Já na Constituição Federal de 1988, o marco legal determina como um dos princípios da educação nacional, o ensino com garantia do padrão de qualidade e que seria estabelecido um plano nacional de educação com diretrizes, objetivos, metas e estratégias que assegurasse a melhoria na qualidade desse ensino (BRASIL, 1988). Conseqüentemente, a Lei de Diretrizes e Bases da Educação, além de reafirmar esse princípio, incumbiu à União o processo de avaliação do rendimento escolar em todos os níveis de ensino, “objetivando a definição de prioridades e a melhoria da qualidade de ensino.” (Art. 9º, inciso VI, LDB, 1996). Ao tratar do título sobre recursos financeiros, a lei os vincula a padrões mínimos de qualidade, sem definir quais são. Do mesmo modo, o Plano Nacional de Educação 2001-2010 apresentou como uma das diretrizes para a educação básica, a indissociabilidade entre o acesso, a permanência e a qualidade da educação escolar, sendo que o direito ao ensino fundamental não se referia

apenas à matrícula, mas ao ensino de qualidade. A meta 26 desse plano previa um programa de monitoramento dos indicadores do Sistema Nacional de Avaliação da Educação Básica, para que fosse assegurada a elevação progressiva dos níveis de desempenho dos alunos, tanto nesse sistema como nos que fossem criados por estados e municípios. Como avalia Dourado (2010), esse plano, a despeito de apresentar algumas metas de longo alcance em face dos grandes desafios de democratização e a melhoria da educação nacional, não se constituiu como base e diretriz para políticas, planejamento e gestão da educação nacional, nem foi acionado como tal pelos diferentes segmentos da sociedade civil ou pelos entes federados.

Estados e municípios já adotavam medidas para regularização do fluxo, tais como a adoção de ciclos de escolarização, promoção continuada, programas de aceleração da aprendizagem, de tal forma que houve uma queda expressiva na taxa de repetência. A questão que surgiu, então, foi se esses mecanismos surtiam o efeito de melhoria da qualidade de ensino, de fato. Oliveira e Araújo (2005, p. 12) consideram que:

Se o combate à reprovação com políticas de aprovação automática, ciclos e progressão continuada incide sobre os índices de “produtividade” dos sistemas, gera-se um novo problema, uma vez que esses mesmos índices deixam de ser uma medida adequada para aferir a qualidade. Se existem políticas e programas que induzem a aprovação, a tarefa de aferir a qualidade num sistema com um índice de conclusão igual ou superior a 70% torna-se mais complexa.

Com a constatação da dificuldade que as políticas de progressão ou aprovação automática estavam gerando sobre os índices de produtividade dos sistemas, um terceiro indicador de qualidade da educação é incorporado, que é a capacidade cognitiva dos estudantes em testes padronizados em larga escala. Assim, a qualidade também passa a ser aferida mediante o estabelecimento de matrizes ou diretrizes que apresentam os conteúdos ou habilidades que os alunos devem dominar nos diferentes níveis ou etapas da escolarização. Segundo esse pressuposto, os testes ou exames são elaborados a partir desses conteúdos prescritos, avaliando se o aluno os aprendeu ou não. Dessa forma, a qualidade é aferida pela medida do desempenho cognitivo dos alunos. Sendo assim, a associação entre desempenho cognitivo e fluxo, que deu origem ao Ideb como índice que sintetiza as duas medidas pode ser vista como um indicador numérico de qualidade e comparabilidade.

Dirce Freitas (2013, p. 73) considera que:

A questão da concepção de “qualidade” em educação escolar básica inscreve-se não só nos discursos justificadores das medidas anunciadas e adotadas, mas na natureza das próprias escolhas feitas e dos processos que as materializam. Na esfera do governo da

educação escolar básica, o Brasil escolheu operar com uma concepção de qualidade baseada no mix “fluxo escolar e desempenho cognitivo”. O primeiro orientado para uma situação de progressão regular, pela redução da reprovação, do abandono e da evasão escolar. O segundo, pelo desenvolvimento de competências e habilidades em leitura (Língua Portuguesa) e em Matemática.

Contudo, o percurso descrito nesses três momentos – i) ênfase na expansão de vagas; ii) ênfase na continuidade e permanência; iii) ênfase no fluxo e desempenho – concretiza-se num contexto supranacional, no qual o Brasil assumiu compromissos, ao lado de outros chefes de Estado, com uma agenda de organismos internacionais de financiamento e cooperação técnica que passaram a interferir gradativamente nos planos educacionais e, conseqüentemente, nas ideias acerca da qualidade da educação. Fonseca (2009) destaca projetos e programas negociados e acordados com o Banco Mundial e a exigência da Organização das Nações Unidas para Educação e Cultura (Unesco) de que os países membros formulassem planos educacionais para a década, embasados na Declaração Mundial de Educação para Todos, elaborada por ocasião da Conferência Internacional de Jomtien, em 1990. O Ministério da Educação elaborou o Plano Decenal de Educação Para Todos (1993-2003) que, para a autora, assumia as orientações doutrinárias e as metas quantitativas do fórum internacional, com a previsão de uma série de ações para qualidade da educação, entre elas, um amplo sistema de avaliação da educação básica.

Dirce Freitas (2011) também destaca os sinais de convergência com as diretrizes políticas supranacionais, citando a participação do Brasil em vários programas de monitoramento, dos quais cita: i) o Programa Internacional de Avaliação do Aluno (PISA), da OCDE; ii) Programa de Indicadores Mundiais de Educação (WEI), que coleta dados anualmente para o levantamento de indicadores estatísticos de países em desenvolvimento; iii) Programa de Indicadores de Educação do Instituto de Estatística, da OCDE (WEI-SPS); iv) Pesquisa Internacional sobre Ensino e Aprendizagem (Talis), da OCDE. Para a autora, nessas esferas, vê-se um alargamento na concepção de qualidade, “embora derivada de políticas que relacionam mercado com gestão, performatividade e transformação na natureza do Estado.” (FREITAS, D., 2011, p. 75). Conclui que a qualidade em educação escolar básica tem se tornado cada vez mais objeto de monitoramento na esfera do governo e da governança.

Dourado, Oliveira e Santos (2007) destacam a participação do Brasil e os compromissos assumidos junto à Cúpula das Américas, num esforço de “identificar as condições, dimensões e fatores fundamentais do entendimento do que seja Educação de Qualidade” (p. 7), considerando a ótica dos países membros e dos organismos multilaterais como a Unesco e o Banco Mundial, tendo em vista a considerável influência que estes

exercem na formulação das políticas educacionais da região. Destacam-se os marcos acordados no Plano de Ação em Educação adotados pelos chefes de Estado e governos com as seguintes metas gerais:

[...] assegurar que até o ano de 2010, todas as crianças concluam a educação primária de qualidade e que, pelo menos, 75% dos jovens tenham acesso à educação secundária de qualidade, com percentagens cada vez maiores de jovens que terminem seus estudos secundários, e oferecer oportunidades de educação ao longo da vida à população em geral. (DOURADO; OLIVEIRA; SANTOS, 2007, p. 8).

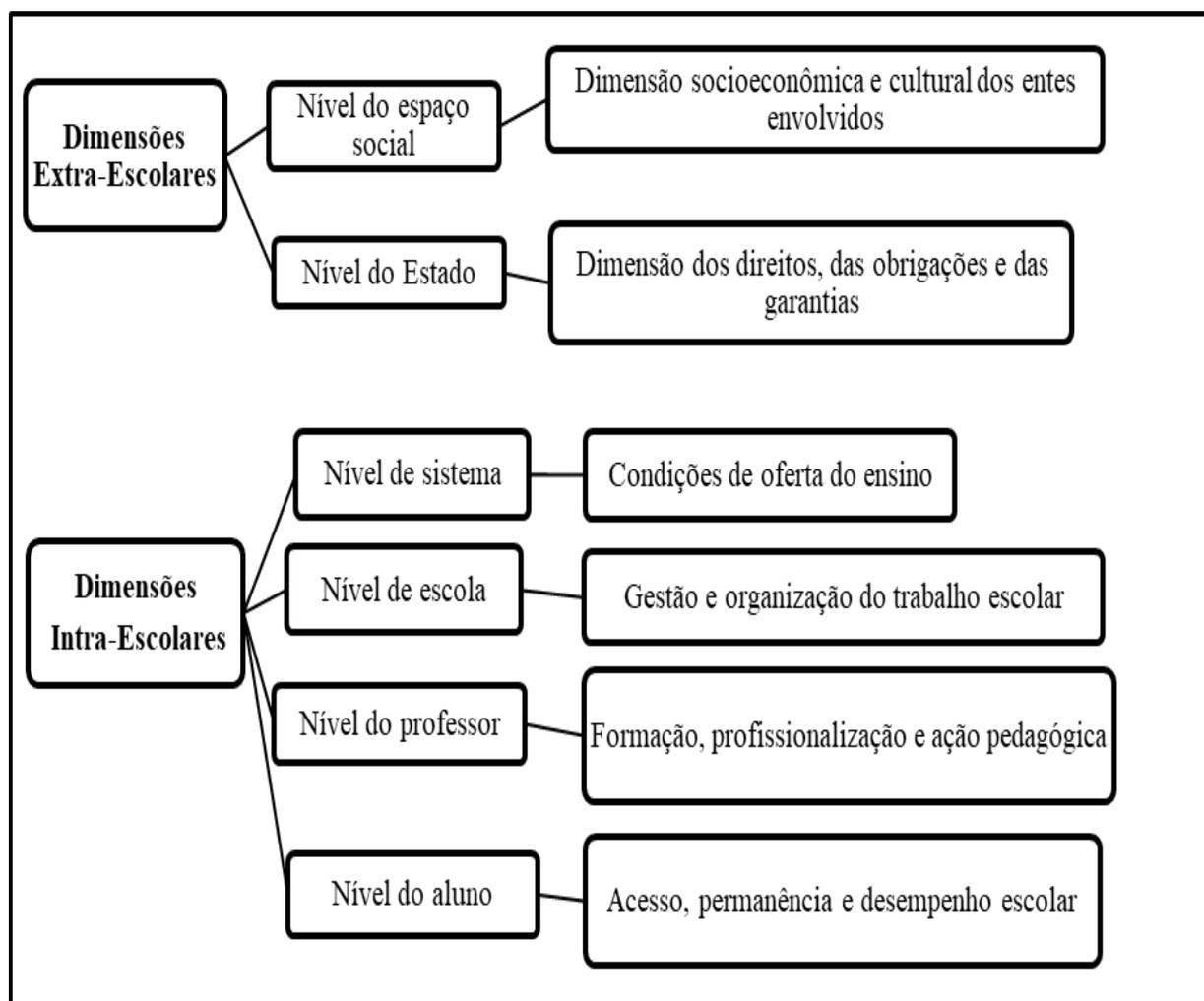
Essas metas estabelecidas nos documentos da Cúpula das Américas perseguindo uma melhoria na qualidade da educação na região refletem as concepções da OCDE e da Unesco que, segundo os autores, utilizam como paradigma “para aproximação da Qualidade da Educação, a relação *insumos-processos-resultados*.” (DOURADO; OLIVEIRA; SANTOS, 2007, p. 9, grifos dos autores). Sobre essas concepções, ainda consideram:

Desse modo, a Qualidade da Educação é definida envolvendo a relação entre os recursos materiais e humanos, bem como a partir da relação que ocorre na escola, na sala de aula, ou seja, os processos de ensino aprendizagem, os currículos, as expectativas de aprendizagem com relação à aprendizagem das crianças etc. Destaca, ainda, que a qualidade pode ser definida a partir de resultados educativos, representados pelo desempenho dos alunos. (DOURADO; OLIVEIRA; SANTOS, 2007, p. 9).

Ao fazerem uma síntese das concepções de qualidade educacional que aparecem nos documentos dos organismos multilaterais, os autores apresentam alguns pontos em comum, “por exemplo, a vinculação do conceito de qualidade a medição, rendimento e a indicação da necessidade da instituição de programas de avaliação da aprendizagem.” (p. 12). A ênfase recai sobre a necessidade de ferramentas e diagnósticos por meio de sistemas de medição que contribuem para a melhoria da qualidade a partir do estabelecimento de parâmetros de referência de aprendizagem e da “indicação de padrões, insumos e processos eficazes que permitam avançar no rendimento escolar, mobilizando a opinião pública a favor da educação.” (p. 12). Os autores, nesse estudo teórico-conceitual das pesquisas e produções sobre a qualidade da educação, traçaram um cenário das dimensões extra e intraescolares, assim como apontaram diferentes atores envolvidos na dinâmica pedagógica que, direta ou indiretamente, interferem nos resultados educativos.

O Quadro 17 representa, de forma sintética, as dimensões mínimas comuns nos documentos analisados pelos autores e que são consideradas como determinantes e devem ser levadas em conta para uma educação de qualidade.

#### **Quadro 17 - Dimensões mínimas comuns da qualidade da educação**



Fonte: DOURADO; OLIVEIRA; SANTOS (2007); elaboração da autora.

Para Dourado, Oliveira e Santos (2007), a amplitude das dimensões<sup>31</sup> que envolvem a garantia mínima de qualidade enseja a busca de parâmetros analíticos mais complexos que possibilitem a avaliação da realidade educacional, além da medição efetivada por meio de provas estandardizadas.

Numa pesquisa sobre as concepções de qualidade de ensino no contexto das avaliações em larga escala em artigos publicados no período de 1995-2012, utilizando a metodologia de análise do conteúdo, Manfio (2013) investigou a produção em periódicos de relevância nacional no período, que compreendeu dois governos de propostas políticas diferentes. O primeiro, de 1995 a 2002, e o segundo, de 2003 a 2012. A autora apresenta duas concepções de qualidade distintas nas publicações dos dois períodos. No primeiro, que corresponde ao

<sup>31</sup> Parte dessas dimensões está presente na proposta original do Saeb de 1990 e foi retomada na proposta do Sinaeb, como vimos no capítulo três (Quadro 9) e que não se estabeleceu, ao ser revogada a portaria que o criou.

governo de Fernando Henrique Cardoso, o conceito de qualidade está aliado à democratização do ensino, para a cidadania e preparo para o mercado de trabalho, porém, os autores analisados são unânimes em afirmar que essa melhoria está relacionada a números e resultados. Assim, o período é descrito da seguinte forma:

Destarte, o conceito de qualidade de ensino implicado nesse período diz respeito à democratização e um ensino para a cidadania e preparo para o mercado de trabalho. Com as alterações no sistema de ensino público vimos que por um lado foi atribuída autonomia para a escola e por outro lado, foram atribuídas avaliações para a regulação dessa autonomia, no entanto, essas avaliações segundo os autores, não estão voltadas para a qualidade de ensino, mas sim deflagram a escola em rankings e apenas números que nada têm a ver com a qualidade de ensino evidenciada pelos autores. (MANFIO, 2013, p. 31339).

No segundo período, que corresponde aos dois mandatos do presidente Luiz Inácio Lula da Silva, a concepção de qualidade de ensino permanece a mesma, porém com uma ênfase na necessidade de participação da comunidade externa e interna das escolas, muitas críticas à capacidade dos sistemas de avaliação em larga escala em promover a melhoria da qualidade da educação e o surgimento do tema da responsabilização vinculado a essas questões. Assim a autora sinaliza seus achados no período:

Parece claro na produção dos artigos nesse período sobre a qualidade na escola que a qualidade de ensino não pode ser algo decretado sem a participação da comunidade. Obtido seu engajamento no processo, os avanços são sobejamente ampliados. O descuido com o envolvimento da comunidade intra e extraescolar acaba deslegitimando o processo de produção de qualidade nas escolas e desperdiçando as evidências obtidas via avaliação. (MANFIO, 2013, p. 31342).

Nesse cenário de mobilização e participação efetiva da sociedade civil, foi aprovado o Plano Nacional de Educação 2014-2024, pela Lei 13.005, de 25 de junho de 2014, que propiciou uma discussão sobre qualidade da educação, associada às bandeiras históricas de luta pela democratização de acesso e permanência, aliadas à garantia de recursos para o financiamento e políticas de gestão mais orgânica para a educação. Apesar da ampla discussão e da proposta de um sistema de avaliação nacional da educação básica que agregasse indicadores de avaliação institucional, de infraestrutura, de recursos pedagógicos, de processos de gestão, além dos indicadores de rendimento escolar e censo, o Plano em vigência na presente década define, na meta 7, “fomentar a qualidade da educação básica em todas as etapas e modalidades, com melhoria do fluxo escolar e da aprendizagem de modo a atingir médias nacionais” (BRASIL, 2014) pré-fixadas para o Ideb. O documento sinaliza dois indicadores de qualidade: o fluxo escolar – indicado pela taxa de aprovação –; e o desempenho dos estudantes – indicado pelas médias de proficiência nas avaliações padronizadas –. Para Dourado (2021), essa visão restritiva da meta ratificou o modelo de

avaliação vigente, reafirmando a centralidade de uma avaliação estandardizada via exames de larga escala e índices. Ainda assim, o autor considera que “o PNE representa um avanço na agenda educacional e, se materializado como política de Estado, contribuirá efetivamente para a melhoria da qualidade da educação no País.” (DOURADO, 2021, p. 182).

Ainda num esforço de revisão bibliográfica sobre a noção de qualidade educacional aplicada à educação básica, Esquinsani e Dametto (2018) procuram evidenciar que a expressão não comporta nenhum sentido propriamente teórico ou científico, como alguns sugerem, mas prioritariamente político: “assemelha-se a um juízo de caráter valorativo moral endereçado a uma experiência educacional, seguindo parâmetros peculiares a cada forma de conceber o atributo ‘qualidade’” (p. 309). Consideram, apesar de carregado de imprecisão, haver um intenso uso político dessa noção. Assim, levantaram oito posições historicamente estabelecidas que situam qualidade educacional como: 1) excelência ou excepcionalidade; 2) oferta de vagas ou acesso universal à educação; 3) fluxo adequado de alunos; 4) presença e diversidade de insumos; 5) investimento desejável por aluno; 6) adequação de processos e do ambiente; 7) relevância social da educação e, por fim, 8) performance em avaliações padronizadas. Após uma análise de cada posição, os autores concluem que:

No momento atual, em que a qualidade se alicerça, predominantemente, em resultados de avaliações em larga escala, tratados estatisticamente e comparados a padrões supostamente desejáveis ou ideais, há uma aparente despolitização do debate acerca dos rumos da educação, emergindo critérios supostamente técnicos de qualidade, mas que elegem, de modo sorrateiro, caminhos politicamente enviesados para a educação básica. (ESQUINSANI; DAMETTO, 2018, p. 309).

Na perspectiva da evolução histórica das concepções de qualidade da educação, encontramos uma nova concepção que, segundo Flach (2013), foi gestada no interior dos movimentos sociais e partidos de oposição a partir de 1980 e que passa a ser defendida em contraposição à ideia de eficiência e produtividade inerente ao modelo considerado transplantado do mundo empresarial, no qual muitos sistemas de ensino adotam o modelo gerencial da qualidade total<sup>32</sup>. Denomina-se qualidade social da educação esse esforço de contraposição, cujos eixos principais são: “educação de qualidade social como direito de cidadania, gestão democrática com participação popular, valorização dos trabalhadores da educação, financiamento em regime de colaboração e integração entre as políticas e outros entes jurídicos e sociais.” (FLACH, 2013, p. 4).

---

<sup>32</sup> Qualidade Total ou Gestão da Qualidade Total refere-se a uma estratégia gerencial que se fundamenta em pontos como: foco no cliente; trabalho em equipe permeando toda a organização; decisões baseadas em fatos e dados; a busca constante da solução de problemas e da diminuição de erros. (LONGO, 1996, p. 10 *apud* GALLANI, 2012)

Silva (2009), sobre a qualidade social, conclui:

A qualidade social da educação escolar não se ajusta, portanto, aos limites, tabelas, estatísticas e fórmulas numéricas que possam medir um resultado de processos tão complexos e subjetivos, como advogam alguns setores empresariais, que esperam da escola a mera formação de trabalhadores e de consumidores para os seus produtos. A escola de qualidade social é aquela que atenta para um conjunto de elementos e dimensões socioeconômicas e culturais que circundam o modo de viver e as expectativas das famílias e de estudantes em relação à educação; que busca compreender as políticas governamentais, os projetos sociais e ambientais em seu sentido político, voltados para o bem comum; que luta por financiamento adequado, pelo reconhecimento social e valorização dos trabalhadores em educação; que transforma todos os espaços físicos em lugar de aprendizagens significativas e de vivências efetivamente democráticas. (SILVA, 2009, p. 225).

Carvalho (2012), fazendo uma análise discursiva da presença do conceito de qualidade social da educação nos documentos oficiais produzidos pelo Ministério da Educação desde a promulgação da Lei de Diretrizes e Bases da Educação, na década de 1990, conclui que “evidenciam-se enunciados da qualidade social da educação na perspectiva da educação republicana em coexistência com elementos do discurso da qualidade total da educação” (p. 1). O discurso da qualidade social opera por meio da colaboração entre os entes federados, associado aos processos de controle e regulação da política educacional do país.

A despeito da polissemia do conceito de qualidade, considera-se que os principais autores referenciados para este estudo a definem como uma construção social que se altera segundo os interesses dos grupos, neste caso, ligados à instituição educativa e que refletem as características da sociedade na qual estão inseridos. Consoante à perspectiva adotada no enunciado da tese proposta, a noção de qualidade educacional que referencia a análise do objeto de pesquisa ampara-se nas ideias de Dourado (2020). Acompanhando o autor, é reconhecido este cenário de tensão entre distintas concepções político-pedagógicas de educação, do papel do Estado, das articulações de organismos multilaterais e de processos e dinâmicas de regulação e avaliação que impactam a qualidade da educação. Dourado (2020) acentua os aspectos intra e extraescolares que precisam ser considerados e que implicam a qualidade da educação.

Uma educação de qualidade implica, portanto, considerar a multiplicidade desses aspectos, como resultado de processos coletivos e democráticos, articulados à concepção de educação e qualidade social, às condições de acesso e permanência, aos sujeitos envolvidos no processo e suas condições concretas, à dinâmica formativa e aos aspectos político-pedagógicos que consubstanciam o ato educativo, envolvendo a aquisição e a produção de conhecimentos e saberes significativos, a avaliação formativa, a definição coletiva de base comum nacional que garanta a unidade na diversidade. É preciso pensar em processos avaliativos mais amplos, vinculados a projetos educativos democráticos e emancipatórios, contrapondo-se à centralidade conferida à avaliação como medida de resultado, que se traduz em instrumento de controle, ranqueamento e competição institucional. (DOURADO, 2020, p. 180).

Portanto, o autor faz uma relação entre o processo avaliativo e a qualidade da educação, mas como um aspecto que pode ser considerado entre outros, e vinculado a projetos educativos democráticos e emancipatórios, não ocupando a centralidade como principal indicador. De fato, faz-se necessário o estabelecimento de padrões referenciais para que seja possível comparar processos educacionais e se emitir um juízo de valor, a partir desses padrões. A avaliação externa e em larga escala pode contribuir como uma fonte de dados para promover a qualidade educacional, desde que articulada a inúmeros outros indicadores observáveis que sejam capazes de refletir as aspirações e expectativas dos atores envolvidos no processo educativo.

Por mais que subsistam diferentes concepções de qualidade educacional, por mais dificuldades em encontrar consensos sobre qual significado adotar em determinado contexto educacional, a busca por uma educação de qualidade tem sido um dos maiores objetivos dos sistemas de ensino. A necessidade de melhoria da qualidade educacional do país tem sido reivindicada por inúmeros segmentos da sociedade, de políticos a pesquisadores da educação, gestores, pais ou alunos. Apesar da complexidade do tema, ainda há um vasto espaço para discussão e investigação em busca de convergências que favoreçam que o direito à educação seja adjetivado pela qualidade, o que abre espaço para questionar: qual qualidade?

## **CAPÍTULO 6 – AVALIAÇÃO DO PROGRAMA DE AVALIAÇÃO EXTERNA DA REDE**

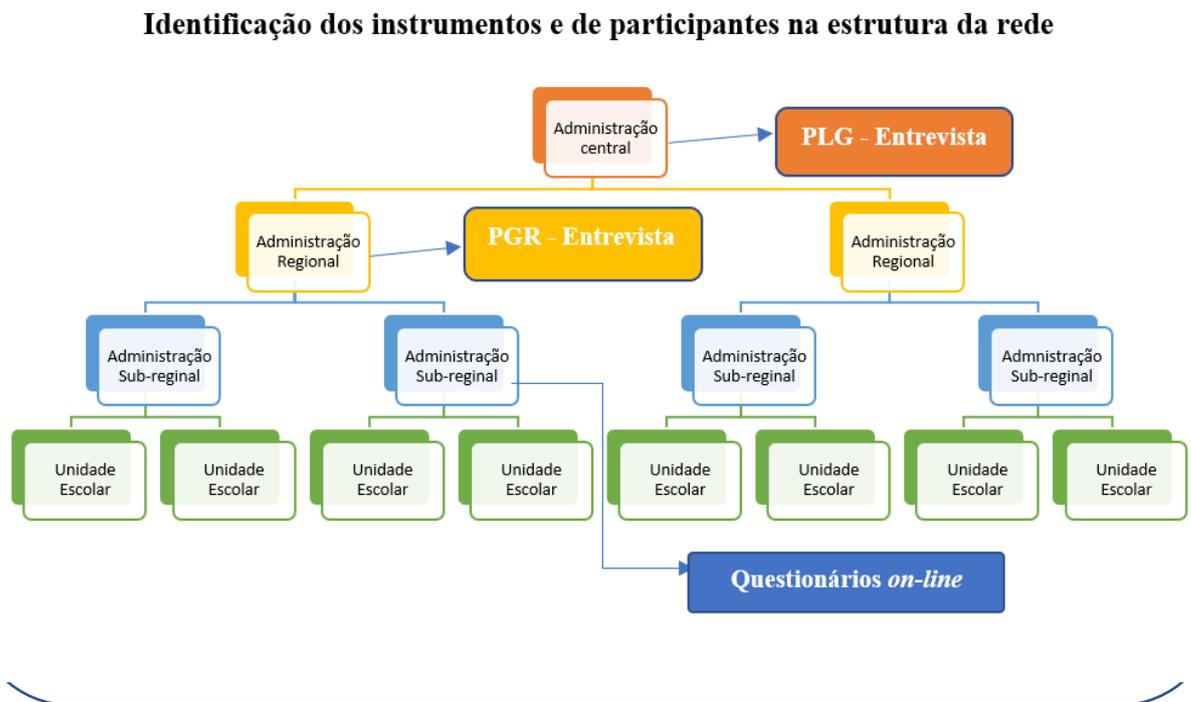
Este capítulo apresenta o resultado das análises dos dados produzidos pelos questionários *on-line*, pelas entrevistas semiestruturadas e documentos disponibilizados pela rede com orientações sobre o PAER para gestores, pais e alunos. No percurso de toda a análise, a indagação inicial que deu origem ao problema de pesquisa, orientou a reflexão analítica, delineando o processo de compreensão e interpretação do material empírico: quais aspectos desenvolvidos e quais a desenvolver podem ser identificados com a implementação de um programa de avaliação externa em larga escala por uma rede privada de ensino de abrangência nacional, com vistas a assegurar qualidade à educação oferecida?

A análise do material empírico, na busca por dar sentido, organizar e categorizar os dados, à luz das referências teóricas apresentadas ao longo do estudo, principalmente no capítulo cinco, possibilita a realização de uma meta-avaliação numa perspectiva formativa e construtivista (GUBA; LINCOLN, 2011). Ao irmos além do aspecto descritivo dos dados coletados, buscando a interpretação destes, com a finalidade de apontar os aspectos desenvolvidos e a desenvolver do programa de avaliação investigado, pretendemos alcançar o propósito de realizar a avaliação de uma avaliação.

O capítulo está estruturado em três seções: na primeira, a análise dos dados produzidos pelas respostas ao questionário *on-line*; na segunda, a análise dos dados produzidos pelas entrevistas semiestruturadas; na terceira, uma síntese integradora de todo o material. Dentro da primeira e da segunda seção, as informações constantes nos documentos escritos e que são atinentes às categorias temáticas se tornam objetos de análise conjuntamente.

Precedendo a análise dos dados, a Figura 7 descreve os instrumentos utilizados nos diferentes níveis da estrutura administrativa da rede. Reiteram-se os critérios adotados para a escolha dos participantes: i) a função que exerce na estrutura administrativa da rede, ocupando posições estratégicas de acesso a dados, tomada de decisão e que pode fazer uso do programa de avaliação como possível propulsor da qualidade educacional pretendida; ii) o tempo de atuação na função de, pelo menos, dois anos ou duas aplicações do PAER, que permita a contribuição com sua percepção e análise do processo.

**Figura 7 - Instrumentos e participantes representados na estrutura organizacional**



Elaboração da autora.

Na Figura 7, indica-se que as entrevistas foram realizadas com participantes dos dois primeiros níveis de gestão: administração para todo o país e administração das oito regiões; e o questionário, para administradores das 50 sub-regiões de todo o território nacional.

### 6.1 Análise dos Questionários

Os questionários mistos - com questões objetivas e perguntas abertas - foram disponibilizados para 105 participantes que ocupam funções de gestão da rede em nível de sub-regiões distribuídas em todo o território nacional. Responderam ao questionário 47 participantes, indicando um percentual de 44% de participação, índice este considerado muito significativo, ante a projeção de Marconi e Lakatos (2017) que indicam um percentual médio de 25% de devolução.

O questionário foi pré-testado e ajustado, atendendo às sugestões dos participantes da testagem. Como foram enviados por correio eletrônico, e com as restrições de acesso a dados pessoais dos servidores da rede sem o consentimento deles, foi necessário que cada gestor regional intermediasse e disponibilizasse os endereços eletrônicos dos respectivos liderados e colaborasse com o envio dos endereços para a pesquisadora. Esse processo durou seis meses,

pois o contato com cada líder regional foi feito por *WhatsApp*, aguardando que respondessem para um número, a princípio, desconhecido de seus contatos. O período de restrições e isolamento social imposto pela pandemia da Covid 19 afastou quase a totalidade dos participantes de seus locais de trabalho, e os telefones disponíveis para a pesquisadora eram os dos escritórios das sedes regionais. Esta foi a razão da dependência de contatos via dispositivo de conversas *WhatsApp*. Esse cenário valoriza o retorno significativo das 47 respostas aos questionários enviados.

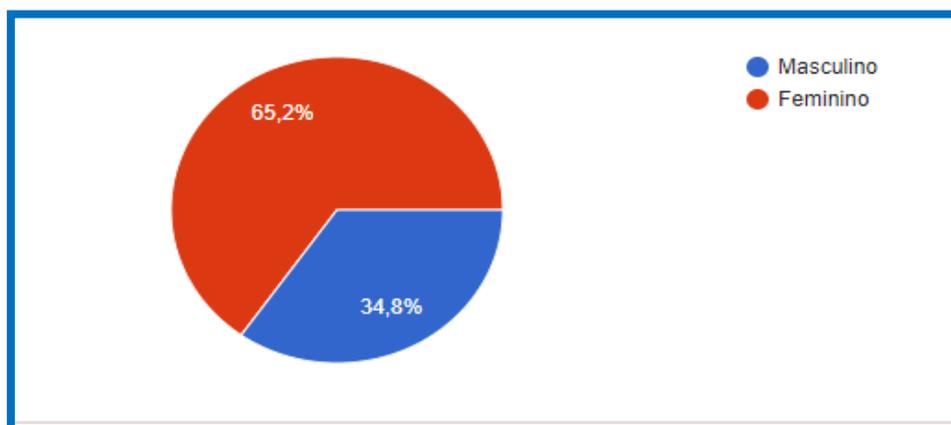
O questionário é composto por sete questões objetivas e cinco questões abertas. O perfil dos participantes é descrito por meio de gráficos que foram gerados pelo próprio formulário do *GoogleDocs* e apresentados a seguir.

### 6.1.1 Perfil dos participantes do questionário

O perfil dos participantes do questionário foi construído a partir de seis itens: i) Sexo; ii) Idade; iii) Tempo de atuação na rede; iv) Tempo de atividade na função atual; v) Tempo de magistério na educação básica; vi) Edições do PAER de que participou na função atual.

O primeiro item sobre o sexo dos participantes é descrito no gráfico a seguir.

**Gráfico 3 - Questão 1: Sexo dos participantes**

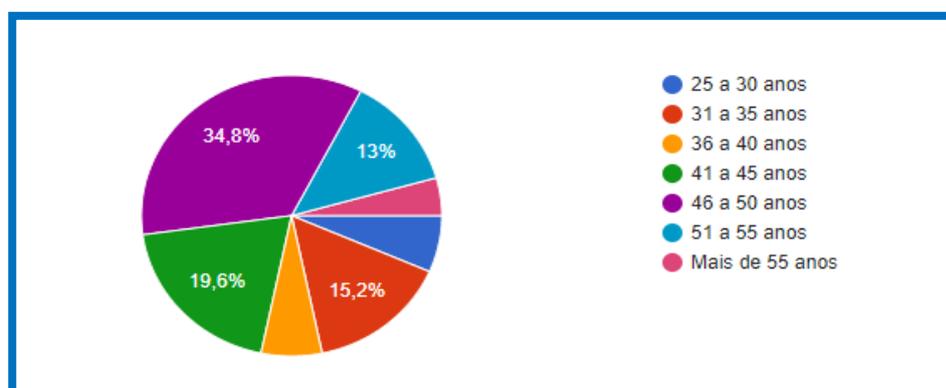


Fonte: *GoogleDocs*

Há o predomínio de mulheres respondentes, e este percentual representa parcialmente a distribuição entre os participantes, pois, dos 105 participantes originais a quem o questionário foi enviado, 53% eram mulheres e 47%, homens. No total de respondentes, 65,2% são mulheres e 34,8% homens.

A idade dos participantes é representada no Gráfico 4 abaixo.

Gráfico 4 - Questão 2: Idade

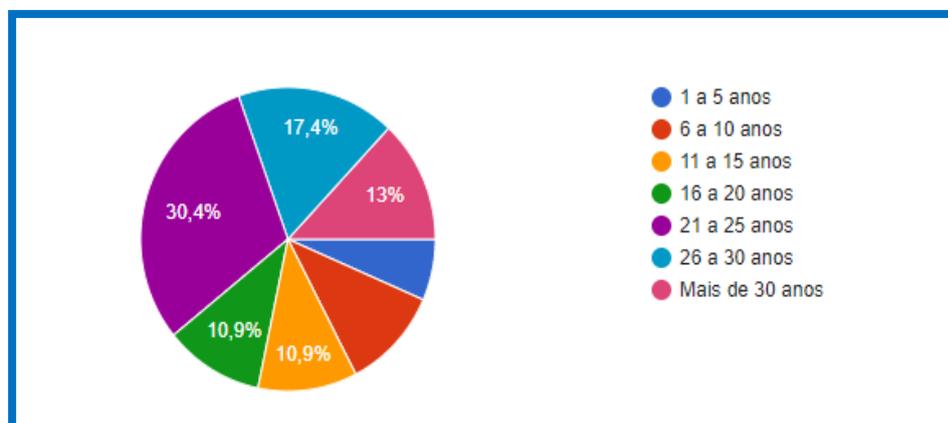


Fonte: *GoogleDocs*

O maior percentual de participantes (34,8%) encontra-se na faixa de 46 a 50 anos, seguido de participantes entre 41 e 45 anos (19,6%) e o menor percentual está na faixa etária acima de 55 anos (4,3%).

O tempo de atuação dos participantes na rede é representado no Gráfico 5.

Gráfico 5 - Questão 3: Tempo de atuação na rede

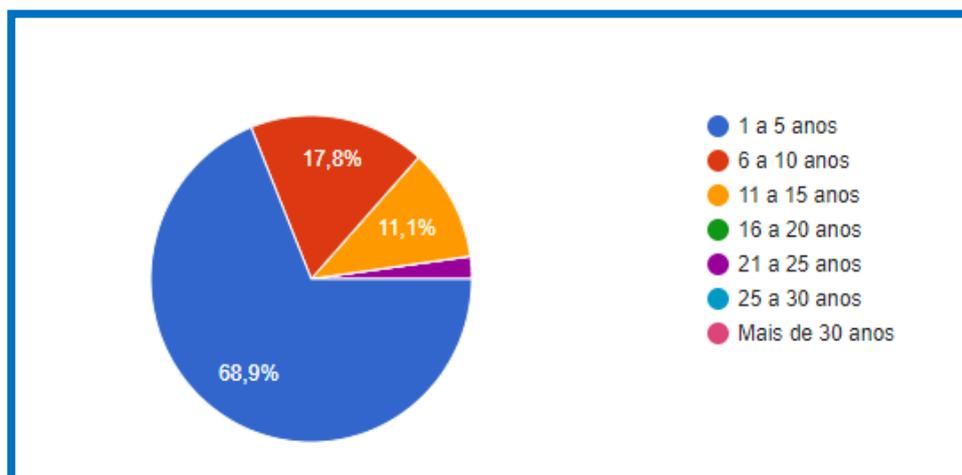


Fonte: *GoogleDocs*

Constata-se que a maior concentração de gestores sub-regionais (30,4%) conta um tempo de atuação na rede entre 21 e 25 anos, seguida pelos que atuam há um período de 26 a 30 anos (17,4), representando mais de 50% dos respondentes, aqueles que estão há mais de 21 anos na rede em função de gestão. Os que têm entre 1 e 5 anos de atuação na rede representam o menor percentual (6,4%).

O tempo que os participantes estão atuando na função é descrito no Gráfico 6.

**Gráfico 6 - Questão 4: Tempo de atividade na função atual**

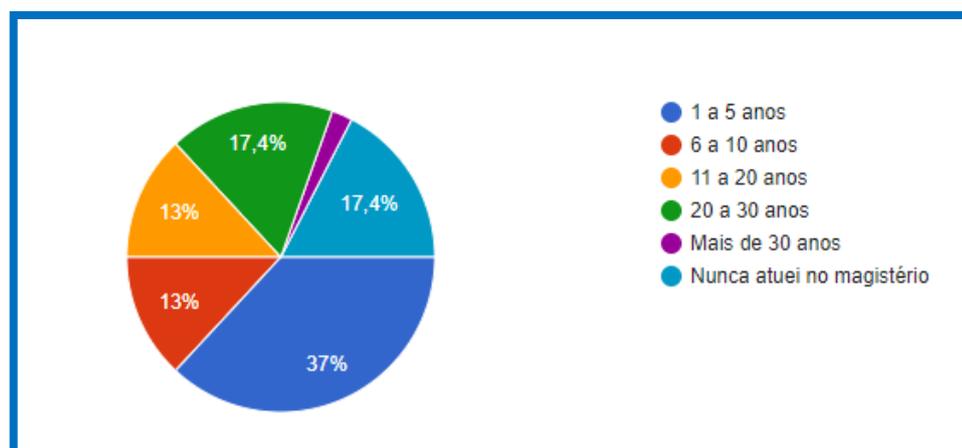


Fonte: *GoogleDocs*

Esta resposta apresenta um percentual significativo de participantes que estão na atual função de gestor sub-regional há um período de 1 a 5 anos (68%). Os respondentes com 21 a 25 anos na função representam o menor percentual (2,2%). Este dado revela que, apesar do maior número de participantes atuantes na rede há mais de 20 anos encontrar-se na faixa acima de 45 anos, esse grupo passou por outras funções antes de ocupar a atual. Dentro da estrutura da rede, podem ter atuado em funções nas unidades escolares, desenvolvendo experiência prévia nas escolas, antes de ocuparem a função atual.

O Gráfico 7 representa o tempo de atuação dos participantes como docentes da educação básica.

**Gráfico 7 - Questão 5: Tempo de magistério na Educação Básica**



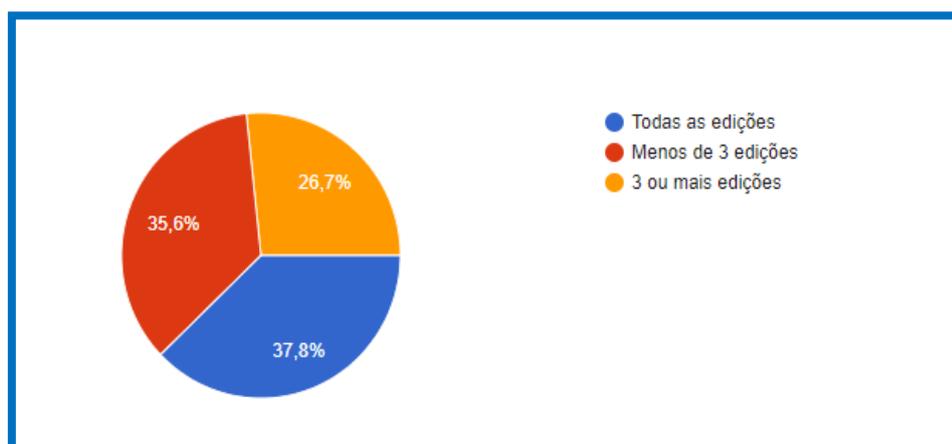
Fonte: *GoogleDocs*

No Gráfico 7, o maior número de participantes atuou no magistério na educação básica por um período de 1 a 5 anos (37%), seguido de um percentual de participantes que atuou de 20 a 30 anos (17,4%), e o mesmo número para aqueles que nunca atuaram (17,4%)

Comparando esses dados com os dados da questão três – Tempo de atuação na rede –, conclui-se que a maior parte dos participantes que atua na rede há mais de 20 anos tem pouco tempo de magistério na educação básica, em relação aos anos de atuação na rede. Pode-se inferir que o grupo com percentual significativo que não exerceu magistério na educação básica, envolveu-se com funções administrativas nas unidades escolares e nas sub-regiões.

O Gráfico 8 representa a participação dos gestores nas edições do PAER.

**Gráfico 8 - Questão 6: Edições do PAER de que já participou na função atual**



Fonte: *GoogleDocs*

O Gráfico 8 revela que a maior parte dos participantes atuou nas nove edições do PAER (2012 a 2020), e isto pode garantir uma percepção pautada pela vivência e experiência dos gestores das sub-regiões, cujas respostas abertas podem traduzir essa realidade experienciada.

A primeira parte do questionário, portanto, nos apresenta como perfil dos participantes, um predomínio de mulheres, com idade entre 45 e 50 anos, com mais de 20 anos de atividade na rede e com cinco anos ou menos, na função atual, mas com vivência nas nove edições do PAER.

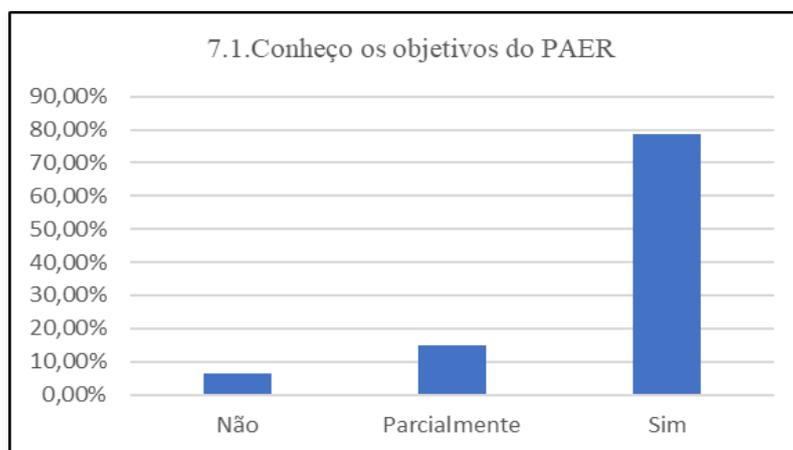
### 6.1.2 Questões objetivas

Os gráficos seguintes correspondem à questão sete que foi construída com 12 itens afirmativos, que o respondente deveria assinalar com: sim; não; parcialmente ou não tenho como responder. As afirmações podem ser organizadas em quatro eixos: i) Objetivos do PAER, com seis itens; ii) Análise dos dados, três itens; iii) Impactos do PAER, dois itens iv) Uso do PAER, um item.

### 6.1.2.1 Objetivos do PAER

No primeiro eixo, sobre o conhecimento dos objetivos do PAER, temos as afirmações: a) Conheço os objetivos do PAER; b) Os diretores conhecem os objetivos do PAER; c) Os professores conhecem os objetivos do PAER; d) Os pais conhecem os objetivos do PAER; e) Os alunos conhecem os objetivos do PAER; e f) Considero que o PAER alcançou seus objetivos.

**Gráfico 9 - Conheço os objetivos do PAER**

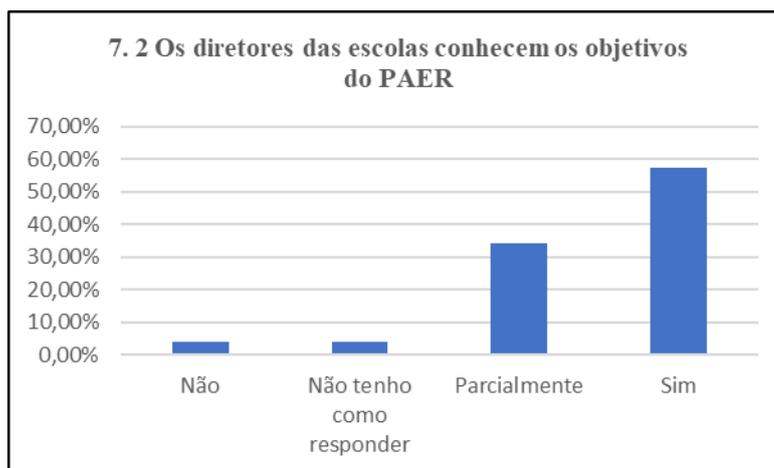


Fonte: Elaboração da autora

Pode-se verificar que dos 47 respondentes, 37 (78%) declaram conhecer os objetivos do PAER; três declaram não conhecer os objetivos do PAER (6%) e sete declaram conhecer parcialmente os objetivos do PAER (14%).

A questão seguinte aponta para os diretores das unidades escolares que são liderados pelos respondentes do questionário. As respostas indicam a percepção que eles têm sobre o conhecimento de seus liderados a respeito dos objetivos do PAER.

**Gráfico 10 - Os diretores conhecem os objetivos do PAER**



Fonte: Elaboração da autora

Observa-se que dos 47 gestores respondentes, 27 (57,45%) consideram que sim, os diretores das unidades escolares conhecem os objetivos do PAER; 16 (34,04%) consideram que conhecem parcialmente; dois (4,26%) consideram que não têm como responder e dois (4,26%) consideram que os diretores não conhecem os objetivos. Nota-se, portanto, que o percentual de diretores que conhecem os objetivos do PAER, na percepção dos gestores, é de cerca de 60%, o que já sinaliza dado importante a ser considerado. Um programa da rede cujos objetivos não são do conhecimento de todos os que vão implementá-lo nas unidades escolares pode ter, nesse fator, um impeditivo para seu êxito. Ao discutir sobre a relação da avaliação e a qualidade da educação, Cappelletti (2015) apresenta o papel do gestor escolar na condução de todo o processo, gerindo as tensões das pressões políticas e conflitos, conduzindo os processos de reflexão, discussão; ele “pode e deve criar um espaço de reflexão no horizonte da política, como requisito legítimo da prática pedagógica envolvendo os profissionais da escola, alunos e pais” (p.102). Na visão da autora, é importante que eles sejam considerados profissionais pedagógicos e não apenas executivos. Portanto, a condução do processo de avaliação externa e todos os desdobramentos são de sua competência no contexto da unidade escolar.

O Gráfico 11, traz a percepção dos respondentes sobre o conhecimento dos professores a respeito dos objetivos do PAER.

### Gráfico 11 - Os professores conhecem os objetivos do PAER



Fonte: Elaboração da autora

Dos 47 gestores respondentes, 26 (55,32%) consideram que os professores conhecem os objetivos do PAER, 18 (38,30%) consideram que conhecem parcialmente, dois (4,26%) consideram que não conhecem, e um (2,13%) considera que não tem como responder. Portanto, na percepção de cerca de 55% dos respondentes, os professores conhecem os objetivos do PAER. Os demais 45% estão entre parcialmente ou não conhecem.

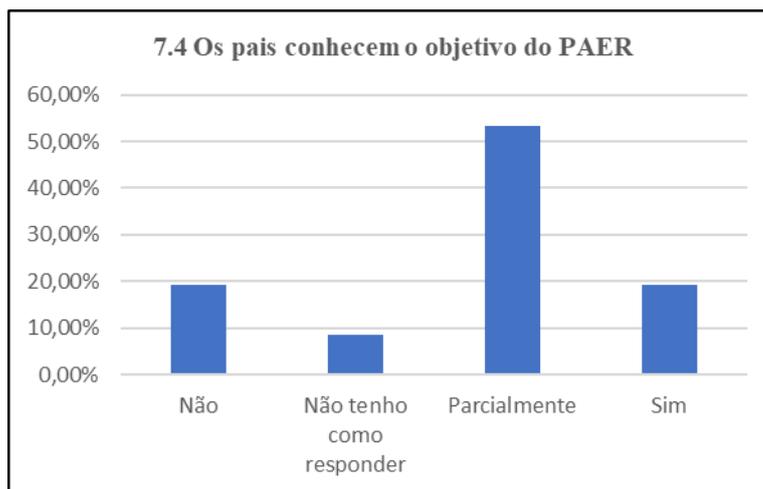
Esta informação é significativa, pois, ainda que os respondentes sejam gestores que não atuam diretamente nas unidades escolares, entende-se que têm a responsabilidade sobre um conjunto de escolas e, conseqüentemente, a função de garantir a identidade da rede com os programas e projetos que são nela desenvolvidos. Desse modo, também compete a eles a disseminação da informação e a construção de uma cultura de avaliação na rede.

Portanto, este índice de professores que conhecem parcialmente os objetivos do PAER, na visão dos respondentes dos questionários, comparado com os dados levantados nas respostas das entrevistas, analisadas a seguir, corrobora com a ideia da necessidade da criação de uma cultura avaliativa. Os professores deveriam ser, em primeira instância, os interessados em conhecer os objetivos de uma avaliação externa na qual o desempenho dos alunos é mensurado, já que a prática docente é associada aos resultados obtidos pelos alunos. Indiretamente, seu trabalho é avaliado e, ainda que, de forma não declarada, são responsabilizados. Este é um aspecto que pode ser indicado como uma fragilidade a ser superada pelo programa pois, ao pretender-se por meio dele a melhoria da qualidade educacional, como declarado no objetivo geral do documento do PAER, o envolvimento dos docentes torna-se necessário pois, como afirma Dourado (2020) os sujeitos envolvidos no processo educativo e suas condições concretas implicam em uma educação de qualidade.

Numa perspectiva democrática e emancipatória, o conhecimento dos objetivos por todos os envolvidos é fundamental para que o processo avaliativo persiga sua finalidade.

As respostas à pergunta 7.4, sobre o conhecimento dos pais a respeito dos objetivos do PAER, estão representadas no Gráfico 12.

**Gráfico 12 - Os pais conhecem os objetivos do PAER**

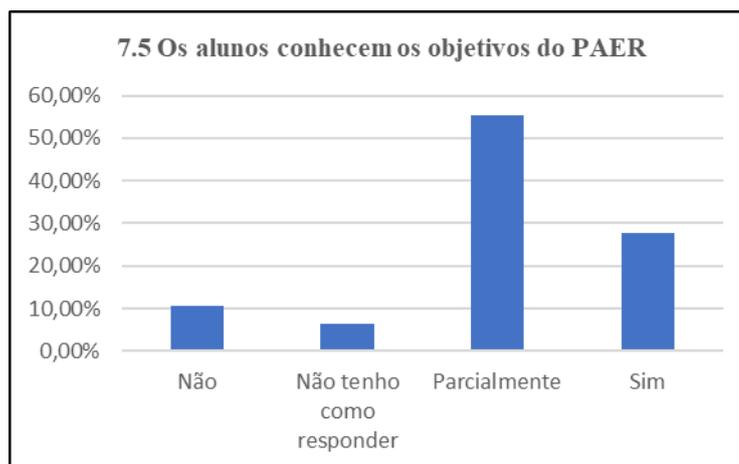


Fonte: Elaboração da autora

Nesta resposta, nove (19,15%) gestores consideram que os pais conhecem os objetivos do PAER; nove (19,15%) consideram que não conhecem; 25 (53,19%) consideram que parcialmente, e quatro (9,52%) não têm como responder. Este, portanto, é um item que apresenta o mesmo número de respostas **sim** e resposta **não**, e com mais de 50% de respostas **parcialmente**. Pode-se considerar que, por esses gestores não atuarem nas unidades escolares em contato com os pais, seja possível que essa percepção não se mostre tão clara e gere esse tipo de divergência. Porém, é um dado a ser considerado, no conjunto das respostas do eixo, como um indicador significativo.

As respostas à afirmação de que os alunos conhecem os objetivos do PAER estão representadas no Gráfico 13.

**Gráfico 13 - Os alunos conhecem os objetivos do PAER**

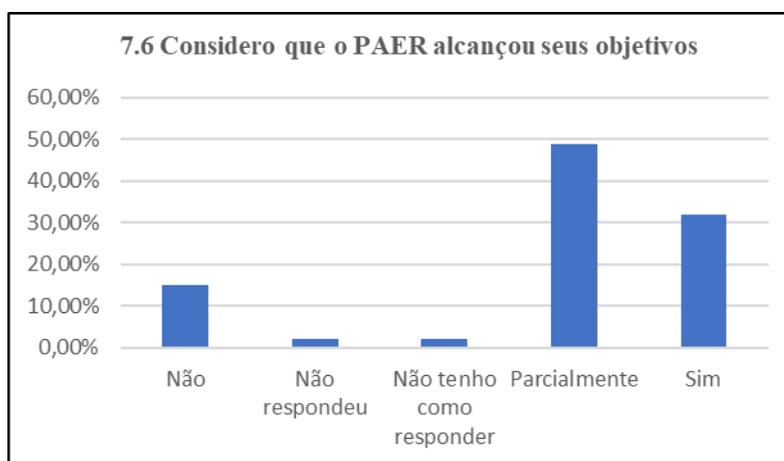


Fonte: Elaboração da autora

Na percepção de 13 (27,66%) gestores, os alunos conhecem os objetivos do PAER, cinco (10,64%) consideram que não conhecem, 26 (55,32%) consideram que os alunos conhecem parcialmente, e três (6,38%) não têm como responder. Verifica-se que, também nesta resposta, o índice de **parcialmente** ultrapassa os cinquenta por cento. É possível que esta percepção também esteja associada ao fato de os respondentes não atuarem como gestores nas unidades escolares. Se, de fato, os alunos não conhecerem os objetivos do programa de avaliação externa, este fator pode contribuir para a não valorização, como descrito por alguns participantes da entrevista e a consequente constatação da necessidade da criação de uma cultura avaliativa na rede, como será demonstrado na síntese ao final do capítulo.

A última resposta do eixo sobre os objetivos do PAER está representada no Gráfico 14.

**Gráfico 14 - Considero que o PAER alcançou seus objetivos**



Fonte: Elaboração da autora

Nesta resposta, dos 47 participantes, 15 (31,91%) consideram que o PAER alcançou seus objetivos; 7 (14,89%) consideram que não, 23 (48,94%) consideram que parcialmente os objetivos foram alcançados. Um participante considerou que não tinha como responder, e um não respondeu.

Portanto, neste eixo sobre os objetivos do PAER, as respostas dos gestores indicam que, nessa esfera administrativa, há conhecimento dos objetivos do PAER, e que eles percebem que os diretores das unidades escolares e os professores conhecem ou conhecem parcialmente os objetivos. De igual modo, na percepção desses gestores, alunos e pais, conhecem parcialmente.

O conhecimento dos objetivos de um programa de avaliação em larga escala, tanto pela comunidade interna de um sistema ou rede, como pela comunidade externa, pode ser um indicativo do grau de envolvimento de todos os interessados, e esse pode ser um fator que desvela qual a lógica ou perspectiva adotada pelo referido programa. Guba e Lincoln (2011) identificam três diferentes grupos de interesse sempre existentes em um processo de avaliação. Para os autores, **interessados** (*stakeholders*) são pessoas ou grupos que, de alguma forma, “são colocados em risco pela avaliação, isto é, pessoas ou grupos que têm algum interesse em jogo.” (GUBA; LINCOLN, 2011, p. 223). Consideram os **representantes** – as pessoas envolvidas na produção, utilização e implementação do objeto de avaliação; os **beneficiários** – as pessoas que ganham com o fato de o objeto de avaliação estar sendo utilizado, que podem ser beneficiários diretos (as pessoas para as quais a avaliação foi criada), e os beneficiários indiretos (pessoas cuja relação com os beneficiários diretos é intermediada ou influenciada positivamente); e, as **vítimas** – as pessoas que são afetadas negativamente pelo uso do objeto de avaliação (GUBA; LINCOLN, 2011).

Para os autores, nas avaliações das três primeiras gerações, cuja tendência ao gerencialismo está fortemente presente, as informações ficam restritas aos **representantes** que são as pessoas que desenvolvem o programa, os que tomam as decisões em nível local, regional ou nacional, e os envolvidos na implementação. Já, numa avaliação de quarta geração, numa perspectiva responsiva e emergente, todos os grupos de envolvidos – neste caso, a comunidade interna e externa: gestores, professores, alunos e pais – que seriam os **beneficiários**, conhecem os objetivos da avaliação, e todos são solicitados a contribuir com suas reivindicações, preocupações e questões. As informações circulam em todos os níveis.

Freitas *et al.* (2017) indicam a necessidade de as redes que possuem sistemas de avaliação em larga escala desenvolverem um processo de comunicação com a comunidade

para que todos tenham acesso às informações e que estas sejam constantemente objeto de análise e de reflexão. Defendem, os autores, a participação dos pais num modelo de conselho gestor de avaliação da rede, entre outras sugestões, para que esse fluxo de informações seja ampliado.

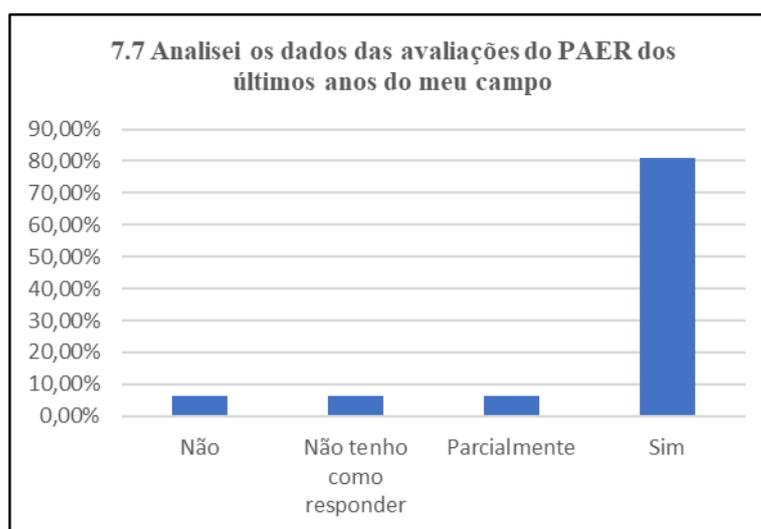
Um dos objetivos específicos do PAER apresentado no capítulo quatro é “divulgar o programa entre os diferentes membros da comunidade educacional, usando diferentes estratégias de comunicação” (DO, 2017, p. 5). Sendo assim, neste primeiro eixo sobre os objetivos do PAER, há que se considerar que a percepção dos gestores indica que há um espaço a ser percorrido para que alunos e pais, como beneficiários do programa, conheçam o programa, seus objetivos, como previsto originalmente.

### 6.1.2.2 Análise dos dados do PAER

O segundo eixo das questões objetivas trata-se da análise dos dados do PAER na percepção dos participantes. Três afirmações foram apresentadas: a) Analisei os dados do PAER nos últimos anos em minha região administrativa. b) Os dados do PAER são analisados pelas equipes administrativas das unidades escolares. c) Os professores analisam os dados do PAER.

O primeiro, Gráfico 15, indica a análise realizada pelo próprio respondente em sua região administrativa.

**Gráfico 15 - Analisei os dados das avaliações do PAER nos últimos anos**

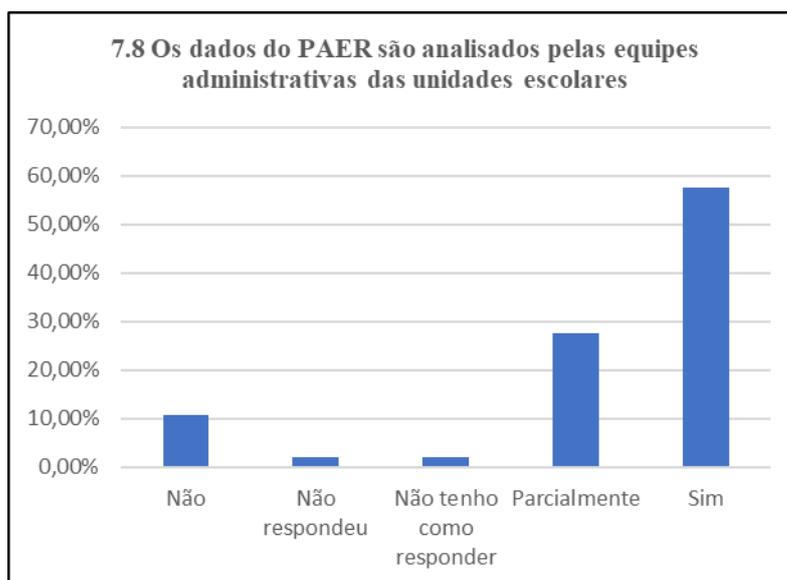


Fonte: Elaboração da autora

Dos 47 respondentes, 38 (80,85%) analisaram os dados produzidos pelas edições do PAER dos últimos anos. Três (6,38%) não analisaram; três (6,38%) analisaram parcialmente e três não tiveram como responder.

Sobre a análise dos dados pelas equipes administrativas das unidades escolares, a percepção dos gestores regionais é representada no Gráfico 16.

**Gráfico 16 - Os dados do PAER são analisados pelas equipes administrativas**

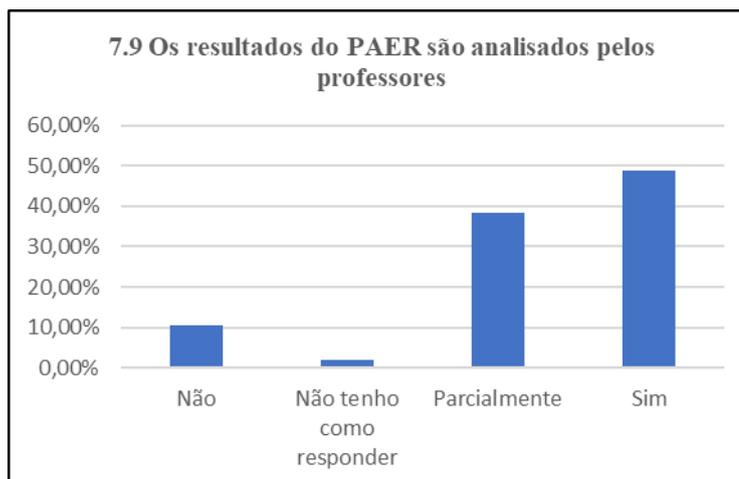


Fonte: Elaboração da autora

Na percepção dos respondentes, 27 (57,45%) consideram que as equipes analisam os dados produzidos pelas avaliações do PAER; 5 (10,64%) consideram que não são analisados; 13 (27,66%) consideram que são analisados parcialmente, um respondente considera que não tem como responder, e um participante não respondeu.

A afirmação seguinte é sobre a análise dos dados pelos professores das unidades escolares. O gráfico 17 representa a percepção dos gestores sobre esta afirmação.

**Gráfico 17 - Os resultados do PAER são analisados pelos professores**



Fonte: Elaboração da autora

Na percepção dos gestores participantes, 23 (48,94%) consideram que os professores analisam os resultados do PAER; cinco (10,64%) consideram que não analisam; 18 (38,30%) consideram que analisam parcialmente, e um participante (2,13%) não teve como responder.

Ao afirmarem sobre sua própria análise dos dados produzidos pelo PAER, os gestores apresentam um alto índice de 80,85%. Ao demonstrarem sua percepção sobre a análise dos dados pelas equipes administrativas e professores, o percentual diminuiu, mas ainda se mantém próximo aos cinquenta por cento. Contudo, mais uma vez, o índice está abaixo de 50% em relação aos professores e análise dos dados gerados pela avaliação externa.

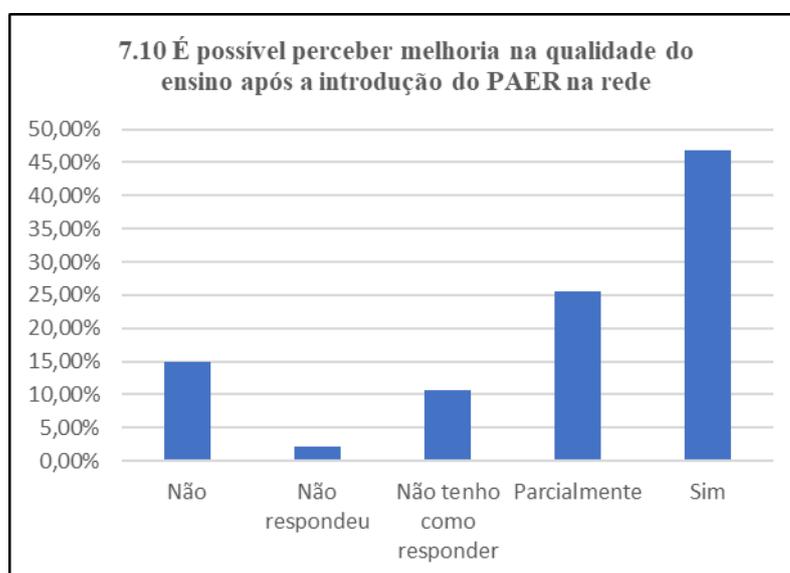
A análise dos dados, por si só, não tem relevância, se não for seguida por tomadas de decisões. Pois a análise não se limita a uma constatação, ela envolve um olhar de contexto, de processo, de causas e efeitos que leva a uma ação. Da mesma forma, os dados gerados pelo PAER não se limitam ao desempenho dos alunos nas provas. Os questionários contextuais oferecem uma gama de dados que podem ser amplamente analisados e correlacionados estatisticamente na própria plataforma do programa.

Para Freitas *et al.* (2017), a análise dos dados da avaliação em larga escala pela comunidade escolar, ancorada na avaliação periódica do projeto político-pedagógico da escola, contribuiu para a formulação e reformulação desse projeto com base nesses dados. Desse modo, revela-se a reflexão que a escola realiza sobre os dados, gerando demandas e compromissos. Portanto, o envolvimento dos professores nesse processo de análise e discussão é fundamental para que as mudanças sejam implementadas, a partir dos resultados obtidos. Sem a participação docente, o sentido da avaliação externa como propulsora de melhoria da qualidade se perde.

### 6.1.2.3 Impactos causados pelo PAER

As duas afirmações seguintes do questionário estão relacionadas ao eixo impactos causados pelo PAER: a) É possível perceber melhoria da qualidade do ensino após a introdução do PAER na rede. b) Houve mudanças significativas na qualidade do ensino após a introdução do PAER na rede. Na percepção dos gestores que responderam ao questionário, a primeira afirmação teve o resultado descrito no Gráfico 18.

**Gráfico 18 - Melhoria da qualidade do ensino após a introdução do PAER na rede**

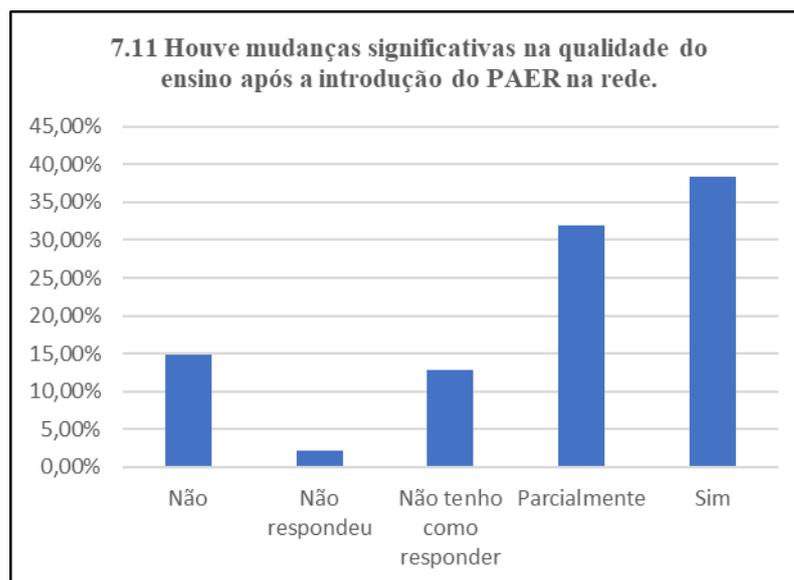


Fonte: Elaboração da autora

Dos 47 respondentes, 22 (46,81%) consideram que é possível perceber melhoria na qualidade do ensino após a introdução do PAER na rede; sete (14,89%) consideram que não; 12 (25,53%) consideram que parcialmente; cinco (10,64%) consideram que não têm como responder, e um não respondeu.

Na afirmação seguinte, a percepção dos gestores é descrita no Gráfico 19.

**Gráfico 19 - Houve mudanças significativas na qualidade do ensino**



Fonte: Elaboração da autora

Na percepção dos gestores das sub-regiões, 18 (38,30%) consideram que houve mudanças significativas na qualidade do ensino após a introdução do PAER na rede; sete (14,89%) consideram que não houve mudanças significativas; 15 (31,91%) consideram que houve parcialmente; seis (12,77%) afirmam que não têm como responder; e um não respondeu.

Em ambas as respostas o percentual que considera que é possível perceber melhora ou se houve melhora significativa da qualidade do ensino após a introdução do PAER é inferior a 50%; com o parcialmente, entre 25 e 30%. O principal objetivo do programa versa sobre a melhoria da qualidade do ensino. Infere-se, portanto, que para esse grupo de gestores, o objetivo está sendo parcialmente alcançado.

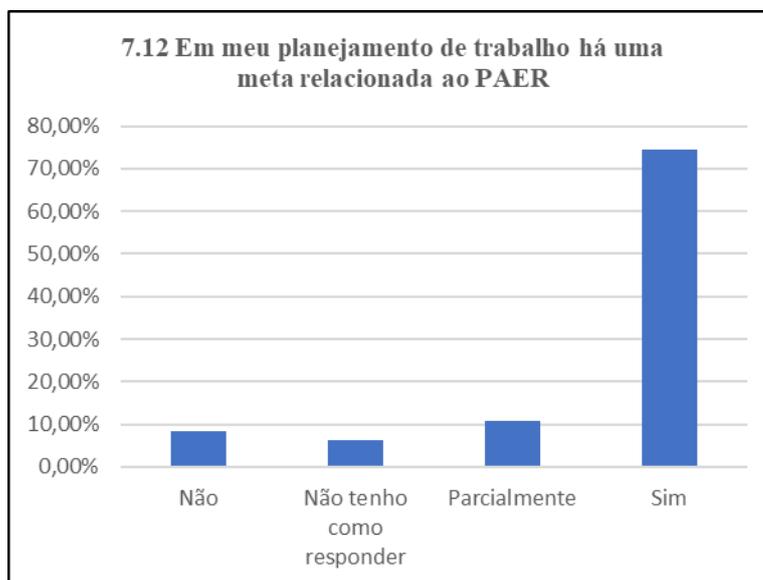
Este é um dado importante a ser considerado, pois está diretamente relacionado ao objetivo da pesquisa em levantar ganhos e dificuldades do programa em questão. Se o objetivo anunciado por todos os sistemas de avaliação externa e em larga escala – quer de sistemas públicos de ensino, quer de redes privadas – é a melhoria da qualidade do ensino; se após um decurso de oito anos de aplicação do PAER, o índice de melhora percebida está abaixo de 50%, pode-se concluir que ainda há o desafio de encontrar caminhos para que o objetivo seja alcançado. Sem anunciar a qual qualidade o objetivo persegue, não há um parâmetro e esta resposta pode significar várias coisas. Nas questões abertas dubiedade ficará aparente.

Em um estudo que considerou a opinião de professores e gestores sobre as políticas de avaliação externa e a efetivação de uma escola de qualidade para todos, Sohn e Eyng (2012)

obtiveram percentuais semelhantes nas respostas das entrevistas, quando perguntaram sobre essa relação, entre os resultados das avaliações e a melhoria da qualidade do processo de ensino e aprendizagem. Apesar de, em muitos casos, ter havido aumento dos indicadores – no caso, o Ideb – cerca de 50% dos professores não consideram que haja, de fato, melhoria ou mudanças positivas. Esta é uma temática que pode ser aprofundada com pesquisas que ampliem o espectro de investigação sobre a percepção de docentes e gestores com relação aos impactos que uma avaliação externa pode causar no processo educativo. Na análise das entrevistas, de que trataremos a seguir, pode-se observar que a percepção de melhoria da qualidade como um ganho do PAER foi indicada apenas por um participante, que é o idealizador do programa. Portanto, é um dado significativo a ser considerado e será melhor analisado na próxima seção.

A última afirmação está relacionada à inclusão do PAER como meta do planejamento do gestor. No Gráfico 20, as respostas são representadas.

**Gráfico 20 - Em meu planejamento há uma meta sobre o PAER**



Fonte: Elaboração da autora

Entre os gestores respondentes do questionário, 35 (74,47%) estabeleceram em seu planejamento uma meta relacionada ao PAER; quatro (8,51%) não estabeleceram meta em seu planejamento; cinco (10,64%), parcialmente; e três (6,38%) não têm como responder. Pode-se inferir por este dado que há um interesse pela continuidade do PAER, que os dados gerados são utilizados como instrumento de gestão. Essa é uma das finalidades da avaliação em larga escala. Cappelletti (2015) considera que, se o gestor conceber a avaliação numa

perspectiva globalizante, “por certo terá à sua disposição informações valiosas sobre o desenvolvimento do trabalho escolar. Terá pistas, alternativas de ação para promover o desenvolvimento curricular em sua escola, assumindo sua função principal de gerir e criar cultura.” (CAPPELLETTI, 2015, p. 102). Sabe-se que os gestores que responderam aos questionários não estão diretamente nas unidades escolares. São gestores educacionais, ou de redes, e não gestores escolares (SOUZA; MARTINS, 2012). Contudo, entendemos que a responsabilidade em gerir e criar cultura também está nesse nível de gestão.

Sousa e Martins (2012), num levantamento da produção científica sobre avaliação educacional e gestão de sistemas e escolas no qual a avaliação em larga escala obteve centralidade na produção analisada, concluem:

Depreende-se que, gradualmente, resultados de desempenho de alunos em testes de larga escala vêm sendo incorporados como indicador relevante de sucesso (ou não) de políticas educacionais e de práticas escolares, o que tende a induzir administradores a colocarem em suas agendas o compromisso com a melhoria do rendimento escolar dos alunos, além de fluxo escolar, tradicionalmente considerado como referência de qualidade. (SOUZA; MARTINS, 2012, p. 23).

Essa inserção do tema como integrante da agenda do gestor, seja do sistema ou rede, seja da unidade escolar, tem sido um movimento que vem se consolidando nos sistemas públicos de ensino e, pelo dado levantado no questionário, também na rede privada aqui pesquisada.

### 6.1.3 Questões abertas

O questionário apresentou quatro questões abertas cujas respostas foram analisadas e categorizadas. Três perguntas são reproduzidas também nas entrevistas. A quarta pergunta foi introduzida a partir do contexto do isolamento social que se estendeu mais do que o esperado em relação ao período no qual as entrevistas foram realizadas, sendo possível incluir o tema no questionário *on-line*. As perguntas foram:

1. Em sua opinião, quais os ganhos (aspectos desenvolvidos) do PAER?
2. Quais os desafios ou dificuldades (aspectos a desenvolver) do PAER?
3. Quais as implicações do PAER no processo educativo?
4. O cenário de 2020 com o ensino remoto e a aplicação do PAER totalmente *on-line* trouxe alguma implicação para a qualidade do processo?

As questões estão alinhadas ao principal objetivo da pesquisa e relacionam-se aos resultados obtidos pelo programa. As categorias de análise que emergiram das respostas foram agrupadas pelo grau de incidência.

### 6.1.3.1 Os ganhos do PAER

Dos 47 respondentes do questionário, 41 responderam à primeira questão aberta. Destes, um respondeu “não sei”. Portanto, 40 respondentes indicaram algum ganho advindo do PAER. Há, na percepção dos participantes, uma relação entre o programa de avaliação externa e um movimento de melhoria da qualidade educacional. Ora voltado à prática docente, ora voltado ao envolvimento dos alunos e à aprendizagem. Essas respostas podem ser agrupadas em quatro categorias de análise: i) Melhoria da qualidade do ensino; ii) Melhoria da qualidade da aprendizagem; iii) Indica pontos fortes e pontos fracos; iv) Mudanças na avaliação interna.

As 40 respostas foram categorizadas, conforme dispostas na Tabela 6, que apresenta o número de respostas para cada categoria e o percentual correspondente:

**Tabela 6 - Pergunta 8: Categorias de análise e incidência**

Categorias	Número	%
Melhoria da qualidade do ensino	18	45%
Melhoria da qualidade da aprendizagem	8	20%
Indica pontos fortes e fracos	7	17,5
Mudanças na avaliação interna	6	15%
Outras respostas	1	2,5

Elaboração: autora.

Sobre a primeira categoria de análise, destaca-se que a percepção de **melhoria na qualidade do ensino** em índice de 45% para os respondentes do questionário é significativamente maior que a percepção dos participantes da entrevista, na qual a proporção foi de um em cinco (20%), que considerou a melhoria da qualidade do ensino como ganho atribuído ao PAER. Ainda que os participantes, de ambos os instrumentos de coleta de dados, sejam gestores, estão em níveis diferentes da estrutura organizacional da rede, o que pode justificar que, perguntas iguais, possam obter respostas com enfoques e percepções diferentes, o que enriquece e amplia as possibilidades de análise dos dados. Em uma pesquisa de abordagem qualitativa, até mesmo o dissenso nas respostas obtidas é um dado importante e enriquece a interpretação e análise do investigador.

A percepção de que há **melhoria da aprendizagem** coloca no âmbito dos alunos a busca por essa melhoria. Há, por parte destes, “a valorização pelo processo avaliativo”, como indicou um dos respondentes. Outro menciona que a avaliação externa “desenvolve nos alunos o desejo de alcançar níveis maiores de aprendizagem”. Ou, ainda, “melhora da qualidade do estudo do aluno”. Esta é uma percepção que, semelhantemente à categoria anterior, não encontrou correlação com as respostas da entrevista. Ambas estão relacionadas e serão analisadas com mais profundidade na próxima seção deste capítulo.

A ideia de que a avaliação externa **indica pontos fortes e pontos fracos**, apresentada por sete respondentes, não tornou clara em todas as respostas qual o objeto considerado que apresenta pontos fortes ou fracos: ora como “avanços ou retrocessos acadêmicos”; ora como “objetivos alcançados ou não alcançados”; ou “conhecer nossas potencialidades e consolidá-las e nossas fragilidades para dirimi-las”; como indicaram alguns respondentes. Contudo, é significativa a percepção de que esse é um dos ganhos de uma avaliação externa em larga escala. Ainda que esta seja circunscrita a uma avaliação do desempenho dos alunos por meio de provas de Português e Matemática para ensino fundamental, e das quatro áreas do conhecimento para o ensino médio, delimitadas por uma matriz de referência. De fato, uma avaliação educacional, seja ela da aprendizagem, institucional, ou externa, potencialmente tem a função de desvelar pontos fortes e pontos fracos. Resta-nos perguntar: esses pontos são fortes ou fracos em relação a quê? Este é um dos riscos do uso dos resultados das avaliações de desempenho como único indicador de qualidade educacional. Quando uso o rendimento do aluno em duas áreas do conhecimento como indicativo do construto qualidade - que deveria ser considerado por um conjunto de indicadores que envolvem elementos intra e extraescolares - tenho clareza de que esses resultados apontam pontos fortes e fracos apenas do desempenho cognitivo dos alunos? Ou considero esses resultados como indicadores de pontos fortes e fracos do processo educativo como um todo? Esta é uma das preocupações e críticas ao reducionismo que as avaliações externas têm causado como condicionante da qualidade de ensino (SOUZA; OLIVEIRA, 2010).

A relação entre a avaliação externa e a avaliação interna é indicada por seis respondentes como um ganho produzido pelo PAER. A **mudança na avaliação interna** é percebida pelos gestores como consequência da iniciativa dos professores que “aprenderam a construir questões mais elaboradas utilizando os descritores”; ou “uma melhoria da qualidade das avaliações”, que permite uma “análise comparativa com os resultados da própria unidade”. Há, portanto, uma busca em estabelecer uma correlação entre a avaliação da

aprendizagem e a avaliação externa em larga escala. Este é um dado que é reconsiderado na análise das entrevistas. Para Souza (2013), essa é uma necessidade que precisa ser enfatizada e compreendida pelos gestores e profissionais da educação. A articulação entre os dois níveis de avaliação também é postulada por Freitas *et al.* (2017), que consideram que a avaliação externa adquire uma relevância e significado quando alcançam essa dimensão mais construtiva, ao ser articulada com a avaliação em sala de aula. O autor postula a existência de três níveis integrados de avaliação da qualidade: avaliação em larga escala de redes de ensino; avaliação institucional da escola, e a avaliação da aprendizagem em sala de aula. Os três níveis necessitam interagir entre si para retratar a realidade educacional.

Se, de fato estas mudanças na avaliação em sala de aula, percebida pelos gestores da rede pesquisada, ocorrem como uma articulação entre ambas as avaliações, aprimorando os saberes docentes e garantindo a formulação de avaliações da aprendizagem de natureza formativa, considera-se que esta mudança atende as características de uma avaliação emancipatória e constitui-se um ganho. Se, ao contrário, a avaliação de sala de aula passou a ser formulada no modelo da avaliação externa como um treino para obtenção de melhores resultados, ou como adverte Freitas (2017) a avaliação externa é feita como um instrumento de avaliação da sala de aula, reduz-se apenas a uma mudança que objetiva a manutenção de uma perspectiva de avaliação convencional.

### 6.1.3.2 Desafios ou dificuldades do PAER

À nona pergunta aberta do questionário, 40 participantes responderam, indicando os desafios ou pontos a desenvolver no PAER. As categorias que emergiram das respostas foram: i) Alunos não entendem a importância da avaliação externa; ii) Uso inadequado dos resultados; iii) Professores e gestores não entendem a importância da avaliação; iv) Não há um programa pós-Paer, que formule políticas e ações para a rede; v) Necessidade de qualificação docente; vi) Acesso às questões da prova pelos professores. A Tabela 7 apresenta as categorias de análise e a incidência de respostas para cada categoria.

**Tabela 7 - Resposta 9: Categorias de análise e Incidência**

<b>Categorias</b>	<b>Número</b>	<b>%</b>
Uso inadequado dos resultados	10	25%
Alunos não entendem a importância da avaliação externa	10	25%
Professores e gestores não entendem a importância da avaliação externa	7	17,5%

Não há um programa pós-PAER, que formule políticas e ações para a rede	6	15%
Necessidade de qualificação docente	5	12,5%
Outras respostas	2	5%

Elaboração: autora.

A primeira categoria, **uso inadequado dos resultados**, deriva da resposta de dez participantes e se desdobra em alguns aspectos já citados anteriormente, como dificuldades advindas da avaliação externa. Observações como: “Os resultados não serem utilizados como ‘*rankeador*’ das unidades, assim como os dados precisam ser analisados num período e não o resultado anual de forma estanque”; “fazer com que os professores não apenas ‘treinem’[...] trabalhar resultados com efetividade de maneira sistemática e global”; “nem sempre é feita uma análise dos resultados”. Esses pontos apresentados encontram eco nas principais críticas e ponderações à avaliação em larga escala, presentes na literatura nacional e internacional, discutidas por Bauer; Alavarse e Oliveira (2015). Os autores apontaram dez argumentos contrários às avaliações padronizadas em larga escala e, entre eles, a questão do *ranking*, incentivando a competição entre as escolas; a substituição do aprendizado pela melhoria do resultado em provas, o que gera o treino para os testes e o afunilamento do currículo; entre outras críticas. Concluem, os autores, que as principais críticas à avaliação em larga escala estão relacionadas ao uso inadequado de seus resultados.

A segunda categoria de análise indicada aponta que os **alunos não entendem a importância da avaliação externa**. Esta categoria pode ser associada à terceira categoria mais indicada pelos respondentes: **professores e gestores não entendem a importância da avaliação externa**. Ambas estão relacionadas. Os alunos poderão ser influenciados pelos professores e gestores a reconhecerem o valor de um processo avaliativo na medida em que estão envolvidos nesse processo, têm acesso às informações e resultados, podem apresentar suas impressões sobre a avaliação, opinar e sugerir mudanças. De igual modo, os professores e gestores escolares que entendem a importância do processo de avaliação externa tendem a influenciar os alunos a reconhecerem a importância que lhe é atribuída. A valorização também depende do desenvolvimento de uma cultura de avaliação pautada nos princípios de uma avaliação emancipatória, na qual é garantida a todos os participantes da comunidade escolar expressar seus pontos de vista, discutir, dialogar, propor alternativas, tendo o gestor como condutor desse processo, que atribui sentido à avaliação (CAPPELLETTI, 2015).

Dias Sobrinho (2004) considera que a avaliação, no paradigma que corresponde a uma epistemologia subjetivista, é uma atribuição de valor e produção de sentidos. Ela não exclui o

controle e a dimensão técnica do processo, mas estes complementam-se, garantindo significado aos dados produzidos e ao produto. Para o autor:

A avaliação, como produção de sentidos, reflexão de valores e significados, tem um grande potencial educativo. Sem deixar de ser objetiva e utilizar instrumentos técnicos, e não se satisfazendo com a mera verificação e checagem de produtos e sua conformidade com uma norma, a avaliação como produção de sentidos alimenta debates, interroga-se sobre significados, as causalidades e os processos, trabalha com a pluralidade e a diversidade, abre possibilidades de emancipação, construção, dinamização. (DIAS SOBRINHO, 2004, p. 724).

Compreender a importância de um processo de avaliação externa e em larga escala torna-se possível, quando o processo adquire sentido e significado para todos os envolvidos que, ainda que não sejam partícipes de todas as suas fases, têm a oportunidade de discutir sobre o produto, os resultados, propor caminhos e alternativas, contribuir para as mudanças e transformações possíveis e necessárias. Sem a participação dos envolvidos, a avaliação é uma intervenção fechada, isolada, que se esgota em si mesma, cujos dados levantados atendem apenas aos interesses dos avaliadores. Dias Sobrinho (2004) a denomina avaliação como controle, fundada numa epistemologia objetivista. Neste modelo, tudo é decidido externamente, sem a participação efetiva dos educadores. Pode produzir importantes efeitos. Porém, como afirma o autor, “tende a produzir mais do mesmo.” (DIAS SOBRINHO, 2004, p. 724).

Em um desdobramento do uso dos resultados, a categoria que indica que **não há um programa pós-PAER, que formule políticas e oriente ações para a rede**, está associada a essa percepção de que o programa está se esgotando com a divulgação dos resultados. O que virá depois da avaliação? O que fazer com os dados levantados? Alguns dos respondentes questionam o “direcionamento de estudos de ações por parte do órgão responsável pela implantação do sistema de avaliação”; ou a criação de um “projeto específico visando melhorar os resultados, através do órgão central, não deixando a responsabilidade exclusivamente com as escolas”. Houve, ainda, a observação de que “deveriam apresentar caminhos a serem trilhados em prol de se ter melhores resultados... enfim, não basta apenas aferir o conhecimento, que orientações são dadas após essa etapa?” E, finalmente, “não existe uma proposta, uma política educacional de gestão pelo órgão responsável pelo PAER, após os resultados”.

Essa percepção de que há algo mais a ser feito além da verificação do rendimento dos alunos pode derivar de alguns sentidos: o primeiro aponta a expectativa de que as diretrizes de como proceder após a divulgação dos resultados sejam emanadas do órgão central que implantou o programa de avaliação externa, tirando da unidade escolar a responsabilidade de

reflexão, discussão, tomada de decisão ou, até mesmo, a não responsabilização pelos resultados. Outro sentido pode ser inferido pela falta de autonomia que pode existir numa estrutura de rede, onde os níveis de gestão são tão diversos e hierarquizados, que a unidade escolar não desenvolve a dinamicidade de tratar com a questão da avaliação externa como um componente a ser agregado ao projeto pedagógico, ao planejamento das atividades educacionais, de natureza formativa e como instrumento de gestão. Desse modo, há uma expectativa de que um programa de avaliação externa gere outro programa para trabalhar os resultados dessa avaliação. Essa percepção demonstrada por seis respondentes pode ser compreendida como uma indicação que função da avaliação se esgota na apresentação dos resultados; não vá além de oferecer dados.

A dificuldade ou desafio apresentado por cinco respondentes sobre a **necessidade de qualificação docente** se diversifica sobre o tipo de formação necessária. Um respondente identifica a “formação para análise dos dados produzidos pelo PAER”; outro considera a necessidade de formação inicial do professor para lidar com a avaliação, pois argumenta sobre “os desafios do recrutamento, seleção e formação dos professores para alcançar os objetivos da avaliação externa”; outro indica a “formação docente para compreenderem a importância do programa como ferramenta para o desenvolvimento de um ensino de qualidade.”.

Souza (2007), ao investigar questões relacionadas à formação inicial do docente em avaliação educacional, constatou como precária essa formação. Para a autora, isso se dá em decorrência da segmentação entre ensino e avaliação como objetos distintos, da dicotomia teoria e prática, da ausência de reflexões sobre temas atuais que possuem implicações diretas com a avaliação. Ainda que a autora tenha pesquisado a avaliação educacional de forma ampla, o aporte teórico é o mesmo, sendo que este tem sido indicado como fator de precariedade na formação docente. Como conclui Cappelletti (2015), associando a fragilidade na formação docente aos conflitos entre a avaliação e qualidade educacional:

A falta de formação e experiência no campo da avaliação impede a oferta aos estudantes, futuros professores, de vivências em processos avaliativos adequados e de acesso a modelos operacionais que possam ser ilustrativos de uma postura crítica em avaliação. (CAPPELLETTI, 2015, p. 104).

A formação continuada pode amenizar essa precariedade, contudo há o risco de essa formação resvalar na mesma dificuldade de associar o fundamento teórico à prática, levando à ausência de reflexão, e a avaliação terminar por ser utilizada apenas para consolidar práticas e fazer ajustes daquilo que já está posto, visando apenas à manutenção de um modelo. Isso não invalida a necessidade de formação de professores para compreenderem a importância da

avaliação externa, ou mesmo a avaliação educacional. Apenas reforça a importância de buscar a coerência entre o arcabouço teórico escolhido e a prática adotada no contexto educacional.

Outrossim, esta percepção dos cinco gestores respondentes pode ser reflexo de uma tendência à culpabilização docente, como sendo o responsável pelo desempenho dos alunos nas avaliações externas e conseqüentemente pela melhoria dos resultados da escola. Uma tendência à responsabilização que, por vezes, não se apresenta de forma declarada, mas faz-se presente no tecido das relações intraescolares.

### 6.1.3.3 Implicações do PAER no processo educativo

Na questão aberta número dez, sobre as implicações do PAER no processo educativo, houve 38 respostas. Um dado relevante é que nove respondentes (23%) não conseguiram identificar qualquer implicação do programa no processo educativo, e oito (21%) apresentaram respostas que não têm relação com a pergunta. Dentre as 21 respostas restantes, as categorias de análise que foram levantadas são: i) Mudança na prática docente; ii) Mudança de atitude dos alunos; iii) Desvelamento a realidade; iv) Mudanças nas disciplinas avaliadas pelo PAER.

A Tabela 8 apresenta as categorias e as incidências das respostas da questão dez.

**Tabela 8 - Resposta 10 – Categorias de análise e incidência**

<b>Categorias</b>	<b>Número</b>	<b>%</b>
Mudanças na prática docente	10	26%
Mudanças na atitude dos alunos	4	10,5%
Desvelamento da realidade	4	10,5%
Mudanças nas disciplinas avaliadas pelo PAER	2	5%
Não responderam ou conteúdo não relacionado à pergunta	18	47%

Elaboração: autora

A mesma pergunta foi apresentada na entrevista semiestruturada e, semelhantemente, dos cinco entrevistados, dois (40%) não conseguiram responder à questão, apresentando conteúdos que não se relacionavam à temática da pergunta: implicações do PAER no processo educativo. Pode-se inferir que a palavra “implicações” não tenha o mesmo significado ou sentido para todos os participantes. Mas é um dado a ser considerado diante da dificuldade de obtenção de respostas em mais de quarenta por cento dos participantes de ambos os instrumentos utilizados no levantamento de dados.

A categoria **mudanças na prática docente** foi a mais relevante da percepção de 26% dos respondentes, indicando que, para estes, o programa tem implicado no processo educativo, com ênfase na atuação do professor. Indicações como: “Muda as estratégias de ensino e a forma de avaliar”; “Uma maior qualidade na sala de aula, no formato da aula/ensino”; ou “Ações de mudanças por parte dos professores”; são algumas das explicações para justificar a possível implicação do PAER no processo educativo. Apesar da quase totalidade das respostas indicarem mudanças positivas, um respondente sinalizou uma implicação que considerou negativa: “Uma busca por ‘treinar’ os alunos para a realização da prova, através de simulados (efeito negativo, pois foge do objetivo de desenvolvimento das habilidades dos alunos)”. De fato, implicar não indica que a consequência ou o encadeamento será necessariamente o desejado. Pode gerar efeito ou acarretar mudanças indesejáveis. Na visão de um dos participantes, portanto, a implicação foi negativa, levando ao treino por meio de simulados.

Essa prática de treinos de alunos por meio de simulados, para a obtenção de melhores resultados, foi apontada em um estudo de Sousa e Arcas (2010), no qual analisaram as implicações das avaliações em larga escala – no caso o Saresp, no sistema público do estado de São Paulo – e seu potencial de controle e conformação do trabalho escolar. Evidências de pesquisas levantadas pelos autores demonstraram que a avaliação da aprendizagem toma como referência a avaliação em larga escala:

Entretanto, o que pode ser constatado pela pesquisa é que o Saresp, ao servir como referência para as práticas avaliativas empreendidas nas escolas, acabou por reforçar práticas tradicionais de avaliação da aprendizagem. Num contexto em que se preconizam práticas de avaliação da aprendizagem fundamentadas na concepção formativa, visando orientar e favorecer a aprendizagem dos alunos, o Saresp reforça a aplicação de provas testes, objetivando, na maioria dos casos, simular a aplicação da avaliação externa, preparando os alunos para este tipo de instrumento de avaliação... (SOUSA; ARCAS, 2010, p. 193).

Outro dado que os autores consideraram, por meio de depoimento de professores, foi a necessidade de os docentes elaborarem as avaliações da aprendizagem nos moldes das provas da avaliação externa, o que exigiu uma capacidade que os professores não tinham e, na ótica dos docentes, implicou num desenvolvimento profissional.

A mesma percepção é demonstrada por alguns respondentes do questionário que apontaram as mudanças na elaboração das avaliações da aprendizagem como uma das implicações do PAER no processo educativo, considerando que “muda [sic] as estratégias de ensino e a forma de avaliar”. De maneira geral, para os participantes, essas mudanças são consideradas de forma positiva, como um crescimento do professor. Das dez respostas que se

enquadram nesta categoria, apenas uma aponta para a mudança como “efeito negativo”, associada à elaboração de provas simuladas com as características das provas da avaliação externa, com a função de treinar. Porém, não é possível saber se as demais respostas que indicam mudanças no ato de avaliar percebidas como algo positivo, também não se refere a treinos e simulações da avaliação externa, valorizados pelos gestores como implicações favoráveis no processo educativo.

Além da mudança nos instrumentos de avaliação da aprendizagem, outras dimensões de mudanças indicadas nas respostas estão associadas à busca por mudanças nas práticas de ensino para que os alunos possam “vencer” os descritores nos quais não obtiveram proficiência na avaliação externa; ou “uma preocupação e maior envolvimento com o aprendizado”.

Já a ideia de que há uma implicação no processo educativo que envolve uma **mudança na atitude dos alunos**, relacionada, portanto, à aprendizagem, é indicada por quatro participantes que consideram que os alunos reconhecem “os pontos a serem melhorados no desempenho acadêmico de cada um” e que “estão mais atentos na realização das provas”. O fato de a avaliação externa desencadear essa atitude dos alunos pode estar associado a alguma pressão por resultados produzida por gestores escolares que, se assim for, pode não estar ocorrendo em relação às avaliações internas. No entanto, entre os que consideraram que não há implicação no processo educativo, as causas apresentadas estão relacionadas, principalmente, aos alunos: “Se existir um programa bem elaborado, apresentando o real objetivo, onde os alunos compreendam o motivo pelo qual existe o programa e sua importância, penso que poderemos ter melhora no processo acadêmico”; ou, “muitos alunos não se dedicam em realizar as questões com responsabilidade, não levam a sério este momento.”. Este é um tema sobre o qual não há consenso, de fato. Numa análise mais ampla, com as respostas a outras perguntas do questionário e comparando com as respostas das entrevistas, pode-se apontar aí, um dos aspectos que necessita de uma atenção especial, pois os alunos são agentes do processo educativo e, conseqüentemente, um dos interessados em compreender a importância da avaliação externa, sua finalidade e propósito.

Semelhantemente, à percepção de quatro participantes do questionário que consideraram que o PAER implica no processo educativo por ser capaz de **desvelar a realidade**, um dos entrevistados também aponta esta categoria. Como está descrito na análise da questão dez da entrevista, é esse desvelar da realidade que permite sua análise, a crítica e potencial transformação. Ideias como: “visão ampla da realidade acadêmica da instituição”;

“mostra da realidade acadêmica do colégio, quais são nossas deficiências para serem trabalhadas”, apresentadas pelos participantes, indicam que reconhecem que essa função da avaliação está sendo desenvolvida pelo programa. Assim como questionado na análise das respostas da pergunta dez, é necessário ir além do desvelar da realidade para, a partir de então, promover sua transformação. E, para isso, questões epistemológicas e axiológicas que emergem do paradigma que norteia a visão de educação que está implícita determinarão o que se fará após o desvelar da realidade.

#### 6.1.3.4 O cenário de 2020 com o ensino remoto e as implicações para o PAER

A última pergunta do questionário foi produzida diante do cenário do isolamento social causado pela pandemia da Covid-19 e o fechamento das escolas num primeiro momento, com a posterior abertura parcial, no segundo semestre do ano 2020. O PAER foi aplicado integralmente *on-line* no mês de outubro. Pelo fato de haver cidades onde não havia a possibilidade do retorno parcial às aulas, em muitas unidades escolares as provas foram realizadas pelos alunos em seus domicílios. Poucas escolas puderam receber alunos para a realização da prova *on-line* nas suas instalações.

A questão aberta que os participantes responderam foi: “Considerando o cenário de 2020 com o ensino remoto e a aplicação da prova totalmente *on-line*, quais foram as implicações para o PAER?”. A esta pergunta, 38 participantes responderam. Dois consideraram que não sabiam como responder; as 36 outras respostas foram categorizadas para a análise em três grupos: i) Infidelidade dos resultados; ii) Dificuldades com a tecnologia; iii) Indicação de forças e fraquezas. A Tabela 9 apresenta a distribuição dessas categorias:

**Tabela 9 - Resposta 11: Cenário de 2020 e implicações para o PAER**

<b>Categorias</b>	<b>Número</b>	<b>%</b>
Infidelidade dos resultados	27	71%
Dificuldades com a tecnologia	4	10%
Indicação de forças e fraquezas	3	7%
Não sei responder	2	5%

A maioria dos participantes considerou que os resultados não são fidedignos, tendo em vista que os alunos realizaram em suas casas, com acesso a várias fontes de informação, entre

livros, a internet e os próprios pais. Para alguns, as dificuldades apresentadas pelo ensino remoto comprometeram tanto o ensino como a aprendizagem, por isso, houve surpresa por parte dos participantes do questionário, diante do aumento significativo das notas dos alunos nesta edição do PAER, de maneira geral. O estatístico responsável pela condução de todo o processo de consolidação dos dados das avaliações e questionários do PAER confirmou essa percepção, ao afirmar que “foi nítida a mudança dos resultados nesta edição de 2020 em relação aos anos anteriores, em todo o Brasil, principalmente no ensino médio. Os resultados surpreenderam na análise dos índices.”. Nas análises estatísticas, os índices ficaram bem melhores, diferente do que ocorria na série histórica. Muitas possibilidades são cogitadas pelo estatístico, que considera que os resultados foram muito diferentes. Não devem ser ignorados, porém, são considerados com reservas. “Não pelo fato de a prova ter sido feita *on-line*, mas pelo fato de ter sido feita em casa.”. Desse modo, a percepção dos 27 participantes do questionário foi ratificada pelo estatístico que trabalha com os dados na mantenedora central do programa, após análise desses resultados em comparação com a curva que se desenhava ao longo do tempo em cada região do Brasil.

Portanto, essa expressiva indicação de 71% das respostas, de que os dados produzidos “não são confiáveis”, que “o resultado foi mascarado”, ou que “não reflete a realidade da escola”, deve-se ao fato de que, ao se depararem com os resultados, pode-se comparar com os anos anteriores e verificar que estavam discrepantes, com níveis de proficiência muito superiores, “exatamente num ano em que o processo de ensino e aprendizagem foi dramaticamente prejudicado neste primeiro ano de pandemia pela Covid-19” e, diante disso, “os resultados foram potencialmente afetados”, consideraram os respondentes.

O segmento da educação básica no contexto da pandemia da Covid-19 necessitou de uma rápida adaptação para o ensino remoto, prática que não encontrava, no país, precedentes legais, nesse nível de escolaridade. As escolas se adaptaram, os professores e os recursos tecnológicos tiveram que passar por adequação a essa realidade. Porém, os alunos dependiam da disponibilidade de equipamentos e internet em seus lares. Tendo em vista que a rede de escolas pesquisada tem centenas de unidades em todas as regiões do país, não somente em capitais e cidades de regiões metropolitanas, mas também em municípios de baixa densidade populacional, em locais onde o acesso a uma rede de computadores com conexão de qualidade para esse tipo de avaliação *on-line* de forma simultânea e sem riscos é mais difícil, quatro participantes indicaram as **dificuldades com a tecnologia**: tanto a falta de

equipamentos disponíveis nos lares, como a dificuldade de conexão com a internet, como implicações geradas para a edição de 2020 do PAER.

Dificuldade semelhante foi indicada por um dos participantes da entrevista, ao discorrer sobre as formas de aplicação das provas em um contexto diferente: excesso de alunos nas salas para o número de computadores disponíveis para a realização das avaliações na modalidade *on-line*. Seja nas condições normais com aulas presenciais, seja na condição de aulas remotas, a igualdade de condições assegurada a todos os alunos por força de norma é um dos requisitos fundamentais em uma avaliação de caráter emancipatório, como já comentado anteriormente. Acesso a equipamentos e a uma conexão de internet com qualidade para todos os alunos gozarem das mesmas oportunidades e condições são direitos fundamentais para que se desenvolva uma educação com qualidade social.

Três participantes consideraram que o cenário de 2020 com o ensino remoto, a necessidade de realizar a prova totalmente *on-line*, pode **indicar fraquezas e forças** dessa forma de aplicar a avaliação externa. Um dos participantes apontou como positiva a rapidez com que os resultados são disponibilizados, sendo possível apresentá-los à comunidade escolar poucos dias após a realização da avaliação. A necessidade de uma rápida adaptação de todos a essa realidade e a possibilidade de usar desse conhecimento em situações futuras de exames *on-line* foram considerados como demonstração de forças da rede de ensino. Ao mesmo tempo, as fragilidades também foram indicadas, e estas não se restringiram à estrutura e à tecnologia. A “falta de visão dos envolvidos” e a impossibilidade de considerar a fidedignidade do resultado para “validá-lo” como expresso nas respostas, indicam que ainda há obstáculos a serem superados e que estes puderam ser detectados nesse contexto de avaliação, também remota.

Essa experiência vivenciada nesses anos de pandemia tem sido objeto de estudos e pesquisas que estão sendo desenvolvidas para conhecermos as implicações educacionais desse período que envolveu, não somente o ensino remoto, mas as formas de relacionamento e os espaços de convivência. Inúmeras variáveis estão associadas a todo o contexto de isolamento social por um lado, e, por outro lado, relacionamento entre filhos e pais ou responsáveis que, por forças de circunstâncias, passaram a conviver mais tempo juntos e em situação de necessidade de continuar com rotinas de estudos, aulas e avaliações. Como bem foi sinalizado na categoria de análise, houve a oportunidade de se identificar a importância da educação formal e, ao mesmo tempo, a urgência de adaptação de todos a novas realidades.

Diante do alto índice de respostas que consideraram não fidedignos os resultados do PAER na edição de 2020, podemos refletir sobre a capacidade da educação, como processo formativo que se desenvolve tanto na vida familiar como nas instituições de ensino, contribuir para o “desenvolvimento pleno do educando, seu preparo para o exercício da cidadania e sua qualificação para o trabalho” (BRASIL, 1996), ao suscitar que, quando sob a necessidade de realizar uma prova remotamente, não se pode validar a veracidade das respostas, considerando a possibilidade de desonestidade na consulta de fontes ou mesmo na ajuda dos pais, como foi apontado por vários participantes. Nessas circunstâncias, o papel da escola e da família na formação moral dos alunos é colocado à prova, e pode-se questionar se os conteúdos abordados nas avaliações em larga escala são, de fato, o que nos sinaliza a qualidade do ensino que está sendo ministrado nas escolas. Ou, se o propósito da educação é levado a cabo com o aumento da proficiência nas disciplinas avaliadas por esse programa de avaliação externa.

A possibilidade de o ensino remoto ou híbrido tornar-se uma prática frequente em novos contextos de emergências sanitárias; por condições sociais, políticas ou mesmo econômicas; sinaliza a necessidade das redes e sistemas de ensino ressignificarem todo o processo educativo, considerando que o aprendizado e a avaliação da aprendizagem ou mesmo a avaliação externa, nestas circunstâncias, demandariam uma nova configuração de vários elementos envolvidos. No caso da rede investigada, em que a avaliação externa adquiriu centralidade; realizá-la de forma remota, sem o controle e vigilância descrito pelos entrevistados que apresentam o processo de aplicação com regras rígidas - como veremos a seguir na análise da questão quatro da entrevista – pode se constituir um desafio que vai muito além de questões técnicas e estruturais. Pode representar uma oportunidade de realinhar finalidades e propósitos não só da avaliação externa, como da própria educação ministrada.

## **6.2 Análise das Entrevistas**

Originalmente, a entrevista fora planejada para ser realizada com os gestores de oito regiões geográficas da estrutura organizacional da rede. Contudo, além do líder gestor que criou o PAER, apenas quatro deles atendiam aos critérios estabelecidos para participação, já descritos na introdução do capítulo. Após a gravação das entrevistas, estas foram enviadas aos participantes para validação, conforme previsto no Termo de Consentimento Livre e Esclarecido (TCLE).

O perfil dos participantes entrevistados considerou: o sexo (masculino – feminino); a idade; a formação; o tempo de atividade na rede e o tempo de atividade na função.

**Tabela 10 - Perfil dos participantes da entrevista**

Identificação	Sexo	Idade/ anos	Formação	Tempo de atividade da rede	Tempo de atividade na função
PLG	M	51	Doutorando	28 anos	10 anos
SGR-1	F	57	Mestre	35 anos	7 anos
SGR-2	F	51	Especialista	20 anos	6 anos
SGR-3	F	40	Mestranda	5 anos	2 anos
SGR-4	F	56	Especialista	39 anos	7 anos

Fonte: entrevistas; elaboração: autora.

As dez questões da entrevista semiestruturada foram organizadas por focos e eixos para facilitar a análise das respostas, conforme representado no Quadro 18.

**Quadro 18 - Questões da entrevista semiestruturada**

Foco	Eixo	Questão
<b>Percepção</b>	Visão	1 - Qual a sua visão sobre o Programa de Avaliação externa da Rede (PAER)?
<b>Processo</b>	Criação/ implantação	2 - Como foi o processo de implantação do PAER na sua região administrativa?
	Aplicação	4 - Como se dá o processo de aplicação do PAER?
	Articulação	6 - Há articulação dos resultados do PAER com outras avaliações (internas ou externas)?
<b>Resultados</b>	Objetivo	3 - O objetivo principal do PAER explicitado no projeto é: “Retratar a realidade da qualidade da educação, conforme os princípios que fundamentam a filosofia de educação da rede, através de informações que incentivarão a adoção de medidas que promovam o crescimento da rede”. Na sua visão, este objetivo foi atingido? Se não, por quê? Se sim, por quê?
	Uso dos resultados	5 -Qual o uso que tem sido feito dos resultados da referida avaliação?
	Ganhos	7 - Na sua visão, quais os ganhos após 7 anos da referida avaliação?
	Dificuldades	8 - Quais as dificuldades ou aspectos a desenvolver relativos ao referido programa de avaliação da rede?
	Implicações	10 - Quais as implicações do PAER no processo educativo?

Elaboração da autora.

A questão nove, por tratar de sugestões dos participantes para o aperfeiçoamento do PAER, está desvinculada deste Quadro 18, com um espaço próprio na análise conclusiva das entrevistas.

Na divisão das questões em focos e eixos temáticos levou-se em conta a natureza do conteúdo das questões e o tipo de abordagem adotada na pesquisa. No primeiro foco, o termo **Percepção** é empregado apoiado na interpretação de Matos e Jardimino (2016) do uso da palavra na literatura educacional para “informar a maneira como as pessoas percebem, avaliam e agem com relação ao fenômeno pesquisado.” (p. 30). O que numa pesquisa de abordagem qualitativa remete à vivência e experiência dos indivíduos em relação ao fenômeno, objeto ou realidade percebida. Neste caso, a visão sobre o PAER.

Nos dois focos seguintes, os conceitos utilizados derivam das ideias de Draibe (2001) que, ao trabalhar por muitos anos com avaliações de políticas, programas e projetos sociais, formulou um instrumental analítico básico para pesquisas de avaliações, propondo não apenas uma metodologia de trabalho, como um modelo analítico para avaliação de processos e de resultados. Para tanto, a autora formulou conceitos instrumentais que são utilizados na estruturação da análise das respostas, entre os quais destacamos: i) **Processo**: refere-se ao ciclo vital dos programas que “nascem, crescem, transformam-se, reformam-se. Eventualmente estagnam, às vezes morrem.” (p. 26). Neste foco, agrupamos as questões ligadas à criação, implementação e desenvolvimento do programa investigado. ii) **Resultado**: quer-se saber se o programa alcançou seus objetivos, o quanto e com qual qualidade. Neste foco, podemos ainda aprofundar a avaliação do programa, levando em conta os resultados no sentido estrito, que são os produtos previstos em suas metas e objetivos propostos, também denominados desempenho; os impactos, que “referem-se às alterações ou mudanças efetivas da realidade sobre a qual o programa intervém e por ele são provocadas.” (DRAIBE, 2001, p. 21); e os efeitos, que são outros impactos, esperados ou não, que afetam o meio institucional e social no qual o programa foi realizado. Desse modo, foram agrupadas as questões relacionadas ao objetivo geral do programa, ao uso dos resultados, às implicações e aos aspectos desenvolvidos e a desenvolver percebidos pelos participantes.

### **6.2.1 – 1º Foco: Percepção sobre o PAER**

No primeiro foco, constituído de um único eixo temático – visão sobre o PAER –, as respostas dos participantes foram organizadas nas seguintes categorias de análise:

**Quadro 19 - Categorias e frequência das respostas do Eixo Visão**

Questão 1 – Visão dos entrevistados sobre o PAER						
Categorias de análise						
Participantes	Avaliação como instrumento de gestão	Apreciação pela avaliação externa	Articulação entre aval. Externa e aval. Interna	Apreciação com ressalvas	Comparação de dados	Propulsor de reformas
PLG	X					
PGR1		X	X			
PGR2				X	X	
PGR3	X				X	
PGR4		X			X	X

Elaboração: autora.

A primeira categoria analisada neste eixo apresenta a percepção do PAER como uma avaliação que pode ser utilizada como **instrumento de gestão**. Vale ressaltar que os cinco respondentes ocupam funções de gestão. O participante PLG assim expressou esta ideia:

PLG – Sem avaliação nós não temos como planejar com assertividade. [...] a importância de um programa de avaliação é fundamental para que, como rede, tenhamos um planejamento mais assertivo e possamos continuar crescendo e aprimorando sempre.

Outro entrevistado também demonstrou esta percepção ao responder:

PGR3 – [...] eu gosto muito de avaliação porque eu gosto de trabalhar com planejamento estratégico. E a avaliação para mim é um instrumento que vai metrificando se eu estou caminhando dentro do objetivo que eu tracei dentro do planejamento. Se eu não avalio o processo, eu não consigo perceber falhas e não consigo fazer nenhum tipo de ajuste.

Worthen, Sanders e Fitzpatrick (2004, p. 151) denominam de “avaliações centradas na administração” aquelas cujo fundamento é oferecer informações para ajudar na tomada de decisões. Apontam que a “informação avaliativa é parte essencial de decisões inteligentes, e o avaliador pode ser eficiente trabalhando para administradores, legisladores, diretorias e outros profissionais que precisam de boas informações avaliativas.” (WORTHEN; SANDERS e FITZPATRICK, 2004, p. 151). Para Calderón e Borges (2020), “este tipo de abordagem baseia-se na concepção da avaliação como um processo para melhoria de programas e das políticas educacionais”. E, conseqüentemente, “radicaria a força positiva dos processos avaliativos.” (CALDERÓN; BORGES, 2020, p 40).

A ideia da avaliação em larga escala como instrumento de gestão está posta nos documentos oficiais como um dos objetivos do sistema de avaliação da educação básica ao longo dos anos, como demonstrado nas várias edições de portarias que o regulamentaram, tal

como a Portaria nº 174, de 13 de junho de 2015, que estabelece como objetivo específico: “fornecer informações sobre as unidades escolares, que sejam úteis aos gestores da rede à qual pertençam as escolas avaliadas” (BRASIL, 2015). De igual modo, o documento normatizador do PAER também estabelece como objetivo específico: “Prover informação para a tomada de decisões na gestão estratégica da educação básica a nível de Divisão Sul Americana.” (D.O., p. 5). Portanto, a percepção se coaduna com um dos objetivos específicos do programa.

Sobre os sistemas de avaliação em larga escala e externos, Souza e Oliveira (2010) apontam que:

Os objetivos da avaliação, declarados nos documentos oficiais, tendem a afirmar como expectativa que a avaliação venha a subsidiar os diferentes níveis do sistema na tomada de decisões com vistas à melhoria da qualidade do ensino. Desse modo, espera-se que os resultados da avaliação venham a ser apropriados pelos gestores e equipes centrais e regionais das Secretarias de Educação, bem como pelas escolas, havendo menção ainda, por alguns estados, quanto à expectativa de que a comunidade escolar como um todo – envolvendo alunos e pais – venha a se inteirar dos resultados obtidos pelas escolas. Ou seja, há referências de que a avaliação deve iluminar e trazer consequências para a formulação e/ ou reformulação das políticas educacionais em relação às suas diretrizes, prioridades e metas, focalizando questões relativas à gestão, ao currículo e à infraestrutura, assim como às propostas e práticas das escolas, particularmente em relação a aspectos que concernem ao desenvolvimento do currículo. (SOUZA; OLIVEIRA, 2010, p. 803).

Contudo, os autores sinalizam que tem se desencadeado um debate sobre o uso dos resultados das avaliações externas como instrumento de gestão e que duas visões estão em disputa. De um lado, esse uso está associado a um processo ligado a uma lógica de forte influência empresarial, na qual torna-se instrumento de controle, de racionalidade instrumental e de busca por maior eficiência. De outro lado, uma busca pela autonomia dos processos educativos no âmbito da própria escola, com ênfase nos processos de autoavaliação e na promoção de ações indutoras de melhoria da qualidade, de fato (SOUZA; OLIVEIRA, 2010).

De igual modo, sobre essa tensão entre dois modelos conflitantes, Dias Sobrinho (2003) considera como paradigmas não exclusivos do campo da avaliação, e sim da ciência, da política, da ideologia e da vida em geral. De um lado, as visões de mundo organizadas por esquemas objetivistas, utilitários, pragmáticos e individualistas, que moldam um modelo de gestão racional que atribui grande importância aos números, aos produtos, aos planejamentos rigorosos. Do outro, as visões de mundo pautadas numa perspectiva que valoriza os processos vividos pelos indivíduos e pelos grupos, orienta-se para os descobrimentos por vias exploratórias, para uma gestão participativa, na qual a negociação faz parte de todo o processo. É fato que, apesar de serem apresentados de forma isolada e como excludentes, o

autor considera que os dois paradigmas “na realidade frequentemente se combinam e se entrecruzam.” (DIAS SOBRINHO, 2003, p. 85).

As principais críticas ao uso dos resultados das avaliações externas como instrumento de gestão estão associadas ao enfoque derivado do paradigma dominante, descrito no capítulo cinco. A gestão educacional – assim denominada pela instância na qual atuam os participantes das entrevistas, distinguindo da gestão escolar<sup>33</sup> – abrange um sistema ou rede de ensino, que se orienta pelos pressupostos deste paradigma, utiliza os resultados e dados como instrumento de controle, de competitividade e de eficácia gerencial que se concretiza, por vezes, na forma de *rankings*, ou do *assessment* que promove a hierarquização e o controle das instituições numa lógica psicométrica, e da *accountability*, tão privilegiada e legitimada nessa cultura gerencial que se instalou nos sistemas educacionais.

Algumas palavras utilizadas pelos participantes da entrevista podem ser interpretadas como indicativos dessa visão, tais como: “planejamento mais assertivo”; “assertividade” (PLG); “metrificando”; “planejamento estratégico” (PGR3); que são termos sugestivos de um modelo gerencialista de busca de eficiência. É, porém, no aprofundamento da análise da percepção dos gestores sobre o uso dos resultados que podemos ter uma compreensão mais ampla da visão que orienta o modelo de gestão da rede.

Ao tratar da temática da avaliação em larga escala, Luckesi (2018) afirma que o ato avaliativo “é parceiro do gestor da ação, revelando-lhe a qualidade dos resultados da ação que administra [...]. A avaliação tem o exclusivo papel de revelar a qualidade da realidade; cabe ao gestor da ação tomar decisões.” (LUCKESI, 2018, p. 201). Desse modo, o uso da avaliação como instrumento de gestão é desejável e necessário. O que pode ser questionável é se será utilizada como instrumento de controle, restrições, sanção e práticas exacerbadas de responsabilização, ou como um empreendimento social, que visa à melhora da qualidade educativa, com a participação dos interessados, que priorize as atividades formativas, envolvendo todos os elementos do sistema, cujos desdobramentos devem concorrer para um foco central: a garantia de um ensino de qualidade e da efetiva aprendizagem.

A segunda categoria de análise que indica a visão dos participantes sobre o PAER é a **apreciação pela avaliação externa**. Analisamos esta categoria em conjunto com a outra categoria que emergiu das respostas, a **apreciação da avaliação externa com ressalva**. Dos

---

<sup>33</sup> Machado (2012) apresenta a distinção entre gestão educacional e gestão escolar, apontando que a primeira operacionaliza e efetiva as políticas no âmbito de sistemas educacionais, portanto, em nível mais abrangente; enquanto a segunda está circunscrita ao contexto da escola.

cinco participantes das entrevistas, dois apreciam a avaliação externa e um aprecia com ressalva. As ideias centrais que expressam esta visão podem ser depreendidas das respostas:

PGR1 – Eu gosto da ideia de ter uma avaliação externa. Porque é um olhar diferente.

PGR4 – Eu acredito que como rede, nós precisávamos ter uma avaliação [...] que o PAER veio trazer esta notícia para a gente, este resultado.

A visão positiva sobre a avaliação em larga escala, principalmente em nível de gestão, pode ser reflexo da compreensão das funções específicas que esta modalidade de avaliação possui. Para Freitas *et al.* (2017), essas avaliações “quando conduzidas com metodologias adequadas podem trazer importantes informações sobre o desempenho dos alunos, dados sobre os professores, condições de trabalho e funcionamento das escolas de uma rede.” (FREITAS, *et al.*, 2017, p. 47).

Domingos Fernandes (2009), após discorrer sobre as características e funções das avaliações externas e em larga escala, considera que esses exames “podem ter certamente um conjunto de vantagens e, talvez em parte, por essa razão, são aplicados em quase todos os países” (FERNANDES, D., 2009, p. 137). O autor ainda considera que, partindo do princípio de que sua concepção e elaboração têm qualidade pedagógica, educacional e formativa, algumas vantagens podem ser listadas:

1. Podem exercer um efeito moderador importante nas avaliações internas.
2. Podem induzir práticas inovadoras de ensino e avaliação.
3. Podem contribuir para avaliar o sistema educacional e ajudar a melhorar a tomada de decisões em todos os níveis.
4. Podem alertar as escolas para as necessidades de melhorarem seus projetos educacionais.
5. Podem dar indicações úteis às escolas, aos professores e aos alunos acerca do que é importante ensinar e aprender. (FERNANDES, D., 2009, p. 137)

Apesar de ser um campo de tensões, como já descrito anteriormente, no qual questões de natureza teórico-epistemológica se confrontam no âmbito da produção científica brasileira, com duras críticas, principalmente aos usos dos resultados, há que se reconhecer, como considera Freitas *et al.* (2017), que é possível ser construída “uma posição menos dicotômica e mais construtiva, que pretende apropriar-se das ferramentas avaliativas sob outro conceito do que seja avaliar.” (FREITAS *et al.*, 2017, p. 10). Desse modo, uma disposição de apreciação ao processo avaliativo externo pode ser a sinalização de uma compreensão dessas possibilidades e vantagens apontadas pelos autores citados.

Outra categoria que emerge das respostas à primeira pergunta está voltada para a **comparação dos dados**. Na percepção de três participantes da entrevista, o PAER possibilita essa comparação. No caso do primeiro respondente, não fica claro o tipo de comparação à

qual se refere: se uma comparação de dados oriunda da análise dos resultados da mesma unidade ao longo da série histórica da avaliação; ou a comparação dos dados de várias unidades escolares ou regiões administrativas. Ao considerar – “Um programa de avaliação externa, sem dúvida nenhuma, é um marco para avaliações, para comparações de dados, enfim, crescimentos saudáveis” (PGR2) – podemos interpretar das duas formas citadas. Já o segundo respondente pontuou sua percepção ao afirmar: “E este específico, que é um de larga escala, eu acho interessante porque, além de eu conseguir ver o resultado *in loco* de uma unidade escolar, eu consigo comparar por campo, regiões administrativas. E também a questão da comparação nacional” (PGR3). Desse modo, considera a comparação de dados nas três instâncias da rede: unidade escolar, sub-região e região geográfica. Por fim, o terceiro participante, à semelhança do segundo, considera a possibilidade de comparação dos dados entre regiões administrativas, ao afirmar: “Eu acredito que, como rede, nós precisávamos ter uma avaliação. Eu sempre trabalhei mais na região sul, a gente fazia algumas avaliações. Mais a nível de [sic] região. Mas a gente nunca teve um parâmetro para ver como estava assim, num todo.” (PGR4).

A Portaria MEC nº 458, de 5 de maio de 2020, que institui normas complementares necessárias ao cumprimento da política nacional de avaliação da educação básica, estabelece como objetivos do Saeb a comparabilidade dos dados:

- I – construir uma cultura avaliativa, ao oferecer à sociedade, de forma transparente, informações sobre o processo de ensino-aprendizagem em cada escola, comparáveis em nível nacional, anualmente e com resultados em tempo hábil, para permitir intervenções pedagógicas de professores e demais integrantes da comunidade escolar;
- II – produzir indicadores educacionais para o Brasil, suas regiões, unidades da federação, municípios e instituições escolares, tendo em vista a manutenção da comparabilidade dos dados, permitindo, assim, o incremento de séries históricas. (BRASIL, MEC, 2020)

Percebe-se que a comparabilidade proposta no inciso I refere-se à possibilidade de comparar os resultados de cada escola, em nível nacional e anualmente, pela comunidade escolar. No inciso II, a comparabilidade dos dados pode ser realizada nos diferentes níveis e sistemas educacionais, valendo-se de séries históricas. A portaria também estabelece que os resultados do Saeb devem possibilitar:

- IV – a implementação gradual de aperfeiçoamentos e inovações nos processos de avaliação e exames, mantendo a comparabilidade entre as distintas pesquisas e a manutenção de séries históricas;
- V – a difusão, com transparência e de fácil acesso, para toda a sociedade, de diagnósticos e pesquisas sobre o sistema de educação básica, em níveis de desagregação distintos, até a instituição escolar, comparáveis em nível nacional. (BRASIL, MEC, 2020).

Portanto, a comparação de dados, seja por meio da análise de uma série histórica da própria instância em questão (escola, sistema de ensino ou país), ou entre diferentes instâncias, faz parte dos objetivos do Saeb – e entende-se, que desejáveis para o PAER –, cuja formulação e criação baseou-se nos pressupostos da avaliação nacional. A possibilidade de comparação de dados não foi arrolada como um objetivo específico do programa, mas o documento oficial, ao tratar do tema Resultados, apresenta as formas pelas quais esses dados podem ser compartilhados: i) o sistema de informações do PAER, no qual cada usuário registrado tem acesso a vários tipos de informações, podendo selecionar dados por ano, por nível de hierarquia (regional ou local), por disciplina etc. ii) por painéis de controle (*dashboards*) construídos e distribuídos como arquivos de Excel. Esses painéis permitem visualizar grande quantidade de dados em forma gráfica e escolher o que se quer visualizar. iii) informes escritos que detalham, principalmente, dados dos questionários contextuais. O documento sinaliza que: “Estas ferramentas são especialmente úteis para apreciar a evolução de certos indicadores, particularmente os relacionados com a avaliação cognitiva.” (D.O. p. 8). Essa “evolução de certos indicadores” é possível por meio da comparação de dados. Desse modo, os próprios dispositivos que dão aos gestores os dados gerados pelo programa oferecem ferramentas que permitem a comparação. A comparação de dados pode ser um nível prévio que pode subsidiar um posterior processo de análise e de reflexão a partir desse exercício.

A categoria **articulação entre avaliação externa e avaliação interna** foi manifestada por uma participante como uma contribuição do PAER para a rede, garantindo a oportunidade de crescimento. Na percepção da entrevistada:

PGR1 – Então, é um olhar que entra na escola e que me ajuda a ver novos horizontes. É claro que nem toda avaliação – aliás, a avaliação externa nunca vai te dar assim: “exatamente este ano, com esta turma, você esteve dessa forma”. Porque nós sabemos das peculiaridades. [...] Então, tendo um olhar maduro que você juntar o teu olhar interno com o olhar externo abre muitas possibilidades de crescimento.

Sobre essa questão, alguns autores (RABELO, 2013; FERNANDES, D., 2019; SOUZA, 2013) têm apontado a necessidade de se buscar uma articulação entre a avaliação externa e a avaliação interna, pois a literatura tem sinalizado um descompasso entre ambas. Souza (2013) assim constata:

Após 25 anos de avaliação externa no Brasil, está na hora de fazer uma espécie de parada para fazer uma reflexão sobre essas avaliações. Uma reflexão cujo intuito é fazer algumas observações e contrapontos com o que deparamos como uso dos resultados de uma avaliação externa e, em consonância, verificar o uso das informações obtidas nas avaliações internas, de aprendizagem, realizadas nas escolas, enfatizando sua articulação com a gestão pedagógica da instituição. É

proveitoso ampliar a compreensão dos profissionais da educação a respeito da importância dessa articulação entre a avaliação interna, que a escola faz em sua sala de aula, e os resultados obtidos nas avaliações externas, que ocorre na educação básica no país. (SOUZA, 2013, p. 163).

Freitas *et al.* (2017) postulam a necessidade da integração entre os três níveis de avaliação da qualidade do ensino: a avaliação em larga escala em redes de ensino; a avaliação institucional; e a avaliação da aprendizagem em sala de aula. A busca pela qualidade educacional requer a interação entre esses três níveis. Nesse contexto, principalmente a avaliação externa vem ganhando uma relevância e significado para a gestão pedagógica, quando adquire uma dimensão mais construtiva e pode ser articulada com a avaliação da aprendizagem em sala de aula.

A última categoria de análise originada na resposta à primeira pergunta refere-se à visão do PAER como um **propulsor de mudanças**. Esta percepção é expressa na resposta de respondente PGR4: “O PAER veio trazer esta notícia para a gente, este resultado, ele veio num momento muito bom, que ele conseguiu provocar algumas mudanças”. Há um pressuposto inerente ao processo avaliativo como propulsor de mudanças. A opção pela realização de uma avaliação externa em um sistema de ensino envolve a ideia de saber-se onde está, para se pensar em um movimento em direção a algum novo patamar. Esse novo patamar é associado à melhoria da qualidade educacional, porém, há implicações geradas pelas avaliações em larga escala que afetam currículos, instituições, alunos, professores, administradores e que são objeto de estudos desde as primeiras iniciativas de avaliação externa e em larga escala até os dias atuais. Sobre essa característica do processo avaliativo, Souza e Oliveira consideram que:

Ao realizar determinado processo avaliativo, espera-se, explicitamente, verificar quão distante se está da situação desejável e, a partir daí, definir elementos para modificar a situação em direção ao padrão desejado. Além disso, a avaliação pode, também, possibilitar a emergência de propostas de redirecionamento ou transformação da situação avaliada. Não é possível pensar a avaliação dissociada da ideia de modificação. Para isso, talvez, a mais importante questão seja como criar um envolvimento tal que impulse os membros da instituição a se engajarem no processo de transformação. (SOUZA; OLIVEIRA, 2010, p. 801).

O potencial de uma avaliação provocar mudanças não significa, necessariamente, que induzirá à situação desejada tão somente porque se faz uso de um programa de avaliação. Há necessidade de clareza da finalidade e propósito da avaliação, com a definição de padrões e critérios consensuais do que se entende por qualidade educacional e todas as dimensões envolvidas no processo, para que seja possível afirmar que as mudanças alcançadas foram as desejadas.

### 6.2.2 – 2º Foco: Processo

O segundo foco abrange três eixos: i) criação/ implementação; ii) aplicação; iii) articulação; que representam os ciclos de desenvolvimento do programa. A primeira questão deste eixo é a segunda pergunta da entrevista dirigida ao líder gestor sobre a criação do PAER e aos demais gestores, sobre a implementação. As categorias de análise que emergem das respostas estão descritas no Quadro 20.

**Quadro 20 - Categorias de análise do Eixo Criação/ Implementação**

Questão 2 – Como foi o processo de criação/ implantação do PAER na sua região administrativa?						
Categorias de análise						
Participantes	Necessidade de aprimoramento do processo de ensino e de aprendizagem	Orientação aos coordenadores e professores	Treinamento e formação só para equipes de liderança	Lacunas em nível conceitual	Ausência de formação para professores	Ausência de participação dos professores
PLG	X					
PGR1		X				
PGR2			X			
PGR3				X		
PGR4					X	X

Elaboração: autora

A razão indicada pelo PLG para a criação do PAER foi a **necessidade de aprimoramento do processo de ensino e de aprendizagem**. Essa ideia está expressa no Documento Oficial do programa:

Tem como propósito desenvolver um sistema de avaliação externa de escola e colégios da rede (nos níveis fundamental e médio) em todo o território nacional, procurando produzir informação que permita assegurar a qualidade da educação oferecida. (D.O., 2017, p. 3)

A palavra utilizada pelo PLG em sua resposta foi “aprimoramento”, que segundo o **Michaelis Dicionário Brasileiro da Língua Portuguesa (2015)**, é a “ação de tornar algo mais perfeito, apuro, esmero”; termo relacionado ao processo de melhoria e, portanto, busca da qualidade do processo de ensino e aprendizagem. Desse modo, o respondente apresentou a causa da criação do programa e não o processo de criação: se houve participação de representantes da comunidade escolar, se houve assessoria externa à rede, quais os níveis da estrutura da rede participaram do processo de criação etc. Por outro lado, na resposta dos

demais entrevistados que versava sobre a implantação do PAER em sua região, estes focaram no processo, com ênfase na orientação, treinamento ou formação das equipes para o conhecimento do programa ou, como se pode identificar nas três últimas categorias, na falta ou ausência dessa formação.

Para o PGR1, na implementação “nós orientamos todos os coordenadores de escola, nós tivemos acesso na capacitação do começo do ano aos professores e nós falamos sobre isso”. Já para o PGR2, que por ocasião da implantação do programa não estava em função de gestão, mas atuava em uma unidade escolar, a lembrança foi que apenas as equipes de liderança receberam treinamentos: “lembro que nesta construção, várias formações chegaram, não para nós diretamente, mas para a nossa liderança e que vinha [sic] inclusive com algumas pessoas do Inep”. De igual modo para o PGR4, os professores não receberam formação sobre o programa e não participaram de sua elaboração. Considera que: “no momento da implantação e até hoje, há um certo descontentamento pelo professor não ter acesso”.

Para o PGR3, que na ocasião da implementação não estava na rede, sua percepção é de que houve **lacunas conceituais** sobre o programa: “Houve algumas lacunas neste processo de implantação, em questão conceitual. E foi criado um conceito bastante equivocado da prova”. Sem esclarecer qual o conceito que foi estabelecido, ou qual a lacuna gerada, a ênfase da resposta foi na falta de valorização da avaliação por parte de todos os envolvidos: professores, alunos e gestores das unidades escolares.

É fato que cada entrevistado estava falando de uma região do país, e que é possível que o processo de implementação tenha sido diferente nas distintas regiões. Contudo, foi possível constatar, em relação à reação dos professores por ocasião da implementação do programa, algumas expressões que reforçam a ideia de não participação ou formação para esse grupo, até mesmo pelo PGR1 que afirmou ter havido formação para os professores em sua região, porém, considera que: “Então, o modo como o professor, a escola recebeu, o modo como foi organizado lá na escola, ainda com muita insegurança, queira ou não queira, a gente ainda precisa falar dos medos também; “E agora, como que vão me olhar? O que vão fazer com estes dados?”. As palavras “susto”, “surpresa” e “suspense” apareceram na fala de alguns entrevistados para caracterizar a reação dos professores nesse processo de implantação.

Analisando a avaliação de larga escala, Luckesi (2018) considera o professor em sala de aula como o primeiro elo do sistema de ensino. Do ponto de vista ascendente, é ele que tem o contato com os alunos que serão avaliados por meio de provas. Portanto, é fundamental que seja considerado como um importante elemento dessa cadeia de componentes do sistema

educacional. Freitas *et al.* (2017), apresentando uma análise sobre a avaliação em larga escala de redes de ensino, de igual modo consideram que “é necessário que os princípios orientadores dos processos de avaliação sejam construídos e conhecidos por todos, de forma a conseguir o maior envolvimento possível no processo.” (FREITAS *et al.*, 2017, p. 78). Entende-se que, no caso de uma rede de extensão nacional, a participação de todos na construção dos princípios orientadores do programa de avaliação pode ser inviável. Porém, fazer tais princípios serem conhecidos de todos passa a ser indispensável para o êxito do próprio programa.

Penna Firme (2017), ao descrever os avanços, nos anos recentes, da avaliação de programas e projetos sociais, educacionais ou culturais, considera como imprescindível “a capacitação de professores, líderes, dirigentes e profissionais, nos vários âmbitos disciplinares, para a melhor utilização da avaliação.” (PENNA FIRME, 2017, p. 1). Sobre a participação de todos os envolvidos no processo, a autora conclui:

Em síntese, porém, um fator sumamente crucial e que está presente neste momento atual de avanços na avaliação é o respeito à participação efetiva de todos os interessados no processo ou *stakeholders* e que independente de sua condição social, econômica acadêmica ou outras, são elementos cruciais que devem atuar no processo, desde sua concepção até seus resultados. Somente assim, a avaliação será útil nas necessárias transformações que os programas e projetos sociais, educacionais e culturais pretendem alcançar. (PENNA FIRME, 2017, p. 4).

Para a autora, a participação dos envolvidos deve ocorrer desde a “concepção” até seus “resultados”. Nessa perspectiva, a participação dos professores, ainda que por representatividade, num processo de criação ou concepção do programa; ou no processo de implementação na unidade escolar, participando ativamente de todas as fases, é imprescindível para que o programa alcance os objetivos declarados de melhoria da qualidade do processo educacional. Essa compreensão se dá, quando a avaliação representa um enfoque mais amplo e amadurecido, que vai além do medir e aferir, do gerar dados que são apenas resultados de testes cognitivos, para um processo de reunião de informações e dados que alimentam e estimulam uma análise reflexiva das práticas e a busca de melhoria de todas as instâncias envolvidas: desde a sala de aula até os gestores da rede, um processo mediante o qual todos têm uma visão clara de que o objeto de atenção da avaliação vai muito além do desempenho dos alunos nos testes e que todos são responsáveis pelo crescimento em qualidade.

Portanto, as categorias de análise que emergiram das respostas indicam que apenas um entrevistado considerou que houve, na implantação, **oportunidade de formação para professores** a respeito do programa, mas também considerou em sua resposta, que houve insegurança e medo por parte deles. O contato com uma avaliação que apresenta o desempenho

dos alunos que estão sob a responsabilidade de um professor, ou a possibilidade de os resultados serem utilizados para algum tipo de comparação ou ranqueamentos, pode gerar essa insegurança, que é até previsível diante do novo. Porém, a insegurança, o medo, a surpresa podem ser agravados com o desconhecimento. E o desconhecimento pode ser evitado com a formação, a discussão e o envolvimento de todos os interessados desde a concepção do programa.

A questão quatro da entrevista abordou a fase de aplicação do PAER. Deu origem ao terceiro eixo de análise, conforme o Quadro 21:

**Quadro 21 - Categorias de análise do Eixo Aplicação**

Questão 4 – Como se dá o processo de aplicação do PAER?					
Categorias de análise					
Participantes	Aplicação impressa e aplicação <i>on-line</i>	Organização estrutural	Regras definidas	Formação técnica para aplicadores e coordenadores	Período e datas institucionalizados
PLG	X			X	
PGR1	X		X	X	
PGR2					X
PGR3	X	X			
PGR4			X		

Elaboração: autora

Três dos cinco entrevistados reportaram a forma como a prova é aplicada: **impressa e *on-line***. Dois dos entrevistados (PLG e PGR1) consideram que as duas modalidades requerem encaminhamentos diferentes, mas não ponderaram se há alguma dificuldade ou vantagem em uma das formas de aplicação. Já o PGR3 indica as duas formas, porém, sendo da região norte do país, considera a forma *on-line* como um grande desafio, em virtude da oscilação da internet e das dificuldades com o volume da rede em algumas sub-regiões (O uso da palavra “volume” pelo participante refere-se à potência da internet disponibilizada para determinada região). Para a aplicação no ano de 2020, com as restrições sanitárias, houve um esforço logístico para que todos os alunos pudessem realizar a prova na unidade escolar, visto que vários não teriam equipamento e uma internet estável para a realizar em suas residências. Essas condições foram assim descritas:

PGR3 - Esse ano, tudo mudou. Da prova física passou a ser digital. Então, nós temos aqui sub-regiões desafiadoras, onde a oscilação de rede é enorme, então, nós fizemos um preparo de direcionamento de volume de rede para a unidade escolar, e nesse dia não houve nenhum outro tipo de aula síncrona. A rede foi destinada especificamente

para as provas. [...] Nós minimizamos, não deixamos a prova na gestão total do aluno ou da família em casa. Trouxemos o aluno para a escola. É claro que alguns alunos preferiram ficar em casa. Mas já é um percentual muito baixo. Praticamente 90% fez a prova digital na escola.

Alguns autores consideram que as políticas de avaliação podem conter possibilidades emancipadoras que conduzam a um crescimento da qualidade do ensino com vistas ao desenvolvimento de uma sociedade mais justa (FREITAS *et al.*, 2017; FERNANDES, D., 2009; SOUZA; ARCAS, 2010). Em contrapartida, é possível que os processos avaliativos conduzam a uma intensificação das desigualdades educacionais e sociais. Por isso, o próprio acesso aos equipamentos eletrônicos e à internet, que ofereça condições para a realização das avaliações externas de forma *on-line* em determinadas regiões do país, pode ser um fator de desigualdade de acesso aos recursos educacionais. Sobre isso, Domingos Fernandes (2009), ao discutir sobre os principais elementos de uma avaliação externa em larga escala, considera a equidade como uma questão a ser refletida e investigada.

Normalmente a discussão tem se circunscrito à necessidade de se criar *igualdade de oportunidades* para todos os alunos e, em particular, para os que pertencem a grupos que carecem de necessidades educacionais especiais e a outros em geral designados por *minoritários*. E isso pode ter relação com a origem socioeconômica, a origem étnica, a filiação religiosa, o gênero, a língua materna ou com qualquer outra característica. Em geral, procura-se assegurar igualdade de oportunidades com sistemas, mais ou menos eficientes, de apoio e complementos educacionais. (FERNANDES, D., 2009, p. 129)

Embora a rede objeto desta pesquisa seja de natureza privada, é filantrópica, atendendo 20 por cento dos alunos com bolsas de 100 ou 50 por cento, oriundos de famílias que comprovam índice de carência estabelecido pela legislação. Sendo assim, a desigualdade socioeconômica pode existir entre esses alunos em relação aos demais, em qualquer parte do país, não apenas na região norte, de onde a entrevistada responde. Porém, ainda que a rede de ensino proveja os meios para reduzir essas desigualdades, é possível que haja fatores externos, como a capacidade de internet disponível na região, que precisam ser levados em conta quando se realiza uma avaliação que considera a equidade como um valor a ser alcançado.

No documento oficial do PAER, a questão da equidade é mencionada como implícita à ideia de qualidade:

Um dos objetivos da nossa rede consiste em garantir uma educação de qualidade para todos os estudantes. Implícito na ideia de qualidade da educação está o conceito de equidade, como princípio que promove justiça, imparcialidade e igualdade de oportunidades para todos. (D.O., 2017, p. 4).

A aplicação das provas na modalidade *on-line* fez-se necessária no ano 2020, diante da necessidade de isolamento social imposta pela pandemia por Covid-19. Porém, nos anos

anteriores, a adesão pela modalidade era optativa para a unidade escolar cujo número de computadores e a velocidade da rede de internet fossem suficientes para atender a todos os alunos avaliados, simultaneamente. Na maioria das regiões brasileiras, a prova de 2020 foi realizada pelos alunos em seus próprios equipamentos e em suas residências. Para os respondentes do questionário *on-line*, que são gestores das sub-regiões e com uma atuação mais próxima às unidades escolares, foi possível incluir uma pergunta sobre a experiência da aplicação do PAER no contexto da pandemia, e a percepção desses participantes está descrita na próxima seção do capítulo.

Três das demais categorias que emergiram das respostas sobre a aplicação estão associadas à organização do processo: **a organização estrutural; as regras definidas e o período e data de aplicação institucionalizados**. Nas respostas, é possível identificar uma percepção comum de que há um modelo estruturado de aplicação a ser seguido por todos e que esse modelo gera confiabilidade no programa. Essas respostas indicam algumas ideias prevalentes:

PGR4 – Ele foi pensado assim, com muito, muito sigilo, a gente tinha que mandar o nome das coordenadoras, das pessoas que estavam responsáveis por receber as caixas. Só podia abrir o envelope – tinha um envelope específico –, só podia ser aberto na frente do aluno. Todo um cuidado. E depois de aplicado, tinha que colocar tudo dentro de um envelope e lacrar na frente dos alunos, não é? Tinha todo um cuidado que deveria ser respeitado.

PGR2 - Tem um espaço reservado, todos sabem que naquele dia vai acontecer uma avaliação determinada, que avaliação é essa, alguns ficam esperando. [...] Mas eu vejo que hoje já é... e tem um espaço institucionalizado de organização. Tanto da equipe pedagógica como professores, gestores.

PGR1 – Então, este processo é feito assim. Existem regras bem rigorosas. E aqui, a gente tenta aplicar todas elas, a gente fica controlando essas maneiras de aplicar. [...] Os alunos recebem isso também no celular e é tudo muito bem organizado, nesse sentido.

A estrutura que envolve o processo de aplicação das provas, em condições distintas, para uma rede que está presente em todos os estados da federação, requer uma gestão e organização sistêmica que controle os fatores que possam causar alguma inconsistência nos resultados. Por isso, o documento oficial sinaliza:

Qualquer processo de investigação ou avaliação que envolve pessoas deve sempre ser realizado cuidando dos aspectos técnico-metodológicos e éticos que correspondam a cada caso. É o que acontece com o PAER e suas diferentes avaliações. Cada uma é planejada com mais de um ano de antecedência, e cada etapa levada a cabo com a rigorosidade científica que demanda a atividade e que o contexto institucional permite. (D.O., 2017, p. 6).

Portanto, o processo de aplicação de uma avaliação externa também precisa ser considerado, ao se levar em conta o grau de confiabilidade que essa avaliação possui. Essa é uma característica a ser considerada em qualquer avaliação. Sobre a confiabilidade, Domingos Fernandes (2009) afirma:

Para diminuir, ou mesmo anular, essas *ameaças* à confiabilidade dos exames, o que em geral se faz é padronizar as condições de administração, os critérios de correção, que devem ser tão detalhados e claros quanto possível e livres de quaisquer ambiguidades. (FERNANDES, D., 2009, p. 135).

Entende-se, portanto, que a rigorosidade proposta no documento oficial do PAER está sendo percebida pelos gestores entrevistados e que esta tem proporcionado uma confiabilidade ao processo.

Uma última categoria que surge e que foi apontada por dois participantes das entrevistas relacionadas à aplicação das avaliações é a **formação técnica para os aplicadores e coordenadoras das unidades escolares**. Sendo que uma das regras é a não participação dos docentes das turmas ou das disciplinas avaliadas, é necessário que outros professores ou funcionários recebam as orientações fundamentais para que o rigor ensejado e descrito há pouco possa ser mantido em todas as edições, dando uniformidade ao processo. Desse modo, essa formação técnica se mostra como uma garantia também de confiabilidade, como demonstram as respostas:

PLG - Os professores que vão aplicar, os educadores que vão aplicar, eles passam por uma formação para saber que existe todo um protocolo, e esse protocolo, ele precisa ser seguido à risca para que a gente possa ter os dados que são recebidos com muita confiabilidade.

PGR1 – [...] as informações que chegam aqui, isso é conversado pessoalmente com cada coordenação de sub-região, os coordenadores fazem formação com os coordenadores de escola, e os coordenadores de escola, às vezes sozinhos, mas muitas vezes, juntos com os coordenadores das sub-regiões, eles chegam até o professor na orientação de tudo que vai acontecer.

Guba e Lincoln (2011), ao apresentarem os preparativos para a condução de uma avaliação, estabelecem como fator primordial para evitar problemas na atividade avaliativa, a seleção e o treinamento de uma equipe de avaliadores. Os autores consideram que existem vantagens em formar e usar uma equipe, mesmo que a avaliação em questão seja pequena e possa ser realizada por uma só pessoa. Da mesma forma, Worthen, Sanders e Fitzpatrick (2004), ao indicarem princípios orientadores para os avaliadores internos, apresentam a necessidade de formação e instrução para a realização das tarefas propostas na avaliação. Assim, essas iniciativas de formação para os responsáveis pela aplicação das provas e

questionários sociodemográficos são necessárias e possibilitam a redução da subjetividade inerente à “utilização de instrumentos humanos.” (GUBA, LINCOLN, 2011, p. 218).

Ainda no foco do processo e desenvolvimento do PAER, a questão seis da entrevista tratou da articulação da avaliação externa com outras avaliações (internas ou externas). No Quadro 22, podemos visualizar as categorias de análise que emergiram das respostas.

**Quadro 22 - Categorias de análise do Eixo Articulação**

Participantes	Comparação com outras Avaliações Externas	Comparação entre escolas	Comparação com a avaliação da aprendizagem
<b>PLG</b>	<b>X</b>		<b>X</b>
<b>PGR1</b>	<b>X</b>	<b>X</b>	
<b>PGR2</b>	<b>X</b>		
<b>PGR3</b>			
<b>PGR4</b>			<b>X</b>

Elaboração da autora.

Pelas respostas dos participantes, verifica-se que há um esforço de comparação dos resultados, porém não de forma sistematizada ou coordenada para toda a rede. Dois participantes responderam que desenvolveram uma planilha, na qual registram uma comparação dos resultados do PAER com as notas do ENEM, enquanto estas eram divulgadas por unidade escolar. Porém, com exceção do PGR3 que não conseguiu responder à pergunta, os demais consideram que há comparação dos resultados com outras avaliações externas e internas.

Para o PLG, as comparações são possíveis porque o PAER usa os mesmos parâmetros de outras avaliações externas realizadas no país. Na sua visão, a comparação não visa ao ranqueamento:

PLG - Não para ranquear, porque nós temos trabalhado muito para evitar o *ranking*, porque o mais importante não é saber se eu estou melhor ou estou pior. Mas é saber, é poder nos posicionar em relação à avaliação externa, à própria avaliação interna; nós evitamos competição entre uma escola nossa com outra.

Essa preocupação em não haver uso dos resultados para *ranking* é percebida em outras falas dos entrevistados. Mas há uma comparação entre escolas com vistas a compartilhar

experiências exitosas daquelas que conseguem obter bons resultados e mantê-los. Assim, o PLG considera:

PLG - O que nós, sim, fazemos é: se a escola tem um resultado melhor, nós vamos procurar saber por que aquela escola está obtendo melhores resultados? E, a partir deste conhecimento, a gente transfere o *know-how* para as demais escolas. O objetivo é o crescimento de todos, e não a competição.

Na comparação dos resultados com outras avaliações externas, a mais citada é o Enem, sendo que o PGR1 e PGR2 mencionam também os resultados PISA, como indicadores que utilizam também para estabelecer essas comparações.

A **comparação dos resultados da avaliação externa com a avaliação interna** ou **avaliação da aprendizagem** foi apresentada pelos entrevistados PLG e PGR4. Fica expresso na resposta da PGR4 que essa comparação ou “confronto” não é uma atividade sistemática de todas as unidades, após a divulgação dos resultados do PAER:

PGR4- Sim. As coordenadoras, elas costumam confrontar resultados [...] confrontar até com as notas do aluno. E as coordenadoras das sub-regiões, tem algumas que fazem o confronto direto destes resultados para ver se é um resultado solto, é só do PAER, ou se ele também aparece – um resultado positivo ou um resultado com dificuldades – nas outras avaliações também.

Essa articulação entre as avaliações externas e as avaliações internas, seja da escola, seja dos alunos, é desejável, na visão de Domingos Fernandes (2009). Para o autor, este é um desafio a ser vencido, e sua prática trará reflexos positivos no processo de ensino e aprendizagem:

Em contrapartida, no que diz respeito à avaliação externa, as escolas devem prever mecanismos de análise e discussão de seus resultados e de resultados dos seus alunos para que, a partir daí, possam retirar as devidas consequências para a eventual reformulação de suas *políticas* de ensino e de avaliação. As escolas não devem ficar indiferentes aos resultados das avaliações externas, quaisquer que eles sejam, pois, de certo modo, é uma forma de se verem no espelho e decidirem se têm que fazer alguma coisa acerca da imagem que é refletida. Além disso, é importante que comparem o currículo que é avaliado pela avaliação externa com o que é avaliado por suas avaliações internas, que analisem os efeitos das avaliações externas nas *políticas* da escola e, sobretudo, nos professores e no seu ensino, nos alunos e nas suas aprendizagens. (FERNANDES, D., 2009, p. 157)

O autor considera que é possível, por meio da articulação entre a avaliação externa e a avaliação interna (neste caso, da aprendizagem), se concretizar a melhoria das aprendizagens. Porém, afirma que, lamentavelmente, essa visão está longe de fazer parte da cultura dos sistemas educacionais em vários países. “Na verdade, a coleta e os usos que se fazem das informações avaliativas são, em geral, insatisfatórias, quer se trate de avaliações internas, quer de avaliações externas.” (FERNANDES, D., 2009, p. 164).

Freitas *et al.* (2017) também consideram a necessidade da articulação entre a avaliação em larga escala e a avaliação de sala de aula. Para os autores, contudo, no Brasil, “não é raro encontrarmos a tendência de se querer fazer da avaliação de sistemas um instrumento de avaliação da sala de aula e da escola.” (FREITAS *et al.*, 2017, p. 65).

Percebe-se, portanto, que na rede há iniciativas de articulação entre o PAER e outras avaliações, mas não são realizadas dentro de uma política de uso dos resultados. Não há uma ação coordenada nem orientações das mantenedoras regionais para que esta seja uma prática contínua ou mesmo necessária. Esta articulação vai muito além da comparação de resultados entre ambas. Envolve a percepção dos domínios cognitivos que são evocados nas avaliações externas cuja ênfase está na resolução de situações problemas, leitura e interpretação (RABELO; MARINHO-ARAÚJO, 2015). Está na ação conjunta de todas as áreas do conhecimento contribuindo para o desenvolvimento destes domínios ou operações mentais, sem responsabilizar apenas Língua Portuguesa e Matemática por tal tarefa. Esta articulação oportunizaria o desenvolvimento de estratégias que redundassem no desenvolvimento de habilidades cognitivas que garantissem o êxito em qualquer tipo de avaliação.

### 6.2.3 – 3º Foco: Resultados

Neste foco, foram agrupadas quatro questões cujos eixos estão relacionados aos resultados obtidos pelo PAER. Essas questões estão relacionadas ao objetivo geral do programa, ao uso dos resultados, às implicações e aos aspectos desenvolvidos e aspectos a desenvolver percebidos pelos participantes. Nestes eixos, o cerne das questões remete à capacidade do PAER de alcançar aquilo que se propôs alcançar e os resultados obtidos.

O primeiro eixo, sobre o objetivo geral do programa e as categorias de análise que as respostas à terceira pergunta indicaram é descrito no Quadro 23:

**Quadro 23 - Categorias de análise do Eixo Objetivo**

Questão 3 - No documento de apresentação do programa, o objetivo principal apresentado é “Retratar a realidade da qualidade da educação, conforme os princípios que fundamentam a filosofia de educação da rede, através de informações que incentivarão a adoção de medidas que promovam o crescimento da rede”. Na sua visão, este objetivo foi atingido? Se não, por quê? Se sim, por quê?						
Categorias de análise						
Participantes	Objetivo atingido	Impulsiona ações em prol da qualidade	Objetivo parcialmente atingido	Crescimento independente da avaliação	Desconhecimento sobre avaliação externa	Rotatividade dos profissionais
PLG	X	X				

<b>PGR1</b>			X		X	X
<b>PGR2</b>			X	X		
<b>PGR3</b>			X			
<b>PGR4</b>	X	X				

Elaboração da autora

Pelo Quadro 23 é possível identificar que dois participantes consideraram que o objetivo geral do PAER foi atingido e ambos justificaram essa ideia com a percepção de que a avaliação externa impulsiona ações em prol da qualidade e do crescimento. Três outros participantes consideraram que o objetivo geral foi parcialmente atingido, e a justificativa do participante PGR1 para essa percepção foi que ainda há um desconhecimento por parte da comunidade educacional sobre a avaliação externa, fator este, que é agravado pela rotatividade de profissionais na rede. Já para o PGR2, o crescimento da rede independe da avaliação externa, como está descrito no texto do objetivo geral.

Um dado a se levar em conta é que nenhum participante que considerou o objetivo como atingido se ateu às primeiras palavras do objetivo geral – “retratar a realidade” da qualidade da educação –, mas a ideia prevalente foi do programa como um “propulsor da qualidade”. Talvez porque essa ideia está associada ao objetivo mais divulgado em relação à avaliação em larga escala, desde a Lei de Diretrizes e Bases da Educação (BRASIL, 1996) e todas as demais normas que regulamentaram o sistema de avaliação da educação básica nacional, indicando o compromisso com a melhoria da qualidade do ensino e melhoria da qualidade da educação.

Para o PLG, que considera que o objetivo foi atingido, sua justificativa se expressa da seguinte forma:

PLG - Então, este processo tem nos ajudado a crescer, não só pelo processo em si, mas o que ele fomenta. Porque a partir do projeto, a partir da avaliação, nós continuamos com reuniões de formação, de adequação ao planejamento, e tudo isso envolve preparação do nosso docente, preparação dos nossos líderes educacionais e, conseqüentemente, um maior acompanhamento daquilo que nós nos propomos a fazer, que é oferecer uma educação de excelência.

Para o PGR4, o objetivo está sendo atingido porque as ações desencadeadas após a divulgação dos resultados contribuem para fomentar a qualidade educacional, sendo assim descritas:

PGR4 - Eu acho que o resultado retrata. E na maioria dos casos, ele impulsiona a coordenação da sub-região a fazer um acompanhamento mais de perto e a fazer um estudo com as escolas, com as coordenadoras das unidades, e chega lá na mão do professor, esse resultado. Então, é onde eles vão analisar os descritores, promover ações – não assim, só para o momento.

Na entrevista, não foi possível identificar o conceito de qualidade educacional subjacente às respostas que estabelecem uma relação entre a avaliação externa e a melhoria da qualidade educacional percebida. Porém, nas ações listadas por ambos os participantes para justificar essa percepção, em nenhum momento o desempenho dos alunos nas provas, ou melhoria da aprendizagem foram mencionados como indicadores de melhoria de qualidade. Ambos consideraram o acompanhamento do processo, a formação, o estudo e o preparo dos professores e gestores, como ações fomentadas a partir dos resultados do PAER. Entende-se que essas ações podem contribuir para a melhoria da aprendizagem, com potencial de refletirem sobre o processo pedagógico. Porém, a preocupação com os desdobramentos dessas iniciativas que afetam o aluno e a aprendizagem não foi expressa.

Para os demais participantes, o objetivo foi parcialmente atingido. O PGR1 expressou em sua resposta dois motivos para não alcance do objetivo geral:

PGR1 - a falta de conhecimento sobre o que significa uma avaliação externa faz com que eu não consiga aplicar os próprios princípios institucionais. [...] Ao mesmo tempo que nós estamos sempre formando os profissionais, nós temos aqui na nossa região, uma rotatividade muito grande. Eu penso que a rotatividade de profissionais é o maior dos motivos pelos quais nós não alcançamos ainda este objetivo.

O PGR2 indicou a seguinte justificativa:

PGR2 - Acredito que parcialmente. A gente está falando de crescimento. Então, é uma avaliação que retrata fatos, tira uma fotografia de onde está e ela então, por si só, ela não vai prever crescimento. [...] A PAER nos diz quem você é, que nota você está, qual é a proficiência do seu aluno. Então, ao olhar para dentro deste boletim eu tenho uma dimensão enorme. Mas crescer não depende da avaliação.

Para o PGR2, a avaliação retrata fatos, e o resultado indica a proficiência dos alunos. Porém, na sua percepção, isso não interfere, necessariamente, no crescimento.

Identificam-se, nas respostas, dimensões distintas da avaliação. Uma dimensão na qual os resultados da avaliação impulsionam ações e propostas que contribuem para a melhoria da qualidade educacional; e a outra que considera a avaliação como verificação do desempenho e da proficiência dos alunos nas provas ou, como afirmou o PGR2; “desvela quem nós somos”, mas, não implica mudanças, necessariamente.

Para Souza e Oliveira (2010), essas duas dimensões não são excludentes, mas precisam se complementar. Consideram que:

O sistema de avaliação tende a se limitar a um sistema de informação educacional. Sem dúvida, a produção de informação é uma etapa do processo avaliativo, mas esse só se realiza quando, a partir de informações, ocorrerem julgamento, decisão e ação. Ou seja, a avaliação ganha sentido quando subsidia intervenções que levem à transformação e à democratização da educação, em suas dimensões de acesso, permanência e qualidade. Não pode estar essencialmente a serviço de apoiar a reiterada denúncia da baixa qualidade do ensino. (SOUZA; OLIVEIRA, 2010, p. 818).

Dias Sobrinho (2008), ao abordar os diversos aspectos e dimensões da avaliação, também considera que a dimensão da verificação, constatação e medidas por meio de notas, exames e testes, é necessária, mas insuficiente. Para o autor, os instrumentos e as práticas de medidas e constatações não bastam a si mesmas, e considera, ainda, que:

Imprescindível é que essas técnicas propiciem reflexão, enriqueçam seus significados com o recurso a análises qualitativas, levem a questionamentos, isto é, façam parte de um conjunto de atividades epistêmicas e valorativas que produzam sentidos complexos e mobilizem os sujeitos para a tomada de decisões e ações de melhoria. (DIAS SOBRINHO, 2008, p. 202)

Sobre a explicação do PGR2, comparando a avaliação externa com uma fotografia que retrata um momento, Dias Sobrinho (2008) utilizou essa mesma analogia para distinguir as dimensões da avaliação – estática e dinâmica –, quando analisa o desempenho do estudante como um fim em si mesmo, e a outra considera a historicidade da pessoa em formação:

Não se trata de fotografia ou medida de retenção de conteúdos num momento dado, mas sim de compreender as mudanças que vão ocorrendo ou os valores que vão se agregando ao longo do percurso. Esta mesma lógica serve para a avaliação de programas, cursos, instituições e assim por diante. Para além da fotografia estática, os movimentos. (DIAS SOBRINHO, 2008, p. 203).

Não é fato estranho que numa rede de extensão nacional e com uma estrutura regional tão abrangente, a diversidade de visões sobre a finalidade da avaliação externa e em larga escala seja encontrada entre os que ocupam funções estratégicas de gestão. Essa diversidade pode ser compreendida e até mesmo aceita como algo favorável, desde que haja espaço para discussões, análises de diferentes pontos de vista, disposição e abertura para uma construção coletiva dos princípios norteadores do programa, em que as dimensões não sejam excludentes, mas complementem-se. O que se observa, desde a forma como o programa foi criado, implementado e a percepção dos entrevistados sobre o uso dos resultados, descrita a seguir, é que não há este espaço para a discussão destas ideias em nenhum dos níveis de gestão da rede.

Quanto ao eixo **objetivo do PAER**, portanto, a maior parte dos entrevistados considera que o objetivo foi parcialmente alcançado, e esta categoria é revisitada após análise das respostas ao questionário *on-line*, sendo possível considerar que o objetivo não está claro para todos, tanto sua enunciação como as ideias de qualidade e crescimento que apresenta.

O segundo eixo derivado da pergunta cinco trabalha o uso dos resultados do PAER. O Quadro 24 apresenta as categorias de análise que emergiram das respostas.

#### **Quadro 24 - Categorias de análise do Eixo Uso dos resultados**

Questão 5 - Qual o uso que tem sido feito dos resultados da referida avaliação?						
Categorias de análise						
Participantes	Processamento de informações	Tabulação dos dados	Estudo dos dados e dos descritores	Socialização com professores	Formação de professores/educadores	Capacitação em elaboração de itens
PLG	X				X	
PGR1			X		X	
PGR2		X	X			
PGR3			X			
PGR4			X	X		X

Elaboração da autora.

Observa-se que, para quatro, dos cinco participantes das entrevistas, os resultados têm propiciado **estudos dos dados** gerados e estudos dos descritores da matriz de referência. Dois participantes identificam que a **formação de professores** e educadores também tem sido desenvolvida a partir dos resultados obtidos e analisados. Para dois participantes, o aspecto aparentemente técnico do **processamento de informações** e **tabulação dos dados** é identificado como uso dos resultados e, para um participante, duas categorias de análise foram identificadas: a possibilidade de **socialização dos resultados** com os professores e a **capacitação em elaboração de itens**.

Da possibilidade de estudo dos dados apontada pelos participantes, dois se referem ao estudo pelos professores, em conjunto com as coordenadoras das unidades escolares, e dois se referem ao estudo dos dados em nível de mantenedora das regiões e sub-regiões. Porém, a oportunidade de estudo ou análise dos dados é um indicativo de que o uso está voltado para uma finalidade pedagógica, podendo causar efeitos<sup>34</sup> na prática educativa por envolver professores, como apontado pelo PLG, ao afirmar que cursos são preparados para superação das deficiências indicadas pelos descritores com menor desempenho.

PLG - E aí, entra um outro departamento que temos chamado Universidade Cooperativa (que hoje está mudando de nome), que prepara formação para os nossos educadores de acordo com a necessidade levantada nestas aplicações dos instrumentos de avaliação.

Estudar os dados pressupõem uma finalidade. Pressupõe uma compreensão da realidade que permita uma ação sobre ela. Se os dados estão sendo “estudados” com esta finalidade, o uso dos resultados poderá promover a transformação ensejada, que produza

<sup>34</sup> Efeito no sentido de alterações, esperadas ou não, provocadas pelo programa, que não as previamente pretendidas e que vão além do grupo-alvo, conforme descrito no início do capítulo. (DRAIBE, 2001)

melhoria da qualidade educacional. Mas este “estudo” pode ser associado a uma análise que desencadeie ações cuja finalidade seja apenas trabalhar descritores cujo desempenho foi insatisfatório, como descrito pelo PLG. Contudo, cotejando estas respostas com os demais dados do questionário, e da própria entrevista, é possível inferir que este estudo dos dados visa a superação das deficiências indicadas pelos descritores com desempenho insatisfatório. Este é um dos motivos, mas não deve ser o único e nem uma finalidade.

Nota-se que somente dois participantes indicaram a categoria **formação de professores** como um dos usos dos resultados. No documento oficial do PAER, há uma indicação de que, uma vez completada a correção das provas e análise das respostas, os resultados são compartilhados com os gestores de várias maneiras e com a Universidade Corporativa (UC), “destinada à formação contínua de professores da rede. O planejamento dos cursos oferecidos pela UC incorpora necessidades detectadas pelo PAER” (D.O., 2017). Assim, a iniciativa da formação para professores originada pelos resultados da avaliação está prevista no documento oficial e mencionada pelo gestor da mantenedora geral para o Brasil, conforme resposta acima. Já a resposta do PGR1 refere-se à formação de professores por iniciativas locais de unidades escolares:

PGR1 – Então, eles traçaram um programa a partir dos estudos dos descritores que apresentavam uma dificuldade. Que as crianças apresentavam mais dificuldades. Depois que eles fizeram esse trabalho, então os professores do 6º ao 9º ano fizeram a formação do professor do 1º ao 5º para desenvolverem, então, um manejo de trabalhar e também conteúdo, repertório para trabalhar com os descritores faltantes.

Embora haja uma iniciativa oficial de formação de professores a partir dos resultados do PAER, gerenciada pela mantenedora central, o Quadro 24 das categorias em análise indica que os entrevistados não consideram essa iniciativa, ou não a percebem como um efeito dos resultados da avaliação externa. Apenas o criador do programa e da “Universidade Corporativa” é quem faz esta associação. Este dado pode indicar que, pode haver a intenção, mas a operacionalização não está concretizando esta intenção de formação dos professores para minimizar os pontos fracos indicados pelo PAER. Se os gestores entrevistados, que estão no primeiro nível de gestão, não tem esta percepção, qual a possibilidade dos professores considerarem esta associação entre a formação continuada e os resultados da avaliação externa?

Portanto, há uma percepção de que os resultados promovem o estudo dos dados gerados, nos níveis de gestão e de docentes, mas a formação desenvolvida em função desses dados não está sendo associada aos resultados do PAER, de acordo com as respostas analisadas.

As categorias **processamento de informações** e **tabulação de dados**, que são atividades, a princípio, de natureza técnica, no desenvolvimento das respostas dos participantes são associadas ao posterior uso pedagógico dos dados:

PLG – Então, nós temos o resultado escalonado, que já é oferecido para a rede, e também nós possibilitamos o cruzamento de informações destes dados levantados, ou estes instrumentos, e também outros que são de cunho socioeconômico etc. Então, nós possibilitamos, dentro de uma base de dados, que é uma *big data*, a utilização destas informações para os planejamentos em nível local da instituição de ensino; ou em nível regional; ou em nível de país; e também favorecemos a preparação de formações.

Para o PGR2, a tabulação de dados com vistas ao estudo, é descrita na resposta:

PGR2 - De dois anos para cá, nós criamos uma outra tabulação para pesquisar descritores fortes e descritores fracos. E aí, este ano, especificamente, nós fizemos um estudo grande sobre os nossos desafios e pegamos esse material e começamos a estudar, então, à luz do próprio ensino.

Em ambas as respostas, nota-se a finalidade pedagógica da ação de tabular e processar dados com vistas a “planejamentos em nível local da instituição de ensino”; “preparação de formações” e “estudar, então, à luz do próprio ensino”. Como já citado alhures, a dimensão métrica pode ter seu valor, desde que seu uso não seja fechado e isolado, sem gerar questionamentos e produção de sentido (DIAS SOBRINHO, 2008).

Para o PGR4, o uso dos resultados está relacionado à **socialização com professores** dos dados levantados e à **capacitação na elaboração de itens** também voltada aos professores. Assim descreve o processo:

PGR4 - Porque é assim: assim que ele chega, os professores já estão acostumados, eles querem saber logo o levantamento, porque daí, as coordenadoras fazem o levantamento por descritor. [...] Então, aqui na nossa região, nós fizemos capacitação de elaboração de itens com todos os professores. E nós percebemos que houve uma mudança no olhar do professor até na maneira de trabalhar os conteúdos. Porque alguns não estavam percebendo que a necessidade vinha bem antes da avaliação. Que era no dia a dia, ao trabalhar com cada componente curricular. Eu acho que eles cresceram bastante.

Seja uma iniciativa pontual em uma única região, seja uma iniciativa mesmo dos gestores em socializar os resultados com os professores – o que parece ser uma prática comum aos demais, pois as respostas que indicam oportunidade de estudo dos dados devem envolver esse compartilhamento de informações previamente –, entende-se que os professores são os principais interessados em ter acesso aos resultados. Se, de fato, enseja-se a melhoria da qualidade do ensino, os professores precisam estar envolvidos integralmente no processo de avaliação, e todas as ações desencadeadas a partir dos resultados necessitam da participação dos docentes, não apenas como executores, mas como propositores. Como afirma

Cappelletti (2015), quando discorre sobre os conflitos entre a avaliação e a qualidade da educação: “Só com a participação dos envolvidos as mudanças serão compatíveis com a identidade da escola, favorecendo a apropriação de saberes e a compreensão da realidade que se quer transformar.” (CAPPELLETTI, 2015, p. 103).

Apesar de a capacitação em elaboração de itens poder ser entendida como uma formação de professores em serviço, a especificidade do tema não deixa margens para alguma inferência sobre os aspectos conceituais ou um quadro teórico que referenciou os estudos sobre a avaliação em larga escala, ou mesmo a avaliação educacional como um todo. Não foi objeto da entrevista, nem houve alguma resposta que indicasse os conteúdos abordados nas iniciativas de formação. Porém, ainda que não tenhamos informações que indiquem o tipo de conteúdo abordado nessas iniciativas, ou mesmo o referencial adotado para o estudo dos dados pelos docentes e gestores, o fato de os resultados gerarem a oportunidade de estudo, reflexão e formação de professores já se constitui um forte indicativo de que a avaliação externa tem o potencial de produzir efeitos na prática pedagógica. Essas reflexões, se pautadas em fundamentos epistemológicos, modelos e pressupostos que orientem a operacionalização das práticas avaliativas, favorecem a mudança desejada com vistas à qualidade do ensino.

As questões sete e oito da entrevista podem ser consideradas como o cerne de uma meta-avaliação, como designou Stufflebeam (2001). Elas tratam dos ganhos (aspectos desenvolvidos) e das dificuldades (aspectos a desenvolver) na percepção dos entrevistados.

No Quadro 25, as principais categorias de análise da questão sete estão representadas.

**Quadro 25 - Categorias de análise do Eixo Ganhos**

Questão 7 - Na sua visão, quais os ganhos após 7 anos da referida avaliação?							
Categorias de análise							
Participantes	Melhoria da qualidade do ensino	Reflexão sobre a prática pedagógica	Desvelamento da realidade	Proposta de Formação de professores	Mudança de olhar sobre Av. Externa/ PAER	Entendimento do valor do PAER	Mudança da prática pedagógica
PLG	X						
PGR1		X					
PGR2			X	X	X		
PGR3						X	
PGR4					X	X	X

Elaboração da autora.

Há sete categorias de análise, sendo que apenas duas delas foram indicadas por dois participantes, simultaneamente: **mudança de olhar sobre a avaliação externa ou sobre o PAER**, e **entendimento do valor do PAER**. As demais foram indicadas apenas uma vez.

Há um dado interessante de ser considerado no Quadro 25. Se o objetivo geral do PAER está relacionado à **melhoria da qualidade de ensino**, apenas o PLG apresenta essa categoria como **ganho**. Vale ressaltar que esse participante também foi um dos dois que considerou o objetivo alcançado e que ele está na condição de liderança mais elevada na estrutura da rede. Não podemos afirmar que os demais não consideram que houve ganho de qualidade no ensino da rede. Porém, eles não expressaram em suas respostas essa percepção. Por outro lado, pode-se considerar que todas as categorias apresentadas nas respostas dos demais podem contribuir para a melhoria da qualidade do ensino. Este dado nos remete novamente à questão: qual a concepção de qualidade que norteia o programa? As mudanças na prática pedagógica são, necessariamente, melhoria da qualidade do ensino?

Ressalta-se, nesse contexto, a posição de alguns autores que consideram duvidoso afirmar que a avaliação externa contribui para a melhoria da qualidade do ensino. Zákia Souza e Oliveira (2010), ao analisarem sistemas de avaliação em larga escala de cinco unidades federativas, buscando explorar como os resultados produzidos contribuíram para a formulação e implementação de políticas educacionais, concluíram que, nos casos estudados, “a relação entre avaliação e melhoria da qualidade do ensino não constitui ainda uma realidade.” (SOUZA; OLIVEIRA, 2010, p. 817). Para os autores, “a ideia de qualidade que vem sendo forjada tem se restringido à apreciação do desempenho do aluno, sem que este seja interpretado à luz de condições contextuais, intra e extraescolares.” (p. 818).

Em uma análise da literatura sobre o tema, Domingos Fernandes (2009) considera que se trata de matéria polêmica para muitos pesquisadores e educadores, a expectativa que se coloca sobre a avaliação externa no papel de contribuir para a melhoria da qualidade do ensino das escolas.

Para Freitas *et al.* (2017):

A avaliação em larga escala, do tipo Saeb, é um instrumento de acompanhamento global de redes de ensino com o objetivo de traçar séries históricas do desempenho dos sistemas, que permitam verificar tendências ao longo do tempo, com a finalidade de reorientar políticas públicas. Quando conduzidas com metodologia adequada, podem trazer importantes informações sobre o desempenho dos alunos, dados sobre os professores, condições de trabalho e funcionamento das escolas de uma rede. (FREITAS *et al.*, 2017, p. 47).

Evidencia-se, nesse parágrafo, que a avaliação em larga escala, quando bem conduzida, pode trazer importantes informações e sua finalidade é reorientar políticas públicas. Essa é uma das finalidades previstas no inciso III do Artigo 9º da Portaria MEC nº 458, de 5 de maio de 2020: “a disponibilidade de informações relevantes para subsidiar a elaboração, o monitoramento e o aprimoramento de políticas públicas;” (BRASIL, 2020). O que se pode inferir dessa finalidade, associada ao parágrafo de Freitas *et al.* (2017), é que as políticas públicas aprimoradas seriam as responsáveis pela qualidade educacional, e não a avaliação externa de *per si*.

Na perspectiva de a avaliação em larga escala trazer importantes informações sobre o desempenho dos alunos e ter a finalidade de reorientar políticas públicas, Tavares Jr. e Neubert (2014), numa pesquisa em que analisaram a associação entre a implantação de sistemas estaduais de avaliação educacional no Brasil e o desenvolvimento da proficiência média entre os alunos das redes estaduais que criaram sistemas próprios de avaliação estáveis e consolidados – Ceará, Minas Gerais e São Paulo –, com uma ampla análise estatística, observaram crescimento nas curvas de desempenho médio dessas redes estaduais de ensino, com tendências gerais de maiores avanços nos anos iniciais do ensino fundamental e apresentação de resultados melhores nas edições do Saeb posteriores a 2005, concluem:

A recuperação da qualidade da educação na primeira década deste século, decorrente do avanço do desempenho médio das redes públicas, esteve associada à adoção de políticas educacionais e reformas que sempre tiveram nas avaliações em larga escala um elemento central. (TAVARES JR; NEUBERT, 2014, p. 43).

A associação da ideia entre evolução do desempenho médio dos alunos com a melhoria da qualidade da educação é retratada no ideário de educadores, assim como na sociedade em geral. Horta Neto (2014), em um estudo sobre a discussão da qualidade educacional na mídia eletrônica escrita, por meio de uma coleta de matérias disponibilizadas na internet no período de janeiro de 2009 a outubro de 2012, observou que dos 2.100 *links* analisados, nenhum apresentava críticas aos testes avaliativos no Brasil e que “tem sido veiculada uma forte mensagem de que aumentar o número de testes é um bom caminho para aprimorar as aprendizagens.” (HORTA NETO, 2014, p. 199). Desse modo, compreende-se que pode ser reforçada pelas mídias eletrônicas essa associação entre desempenho dos alunos e qualidade da educação. Evidencia-se, portanto, que a ideia de qualidade está relacionada ao desempenho dos alunos nas avaliações externas.

Como constatado, há um ideário que associa o desempenho dos alunos nos testes avaliativos à qualidade educacional. Esta informação acentua a importância do fato de que,

entre os entrevistados, apenas o PLG considerou que houve ganho na qualidade do ensino por meio da avaliação externa. Qual seria a razão pela qual os demais entrevistados não percebem, ou não indicam, que houve melhoria da qualidade? Isto implicaria na visão desta relação entre avaliação e qualidade educacional? São questões que se reproduzem em outros sistemas de ensino?

De acordo com Rios (2021), avaliação educacional e qualidade da educação e do ensino estão intrinsecamente relacionadas. A autora, contudo, questiona: “qual avaliação? Qual qualidade? Esses questionamentos emergem das concepções que coexistem acerca da avaliação e da polissemia do termo qualidade.” (RIOS, 2021, p. 123). Portanto, esse dissenso entre os autores sobre a relação entre a avaliação externa e a melhoria da qualidade educacional pode ser fruto da matriz sobre a qual se assentam as ideias e concepções sobre a temática.

A avaliação pautada numa perspectiva transformadora e emancipatória considerará o uso dos dados e resultados de uma avaliação externa como um importante instrumento propulsor de mudanças. Porém, serão as políticas desencadeadas a partir desses dados – numa lógica não classificatória e excludente; buscando o diálogo entre os interessados: avaliados e avaliador; compreendendo que a qualidade educacional vai muito além do desempenho em testes cognitivos; atuando sobre fatores intra e extraescolares – que poderão garantir a qualidade almejada.

A análise da mesma pergunta feita no questionário *on-line* para os demais gestores em posições sub-regionais mais próximas das unidades escolares pode desvelar, com mais clareza, uma percepção coletiva com respeito a esse tema por uma representatividade mais significativa de gestores. De fato, infere-se que não há, na rede pesquisada, uma compreensão ou um consenso sobre a concepção de qualidade ou sua relação com a avaliação externa.

A categoria **reflexão sobre a prática pedagógica**, apresentada pelo PGRI, indica a compreensão de uma das principais finalidades da avaliação educacional na perspectiva emancipatória. Para Freitas *et al.* (2017), qualquer dos três níveis avaliativos – avaliação da aprendizagem, avaliação institucional ou avaliação externa em larga escala – trata-se de um processo de reflexão coletivo. Consideram que: “Há que se medir, mas esta não é a parte mais importante; há que se avaliar – esta sim é fundamental. Avaliar é promover no coletivo a permanente reflexão sobre os processos, seus resultados, em função de objetivos a serem superados” (FREITAS *et al.*, 2017, p. 78). Desse modo, podemos depreender da resposta do participante a compreensão da relevância da reflexão como um dos efeitos da avaliação

externa. PGR1 – “Acho que reflexão, essa é a palavra. É o maior ganho que a gente tem, é a reflexão sobre as nossas ações mais pedagógicas”.

Dias Sobrinho (2003) distingue dois tipos de efeitos das avaliações realizadas no campo educacional: políticos e pedagógicos ou educativos. Sobre os efeitos pedagógicos – que para o autor, também têm significação política – os mais destacados aos quais se refere são: “as dimensões mais específicas do currículo, da produção e reprodução do conhecimento, da organização do processo ensino-aprendizagem, das relações interpessoais etc.” (DIAS SOBRINHO, 2003, p. 95). Conclui que, no caso dos efeitos pedagógicos, a avaliação tem por finalidade a melhora das funções propriamente educativas. Para que esses efeitos possam concorrer para a melhora das funções educativas de uma instituição de ensino, com a potência de iluminar e instrumentalizar as reformas educacionais, a reflexão sobre a prática pedagógica à luz dos dados produzidos pela avaliação é indispensável, pois ela possibilita a construção de sentidos amparados no contexto, valorizando a experiência dos envolvidos na prática pedagógica, com uma intencionalidade educativa e procedimentos fundados em princípios democráticos.

Portanto, a reflexão sobre a prática pedagógica visa à **mudança na prática pedagógica**, categoria que emergiu da resposta do PGR4. Porém, a mudança na prática pedagógica pode ser consequência de uma avaliação externa tanto numa perspectiva objetivista, pautada no paradigma tradicional, convencional ou racionalista, como na perspectiva subjetivista, pautada no paradigma emergente, de caráter formativo e emancipatório. Na primeira possibilidade, estudos têm demonstrado uma redução dos currículos com ênfase na matriz de referência das avaliações externas; a seleção de conteúdos visando preparar e treinar os alunos para as provas em virtude da necessidade do cumprimento de metas; e forte pressão sobre os professores (CAPPELLETTI, 2015). Nesse modelo, a avaliação “induz mudanças nos programas pedagógicos, transformando-os em programas de exames.” (DIAS SOBRINHO, 2003, p. 174). Na segunda possibilidade, as mudanças ocorrem, não sem tensões e desafios, pois ao envolver o coletivo, diferentes concepções de avaliação e educação podem emergir e, num movimento participativo, a realidade complexa da educação é evidenciada. Porém, as mudanças produzidas por esse processo redundam em ganho qualitativo, quando, por exemplo, o currículo é considerado por seu valor de formação, e o conhecimento ultrapassa a dimensão técnica (DIAS SOBRINHO, 2003).

Um levantamento bibliográfico desenvolvido por Bauer (2016) apontou trabalhos acadêmicos que procuram compreender as interações entre avaliação externa, currículo e a prática pedagógica. A partir da implementação da avaliação censitária e dos dados produzidos pela Prova Brasil, a autora identificou duas linhas de conclusões dos estudos que discutem essa interação: um grupo de autores que levantam críticas e apontam uma simplificação dos saberes oferecidos aos alunos, a indução curricular, um ataque à autonomia dos professores e direcionamento das práticas pedagógicas, a partir das avaliações externas; e outro grupo que assinala aspectos positivos nessa relação, reportando à reflexão didática, à reorganização do trabalho pedagógico, ressaltando as potencialidades da avaliação externa para a organização do trabalho educativo e para a garantia do acesso aos alunos a determinados conteúdos. Portanto, a autora conclui que as pesquisas identificam que tem havido uma forte influência das avaliações externas no currículo escolar e nas práticas pedagógicas.

Sobre esses dois polos ou perspectivas que atendem às duas concepções distintas de avaliação, mas que ambos servem ao propósito de produzir mudanças na prática pedagógica, Dias Sobrinho (2003) conclui:

É perfeitamente admissível que ambas as intencionalidades podem levar a ações e transformações. Entretanto, também essas transformações são de natureza diferentes e podem estar servindo a interesses muito distintos. O controle tende a levar a mais do mesmo: conformação, ajuste, repetição, modelagem, heteronomia. A avaliação educativa, ao contrário, sem deixar de ser objetiva, de controlar os processos científicos, pedagógicos e administrativos, tende a produzir os debates, a reconhecer a diversidade de ideias, a interpretar a pluralidade, a construir novos sentidos, a questionar a razão dos projetos e currículos, a valorar a inserção crítica e produtiva na sociedade, a dinamizar a construção da autonomia. (DIAS SOBRINHO, 2003, p. 189)

As mudanças na prática pedagógica, portanto, podem ocorrer em ambas as concepções subjacentes à avaliação. Na resposta do PGR4, que apontou esta categoria de análise, podemos ver indicativos de uma concepção de avaliação num caráter formativo e participativo, com ênfase na figura do professor como protagonista no processo:

PGR4 - O primeiro ganho que eu vejo é o amadurecimento das pessoas envolvidas, para enxergar esse resultado como uma mola para a gente crescer [...] para o professor, ele também passou pelo amadurecimento e ele hoje, já consegue ver os resultados como indicadores de conteúdo, de descritores, de novas possibilidades que ele precisa trabalhar com o aluno, para que o estudante, de fato, aprenda.

Esse olhar direcionado para o professor se revela também nas respostas de dois outros participantes que apontaram uma **mudança de olhar sobre a avaliação externa/ PAER** e um **entendimento do valor do PAER** por parte dos professores. PGR3: “Então, eu percebo que o grupo começou a amadurecer nos últimos três anos. Dos últimos três anos para cá. Então,

considerando este último triênio, o ganho é o entendimento do valor que esta prova tem”. Ou, como respondeu o outro participante:

PGR2: Hoje, você vê professores discutindo sobre a elaboração de itens, conversando sobre matriz de referência, e isso não era uma realidade antes disso. [...] Então, uma compreensão de modo geral do que esta avaliação simboliza, e redimensionar este trabalho com as iniciativas públicas da educação, acho que isso é algo bastante significativo.

Essa compreensão de que um programa de avaliação externa pode agregar valor à prática docente e que, bem direcionado, pode trazer desenvolvimento profissional é apresentada por Dias Sobrinho (2003) como uma das possibilidades desse modelo de avaliação. Sobre a avaliação externa, o autor considera que:

Tomar distância e refletir sobre as próprias atividades é uma prática necessária ao desenvolvimento profissional. Este processo de “descentração” permite que o sujeito observe suas ações com critério de exterioridade, como se fosse um olhar de alteridade ou sendo realmente um olhar de um outro. É importante assegurar que a avaliação externa tenha a mesma intencionalidade formativa e não-controladora e seja sensível à multidimensionalidade da realidade humana e à polissemia da educação. (DIAS SOBRINHO, 2003, p. 173)

Se os professores conseguem identificar na avaliação externa essa possibilidade e a ela atribuírem valor, esse é um indicativo que o caráter formativo e não-controlador esteja sendo alcançado pelo programa investigado. Ainda que os participantes da investigação estejam em níveis distintos da gestão da rede e que não se tenha levantado a percepção de professores por meio dos instrumentos utilizados, o que se pode inferir é que, na percepção dos gestores, os professores estão mudando o olhar para a avaliação externa, ao longo das edições, e entendendo o valor desse modelo de avaliação.

Ao indicarem como um dos ganhos do programa as iniciativas de **formação de professores** – PGR2: “Agora, claro, em termos de rede, os avanços foram muitos. A partir desses dados, propostas de formação de professores puderam ser iniciadas” –, pode-se considerar que há uma preocupação com a melhoria da qualidade do ensino e com a dimensão pedagógica, a partir dos resultados da avaliação externa. Como foi mencionado por mais de um participante da entrevista, há uma instância de formação na rede denominada “Universidade Corporativa”, que desenvolve cursos a partir das fragilidades apontadas pelos resultados do PAER. Porém, até mesmo a formação docente pode ser pautada em duas concepções de avaliação – e conseqüentemente – de educação. O professor “como instrutor e transmissor de um pacote pedagógico alheio, ou, por outro lado, como um profissional que reflete crítica e criativamente sobre sua prática e os efeitos de suas atividades.” (DIAS SOBRINHO, 2003, p. 172).

Portanto, a formação de professores, em si, não é um dado que pode ser considerado isoladamente. Ela pode ser identificada como uma valorização da titularidade docente no processo ou como um mecanismo de reprodução de um pacote pedagógico alheio, por intermédio do qual apenas se reforça uma expectativa de desempenho.

A oitava pergunta de entrevista busca levantar as dificuldades ou aspectos a desenvolver no PAER, a partir da perspectiva dos gestores das regiões do país.

O Quadro 26 apresenta as categorias de análise que emergiram das respostas.

**Quadro 26 - Categorias de análise do Eixo Dificuldades**

Questão 8 - Quais as dificuldades ou aspectos a desenvolver relativos ao referido programa de avaliação da rede?				
Categorias de análise				
Participantes	Desenvolvimento de uma cultura de avaliação	Aprimoramento da tecnologia	Entendimento do valor do PAER	Formação para o uso dos resultados
PLG				
PGR1	X	X	X	
PGR2				X
PGR3			X	
PGR4	X			

Elaboração: autora.

Para os participantes da entrevista, quatro aspectos foram indicados como dificuldades ou pontos a desenvolver no PAER: **desenvolvimento de uma cultura de avaliação; aprimoramento da tecnologia; entendimento do valor do PAER; e formação para uso dos resultados.** Um destaque no Quadro 26 pode ser dado ao fato do PLG não ter indicado nenhum aspecto que considere necessário desenvolver ou que represente alguma dificuldade. Questiona-se se isso se deve ao fato de ser o gestor responsável pela criação do PAER? As possíveis dificuldades indicariam algum equívoco na implementação ou acompanhamento do programa sob sua responsabilidade? Se levarmos em conta que ele considera que o objetivo está sendo alcançado, pode-se também interpretar que, de fato, para o entrevistado, não há dificuldades a serem superadas no programa, pois o objetivo proposto foi alcançado. O fato é que não apresentou nenhuma dificuldade a ser superada. Outro destaque pode ser feito ao fato de que uma das categorias que foi apontada como ganhos pelos PGR3 e PGR4, no Quadro 25,

é apresentada como dificuldade pelo PGR2, no Quadro 26, que aponta as dificuldades do PAER.

Dois participantes indicaram como uma das dificuldades o desenvolvimento de uma cultura de avaliação. Para o PGR1, ainda não há uma cultura de avaliação estabelecida em todos os níveis da rede, inclusive, no de gestores: PGR1: “Eu acho que, como cultura, ainda não está bem estabelecida a nossa avaliação externa. O que existe é um esforço para o grupo que participa. Quer seja aqui, entre nós mesmos, na mantenedora geral, ou nos campos (sub-regiões) ou na escola.”

Para o PGR4, essa cultura não está desenvolvida entre os alunos. Isso impede uma participação com compromisso de fazer seu melhor. PGR4: “Então, eu acho que a gente ainda precisa [...] achar uma maneira do aluno se envolver de modo responsável, para a gente ter o feedback real, para ver onde... [...] o aluno só se envolve se tem nota.”

A percepção dos participantes sobre a necessidade do desenvolvimento de uma cultura avaliativa é um dado relevante a ser analisado e está relacionado às concepções apresentadas no capítulo cinco. Defendemos a necessidade de criação de uma cultura avaliativa apoiados nas ideias de Luckesi (2005) que apresenta duas culturas distintas associadas às duas concepções de avaliação educacional existentes: a primeira, a cultura do exame, que herdamos do modo “examinatório de agir na prática escolar” (p.3), desde as pedagogias jesuíta e comeniana, discutidas. Esta cultura “configura os exames como um modo adequado e satisfatório, pedagógica e disciplinarmente, de controlar a aprendizagem dos educandos” (p.3). Os exames selecionam, marginalizam, classificam, excluem e reproduzem a lógica do paradigma dominante. Para o autor, esta cultura é uma herança poderosa e consistente. Confundimos examinar com avaliar e acreditamos que estamos avaliando, quando estamos apenas reproduzindo a cultura do exame. A segunda cultura, a avaliativa, se contrapõe à cultura do exame, “exige uma posição democrática do sistema de ensino e do professor” (p.2), tem o foco na aprendizagem, olha para o ser humano como um ser em desenvolvimento, em construção permanente, é inclusiva e construtiva. Portanto, a instituição da cultura avaliativa se dá mediante um estado de consciência que se abre para a acolhida dos resultados, numa perspectiva qualificadora, que subsidie a compreensão dos problemas identificados e a tomada de decisões que possibilitem o desenvolvimento constante da realidade escolar (LUCKESI, 2005).

A formação de uma cultura de avaliação é um dos objetivos do Saeb replicado em sistemas de avaliação em larga escala desenvolvidos em estados e municípios. Desde a

institucionalização do sistema em 1994 aos dias atuais, tem-se mantido o objetivo, como consta na Portaria MEC nº 458, de 5 de maio de 2020, no artigo 6º:

I - construir uma cultura avaliativa, ao oferecer à sociedade, de forma transparente, informações sobre o processo de ensino-aprendizagem em cada escola, comparáveis em nível nacional, anualmente e com resultados em tempo hábil, para permitir intervenções pedagógicas de professores e demais integrantes da comunidade escolar; (BRASIL, 2020).

Ao analisar a criação de sistemas de avaliação em larga escala nos estados e municípios da federação, Souza e Oliveira (2010) observaram que é usual que esse objetivo seja tido como referência na criação de tais sistemas. Porém, os autores consideram que esse conceito não é explicitado ou discutido de modo a permitir uma apreensão clara de seu significado. O crescimento do número de sistemas de ensino que implementaram suas próprias avaliações externas em larga escala denota a importância que os processos avaliativos têm representado, principalmente, para os gestores desses sistemas. Seria esse crescimento de sistemas de avaliação externa um indicativo de uma cultura avaliativa? Os autores concluem que isso não configura plenamente a construção de uma cultura de avaliação, embora a literatura proponente indique uma sedimentação da importância dos sistemas de avaliação.

Já nas primeiras edições do Saeb, ao escrever sobre as avaliações em larga escala em andamento no país, Vianna (2003) constrói um panorama da cultura avaliativa e considera a necessidade da criação de uma nova cultura de avaliação na qual, além da divulgação dos resultados, fossem definidas diretrizes de como usá-los, a fim de que resultassem em melhorias no processo de educação. O autor considera que:

A expressão “cultura da avaliação” integra, atualmente, a constelação de palavras técnicas no âmbito da comunidade educacional e aos poucos se vai tornando verdadeiro lugar comum, quase que simples figura de retórica; no entanto, é preciso que essa expressão se liberte do seu caráter de mero truismo e se transforme numa efetiva política de ação. (VIANNA, 2003, p. 34).

Para que a expressão não se transforme numa mera figura retórica e se transforme numa política de ação, necessita estar inserida num contexto em que haja espaço para isso ocorra. Para Cappelletti (2015), cabe ao gestor a função principal de gerir e criar essa cultura. Porém, também a cultura será instituída a partir das concepções adotadas pelo gestor: quer numa perspectiva de seleção, punição, *ranking* ou controle (cultura do exame); quer numa perspectiva de reflexão, transformação, respeito às diversidades e construção coletiva (cultura avaliativa). Em qualquer perspectiva, a criação de uma cultura de avaliação esbarra nas distintas e contraditórias visões de avaliação que as pessoas possuem. Como afirma Dias

Sobrinho (2008), porque as pessoas a serem envolvidas na construção dessa cultura “têm distintas e muitas vezes contraditórias visões de mundo, incluindo concepções sobre educação.” (p. 205). Para o autor, “a avaliação é, muito mais uma expressão da filosofia educativa que um instrumento burocrático.” (DIAS SOBRINHO, 2003a, p. 180). Desse modo, a criação de uma cultura de avaliação requer tempo e a produção de sentidos para todos os envolvidos no processo: gestores, professores, alunos, pais e comunidade em geral.

Diante do exposto, questionamos: Os participantes que percebem a necessidade da criação desta cultura de avaliação estão se referindo a esta cultura avaliativa defendida pelos autores aqui citados? Seria, de fato, uma cultura avaliativa que sentem necessidade? Há espaço para o desenvolvimento desta cultura na estrutura de gestão da rede?

A segunda categoria de análise destaca a necessidade de **aprimoramento da tecnologia**. O PGR1 que aponta essa dificuldade encontra-se em uma região do país onde se espera que o acesso a equipamentos e à internet seja mais fácil que em outras regiões, cujos gestores não apontaram tal dificuldade. Porém, o contexto da resposta indica a dificuldade das provas *online* para os alunos, em virtude da pandemia. A falta de equipamentos para o número de alunos nas escolas e a qualidade dos laboratórios de informática são objeto dessa preocupação:

PGR1 - Com as dificuldades que nós temos, na questão de demanda de tempos, da cultura não estabilizada, da falta de tecnologia, eu acho que isso nos impede de crescermos um pouquinho mais. [...] Por exemplo, este ano, os alunos estão fazendo em casa. Mas tem um celular para a família inteira. Como que ele vai fazer isso? Nos outros anos, que tem na escola, como estão os laboratórios? Será que a gente já conseguiu organizar laboratórios bons em todas as escolas? Quer seja móvel, quer seja mais antigo, nós temos máquinas que possam trabalhar com todos os alunos? Então, nós temos, sim, algumas dificuldades para serem desenvolvidas que o próprio resultado mostra esse retrato, não é?

Nota-se que a preocupação é com a estrutura nas unidades escolares para a realização das avaliações na modalidade *online*. Para o entrevistado, esse fator interfere no resultado. Ao ser abordada, na pergunta quatro, a questão da aplicação da prova, essa temática foi comentada por um participante de outra região, como um elemento complicador que demandou iniciativas locais de logística para garantir acesso de todos os alunos aos equipamentos das unidades escolares e internet de boa qualidade.

A infraestrutura das unidades escolares, como uma das dimensões de qualidade da educação básica a ser avaliada, está estabelecida na sétima meta do PNE – Lei 13.005, de 25 de junho de 2014, que propôs como estratégia:

7.3) constituir, em colaboração entre a União, os Estados, o Distrito Federal e os Municípios, um conjunto nacional de indicadores de avaliação institucional com

base no perfil do alunado e do corpo de profissionais da educação, nas condições de infraestrutura das escolas, nos recursos pedagógicos disponíveis, nas características da gestão e em outras dimensões relevantes, considerando as especificidades das modalidades de ensino. (BRASIL, 2014)

Albernaz, Ferreira e Franco (2002), a partir dos microdados produzidos pelo Saeb, realizaram uma investigação dos determinantes das medidas de desempenho escolar com base em rendimento dos alunos nos testes padronizados e, aplicando modelos estatísticos hierárquicos lineares aos microdados, puderam estimar a contribuição de diferentes variáveis escolares sobre a qualidade do ensino fundamental no Brasil. Na categoria que refere-se ao papel das variáveis escolares, os autores concluem que a qualidade da infraestrutura da escola contribui para um melhor desempenho de seus alunos nas avaliações. Tanto a qualidade dos ambientes: salas arejadas, silenciosas; quanto o estado e conservação dos equipamentos e recursos disponíveis contribuem para um melhor resultado.

Dourado, Oliveira e Santos (2007), em um estudo que visou contribuir com a identificação de condições, dimensões e fatores fundamentais do entendimento do que seja educação de qualidade, considerando a ótica dos países membros da Cúpula das Américas e documentos de organismos multilaterais, apresentaram dentre as dimensões da qualidade da educação, em nível de escola: gestão e organização do trabalho escolar, que um dos aspectos considerados impactantes da qualidade, é “a existência e utilização adequada de tecnologias educacionais e recursos pedagógicos apropriados ao processo de aprendizagem.” (DOURADO; OLIVEIRA; SANTOS, 2007, p. 21). Nesse caso, o uso da tecnologia é apresentado como apoio ao processo de aprendizagem. Porém, na percepção do participante da entrevista, as dificuldades com os recursos tecnológicos interferem nos resultados da avaliação. O que não configura, necessariamente, interferência na aprendizagem. Contudo, é um dado relevante, na medida em que há essa relação apontada nos estudos sobre as condições de infraestrutura e qualidade educacional.

A categoria **entendimento do valor do PAER** foi indicada na resposta de dois participantes – PGR1 e PGR3 – como dificuldade ou ponto a ser desenvolvido. Porém, há que se destacar que esta categoria foi indicada como ganho na resposta sete pelo mesmo sujeito – PGR3. Seria este um paradoxo nas respostas do participante? Ele mesmo explica, ao argumentar:

PGR3 - Para mim, ainda é ampliar esse entendimento. Dos professores, agora. Nós já temos o entendimento como regional, dessa prova; eu percebo claramente que há um entendimento e um consenso já, de perspectiva dessa prova, nas coordenadoras das sub-regiões; mas ainda é preciso ampliar na coordenação das escolas – até porque nós temos uma rotatividade desses profissionais.

Para o participante, o entendimento do valor da avaliação externa precisa ser ampliado junto aos docentes e, ao responder na questão anterior que este é um ganho, refere-se à compreensão dos gestores e coordenadoras dos diversos níveis administrativos da rede. Portanto, ainda considera que os docentes não entendem o valor do programa.

O participante PGR1 também considera que há falta de conhecimento sobre a importância e valor da avaliação externa, não especificando em quais níveis da estrutura da rede isso ocorre.

Atribuir valor para a avaliação externa requer que o propósito e a finalidade da avaliação sejam conhecidos por todos os envolvidos. Becker (2010), ao apresentar um estudo sobre a experiência brasileira na área de avaliação educacional em larga escala, destaca a necessidade de que, ao se implantar um sistema de avaliação em larga escala, os propósitos desse sistema sejam previamente debatidos e que suas metas e objetivos sejam conhecidos por todos: professores, diretores e sociedade em geral. Nesse caso, podemos estabelecer uma relação entre o modelo de criação e implementação de um sistema de avaliação externa e a sua valorização por parte dos envolvidos. Caso não haja esse debate prévio, nem a participação dos principais interessados nos dados que a avaliação externa pode oferecer – neste caso a verificação do desempenho estudantil com vistas a promover ações que concorram para a melhoria do processo educativo como um todo - dificilmente será atribuído valor à avaliação externa.

Como última categoria indicada, o PGR2 levanta a necessidade de **formação para o uso dos resultados do PAER**, especificando que essa formação se destina aos professores que precisam “olhar estes resultados e, a partir dos resultados da prova, olhar as perspectivas pós-resultados”. Ir além dos resultados é compreender que a avaliação tem um caráter dinâmico que gera ações, mudanças, promove questionamentos, transformações e, como afirmam Freitas *et al.* (2017, p. 78), “um processo destinado a promover o permanente crescimento.”. Para que isso ocorra, “os dados são produzidos nos vários níveis com o objetivo de serem usados pelos interessados na geração de processos de reflexão local e melhoria da escola.” (FREITAS *et al.*, 2017, p. 80). Contudo, esse processo depende de uma ação coletiva, não restrita a professores, sendo necessária uma intenção política da instituição na qual ocorre, neste caso, a rede de ensino. O uso que será feito dos resultados dependerá das concepções de avaliação que estão subjacentes a todo o processo: desde os motivos pelos quais foi criado o sistema de avaliação externa, seus objetivos, a forma de aplicação, a participação dos interessados, a comunicação dos resultados, até os possíveis efeitos ou

impactos na organização do trabalho pedagógico. Com essas considerações, entende-se que não seria uma formação restrita aos professores sobre o uso dos resultados da avaliação externa que garantiria a melhoria da qualidade objetivada.

A décima pergunta da entrevista é retratada no Quadro 27, com as categorias de análise que as respostas geraram.

**Quadro 27 - Categorias de análise Eixo Implicações**

Questão 10 - Quais as implicações do PAER no processo educativo?			
Categorias de análise			
Participantes	Desvelamento da realidade	Direcionamento do processo educativo	Problematização a partir dos resultados do PAER
PLG	X		
PGR1			
PGR2		X	
PGR3			
PGR4			X

Elaboração: autora.

Nesta pergunta, dois participantes não conseguiram indicar alguma implicação da avaliação externa da rede no processo educativo. Os demais indicaram o **desvelamento da realidade**, o **direcionamento do processo educativo**, e a **problematização a partir dos resultados do PAER**.

A primeira categoria de análise emerge da visão do PLG, ao afirmar:

Seja qualquer modelo de avaliação, ela tem que acontecer para que a educação possa ser percebida. A implicação do PAER é, praticamente, ela se funde com a própria educação. Ou seja, ela é tão importante quanto a educação em si. A educação não se resume só à avaliação, mas educação sem avaliação é inconcebível. Então, todo este programa de avaliação ele é tão relevante quanto a própria educação.

A ideia de que a avaliação precisa acontecer para que a “educação possa ser percebida”, assim como a vinculação entre ambas, demonstra-nos que o participante tem uma percepção clara da potencialidade da avaliação em **desvelar a realidade**. Porém, qual visão de realidade? Sob a ótica do paradigma dominante, descrito no capítulo cinco, a realidade é estática, desprovida de historicidade, de movimento, alheia ao sujeito que busca conhecê-la e

que, para isso, precisa distanciar-se dela e fragmentá-la. Já, na perspectiva do paradigma emergente, a realidade é dinâmica, sistêmica, interacionista, sendo impossível desvincular o indivíduo e suas relações com a realidade, quando ele busca compreendê-la e transformá-la. Na resposta do PLG pode-se observar indícios das noções da avaliação numa perspectiva do paradigma emergente, no qual há uma integralidade entre a educação e a avaliação, quando a função dessa é trazer à luz a totalidade das situações educativas, não restrita a resultados de exames. Nesse contexto, o processo de descrição da realidade faz-se necessário para que possa ocorrer a análise, a crítica dessa realidade, visando transformá-la. É o desvelar da realidade que pode iluminar o caminho da transformação (SAUL, 2010).

É possível, portanto, que a avaliação externa contribua para esse desvelar da realidade, visando transformá-la? Sem dúvida, ela abriga esse potencial. Porém, como afirma Saul (2010): “A *emancipação* prevê que a consciência crítica da situação e a proposição de alternativas de solução para a mesma constituam-se em elementos de luta transformadora para os diferentes participantes da avaliação” (SAUL, 2010, p. 66, grifo da autora). Faz-se necessário que após o desvelar, haja o envolvimento de todos os participantes na indicação de alternativas e na corresponsabilização pelas mudanças desejáveis.

De igual modo, o **direcionamento do processo educativo**, indicado na resposta do participante PGR2 pode ser interpretado à luz de ambos os paradigmas. Há um reconhecimento de que as implicações no processo educativo ocorrem em vários níveis e aspectos. Porém, como indicado anteriormente, o uso da avaliação pode ser determinado por uma perspectiva impositiva ou participativa; coercitiva ou colaborativa. Contudo, há uma percepção das implicações, ao responder:

PGR2 – Então, é uma avaliação que está avaliando desde a educação infantil. Currículo, a maneira como o professor aborda esse currículo, a formação do próprio professor, a base dele. Então, são vários elementos que estão aí implicados. [...] Tudo reporta ao fato, enfim, ao desempenho que este aluno vai ter, quando por fim, ele chegar na prova. Então, acho que a implicação é constante e total dentro do ensino.

São apresentadas as dimensões do currículo, da formação do professor, da prática docente, como elementos envolvidos no processo do ensino e, conseqüentemente, da aprendizagem.

Como afirma Domingos Fernandes (2009), as avaliações externas podem trazer um conjunto de vantagens ao sistema de ensino, razão esta, por serem adotadas em tantos países. Dentre as vantagens citadas, está a possibilidade de “alertar as escolas para a necessidade de melhorarem seus projetos educacionais” e “dar indicações úteis às escolas, aos professores e

aos alunos acerca do que é importante ensinar e aprender.” (FERNANDES, D., 2009, p. 137). Mas o autor considera que essas vantagens podem ser associadas às avaliações externas, partindo do princípio de que elas são concebidas e elaboradas com qualidade pedagógica, educacional e formativa.

De igual modo, podem apresentar desvantagens ao imprimirem um direcionamento do processo educativo que gere estreitamento curricular; reducionismo das avaliações internas visando ao treino para as avaliações externas; substituição de alunos que deveriam realizar os exames por aqueles que têm possibilidades de sucessos nas provas; discriminação e práticas fraudulentas para obtenção de bons resultados. O direcionamento do processo educativo como consequência da avaliação externa, portanto, será determinado pelos princípios que motivaram a concepção de um sistema de avaliação externa e o uso que se faz de seus resultados. Em ambos os casos, o papel dos gestores é determinante na condução de todo o processo. Pois é na unidade escolar que se dá o processo educativo. Nela, não há o processo decisório sobre a existência, ou não, da avaliação externa. Mas é nela que os envolvidos e interessados em uma prática educativa de qualidade têm condições de potencializar as vantagens de uma avaliação externa e reduzir suas possíveis desvantagens.

A última categoria levantada nas respostas à pergunta dez foi a **problematização a partir dos dados do PAER**. Na percepção do participante, há implicações no processo educativo, e a mais importante é a possibilidade de problematizar e discutir a partir dos resultados obtidos: PGR4 - “Eu tenho este resultado, e o que nós vamos fazer? Vamos buscar soluções. Eu acho que ele é um indicador muito bom, hoje, dentro da rede.”.

Para Cappelletti (2015), em uma avaliação concebida por um processo de emancipação, compete ao gestor escolar investigar criticamente as situações, problematizá-las para ressignificá-las e transformá-las. Esse não é um percurso simples. Demanda conhecimento teórico e prático, lidar com concordâncias e discordâncias, dar liberdade para a expressão dos pontos de vista dos envolvidos e clareza de onde se quer chegar e quais mudanças precisam ser feitas. Para isso, a autora considera que:

Múltiplos conhecimentos, competências e habilidades para o exercício da gestão têm sido fartamente explicitados na literatura: competências de cunho administrativo, pedagógico, social, político e cultural. O mais importante, porém, é o gestor ter clareza do espaço ideológico que “habita”, do contexto socioeconômico e político como cenário inevitável de suas ações, sem o que não terá como conquistar a tão desejada qualidade educacional em seu âmbito de trabalho. (CAPPELLETTI, 2015, p. 103).

Apesar de um percurso desafiador, esse é o modo de produzir sentido à avaliação externa que adquire significado e importância para a comunidade escolar. O contrário disso é a reprodução de um modelo no qual a avaliação atende a funções classificatórias, seletivas, de controle e responsabilização, apenas.

### 6.3 Síntese integradora dos ganhos e das dificuldades

Como demonstrado no início do capítulo, os instrumentos de coletas de dados foram aplicados em gestores de dois níveis da estrutura organizacional de todas as regiões do país. Portanto, é possível afirmar que, ainda que seja uma rede com características específicas definidas por sua natureza confessional e filantrópica, a abrangência geográfica produz contextos distintos, o que diversifica a análise de elementos envolvidos no processo avaliativo, que são comuns a todos. As características socioculturais precisam ser levadas em conta, ao se avaliarem as expectativas da comunidade onde a unidade escolar está inserida. De igual modo, as expectativas dos gestores e suas percepções também são moldadas pelo contexto no qual exercem suas atividades. Essa diversidade, em um país de extensão geográfica continental, pode oferecer uma amplitude de visões, que se refletem nas respostas apresentadas.

Por outro lado, uma das premissas de uma pesquisa de abordagem qualitativa é a subjetividade advinda da singularidade dos indivíduos: suas opiniões, representações, experiências e vivências são únicas, ao passo que sua visão da realidade é interpretada à luz de suas crenças e valores. Minayo (2012) apresenta o termo dialetizar como a capacidade do pesquisador de exercitar o entendimento das contradições, pois tanto as ações como as palavras possuem características contraditórias e conflituosas.

Por vezes, a diversidade e a singularidade podem produzir contradições ou dissensos nas perspectivas apresentadas pelos participantes, o que, em uma abordagem qualitativa, só enriquece a análise e a interpretação dos dados. Um exemplo dessa aparente contradição está na categoria de análise: **entendimento do valor do PAER**, que foi indicada como um ganho que o programa obteve, em respostas à pergunta 7 da entrevista e, ao mesmo tempo, foi indicada como dificuldade a ser superada, em respostas à pergunta 8. Compreender que são visões diferentes para um mesmo fenômeno que ocorre em contextos distintos, reforça a necessidade de o pesquisador ter uma atitude atenta, aberta e interpretativa. Este é o desafio com o qual se depara, igualmente, um meta-avaliador que deve emitir um juízo sobre a própria avaliação.

A análise efetuada oferece suporte para a realização de uma síntese. Para consolidá-la, sem qualquer pretensão de esgotá-la, mas com a humildade de reconhecermos que é um olhar externo também eivado de subjetividade, apoiamo-nos no conhecimento adquirido pelo suporte teórico de autores referenciados que lançam luz sobre a análise. Sendo assim, organizamos o Quadro 28, a título de síntese, dos ganhos (aspectos desenvolvidos) e as dificuldades (aspectos a desenvolver), que pudemos depreender das respostas às entrevistas, ao questionário e no documento de criação do PAER.

**Quadro 28 - Ganhos e dificuldades do PAER**

<b>Fonte</b>	<b>Ganhos</b>	<b>Fonte</b>	<b>Dificuldades</b>
<b>E</b>	Reflexão sobre a prática pedagógica	<b>E</b>	Desenvolver uma cultura de avaliação
<b>E/Q</b>	Desvelamento da realidade	<b>E/Q</b>	Aprimoramento da tecnologia
<b>E</b>	Entendimento do valor do PAER	<b>E</b>	Entendimento do valor do PAER
<b>E</b>	Mudança de olhar sobre a avaliação externa	<b>Q</b>	Alunos não entendem a importância da avaliação externa
<b>E</b>	Proposta de formação de professores	<b>E</b>	Formação para uso dos resultados
<b>E/Q</b>	Melhora na qualidade do ensino	<b>Q</b>	Não há políticas ou ações pós-PAER
<b>Q</b>	Melhora na qualidade da aprendizagem	<b>Q</b>	Professores e gestores não entendem a importância da avaliação externa
<b>Q</b>	Indicador de pontos fortes e fracos	<b>Q</b>	Necessidade de qualificação docente
<b>Q</b>	Mudanças na avaliação interna	<b>Q</b>	Professores não conhecem o objetivo do PAER
<b>Q</b>	Análise dos resultados pelos gestores	<b>Q</b>	Pais e alunos conhecerem os objetivos do PAER
<b>D</b>	A forma de divulgação dos resultados para a rede		
Legenda: <b>E</b> – Entrevista; <b>Q</b> – Questionário; <b>D</b> – Documento do PAER			

Elaboração: autora

A análise das respostas ao questionário e entrevistas, desenvolvida nas duas primeiras partes do capítulo, nos indicam que o PAER é visto potencialmente como um propulsor de mudanças pelos gestores participantes. Sua implantação, a princípio, gerou surpresa e um certo desconforto que se estendeu por algum tempo, como sinalizado por alguns entrevistados. Após oito anos de implementação, o programa é conhecido da comunidade interna das unidades escolares, porém, os objetivos que se propõe atingir não são do conhecimento de todos os envolvidos, incluindo professores, alunos e pais, como apontado pelas respostas ao questionário. Há uma crescente percepção da importância da avaliação

externa, mas ainda não há uma cultura de avaliação instituída na rede, na percepção de alguns entrevistados. Como discutido na seção de análise das entrevistas, esta cultura avaliativa só se desenvolve a partir de uma concepção de avaliação pautada numa perspectiva emancipatória.

Há uma ideia, não generalizada entre os participantes, de que o programa contribui para a melhoria da qualidade educacional. Apesar de que, nas respostas ao questionário, essa melhora está vinculada à qualidade do ensino pelas iniciativas dos docentes em promoverem ações para o preparo dos alunos para as provas, quer por meio de mudanças no estilo de suas avaliações da aprendizagem, quer por meio de simulados e treinos. Desse modo, pode-se inferir que há uma percepção de que a qualidade educacional está relacionada ao desempenho docente ou à melhora das notas dos alunos nas provas. Como já analisado, esta é apenas uma das várias dimensões da qualidade e, por vezes, tem gerado uma visão reducionista que associa os índices produzidos pelas avaliações externas de sistemas e redes como um sinônimo de qualidade educacional. O suposto desenvolvimento cognitivo certamente compõe o processo, porém, não o resume (ESQUINSANE; DAMETTO, 2018).

As iniciativas de formação docente, de cursos organizados pela mantenedora, ministrados pela Universidade Corporativa para atender aos descritores cujos índices estão abaixo do esperado, ao longo das edições do PAER, assim como as iniciativas regionais de formação em serviço, são reconhecidas como avanços e oportunidade de crescimento para os professores. A formação docente voltada para uma compreensão da avaliação e sua importância no processo educativo é necessária, como já discutido no capítulo. Contudo, reafirmamos a necessidade de uma formação em que os aspectos teóricos e práticos se entrelacem, dando sentido e significado, de fato, a todos os processos avaliativos que ocorram no contexto educacional. Não apenas ensinar a elaboração de itens de prova, que preparem os alunos para as avaliações internas, nem amoldar as avaliações da aprendizagem ao padrão da avaliação externa, para que os alunos se acostumem com o modelo. A compreensão dos fundamentos teóricos que sustentam a prática e lhe dão sentido é que pode produzir mudanças significativas.

A questão da formação de professores também se constitui uma aparente contradição, pois está presente como “ganho” e como “dificuldade”. A necessidade de formação de professores para o uso dos resultados está presente nas respostas aos questionários. Foi apontada por alguns participantes, a expectativa de que o mesmo nível administrativo que criou o PAER (mantenedora para toda a América do Sul) se responsabilize por formular políticas de formação e programas “pós-PAER”, para capacitar os professores para o uso dos

resultados. A formação em serviço que propicie reflexão, discussão, uma construção coletiva de alternativas e propostas que se materializem em mudanças e transformações no processo educativo, necessita ocorrer no *locus* de atuação do docente. Ainda que a iniciativa de formação ocorra em níveis de gestão mais amplo que a unidade escolar, é na escola que se dá a concretização do processo. Na discussão com os pares, na interação, na liberdade de expressão de pontos de vista, dúvidas e inseguranças, sem receios de cerceamento, numa relação de respeito; orquestrada pelo gestor escolar, que traduz para a comunidade escolar os valores subjacentes aos programas da rede ou sistema de ensino (CAPPELLETTI, 2015).

Como apontado pelos entrevistados, não há iniciativas de *rankings* após a divulgação dos resultados e nem uma política de responsabilização. Nas primeiras edições, os resultados eram encaminhados aos alunos por meio de correspondência enviada às unidades escolares, com uma divulgação ampla para a comunidade interna e para os pais. Desde 2016, os resultados são disponibilizados em uma plataforma, e o gestor escolar pode divulgá-los, ou não. Sendo assim, a comunicação dos resultados para a comunidade interna e externa está restrita à atuação do gestor. Levando esses aspectos em conta, poderíamos afirmar que se trata de uma avaliação externa de primeira geração (BONAMINO; SOUZA, 2012) e, como percebido por alguns participantes, com fracas consequências para o currículo escolar.

Quanto às implicações no processo educativo, os participantes de ambos os instrumentos apresentaram alguma dificuldade em responder a essa temática. Há uma ideia prevalente de que há implicações, mas há também participantes que não as consideram significativas. Essas implicações também estão associadas à atuação docente e mudanças nas práticas pedagógicas que visam ao alcance de melhores resultados.

Um dos pontos indicados como ganho é a oportunidade de refletir sobre os dados produzidos pelas avaliações. Como indicado por um participante da entrevista, “reflexão é o nosso maior ganho” (PGR1). Numa perspectiva de uma educação emancipatória, que busca a transformação e a qualidade como um direito social, mediante a qual a análise da realidade se faz necessária para que, coletivamente se promovam as mudanças, todo movimento precisa iniciar no desvelar da realidade e na reflexão sobre essa descrição que o processo avaliativo nos apresenta. Pelos dados apresentados pelos gestores entrevistados e pelos questionários, há uma análise dos resultados no nível de gestão. Porém, o mesmo não ocorre em nível de docência. O percentual de professores que analisam os dados está abaixo de cinquenta por cento. A análise e a reflexão devem se complementar. E, sem a participação de todos os envolvidos, os esforços para transformar a realidade tonam-se estéreis.

Ao confrontarmos os dados sobre a participação dos professores na análise dos dados e o conhecimento que possuem sobre os objetivos do PAER, ambos os aspectos indicados pelos gestores abaixo de cinquenta por cento, é possível afirmar que ainda se faz necessário um esforço de envolvimento dos docentes como partícipes e interessados nas contribuições que uma avaliação externa é capaz de proporcionar. Não os envolver pode ser considerada uma indicação de que a finalidade do programa de avaliação restringe-se à verificação de desempenho cognitivo, com uma branda responsabilização dos docentes pelos resultados. Como consequência da não responsabilização, pode haver, por parte dos professores, uma dissociação da avaliação externa com sua prática pedagógica, um desinteresse pelos dados produzidos, considerando como um elemento alheio ao processo educativo. Essa não é a finalidade declarada no documento do PAER, mas pode ser a finalidade derivada de seu uso.

Tanto nas respostas às entrevistas, quanto nas respostas abertas do questionário *on-line* há uma expressão de reconhecimento e “orgulho” pelo fato da rede ter um programa de avaliação externa em larga escala com um grau de confiabilidade, segurança, estrutura e organização que o PAER tem demonstrado. Pelos dados levantados pelos instrumentos, é possível concluir que ainda que gestores valorizem o programa, faz-se necessário que docentes e alunos reconheçam esse valor. Para tanto, é imprescindível o conhecimento dos objetivos, finalidades e usos da avaliação externa por toda a comunidade acadêmica. A descrição de como o programa foi implementado, em resposta à pergunta dois da entrevista, pode ser uma das explicações para essa dificuldade apontada pelos participantes. As informações ficaram restritas às equipes gestoras de níveis superiores da estrutura da rede. Os professores ficaram “surpresos” com a implantação, criando um certo “descontentamento”, quando o processo iniciou. Porém, o programa é reconhecido como algo de valor para a rede. A avaliação, por si só, não é responsável por gerar valor ou produzir mudanças. O fato de gerar mudanças não configura, necessariamente, que seja uma mudança para melhoria do processo ou como sinônimo de qualidade. A mudança pode ser direcionada para um maior controle, uma mobilização para a obtenção de resultados, sem que haja, de fato, uma reflexão crítica sobre os dados com vistas à melhoria da qualidade educacional. Como afirma Luckesi (2018), o ato de avaliar investiga a qualidade da realidade e “revela” a qualidade do objeto avaliado. O uso que se faz dessa “revelação” ou dos resultados é que tipifica o fundo filosófico e epistemológico que subsidia as decisões envolvidas no ato de avaliar. Para o autor, o ato de avaliar se encerra em si mesmo. Porém, “é subsidiário de decisões, por isso está sempre “a serviço de”, a serviço de quem decide.” (LUCKESI, 2018, p. 188).

Além de subsidiar gestores na tomada de decisões, a avaliação em larga escala deve ter sentido para professores e estudantes. É imprescindível que todos se sintam parte do processo. Afinal, são eles que concretizam o processo educativo. É por meio deles e para eles que se busca uma educação de qualidade. Contudo, este ainda é um aspecto que precisa ser desenvolvido para que o PAER se torne significativo e relevante para alunos e professores. O modelo centrado no desempenho dos estudantes pode, potencialmente, oferecer um obstáculo para o desenvolvimento da cultura de avaliação; pois apenas um dos elementos envolvidos – o aluno – é responsabilizado pelos resultados; o que caracterizaria a cultura do exame. Por meio da reflexão e do diálogo, que congregue a todos: docentes, discentes e gestores, é possível a construir a produção de sentidos a todos os envolvidos.

Outrossim, diante da síntese dos ganhos e dificuldades que foram levantados pela análise desenvolvida no capítulo, é possível considerar que ainda se faz necessário um esforço para que o PAER concretize seu objetivo de retratar a qualidade da realidade da rede, limitado como está a uma prova de desempenho cognitivo e um questionário contextual. Como descrito no capítulo cinco, sobre as concepções de qualidade, a realidade educacional não se restringe a estes aspectos. Há inúmeras dimensões intra e extraescolares a serem consideradas e avaliadas, a partir de critérios definidos; a fim de se retratar a qualidade de uma realidade educacional. Em se tratando de um programa de avaliação, a possibilidade de construção de uma avaliação interna em nível institucional, que envolva professores, alunos, funcionários e até mesmo participantes externos, como pais e representantes da comunidade, e que seja articulada com uma autoavaliação, permitiria um amplo processo avaliativo que viabilizaria o retratar da qualidade da rede envolvendo as dimensões que a compõem. Desse modo, os dados e as informações levantadas, o estímulo a uma análise reflexiva de todos os envolvidos, poderiam assegurar “um processo de descrição, análise e crítica de uma dada realidade, visando transformá-la.” (SAUL, 2010, p. 65).

Franco (2004, p.45), na introdução de um artigo que visa refletir sobre as contribuições da avaliação em larga escala para as políticas educacionais, afirma que “há razões para vislumbrar que estas avaliações podem dar contribuições relevantes às políticas educacionais”. Sem desprezar as críticas que são apresentadas a ela, a permanência desse tipo de avaliação ao longo das últimas três décadas e sua expansão em sistemas públicos e privados nos revelam a necessidade de mais pesquisas que nos permitam compreender como esses resultados estão repercutindo, onde e com quem, de fato, a educação ocorre.

Por certo, há muitas informações disponíveis nos dados coletados que permitem inúmeras e variadas interpretações sob o olhar atento de outro pesquisador, que poderia estabelecer outras relações e considerações. Afinal, interpretar é um ato que sucede a compreensão, é a apropriação do que foi compreendido para elaborar possibilidades. Somos limitados no que compreendemos e, conseqüentemente, no que interpretamos. Esse reconhecimento nos conduz a uma atitude de busca constante de interpretar à luz dos fundamentos teóricos que nos orientaram nas indagações que ocorrem durante todo o percurso investigativo. Essas indagações apenas iniciaram, quando se estabeleceu o problema de pesquisa. Novas possibilidades se abrem após procurarmos, ao longo deste capítulo, respondê-lo.

## CONSIDERAÇÕES FINAIS

O percurso investigativo é norteado pela questão ou problema de pesquisa que o motivou. Espera-se que, ao final desse percurso, respostas sejam obtidas como resultado da investigação. Contudo, por vezes, além de respostas, novas questões surgem, abrindo expectativas de buscas por novas respostas a novos problemas a serem desbravados. Assim tem sido na área da avaliação de sistemas, considerada nas últimas décadas como eixo estruturante das políticas educacionais, constituindo-se num campo de pesquisas com muitos desafios e complexidades a serem exploradas. Sua abrangência: alcançando todos os níveis de ensino; sua extensão: expandindo-se nos sistemas federal, estaduais e municipais; e sua permanência: preservada por governos de diferentes matrizes políticas, atribuem uma relevância aos estudos e, por vezes, justificam essa complexidade. É com a consciência dessa complexidade e amplitude que delimitamos esta investigação ao problema que a originou: quais ganhos (aspectos desenvolvidos) e quais dificuldades (aspectos a desenvolver) podem ser identificados com a implementação do programa de avaliação externa em larga escala por uma rede privada de ensino de abrangência nacional, com vistas a assegurar qualidade à educação oferecida?

Em um cenário de crescente interesse pelos índices produzidos pelas avaliações em larga escala em diferentes entes federativos, a rede privada, *locus* desta pesquisa, desenvolveu seu programa de avaliação externa em larga escala, pautado nos princípios da Prova Brasil (Saeb), oportunizando que a investigação de que trata a presente tese se desenvolvesse com uma abordagem meta-avaliativa, com dados coletados em diferentes níveis de gestão e em toda a extensão regional do país.

A revisão da literatura sinalizou a carência de estudos pautados em procedimentos meta-avaliativos voltados para a área educacional. Desse modo, compreende-se que esta pesquisa oferece um contributo como um exercício de meta-avaliação em um contexto singular, porém, com aplicabilidade para outros contextos em que sistemas de avaliação em larga escala para a educação básica sejam desenvolvidos. Com o crescimento de sistemas de avaliação nos últimos anos, há que se ampliar os esforços de meta-avaliações capazes de aprimorar os processos de avaliação e de acreditação de sistemas e redes, em todos os níveis de ensino.

Para a consecução dos objetivos da pesquisa, derivados do problema, a contextualização e o reconhecimento da trajetória das políticas públicas voltadas à avaliação

da educação básica foram delineados, confirmando o crescimento de iniciativas dessa natureza em estados e municípios ao longo dos últimos anos, com a vinculação dessas iniciativas à ideia de melhoria da qualidade educacional por meio das avaliações externas. Desde a concepção original do Saeb, na última década do século XX, até o modelo atual, passando por algumas mudanças que o distanciaram de sua concepção original ao longo dessa trajetória, é preservado o objetivo de contribuir para a melhoria da qualidade educacional, objetivo este confirmado em todos os atos normativos que asseguraram a sua continuidade.

Da mesma forma, buscou-se reconhecer a trajetória do PAER, tendo como ponto de partida o processo histórico de consolidação da educação privada no Brasil, sua caracterização no cenário nacional, assim como a origem da rede privada de extensão nacional que desenvolveu o programa de avaliação externa. A descrição do contexto, dos marcos normativos que garantem à educação privada uma relativa autonomia, inclusive de adotar, ou não, o modelo de sistema de avaliação desenvolvido pelo poder público, caracterizam o ensino privado e sinalizam que a rede, ao criar e implementar um sistema próprio de avaliação em larga escala, com a complexidade e os custos que um programa dessa envergadura requer, retrata uma compreensão de que tal iniciativa contribui para a melhoria da qualidade educacional da rede e seu respectivo crescimento, como descrito no objetivo do programa investigado.

A relação entre as ideias de avaliação educacional e, mais especificamente, avaliação externa, e a qualidade educacional, presentes nos documentos oficiais, assim como no ideário coletivo, foram objeto de um amplo levantamento das concepções subjacentes a ambos os termos, ao longo do capítulo cinco. As diferentes concepções de avaliação, pautadas quer no paradigma dominante, quer no paradigma emergente, coexistem em todas as esferas e níveis dos sistemas educacionais de forma explícita ou implícita, gerando tensões ou convivendo ecleticamente, ora com avaliações com funções classificatórias e seletivas, ora com avaliações de caráter formativo, visando à transformação da realidade educacional. Igualmente, as ideias de qualidade educacional associadas ao longo do tempo às ideias de expansão, de garantia de acesso e permanência, eficiência de processos ou, mais recentemente, às ideias de qualidade social, com igualdade de oportunidades e direitos, gestão democrática e participativa, visando ao bem comum, são associadas nos discursos que justificam o desenvolvimento de avaliações de sistemas de ensino. Contudo, ainda há que se percorrer um caminho de discussão sobre as concepções de qualidade educacional preconizadas nos documentos normativos que regulamentam esses processos avaliativos. Ainda há a necessidade de se explicitarem os

indicadores de qualidade que contemplem as dimensões intra e extraescolares, rompendo com a unicidade dos índices limitados ao desempenho dos alunos e fluxos escolares. E, no contexto da educação básica, a possibilidade de se ampliar a avaliação educacional, articulando a avaliação externa com a avaliação institucional, envolvendo a autoavaliação e a contribuição de participantes internos e externos da comunidade escolar nesse processo, viabilizaria a avaliação das diferentes dimensões determinantes de uma educação de qualidade.

Ao longo da análise dos dados ficou evidente que as ideias acerca da qualidade educacional para os gestores da rede investigada ainda são ambíguas e confusas. Ora consideram a qualidade do ensino ou o desempenho docente como indicador de qualidade educacional; ora consideram o desempenho dos alunos na avaliação externa como sinônimo de qualidade educacional. A rede reflete a ambiguidade e a falta de clareza que coexiste de maneira geral nos sistemas educacionais.

A ideia de avaliação externa em larga escala como uma iniciativa do Estado, com vistas a atender as necessidades ou demandas sociais – neste caso, a qualidade educacional – a vincula às ações de políticas públicas. Como analisado na introdução desta tese, uma ação política que se concretiza numa prática educativa. Porém, ela não se constitui de *per se* numa política pública, pois, como demonstrado ao longo da pesquisa, a avaliação é capaz de desvelar a realidade e gerar dados. O uso desses dados é que pode desencadear ações, programas ou projetos que, implementados, têm potencial para promover as mudanças e transformações necessárias com vistas a atender aos interesses sociais. Trata-se, portanto, de um instrumento valioso para que políticas públicas possam ser instituídas, a partir das informações geradas pelos resultados das avaliações. Por ser confundida com uma política pública, a avaliação de sistemas, por vezes, adquire um caráter finalístico, de constatação e verificação, como um fim em si mesma, esgotando-se na produção de índices e dados. Assim como um dos participantes da pesquisa questionou: “O que está proposto para depois do PAER?” Ou: “Como continuar a mover a roda?” questionamos: o que está proposto para além das avaliações externas em larga escala, com vistas à qualidade educacional? É certo que os índices desencadearam ações nos diferentes níveis de gestão dos sistemas educacionais. Ações que são pautadas em concepções e perspectivas políticas e ideológicas distintas; desde o uso dos resultados para diagnóstico, planejamento e intervenção com vistas a monitorar pontos fortes e fracos da educação escolar; ao uso para ranqueamentos, responsabilização

vertical de professores e gestores escolares, e estigmatização. Daí, considera-se a importância de se discutir o uso dos resultados dessas avaliações, seus limites e suas potencialidades.

O referencial teórico apresentado ao longo da tese, especificamente, no quinto capítulo e no diálogo com os autores na análise dos dados no percurso do sexto capítulo, permitiu que o objetivo de identificar os ganhos e as dificuldades na implementação do programa de avaliação em larga escala da rede fosse alcançado. Contudo, a partir desse referencial, podemos questionar se, de fato, o que se considerou ganho, na percepção dos participantes, é compatível com a visão de qualidade educacional ensejada. Se as mudanças introduzidas após a análise dos resultados do desempenho cognitivo dos alunos, de fato, são mudanças que se concretizam em ganho de qualidade. Um exemplo dessa questão está no depoimento de vários participantes dos questionários *on-line* que indicam como ganho obtido com a implementação do PAER as mudanças na prática de docentes que passaram a treinar os alunos para as avaliações externas, a realização de simulados pelas unidades escolares e a ênfase na organização curricular a partir dos descritores das áreas do conhecimento, objeto das provas. Essas ações são percebidas pelos gestores participantes como positivas e favoráveis ao processo educativo. Diante dessa compreensão, identifica-se uma lógica performática de resultados, lógica esta, alinhada ao paradigma dominante. Assim, pode-se considerar que alguns aspectos percebidos como ganhos pelos participantes não o são, se considerados numa perspectiva emancipatória e transformadora.

De igual modo, ao longo da investigação, pode-se identificar que o processo avaliativo possui um potencial de produzir mudanças. Contudo, questionamos: todas as mudanças contribuem para a melhoria da qualidade educacional? Como já citado, à luz da discussão realizada, podemos afirmar que as mudanças não promovem, necessariamente, a melhoria da qualidade educacional. Ainda que essas mudanças envolvam a mobilização dos professores interessados na melhora dos índices e resultados, por sentirem indiretamente que seu desempenho é avaliado por meio desses resultados. Essa mobilização não significa, necessariamente, uma prática pedagógica que redunde em qualidade de ensino. Mudanças, portanto, não são sinônimo de melhoria.

Ainda que a avaliação externa em larga escala seja planejada e desenvolvida em níveis de gestão de sistemas e redes que estão distantes da realidade das unidades escolares, sem que seja oferecida a oportunidade de participação da comunidade escolar na concepção e implementação do processo avaliativo, esta pode ser utilizada como um instrumento de gestão e promoção de transformação da realidade, quando os resultados são interpretados como indicadores de uma

fração dessa realidade e, a partir de um esforço de reflexão e discussão coletiva, envolvendo e dando voz aos atores do processo educativo, o papel da avaliação seja compreendido e valorizado na justa medida. Esse complexo movimento é que permite que seja desenvolvida uma cultura de avaliação no ambiente escolar. O uso dos resultados em uma perspectiva emancipatória favorece, portanto, a formação dessa cultura. Os gestores da rede identificaram como uma das fragilidades do PAER: a necessidade da formação de uma cultura de avaliação. Reafirma-se que, a partir do modo como os resultados são considerados e trabalhados, essa realidade pode ser mantida ou transformada. Tanto para a criação de uma cultura de avaliação no ambiente escolar, como para que haja uma atmosfera propícia para as transformações necessárias, visando à qualidade educacional, o caminho a ser percorrido requer um esforço coletivo, com o envolvimento de todos os interessados. Não é um caminho fácil, mas é seguro.

A investigação permitiu um exercício de análise dos ganhos e das dificuldades advindos da percepção dos participantes, sobre o programa de avaliação externa e, conseqüentemente, tornou possível a construção de uma síntese destes aspectos, que se constituíram o objeto da meta-avaliação proposta. Neste esforço, foi possível reunir elementos cuja relevância já tem sido apontada na literatura sobre o tema, tais como: i) a importância da formação dos professores para a avaliação educacional; ii) o papel dos gestores na condução dos processos avaliativos; iii) a importância do envolvimento da comunidade interna e externa na discussão dos resultados e seus usos; iv) a necessidade de ir além da avaliação de desempenho cognitivo, ampliando o processo com a avaliação institucional e autoavaliação; entre outros. A síntese proposta também indicou aparentes contradições na percepção dos gestores, ao indicarem dados que, ora foram apontados como ganhos, ora apontados como dificuldades. Esses elementos corroboraram a visão de que as experiências e vivências dos participantes são únicas, singulares, enriquecem o material empírico e, retratá-las com fidedignidade exige do investigador uma compreensão do contexto, a habilidade de confrontar a opinião dos participantes com o quadro de referência e, especialmente, com a complexidade do fenômeno investigado – características de uma pesquisa de abordagem qualitativa.

No decurso e desenvolvimento da pesquisa, a tese enunciada de que, “de fato, para que o programa de avaliação em larga escala promova a melhoria da qualidade ensejada é necessário atenção ao uso dos resultados, em uma perspectiva emancipatória. De igual modo, há necessidade de articulação entre a avaliação externa, foco da investigação, e avaliação da aprendizagem, na qual se oportunize e promova a reflexão para a transformação” pode ser demonstrada e evidenciada. Reafirma-se que, à luz da investigação realizada, considerar com

exclusividade os resultados da avaliação externa, centrada em avaliação de desempenho, torna-se insuficiente para expressar a qualidade educacional.

Uma das proposições consideradas pela pesquisadora na elaboração da investigação e expressa no primeiro capítulo está relacionada à crença de que as pesquisas educacionais devam estar a serviço da vida das escolas e da compreensão de seus problemas práticos. Portanto, espera-se que as discussões e reflexões desencadeadas pela pesquisa e seus possíveis achados constituam-se um contributo para a área da avaliação de sistemas de ensino, na medida em que a realidade da rede de ensino investigada possa ser identificada com contextos semelhantes, nos quais a avaliação externa e em larga escala ocorra. Espera-se, também, que contribua para sensibilizar para a importância da meta-avaliação de programas de avaliação na área educacional, uma vez que essa prática pode contribuir para análises consistentes das políticas públicas adotadas para essa área. A avaliação externa em larga escala na educação básica já possui uma trajetória que justifica a consideração da necessidade de realização de avaliação dos sistemas de avaliação. Talvez, com a realização da meta-avaliação dos sistemas avaliativos, poder-se-iam obter repostas para as perguntas: Por que avaliamos? O que, de fato, avaliamos?

Outra proposição que orientou a pesquisa relaciona-se à consciência da complexidade da educação como uma prática social humana, um processo histórico e, portanto, inconcluso e em transformação, o que faz com que se compreenda que a pesquisa é capaz de refletir um recorte de uma prática social a ser analisada por indivíduos também em processo de desenvolvimento e transformação. O estudo promove alterações em quem se debruça sobre ele, permitindo a ressignificação do fenômeno estudado. Com esta consciência, compreende-se que o papel da pesquisa não é produzir certezas, mas indicar respostas para problemas reais, a partir de uma visão específica de mundo, de homem e de educação.

## REFERÊNCIAS

AFONSO, Almerindo J. **Avaliação Educacional: regulação e emancipação: para uma sociologia das políticas avaliativas contemporâneas.** São Paulo, Cortez, 2000.

\_\_\_\_\_. Estado, políticas educacionais e obsessão avaliativa. **Contrapontos**, Itajaí, v. 7, n. 1, p. 11-22 - Itajaí, jan./abr., 2007.

\_\_\_\_\_. **Para uma concetualizacao alternativa de *accountability* em educação.** Educação e Sociedade, Campinas, v. 33, n. 119, p. 471-484, abr./jun., 2012.

\_\_\_\_\_. **Políticas de responsabilização: equívocos semânticos ou ambiguidades político-ideológicas?** Revista Educação PUC-Camp., Campinas, v. 23(1), p. 8-18, jan./abr., 2018.

ALBERNAZ, Ângela; FERREIRA, Francisco.; FRANCO, Creso. **Qualidade e equidade no ensino fundamental brasileiro.** Pesquisa e Planejamento Econômico, Rio de Janeiro: IPEA, v. 32, n. 3, p. 453 - 475, dez. 2002.

ALVES, Manoel. **A histórica contribuição do ensino privado no Brasil.** Educação, Porto Alegre, v. 32, n. 1, p. 71-78, jan./abr. 2009.

AZEVEDO, Janete M. L. **Notas sobre a análise da gestão da educação e da qualidade do ensino no contexto das políticas educativas.** Revista Brasileira de Política e Avaliação da Educação, v.27, n.3, p. 361-588, set./dez. 2011.

BARDIN, Laurence. **Análise do Conteúdo.** Lisboa, Edições 70, 1979.

BARROSO, João. **Autonomie et modes de régulation locale dans le système éducatif.** In: Revue française de pédagogie. Volume 130, 2000. p. 57-71 CNE. Comité national d'évaluation. Disponível em [https://www.persee.fr/doc/rfp\\_0556-7807\\_2000\\_num\\_130\\_1\\_1053](https://www.persee.fr/doc/rfp_0556-7807_2000_num_130_1_1053).

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.

\_\_\_\_\_. **Estudos sobre sistemas de avaliação no Brasil: um retrato em branco e preto.** Revista @mbienteeducação, São Paulo, v. 5(1), p. 7-31, jan./jun., 2012.

\_\_\_\_\_. **Avaliação e Currículo: sistematizando o debate.** In. **XII COLÓQUIO SOBRE QUESTÕES CURRICULARES/VIII COLÓQUIO LUSO-BRASILEIRO DE CURRÍCULO/II COLÓQUIO LUSO-AFRO-BRASILEIRO DE QUESTÕES CURRICULARES**, 2016, na Universidade Federal de Pernambuco, Recife. Currículo, tecnologias, conhecimento escolar, gestão da escola e história social das disciplinas. Recife, ANPAE, 2016, p. 729 – 741. Disponível em: <https://www.seminariosregionaisnaepae.net.br/BibliotecaVirtual/2- Colóquio/Serie2 pdf>.

BAUER, A.; ALAVARSE, O. M.; OLIVEIRA, R. P. **Avaliação em larga escala: uma sistematização do debate.** Educação e Pesquisa, São Paulo, v. 41, n. especial, p. 1367-1382, dez., 2015.

BAUER, A.; PIMENTA, C. O.; HORTA NETO; J. L.; SOUZA; S. Z. L. **Avaliação em Larga Escala em Municípios Brasileiros: o que dizem os números?**. Est. Aval. Educ., São Paulo, v. 26, n. 62, p. 326-352, maio/ago. 2015.

BECKER, Fernanda R. **Avaliação Educacional em Larga Escala: a experiência brasileira**. Revista Ibero-americana de Educação. Organização dos estados Ibero-americanos para a Educação, a Ciência e a Cultura, n. 53/1, p. 1-11, jul. 2010.

BERLINER, M. R. **Avaliação do Indicador Nacional de Alfabetismo Funcional**. 2010. 67 f. Dissertação (Mestrado Profissional em Avaliação) – Fundação Cesgranrio, Rio de Janeiro, 2010. Orientadora: Profa. Dra. Ligia Gomes Elliot.

BONAMINO, Alícia. **A evolução do Saeb: desafios para o futuro**. Em aberto, Brasília, v. 29, n. 96, p. 113 – 126, maio/ago. 2016.

BONAMINO, Alícia; FRANCO, Creso. **Avaliação e política educacional: o processo de institucionalização do SAEB**. Cadernos de Pesquisa, nº 108, p. 101-132, novembro, 1999.

BONAMINO, Alícia; SOUZA, Sandra Z. **Três gerações de avaliação da educação básica no Brasil: interfaces com o currículo da/na escola**. Educação e Pesquisa, São Paulo, v. 38, n. 2, p. 373-388, abr./jun. 2012.

BORDINI, Débora Valim C. **Um estudo meta-avaliativo da avaliação da docência**. 2017, Dissertação (Mestrado Profissional em Educação), Pontifícia Universidade Católica de São Paulo, 2017. Orientador: Prof. Dr. Nelson Antônio Simão Gimenes.

BRASIL. **Constituição da República Federativa do Brasil**. Diário Oficial da União, Brasília, DF, 05 out. 1988. Disponível em [http://www.planalto.gov.br/ccivil\\_03/constituicao/constituicaocompilado.htm](http://www.planalto.gov.br/ccivil_03/constituicao/constituicaocompilado.htm). > Acesso em 12.07.2019

\_\_\_\_\_. Instituto Nacional de Estudos e Pesquisas Educacionais, Ministério da Educação. **Saeb - Sistema Nacional de Avaliação da Educação Básica: objetivos, diretrizes, produtos e resultados**. Brasília, junho, 1994.

\_\_\_\_\_. Instituto Nacional de Estudos e Pesquisas Educacionais, Ministério da Educação. **Saeb: Primeiros resultados**. Brasília, 1997. Disponível em < [https://download.inep.gov.br/publicacoes/institucionais/avaliacoes\\_e\\_examens\\_da\\_educacao\\_basica/Saeb\\_97\\_primeiros\\_resultados.pdf](https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_examens_da_educacao_basica/Saeb_97_primeiros_resultados.pdf) > Acesso em 06.08.2019

\_\_\_\_\_. Instituto Nacional de Estudos e Pesquisas Educacionais, Ministério da Educação. **Relatório: Saeb – Sistema Nacional de Avaliação da Educação Básica – Resultados 99**. Brasília, 2000. Disponível em < [http://portal.inep.gov.br/informacao-da-publicacao/-/asset\\_publisher/6JYIsGMAMkW1/document/id/485776](http://portal.inep.gov.br/informacao-da-publicacao/-/asset_publisher/6JYIsGMAMkW1/document/id/485776) > Acesso em 06. 08.2019.

\_\_\_\_\_. Instituto Nacional de Estudos e Pesquisas Educacionais, Ministério da Educação. **Saeb 2001 Novas Perspectivas**. Brasília, 2002. Disponível em < <http://200.130.24.54/documents/186968/484421/SAEB+2001+novas+perspectivas/cd69d523-8fc6-48f4-8007-7a2aa4bb99bd?version=1.2> > Acesso em 04.08.2019.

\_\_\_\_\_. Instituto Nacional de Estudos e Pesquisas Educacionais, Ministério da Educação. **Qualidade da Educação: uma nova leitura do desempenho dos estudantes da 4ª série do ensino fundamental.** Brasília, abril, 2003. [http://portal.inep.gov.br/informacao-da-publicacao/-/asset\\_publisher/6JYIsGMAMkW1/document/id/485880](http://portal.inep.gov.br/informacao-da-publicacao/-/asset_publisher/6JYIsGMAMkW1/document/id/485880) > Acesso em 06.08.2019.

\_\_\_\_\_. **Portaria Inep/Mec nº 47** de 3 de maio de 2007. Diário Oficial da União, Brasília, Seção 1, p. 11, publicada em 07 de maio de 2007. Disponível em [http://download.inep.gov.br/educacao\\_basica/prova\\_brasil\\_Saeb/legislacao/Port47\\_3MAI07.pdf](http://download.inep.gov.br/educacao_basica/prova_brasil_Saeb/legislacao/Port47_3MAI07.pdf) > Acesso em 18.09.2019.

\_\_\_\_\_. **Relatório SAEB (ANEB e ANRESC) 2005-2015: panorama da década.** – Brasília, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2018. Disponível em < <http://portal.inep.gov.br/documents/186968/484421> > Acesso em 14.10.2019.

\_\_\_\_\_. **Portaria Inep/Mec 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, Seção 1, p. 47, publicado em 02 de maio de 2019. Disponível em <https://www.in.gov.br/web/dou/-/portaria-n%C2%BA-366-de-29-de-abril-de-2019-86232542> > Acesso em 20.01.2020.

\_\_\_\_\_. **Censo Escolar 2019: Resultados e Resumos.** Disponível em < <http://portal.inep.gov.br/web/guest/resultados-e-resumos> > Acessado em 19.01.2020.

\_\_\_\_\_. **Lei nº 9394**, de 20 de dezembro de 1996. Estabelece as diretrizes e bases da educação nacional. Diário oficial da União, Brasília, DF, 23 dez. 1996. Seção 1. Disponível em < [http://www.planalto.gov.br/ccivil\\_03/leis/19394.htm](http://www.planalto.gov.br/ccivil_03/leis/19394.htm). > Acesso em 19.05.2019.

\_\_\_\_\_. **Portaria MEC nº 1.795** de 27 de dezembro de 1994. Diário Oficial da União, Brasília, DF, seção 1, n. 246, p. 20767-68, 28 de dezembro de 1994. Disponível em < <http://pesquisa.in.gov.br/imprensa/jsp/visualiza/index.jsp?data=28/12/1994&jornal=1&pagina=91&totalArquivos=152> > Acesso em 19.08.2019.

\_\_\_\_\_. **Decreto 2.146** de 14 de fevereiro de 1997: Aprova a Estrutura Regimental e o Quadro Demonstrativo de Cargos em Comissão e Funções Gratificadas do Instituto Nacional de Estudos e Pesquisas Educacionais – Inep e dá outras providências. Diário Oficial da União, 14 de fevereiro de 1997.

\_\_\_\_\_. **Portaria MEC 839** de 26 de maio de 1999. Diário Oficial da União, Brasília, DF, seção 1, n. 100-E, p. 11, 27 de Maio de 1999. Disponível em < <https://pesquisa.in.gov.br/imprensa/jsp/visualiza/index.jsp?jornal=1&pagina=11&data=27/05/1999> > Acesso em 19.08.2019.

\_\_\_\_\_. Conselho Nacional de Educação. Câmara de Educação Básica. **Parecer CNE/CEB nº 30.** Responde à consulta da União dos Conselhos Municipais de Educação esclarecendo o conceito de sistemas de educação. Diário Oficial da União, Brasília, DF, 06 de out. de 2000, seção 1, n. 194-E, p. 25, 06 de Outubro de 2000. Disponível em < <https://pesquisa.in.gov.br/imprensa/jsp/visualiza/index.jsp?jornal=1&pagina=25&data=06/10/2000> > Acesso em 22.08.2019.

\_\_\_\_\_. **Portaria MEC 931** de 21 de março de 2005. Diário Oficial da União, Brasília, DF, seção 1, n. 55, p. 17, 22 de março de 2005. Disponível em <

[http://download.inep.gov.br/educacao\\_basica/prova\\_brasil\\_Saeb/legislacao/Portaria931\\_Novo\\_Saeb.pdf](http://download.inep.gov.br/educacao_basica/prova_brasil_Saeb/legislacao/Portaria931_Novo_Saeb.pdf)> Acesso em 22.08.2019.

\_\_\_\_\_. **Decreto 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. Disponível em <[http://www.planalto.gov.br/ccivil\\_03/\\_Ato2007-2010/2007/Decreto/D6094.htm](http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2007/Decreto/D6094.htm)> Acesso em 24.08.2019.

\_\_\_\_\_. **Portaria MEC 482** de 7 de junho de 2013. Dispõem sobre o Sistema de Avaliação da Educação Básica – SAEB. Diário Oficial da União, Brasília, DF, seção 1, p. 17, 10 de junho de 2013. Disponível em [https://www.in.gov.br/materia/-/asset\\_publisher/Kujrw0TZC2Mb/content/id/30677986/do1-2013-06-10-portaria-n-482-de-7-de-junho-de-2013-30677978](https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/30677986/do1-2013-06-10-portaria-n-482-de-7-de-junho-de-2013-30677978) > Acesso em 24.08.2019.

\_\_\_\_\_. **Lei 13.005** de 25 de junho de 2014. Aprova o Plano Nacional de Educação – PNE e dá outras providências. Disponível em [http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2014/lei/113005.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/113005.htm) > Acesso em 30.08.2019.

\_\_\_\_\_. **Portaria Inep nº 174** de 13 de junho de 2015. Dispõem sobre o Sistema de Avaliação da Educação Básica – SAEB. Diário Oficial da União, Brasília, DF, seção 1, nº 90, publicada em 14 de maio de 2015. Disponível em [https://download.inep.gov.br/educacao\\_basica/prova\\_brasil\\_Saeb/legislacao/2015/portaria\\_n174\\_13052015.pdf](https://download.inep.gov.br/educacao_basica/prova_brasil_Saeb/legislacao/2015/portaria_n174_13052015.pdf) > Acesso em 10.09.2021

\_\_\_\_\_. **Portaria MEC 369** de 05 de maio de 2016. Institui o Sistema Nacional de Avaliação da Educação Básica. Diário Oficial da União, Brasília, DF, seção 1, p. 26, publicado em 06 de maio de 2016. Disponível em [https://www.in.gov.br/materia/-/asset\\_publisher/Kujrw0TZC2Mb/content/id/22793545/do1-2016-05-06-portaria-no-369-de-5-de-maio-de-2016-22793435](https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/22793545/do1-2016-05-06-portaria-no-369-de-5-de-maio-de-2016-22793435) > Acesso em 04.09.2019.

\_\_\_\_\_. **Portaria MEC 981** de 25 de agosto de 2016. Revoga a Portaria MEC nº 369, de 05 de maio de 2016, e dá outras providências. Diário Oficial da União, seção 1, p. 16, publicado em 26 de agosto de 2016. Disponível em [http://portal.mec.gov.br/index.php?option=com\\_docman&view=download&alias=47621-portaria-981-02set-pdf&category\\_slug=agosto-2016-pdf&Itemid=30192](http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=47621-portaria-981-02set-pdf&category_slug=agosto-2016-pdf&Itemid=30192) > Acesso em 04.09.2019.

\_\_\_\_\_. **Portaria MEC 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, seção 1, p.23, publicado em 20 de abril de 2017. Disponível em [https://www.in.gov.br/materia/-/asset\\_publisher/Kujrw0TZC2Mb/content/id/20173409/do1-2017-04-20-portaria-n-564-de-19-de-abril-de-2017-20173356](https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/20173409/do1-2017-04-20-portaria-n-564-de-19-de-abril-de-2017-20173356) > Acesso em 04.09.2019.

\_\_\_\_\_. **Resolução CNE nº 2** de 22 dezembro de 2017. Institui e orienta a implantação da Base Nacional Comum Curricular, a ser respeitada obrigatoriamente ao longo das etapas e respectivas modalidades no âmbito da Educação Básica. Diário Oficial da União, seção 1, p.

41-44 publicado em 22 de dezembro de 2017. Disponível em < [http://basenacionalcomum.mec.gov.br/images/historico/RESOLUCAOCNE\\_CP222DEDEZE\\_MBRODE2017.pdf](http://basenacionalcomum.mec.gov.br/images/historico/RESOLUCAOCNE_CP222DEDEZE_MBRODE2017.pdf)> Acesso em 10.10.2019.

\_\_\_\_\_. **Decreto 9.432** de 29 de junho de 2018. Regulamenta a Política Nacional de Avaliação e Exames da Educação Básica. Disponível em < [http://www.planalto.gov.br/ccivil\\_03/\\_Ato2015-2018/2018/Decreto/D9432.htm](http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Decreto/D9432.htm) > Acesso em 20.10.2019.

\_\_\_\_\_. **Portaria MEC 1.100** de 26 de dezembro de 2018. Estabelece as diretrizes para a realização do Sistema de Avaliação da Educação Básica - Saeb no ano de 2019. Diário Oficial da União, seção 1, p. 132, publicado em 28 de dezembro de 2018. Disponível em [https://www.in.gov.br/materia/-/asset\\_publisher/Kujrw0TZC2Mb/content/id/57217372](https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/57217372)> Acesso em 23.10.2019.

\_\_\_\_\_. **Portaria MEC 458** de 5 de maio de 2020. Institui normas complementares necessárias ao cumprimento da Política Nacional de Avaliação da Educação Básica. Diário Oficial da União, seção 1, p. 57, publicado em 06 de maio de 2020. Disponível em <https://www.in.gov.br/web/dou/-/portaria-n-458-de-5-de-maio-de-2020-255378342> > Acesso em 03.08.2020.

BROOKE, Nigel; ALVES, Maria Teresa G.; OLIVEIRA, Lina Kátia M. (orgs). **A Avaliação da Educação Básica: a experiência brasileira**. Belo Horizonte, Fino Traço, 2015.

BROOKE, Nigel; SOARES, Francisco (orgs.). **Pesquisa em eficácia escolar: origem e trajetórias**. Belo Horizonte: Editora UFMG, 2008.

BRUNS, Barbara; LUQUE, Javier. **Professores Excelentes; como melhorar a aprendizagem dos estudantes na América Latina e no Caribe**. Washington D.C., Grupo Banco Mundial, 2014.

BRZEZINSKI, Iria. **Tramitação e desdobramento da LDB/1996: embates entre projetos antagônicos de sociedade e de educação**. Trabalho, Educação e Saúde, Rio de Janeiro, v. 8, n. 2, p. 185-206, jul./out. 2010.

BUSTELO, M. **La evaluación y la metaevaluación: dos procesos de la estrategia de gestión de la calidad de la educación superior centroamericana**. Disponível em < [www.gc21.de/ibt/site/csuca/.../lametaevaluaciondel%20SICEVAES](http://www.gc21.de/ibt/site/csuca/.../lametaevaluaciondel%20SICEVAES) > Acesso em 28.04.2020.

CALDERÓN, Adolfo I.; BORGES, Regilson M. **Avaliação Educacional no Brasil: da transferência cultural à avaliação emancipatória**. Brazilian Geographical Journal: Geosciences and Humanities research médium, Ituiutaba, v. 4, Special Issue 1, p. 259-274, jul./dez. 2013.

CAMELO, Rafael. **A educação privada em São Paulo: expansão e perspectivas**. 1ª Análise, SEAD, Fundação Sistema Estadual de Análise de Dados, São Paulo, n. 19, out. 2014.

CAPPELLETTI, Isabel F. **Os conflitos na relação avaliação e qualidade da educação**. Educar em Revista, Curitiba, Brasil, Edição Especial n. 1/2015, p. 93-107.

CARVALHO, Rosângela T. **A qualidade social da educação básica no discurso do ministério da Educação.** Revista Estudos de Sociologia, v. 2, n. 18, p.1-18 mar. 2012.

CASTRO, Maria Helena G. **A consolidação da política de avaliação básica no Brasil.** Meta: Avaliação, Rio de Janeiro, v. 1, n. 3, p.271-296, set./dez. 2009.

CASTRO, Maria Helena G. **O Saeb e a agenda de reformas educacionais: 1995 a 2002.** Revista Em Aberto, Brasília, v. 26, n. 96, p. 85-98, 2016.

CHIZZOTTI, Antonio. **Pesquisa em Ciências Humanas e Sociais.** 11 ed., São Paulo, Cortez, 2010.

CHAER, Galdino; DINIZ, Rafael Rosa P.; RIBEIRO, Elisa A. **A técnica do questionário na pesquisa educacional.** Revista Evidência, Araxá, v. 7, n. 7, p. 251-266, 2011.

COELHO, Maria Inês de M. **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, abr./jun. 2008.

COTERA, Aurora T.; MATAMOROS, Ariana A. **Proposta para implementar la Metaevaluación em la Escuela de Ciencias de la Educación de la Universidad Estatal a Distancia.** Revista Calidad en la Educación Superior, Costa Rica, 3ª edición, v. 2, n. 1, p. 247-265, Noviembre, 2011.

CURY, Carlos Roberto J. **Do Público e do Privado na Constituição de 1988 e nas leis educacionais.** Educação e Sociedade, Campinas, v. 39, n. 145, p. 870-889, out./dez., 2018.

DANTAS, Lys Maria V. **As contribuições das políticas de avaliação educacional em larga escala: o caso da avaliação da aprendizagem na Bahia.** 2009. Tese. Programa de Pesquisa e Pós-Graduação, Universidade Federal da Bahia, 2009. Orientador: Prof. Dr. Robert E. Verhine.

DAVIDSON, E. Jane. **Principios básicos de la evaluación para la acción.** Obteniendo respuestas sucintas a las preguntas más importantes. Real Evaluacion Ltd. E-Book, 2013.

DAVOK, Delsi F. **Modelo de meta-avaliação de processos de avaliação da qualidade de cursos de graduação.** Tese, Programa de Pós-Graduação em Engenharia da Produção, Universidade Federal de Santa Catarina, 2006. Orientador: Prof. Dr. Jair dos Santos Lapa.

DEMO, Pedro. **Pesquisa: princípio científico e educativo.** 14 ed. São Paulo, Cortez, 2011.

DIAS SOBRINHO, J. **Avaliação: políticas educacionais e reformas da educação superior.** São Paulo, Cortez, 2003.

\_\_\_\_\_. **Avaliação ética e política em função da educação como direito público ou como mercadoria?** Revista Educação & Sociedade, Campinas, vol. 25, n. 88, p. 703-725, out. 2004.

DIAS SOBRINHO, J. **Avaliação educativa: produção de sentidos com formação de valor.** Revista Avaliação, Campinas; Sorocaba, v. 13, n. 1, p. 193-207, mar. 2008.

DIAS SOBRINHO, J.; RISTOFF, Dilvo I.(org.). **Avaliação e compromisso público: a educação superior em debate.** Florianópolis, Insular, 2003.

DÍAZ, Luis F. **La Metaevaluación y su método.** Revista Ciencias Sociales, 92-93, p. 171-193, 2001.

DOURADO, Luiz F. **Avaliação do Plano Nacional de Educação 2001-2009:** questões estruturais e conjunturais de uma política. Educação & Sociedade, Campinas, v. 31, n. 112, p. 677-705, jul./set. 2010.

\_\_\_\_\_. **Educação de Qualidade.** Em Aberto, Brasília, v. 33, n. 109, p. 179-187, set./dez. 2020.

DOURADO, Luiz F.; OLIVEIRA, João F. **Qualidade da educação:** perspectivas e desafios. Cadernos Cedes, Campinas, v. 29, n. 78, p. 201-215, mai./ago. 2009.

DOURADO, Luiz F.; OLIVEIRA, João F.; SANTOS, Catarina de A. **A Qualidade da Educação:** conceitos e definições. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, Série Documental Textos para Discussão, v. 24, Brasília, 2007.

DRAIBE, Sônia Maria. **Avaliação de implementação:** esboço de uma metodologia de trabalho em políticas públicas. In: BARREIRA, Maria Cecília R.; CARVALHO, Maria do Carmo de. (Orgs). Tendências e Perspectivas na Avaliação de Programas Sociais. São Paulo: IEI/ PUC-SP, 2001.

DUTRA, Fillippe R. **O SIMAVE e sua relação com o trabalho docente:** percepções de diferentes atores do cotidiano escolar. Dissertação – Programa de Pós-graduação em Educação Universidade Federal de Viçosa. Viçosa, 2018. Orientadora: Alvanize Valente Fernandes Ferenc.

ELLIOT, Ligia G. **Meta-avaliação:** das abordagens às possibilidades de aplicação. Ensaio: avaliação e políticas públicas em educação, Rio de Janeiro, v. 19, n. 73, p. 941-964, out./dez. 2011.

ESQUINSANI, Rosimar S. S.; DAMETTO, Jarbas. **A multiplicidade de sentidos e o condicionamento político da noção de qualidade em educação básica.** Revista brasileira de estudos pedagógicos, Brasília, v. 99, n. 252, p. 294-312, mai./ago. 2018.

FERNANDES, Domingos. **Algumas reflexões acerca dos saberes dos alunos em Portugal.** Educação & Sociedade, Campinas, v. 29, n. 102, p. 275-296, jan./abr. 2008.

FERNANDES, Domingos. **Avaliar para aprender:** fundamentos, práticas e políticas. São Paulo, Editora UNESP, 2009.

FERNANDES, Domingos. **Para uma compreensão das relações entre avaliação, ética e política pública.** Revista Educação PUC-Camp., Campinas, v. 23, p. 19-36, jan./abr., 2018.

FERNANDES, Margarida R. **A mudança de paradigma na avaliação educacional.** Educação, Sociedade & Culturas, Porto, n. 9, p. 7-32, 1998.

FERNANDES, Reynaldo. **A universalização da avaliação e a criação do Ideb:** pressupostos e perspectivas. Em aberto, Brasília, v. 29, n. 96, p. 99-111, maio/ago., 2016.

FERNANDES, Reynaldo; GREMAUD, Amaury. **Qualidade da Educação:** avaliação, indicadores e metas. In: VELOSO, Fernando *et al.* (Org.). Educação básica no Brasil: construindo o país do futuro. Rio de Janeiro: Elsevier, 2009, p.213-238. Disponível em: <[www.cps.fgv.br/ibrecps/rede/seminario/reynaldo\\_paper.pdf](http://www.cps.fgv.br/ibrecps/rede/seminario/reynaldo_paper.pdf)>. Acesso em: 20 ago. 2018.

FIGUEIRÓ, A. C.; THULER, L. C.; DIAS, A. L. F. **Padrões internacionais dos estudos de linha de base.** In: HARTZ, Z. M. A.; FELIBERTO, E.; SILVA, L. M. V. (Org.). Meta-avaliação da atenção básica à saúde: teoria e prática. Rio de Janeiro: Ed. Fiocruz, 2008.

FLACH, Simone F. **Qualidade social da educação:** uma concepção na perspectiva democrática. In: 6ª Jornada da Faculdade de Educação, UNICAMP, Campinas, 2013.

FLICK, Uwe. **Desenho da pesquisa qualitativa.** Porto Alegre, Artmed, 2009.

FORTUNATO, Raquel P. **Ensino Fundamental de nove anos: para além do ajuste legal.** Dissertação. Programa de Pós-graduação em Educação, Universidade de Passo Fundo, 2014. Orientador: Júlio Cesar Godoy Bertolin.

FRANCO, Creso. **Quais as contribuições da avaliação para as políticas educacionais?** In: FRANCO, Creso; BONAMINO, Alicia; BESSA, Nícia (orgs). Avaliação da educação básica: pesquisa e gestão. Rio de Janeiro: Ed. PUC-Rio; São Paulo: Loyola, 2004, p.45-63.

FRANCO, Maria Laura P.B. **Pressupostos epistemológicos da avaliação educacional.** Cadernos de Pesquisa, v. 73, p. 63-69, 1990.

FREIRE, Paulo. **Pedagogia da Autonomia:** Saberes Necessários à Prática Educativa. 36ª ed., Paz e Terra, Coleção Saberes, São Paulo, 1996.

FREITAS, Dirce Nei T. **Avaliação da Educação Básica no Brasil:** características e pressupostos. In: BAUER, Adriana; GATTI, Bernadete A.; TAVARES, Marialva R. (org). Ciclo de Debates: Vinte e cinco anos de avaliação de sistemas educacionais no Brasil – origens e pressupostos. Florianópolis, Insular, 2013.

FREITAS, L. C.; DE SORDI; M. R. L.; MALAVASI, M. M. S., FREITAS, H. C. L. **Avaliação Educacional:** caminhando na contramão. 7 ed., Petrópolis, Vozes, 2017.

FURTADO, Juarez P. **A avaliação de programas e serviços.** In: CAMPOS, Gastão W.S.; MINAYO, Maria C. S., (org). Tratado de Saúde Coletiva. São Paulo: Hucitec; 2006.

FURTADO, Juarez P.; LAPERRIÈRE, Hélène. **Parâmetros e paradigmas em meta-avaliação:** uma revisão exploratória e reflexiva. Ciência & Saúde Coletiva, v. 17, p. 695-705, 2012.

GADELHA, M. G. **Elicitação de requisitos baseada em meta-avaliação:** o caso dos centros de referência social. 2010. 99 f. Dissertação (Mestrado em Gestão do Conhecimento e Tecnologia da Informação) – Universidade Católica de Brasília, Brasília, DF, 2010.

GADOTTI, Moacir. **Qualidade na Educação: uma nova abordagem.** Texto apresentado ao Congresso de Educação Básica: Qualidade na Aprendizagem. Rede Municipal de Ensino de Florianópolis, 2013.

GALLANI, Mariângela R. B. **Avaliação da Qualidade da Educação: aspectos críticos para a administração escolar.** Dissertação: Faculdade de Educação da Universidade de São Paulo. São Paulo, 2012. Orientador: Vitor Henrique Paro.

GARCIA, Regina L. (org.). **Para quem pesquisamos para quem escrevemos: o impasse dos intelectuais.** São Paulo, Cortez, 2011.

GATTI, Bernardete A. **Avaliação Educacional no Brasil: pontuando uma história de ações.** EccoS Revista Científica, São Paulo, v. 4, n. 1, p. 17-41, 2002.

GATTI, Bernardete A. **Avaliação de sistemas educacionais no Brasil.** Sisífo. Revista de Ciências da Educação, n. 9, pp 7-18, mai-ago, 2009.

GATTI, Bernardete A. **Possibilidades e fundamentos de avaliações em larga escala: primórdios e perspectivas contemporâneas.** In: BAUER, A. GATTI, B.A., TAVARES, M. Ciclo de Debates: Vinte e Cinco Anos de Avaliações de Sistemas Educacionais no Brasil. São Paulo: Editora Insular/ FCC, 2013.

GATTI, Bernardete A. **Avaliação de sistemas educacionais no Brasil.** In: BROKE, Nigel, ALVES, Maria Teresa G.; OLIVEIRA, Lina Kátia M.(org.). Avaliação da educação básica: a experiência brasileira. Belo Horizonte, Fino Traço, 2015.

GATTI, Bernardete A; VIANNA, Heraldo; DAVIS, Cláudia. **Problemas e Impasses da Avaliação de Projetos e Sistemas Educacionais: Dois Casos Brasileiros.** Estudos em Avaliação Educacional, n 4, p. 7-26, 1991.

GHEDIN, Evandro; FRANCO, Maria Amélia S. **Questões de método na construção da pesquisa em educação.** 2 ed., São Paulo, Cortez, 2011.

GIMENES, Nelson Antônio S. **Um estudo meta-avaliativo de uma experiência de autoavaliação em uma Instituição de Ensino Superior.** Dissertação. Pontifícia Universidade Católica de São Paulo, 2006. Orientadora: Bernardete Angelina Gatti.

GIMENES, Nelson Antônio S. **Estudo meta-avaliativo de um programa de autoavaliação de uma Instituição de Educação Superior no Brasil.** Estudos em Avaliação Educacional, v. 18, n. 37, p. 217-243, 2007.

GIMENES, N.; SILVA, V.; PRÍNCIPE, L.; LOUZANO P.; MORICONI, G. **Além da Prova Brasil: Investimento em Sistemas próprios de avaliação externa.** Estudos em Avaliação Educacional, São Paulo, v. 4, n. 55, p. 12-32, abr./ago. 2013.

GOMES, Candido A. **O tema da avaliação educacional na Constituinte de 1988 e na Lei de Diretrizes e Bases da Educação Nacional de 1996.** Em Aberto, Brasília, v. 29, n. 96, p. 53-69, maio/ago. 2016.

GOMES, Romeu. Análise e interpretação de dados em pesquisa qualitativa. *In*: MINAYO, Maria C.(org.) **Pesquisa Social: Teoria, método e criatividade**. 28 ed. Petrópolis, Vozes, 2009.

GUBA, Egon G.; LINCOLN, Yvonna S. **Avaliação de quarta geração**. Tradução Beth Honorato. Campinas, Editora da Unicamp, 2011.

GUERRA, Isabel C. **Pesquisa Qualitativa e Análise de Conteúdo** – sentidos e formas de uso. 5 ed., Parede, Portugal, Príncipeia, 2014.

GREGO, Sônia Maria D. A meta-avaliação como referencial de análise e de reflexão. *In*: SGUISSARDI, Valdemar (org). **Avaliação Universitária em Questão: reformas do Estado e da Educação Superior**. Campinas, Autores Associados, 1997.

HÖFLING, Eloisa de M. Estado e Políticas (Públicas) sociais. **Cadernos Cedes**, ano XXI, n. 55, p. 30-41, 2001.

HARTZ, Zulmira Maria de A. (org). **Avaliação em Saúde: dos modelos conceituais à prática na análise da implantação de programas** [online]. Rio de Janeiro: Editora FIOCRUZ, 1997. Disponível em SciELO Books.

HARTZ, Z. M. A.; FELIBERTO, E.; SILVA, L. M. V. (Org.). **Meta-avaliação da atenção básica à saúde: teoria e prática**. Rio de Janeiro: Ed. Fiocruz, 2008.

HEDLER, H. C.; TORRES, C. V. **Meta-avaliação de auditorias de natureza operacional do Tribunal de Contas da União**. *RAC*, Curitiba, v. 13, n. 3, p. 469-486, 2009. Disponível em: <<http://www.scielo.br/pdf/rac/v13n3/v13n3a08.pdf>>.

HORTA NETO, João L. **Um olhar retrospectivo sobre a avaliação externa no Brasil: das primeiras medições em educação até o SAEB de 2005**. *Revista Iberoamericana de Educación, Organización de Estados Iberoamericanos para la Educación la Ciencia y la Cultura*, n. 42/5, abril, 2007.

\_\_\_\_\_. **Avaliação externa de escolas e sistemas: questões presentes no debate sobre o tema**. *Revista Brasileira de Estudos Pedagógicos*, Brasília, v. 91, n. 227, p. 84-104, jan./abr., 2010.

\_\_\_\_\_. **Avaliações Educacionais e seus reflexos em ações federais e na mídia eletrônica**. *Estudos em Avaliação Educacional*, Fundação Carlos Chagas, São Paulo, v. 25, n. 58, p. 172-201, set/dez 2014.

\_\_\_\_\_. **Avaliação Educacional no Brasil para além dos testes cognitivos**. *Revista Educação PUC-Camp*, Campinas, 23 (1) p. 37-53, jan./abr. 2018.

HORTA NETO, João L.; JUNQUEIRA, Rogério D.; OLIVEIRA, Adolfo S. **Do Saeb ao Sinaeb: prolongamentos críticos da avaliação da educação básica**. *Em Aberto*, Brasília, v. 29, n. 96, p. 21-37, maio/ago. 2016.

IBGE, Instituto Brasileiro de Geografia e Estatística. **Síntese dos Indicadores Sociais – SIS 2019**. Disponível em <https://www.ibge.gov.br/estatisticas/sociais/populacao.html> > Acesso em 19.05.2020.

KLEIN, Ruben. **Utilização da Teoria de Resposta ao Item no Sistema Nacional de Avaliação da Educação Básica (SAEB)**. In: BROOKE, Nigel; ALVES, Maria Teresa G.; OLIVEIRA, Lina Kátia M. (orgs). *A Avaliação da Educação Básica: a experiência brasileira*. Belo Horizonte, Fino Traço, 2015.

LETICHEVSKY, Ana C.; VELLASCO, Marley, M. B. R; SOUZA, Reinaldo C. **A categoria precisão na avaliação e na meta-avaliação: aspectos práticos e teóricos em um novo enfoque**. *Ensaio: Avaliação e Políticas Públicas em Educação*, v. 19, n. 73, pp. 941-964, 2005.

LETICHEVSKY, Ana C.; VELLASCO, Marley, M. B. R; TANSCHKEIT, Ricardo. **Um sistema fuzzy de suporte à decisão para meta-avaliação uma abordagem e um estudo de caso desenvolvido no Brasil**. *Ensaio: Avaliação e Políticas Públicas em Educação*, v. 15, n. 56, pp. 447-462, 2007.

LÜDKE, Menga; ANDRÉ, Marli E.D.A. **Pesquisa em Educação: abordagens qualitativas**. São Paulo, E.P.U., 1986.

LUCKESI, Cipriano C. **Avaliação da aprendizagem na escola: reelaborando conceitos e recriando a prática**. 2. ed. Salvador: Malabares Comunicação e Eventos, 2005.

LUCKESI, Cipriano C. **Avaliação da aprendizagem: visão geral**. Entrevista concedida ao Jornalista Paulo Camargo, publicado no caderno do Colégio Uirapuru, Sorocaba, estado de São Paulo, por ocasião da Conferência: Avaliação da Aprendizagem na Escola, Colégio Uirapuru, Sorocaba, SP, 8 de outubro de 2005. Disponível em: <http://www.intaead.com.br/webinterativo/didatica/arq/19.cipriano.htm>. Acesso em 3 de março de 2022.

LUCKESI, Cipriano C. **Avaliação em Educação: questões epistemológicas e práticas**. São Paulo, Cortez, 2018.

LUÍS, Suzana M. B. **Escrevendo a avaliação: a escrita de diários como exercício avaliativo**. Tese – Programa de Pós-graduação em Educação Conhecimento e Inclusão Social da Universidade Federal de Minas Gerais. Belo Horizonte, 2007. Orientadora: Ana Lúcia Amaral.

KUHN, Thomas S. **A Estrutura das Revoluções Científicas**. 5 ed., São Paulo, Editora Perspectiva, 1998.

MACEDO-SILVA, Rosilania; SILVA, Josiete G. F.; SANTOS, Maria R.F. **A organização dos sistemas de ensino no Brasil**. 2017. Disponível em <<http://www.maceio.al.gov.br/wp-content/uploads/2017/09/pdf/2017/09/9-A-ORGANIZA%C3%87%C3%83O-DOS-SISTEMAS-DE-ENSINO-NO-BRASIL.pdf>> Acesso em 13. 01. 2020.

MACHADO, Cristiane. **Possíveis contribuições da avaliação externa para a gestão educacional**. Revista on line de Política e Gestão Educacional, Araraquara, n. 12, p. 22–34, 2012. DOI: 10.22633/rpge.v0i12.9327. Disponível em: <https://periodicos.fclar.unesp.br/rpge/article/view/9327>. Acesso em: 12 set. 2021.

MALUF, Mônica M. B. **Sistema Nacional de Avaliação da Educação Básica no Brasil: análise e proposições**. Texto apresentado como exigência do curso “Formación de Analistas de Políticas Educativas” Projeto REDUC/BID, maio 1996. Disponível em <<http://publicacoes.fcc.org.br/ojs/index.php/eae/article/view/2280>> Acesso em 25.10.2019.

MANFIO, Aline. **Avaliação em larga escala e qualidade de ensino: análise da produção em periódicos qualificados (1995-2012)**. Texto apresentado no XI Congresso Nacional de Educação EDUCERE, Pontifícia Universidade Católica do Paraná, set. 2013. Disponível em <[https://educere.bruc.com.br/CD2013/pdf/11773\\_6427.pdf](https://educere.bruc.com.br/CD2013/pdf/11773_6427.pdf).> Acesso em 23.03.2020.

MARCONI, Marina de A.; LAKATOS, Eva M. **Fundamentos de Metodologia Científica**. 8 ed., São Paulo, Atlas, 2017.

MARÇAL RIBEIRO, Paulo R. **História da Educação Escolar no Brasil: notas para uma reflexão**. Paidéia, FFCLRP – USP, Ribeirão Preto, n. 4, fev./jul., 1993.

MATOS, Daniel A.S.; JARDILINO, José R. L. **Os conceitos de concepção, percepção, representação e crença no campo educacional: similaridades, diferenças e implicações para a pesquisa**. Educação & Formação, Fortaleza, v. 1, n. 3, p. 20-31, set./dez. 2016.

MEDEIROS, Jonas; JANUÁRIO, Adriano. **A nova classe trabalhadora e a expansão da escola privada nas periferias da cidade de São Paulo**. In: REUNIÃO NACIONAL DA ANPOCS, 38., 2014. Anais... Caxambu. Caxambu: Anpocs, p. 1-30, 2014.

MICHAELIS. **Dicionário Brasileiro da Língua Portuguesa**. Disponível em <https://michaelis.uol.com.br/moderno-portugues/> Acesso em 03.10.2021

MINAYO, Maria C. **O desafio do conhecimento: pesquisa qualitativa em saúde**. 12 ed. São Paulo, Hucitec, 2010.

\_\_\_\_\_. **Análise qualitativa: teoria, passos e fidedignidade**. Ciência & Saúde Coletiva, v. 17, p. 621-626, 2012.

MINAYO, Maria C.(org.) **Pesquisa Social: Teoria, método e criatividade**. 28 ed. Petrópolis, Vozes, 2009.

MINAYO, Maria C.; SANCHES, Odécio. **Quantitativo-Qualitativo: oposição ou complementaridade**. Caderno de Saúde Pública, Rio de Janeiro, n. 9 (3), P. 239-262, jul./set., 1993.

MINHOTO, M. A. **Política de Avaliação da Educação Brasileira: Limites e Perspectivas**. In: SOUZA, R. A.; GOUVEIA, A. B.; TAVARES. T. M. (Org.) Políticas Educacionais: Conceitos e Debates, Curitiba: Appris, 2016.

MORAES, Maria C. **O paradigma educacional emergente: implicações na formação do professor e nas práticas pedagógicas.** Em Aberto, Brasília, ano 16, n. 70, p. 57-69, abr./jun., 1996.

\_\_\_\_\_. **O paradigma educacional emergente.** Campinas, Papirus, 1997.

MOREIRA, Sebastiana Ap. **Avaliação do Programa Mais Educação para as escolas municipais de Rio Verde – GO: dos números do Ideb à percepção qualitativa dos sujeitos.** Dissertação, Programa de Pós-Graduação em educação da Universidade de Passo Fundo, 2013. Orientador: Prof. Dr. Júlio César Godoy Bertolin.

MOROSINI, Marília C. **Qualidade na Educação Superior: tendências do século.** Estudos em Avaliação Educacional, São Paulo, v. 20, n. 43, maio/ago. 2009.

NAGLE, Jorge. **Educação e Sociedade na Primeira República.** São Paulo: EPU; Rio de Janeiro: Fundação Nacional de Material Escolar, 1974.

NARDI, Elton L.; SCHNEIDER, Marilda P.; RIOS, Mônica P. G. **Qualidade na Educação Básica: ações e estratégias dinamizadoras.** Educação & Realidade, Porto Alegre, v. 39, n. 2, p. 359-390, abr./jun., 2014.

OLIVEIRA, Romualdo P.; ARAÚJO, Gilda C. **Qualidade do ensino: uma nova dimensão da luta pelo direito à educação.** Revista Brasileira de Educação, Rio de Janeiro, n. 28, p. 5-24, jan./abr., 2005.

OLIVEIRA, Romualdo P. et al. **Análise das desigualdades intraescolares no Brasil.** Estudos e Pesquisas Educacionais, v. 4, p. 19, 2013.

OLIVEIRA, Cleide P. **Avaliação institucional na educação básica: limitações e possibilidades.** Dissertação (mestrado profissional) Programa de Pós-graduação Gestão e Tecnologias Aplicadas à Educação – GESTEC da Universidade do Estado da Bahia, Salvador, 2013. Orientadora: Prof<sup>ª</sup> Dr<sup>ª</sup> Elisabete Conceição Santana.

PENNA FIRME, Thereza. **Os avanços da avaliação no século XXI.** Revista Educação Geográfica em Foco, [S.l.], v. 1, n. 1, dec. 2018. ISSN 2526-6276. Disponível em: <http://periodicos.puc-rio.br/index.php/revistaeducacaogeograficaemfoco/article/view/810>  
Acesso em: 04 oct. 2021.

PENNA FIRME, Thereza; LETICHEVSKY, Ana C. **O Desenvolvimento da Capacidade de Avaliação no Século XXI: enfrentando o desafio através da meta-avaliação.** Meta: Avaliação, Rio de Janeiro, v. 2, n. 5, p. 180-195, mai./ago. 2010.

PEREIRA, T. L. P. C. **Avaliação do edital de seleção de professores de primeiro segmento do Colégio Pedro II.** 2010. Dissertação (Mestrado Profissional em Avaliação) – Programa de Pós-Graduação em Avaliação, Fundação Cesgranrio, Rio de Janeiro, 2010. Orientadora: Profa. Dra. Ligia Gomes Elliot.

PEREZ, Marinês O. **Meta-avaliação do sistema de avaliação institucional da Escola Técnica Estadual Parque da Juventude do Centro Paula Souza.** Dissertação – Programa de Mestrado em Gestão e Prática Educacionais da Universidade Nove de Julho, São Paulo, 2014. Orientadora: Prof<sup>ª</sup> Dr<sup>ª</sup> Amélia Silveira.

PILATI, Orlando. **Sistema Nacional de Avaliação da Educação Básica (SAEB)**. In: BROOKE, Nigel; ALVES, Maria T. G.; OLIVEIRA, Lina K. M.(org.) A avaliação da Educação Básica: a experiência brasileira. Belo Horizonte, Fino Traço, 2015.

PINTO, Rodrigo S. **Meta-avaliação: uma década do processo de avaliação institucional do SINAES**. Tese – Programa de Pós-Graduação em Administração da Universidade Federal de Santa Catarina. Florianópolis, 2015. Orientador: Prof. Dr. Pedro Antônio de Melo.

PESTANA, Maria I. **O sistema de avaliação brasileiro**. Revista Brasileira de Estudos Pedagógicos, Brasília, v. 79, n. 191, p. 65-73, jan./abr. 1998.

\_\_\_\_\_. **Trajetória do Saeb: criação, amadurecimento e desafios**. Em Aberto, Brasília, v. 29, n. 96, p. 71-84, maio/ago. 2016.

RABELO, Mauro. **Avaliação educacional: fundamentos, metodologia e aplicações no contexto brasileiro**. Rio de Janeiro: SBM, 2013.

RABELO, Mauro L.; MARINHO-ARAÚJO, Claisy M. **Avaliação educacional: a abordagem por competências**. Avaliação, Campinas, Sorocaba, SP, v.20, n.2, p. 443-466, jul. 2015.

RIBEIRO, Alessandro P.; GESSINGER, Rosana M. **Instrumentos de coletas de dados em pesquisas: questionamentos e reflexões**. In: LIMA, Valdeez M. R.; HARRES, João B. S.; PAULA, Marlúbia C. Caminhos da pesquisa qualitativa no campo da educação em Ciências: pressupostos, abordagens e possibilidades. Porto Alegre, EdiPUCRS, 2018.

RIOS, Mônica P. G. **Avaliação Educacional: uma conversa com inspiração em Paulo Freire**. In: DICKMANN, Ivo; DICKMANN, Ivanio, (Orgs.). 100 anos com Paulo Freire. 1 ed. Chapecó, Livrologia, v. 3, 2021.

RISCAL, José R.; LUIZ, Maria C. **Gestão democrática e a análise da avaliação em larga escala: o desempenho de escolas públicas no Brasil**, São Carlos, Pixel, 2016.

RISTOFF, Dilvo I. **Avaliação de Programas Educacionais: discutindo padrões**. Avaliação: Revista da Avaliação da Educação Superior, 5(4). 2000. Recuperado de <<http://periodicos.uniso.br/ojs/index.php/avaliacao/article/view/1124>> Acesso em 23.04.2020.

RODRIGUES, Maria S. S. **Autoavaliação e planejamento participativo no âmbito da Universidade Federal do Ceará: em busca de saberes e da mudança de cultura institucional**. Tese – Programa de Pós-Graduação da Faculdade de Educação da Universidade Federal do Ceará, 2007. Orientador; Dr. Wagner Bandeira Andriola.

ROTHEN, José C. **Uma pequena história da avaliação da educação a partir do caso brasileiro e francês**. In: ROTHEN, José C. SANTANA, Andréia da C. M. (org.). Avaliação da Educação: referências para uma primeira conversa. São Carlos, EduFSCar, 2018.

SANTANA, Andréia C. M. **A constituição do Estado Avaliativo e o aumento das avaliações externas: propagando um ensino desigual para todos**. In: ROTHEN, José C. SANTANA, Andréia da C. M. (org.). Avaliação da Educação: referências para uma primeira conversa. São Carlos, EduFSCar, 2018.

SANTOS, Alexandre A.; HORTA NETO, João L.; JUNQUEIRA, Rogério D. **O Sistema Nacional de Avaliação da Educação Básica (Sinaeb):** proposta para atender o disposto no Plano Nacional de Educação. Brasília, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2017.

SAUL, Ana M. **Avaliação Emancipatória:** desafio à teoria e à prática de avaliação e reformulação de currículo. 8 ed. São Paulo, Cortez, 2010.

SAVIANI, Demerval. **Educação Brasileira: Estrutura e Sistema.** São Paulo, Editora Saraiva, 1999.

SAVIANI, Demerval. **A política educacional no Brasil.** In: STEPHANOU, Maria; BASTOS, Maria H. C. (org.). *Histórias e memórias da Educação no Brasil.* Vol. III, Petrópolis, RJ, Vozes, 2005.

\_\_\_\_\_. **História da história da Educação no Brasil:** um balanço prévio e necessário. *EccoS – Revista Científica*, São Paulo, v. 10, n. especial, p. 147-167, 2008.

\_\_\_\_\_. **História das Ideias Pedagógicas no Brasil.** 3 ed. rev., Campinas, Autores Associados, 2010.

SCARINCI, Anne L. **Uma proposta para caracterizar a atuação do professor na sala de aula.** Dissertação – Instituto de Física e Faculdade de Educação da Universidade de São Paulo. São Paulo, 2006. Orientadora: Jesuína Lopes de Almeida Pacca.

SCHNEIDER, Marilda P.; NARDI, Elton L. **O IDEB e a construção de um modelo de *accountability* na educação básica brasileira.** *Revista Portuguesa de Educação*, v. 27, p. 7-28, 2014. Disponível em < <https://revistas.rcaap.pt/rpe/article/view/4295> > Acesso em 15.09.2019.

SCHULZ, Almiro. **Educação Superior Protestante no Brasil.** Engenheiro Coelho, UNASPRESS, 2003.

SCRIVEN, Michael. **Meta-evaluation revisited.** *Journal of MultiDisciplinary Evaluation*, v. 6, n. 11, p. iii-viii, january, 2009. [https://journals.sfu.ca/jmde/index.php/jmde\\_1/article/view/220](https://journals.sfu.ca/jmde/index.php/jmde_1/article/view/220) Acesso 09.06.2020.

SCRIVEN, Michael, **The meta-evaluation Checklist (new version).** Disponível em <<http://www.michaelscriven.info/papersandpublications.html>> Assesso em 27.06.2020.

SEROUR, M. G. **Meta-avaliação da avaliação institucional externa da Universidade do Grande Rio Prof. José de Souza Herdy:** Unigranrio. Dissertação (Mestrado Profissional em Avaliação) – Fundação Cesgranrio, Rio de Janeiro, 2010. Orientadora: Profa. Dra. Angela Carrancho da Silva.

SILVA, José A. **Inteligência Humana:** abordagens biológicas e cognitivas. São Paulo, Lovise, 2003.

SILVA, Maria A. **Qualidade Social da Educação Pública:** algumas aproximações. *Cadernos Cedes*, Campinas, vol. 29, n. 78, p. 216-226, maio/ago., 2009.

SILVA, Fernanda G. **A inclusão de Ciências Naturais nos sistemas de avaliação da educação: o caso do Avalia BH.** Dissertação – Mestrado Profissional em Gestão e Avaliação da Educação Pública da Universidade Federal de Juiz de Fora, Juiz de Fora, 2014. Orientadora: Prof<sup>a</sup> Dr<sup>a</sup> Lina Kátia Mesquita de Oliveira.

SICEVAES – Sistema Centroamericano de Evaluación de la calidad de la Educación Superior. **La evaluación y la metaevaluación: dos procesos de la estrategia de estión de la calidad de la educación superior centroamericana.** Febrero,, 2005 Disponível < [https://gc21.giz.de/ibt/site/csuca/ibt/cat08/web/Ejes/Ejes%20Cultura/lametaevaluaciondel\\_1%20SICEVAES.pdf](https://gc21.giz.de/ibt/site/csuca/ibt/cat08/web/Ejes/Ejes%20Cultura/lametaevaluaciondel_1%20SICEVAES.pdf) > Acesso em 01.06.2020

SIQUEIRA, Ana R. **Um segmento menos elitizado da rede privada de ensino: três escolas no município de Contagem.** 2016. Dissertação (Mestrado em Educação) – Faculdade de Educação da Universidade Federal de Minas Gerais. Belo Horizonte, 2016. Orientadora: Profa. Dra. Maria Alice Nogueira.

SOHN, Cleide Eurich; EYNG, Ana Maria. **As Políticas de Avaliação: desafios para uma educação básica de qualidade para todos.** XVI ENDIPE - Encontro Nacional de Didática e Práticas de Ensino - UNICAMP – Campinas, Junqueira & Marin Editores, Livro 3 - p.003955, 2012.

SOUTO, Adilour N. **Do ensino público ao ensino privado: uma análise da Escola Santa Terezinha em Ibiá – MG (1937-1956).** Dissertação – Programa de Pós-graduação em Educação, Universidade Federal de Uberlândia, 2012. Orientadora: Prof<sup>a</sup> Dr<sup>a</sup> Betânia de Oliveira Laterza Ribeiro.

SOUZA, E. D. **A formação dos formadores em avaliação da aprendizagem: o processo de formação inicial em debate.** 158 f. Dissertação (Mestrado em Educação: Currículo) – Pontifícia Universidade Católica, São Paulo, 2007.

SOUZA, Maria A. **O Uso dos Resultados da Avaliação Externa da Escola: relação entre os Resultados da Avaliação Externa e a Avaliação Interna dos Alunos.** In: BAUER, Adriana; GATTI, Bernadete A.; TAVARES, Marialva R. (org). Ciclo de Debates: Vinte e cinco anos de avaliação de sistemas educacionais no Brasil – origens e pressupostos. Florianópolis, Insular, 2013.

SOUSA, Sandra Zákia; ARCAS, Paulo Henrique. **Implicações da Avaliação em Larga Escala no Currículo: revelações de escolas estaduais de São Paulo.** Educação: Teoria e Prática, v. 20, n. 35, jul./dez. 2010, p. 181-199.

SOUZA, Sandra Z. **Concepções de qualidade da educação básica forjadas por meio de avaliações em larga escala.** Avaliação, Campinas; Sorocaba, v. 19, n. 2, p. 407-420, jul. 2014.

SOUZA, Sandra Z.; OLIVEIRA, Romualdo P. **Sistemas estaduais de avaliação: uso dos resultados, implicações e tendências.** Cadernos de Pesquisa, v.40, n.141, p.793-822, set./dez. 2010.

SOUZA SANTOS, B. **Um discurso sobre as ciências.** 8 ed. São Paulo, Cortez, 2018.

STAKE, Robert E. **Estudos de caso em pesquisa e avaliação educacional**. Educação e Seleção, São Paulo, n. 7, p. 5-14, 1983.

STUFFLEBEAM, Daniel L. **The Meta-evaluation Imperative**. American Journal of Evaluation, v. 22, N. 2, p. 183–209, 2001. Disponível em <<http://aje.sagepub.com/cgi/content/abstract/22/2/183>> . Acesso em 15.06.2020.

STUFFLEBEAM, Daniel L. **Meta-evaluation**. Journal of MultiDisciplinary Evaluation, v. 7, n. 15, p. 99-158, february, 2011.

SUFFLEBEAM, Daniel L.; SCHINKFIELD, Anthony J. **Evaluación Sistemática: Guia teórica y práctica**. Colección Temas de Educación. Barcelona, Paidós, 1985. Disponível <http://pdfhumanidades.com/sites/default/files/apuntes/STUFFLEBEAM%20%20evaluacion%20sistemica%20cap1>> Acesso em 12.07.2020.

TAVARES, JR. Fernando; NEUBERT, Luiz F. **A qualidade da educação e a disseminação de sistemas de avaliação**. Estudos em Avaliação Educacional, Fundação Carlos Chagas, São Paulo, v. 25, n. 58, p. 22-47, set/dez 2014.

TRAVITZKI, Rodrigo. **Enem: limites e possibilidades do Exame Nacional do Ensino Médio enquanto indicador de qualidade**. Tese – Faculdade de Educação da Universidade de São Paulo, São Paulo, 2013. Orientadora: Prof<sup>a</sup> Dr<sup>a</sup> Carlota Boto.

TREVISOL, Maria T. C.; RIOS, M. P. G.; GABRIEL, N.; DE BASTIANI, S. C. **Ações e Estratégias com incidência nas práticas pedagógicas: implicações na qualidade da educação**. Roteiro, Joaçaba, Edição Especial, p. 193-214, 2015.

TOBIAS, José A. **História da educação brasileira**. 5<sup>a</sup> ed. São Paulo, Juriscredi, 1986.

VÍCTORA, Ceres G.; KNAUTH, Daniela R.; HASSEN, Maria N. A. **Pesquisa Qualitativa em Saúde: uma introdução ao tema**. Porto Alegre, Tomo Editorial, 2000.

VIANNA, Heraldo M. **Avaliação Educacional: uma perspectiva histórica**. Estudos em Avaliação Educacional, n. 12, p. 07-24, 1995.

VIANNA, Heraldo M. **Fundamentos de um Programa de Avaliação Educacional**. Estudos em Avaliação Educacional, v.3, p. 23-38, jul./dez., 2003.

WORTHEN, Blaine R.; SANDERS, James R.; FITZPATRICK, Jody L. **Avaliação de Programas: concepções e práticas**. Tradução Dinah de Abreu Azevedo, São Paulo, Editora Gente, 2004.

## APÊNDICE

### Questões norteadoras para entrevista online semiestruturada de gestores regionais da rede.

- 1) Qual a sua visão sobre o Programa de Avaliação Externa da Rede (PAER)?
- 2) Como foi o processo de implantação do PAER na sua região administrativa?
- 3) O objetivo principal do PAER explicitados no projeto é: *“Retratar a realidade da qualidade da educação, conforme os princípios que fundamentam a filosofia de educação da rede, através de informações que incentivarão a adoção de medidas que promovam o crescimento da rede.”* Na sua visão, este objetivo foi atingido? Se não, por quê? Se sim, por quê?
- 4) Como se dá o processo de aplicação do PAER?
- 5) Qual o uso que tem sido feito dos resultados da referida avaliação?
- 6) Há articulação dos resultados do PAER com outras avaliações (internas ou externas)?
- 7) Na sua visão, quais os ganhos após 7 anos da referida avaliação?
- 8) Quais as dificuldades ou aspectos a desenvolver relativos ao referido programa de avaliação da rede?
- 9) Quais as suas sugestões para o aperfeiçoamento do PAER?
- 10) Quais as implicações do PAER no processo educativo?
- 11) Em face do propósito do estudo, há algo que você gostaria de acrescentar?

## Questões do questionário on-line para gestores das sub-regiões da rede.

5. **SEXO:** F ( ) M ( )

6. **IDADE**

- 25 a 30 anos ( )
- 31 a 35 anos ( )
- 36 a 40 anos ( )
- 41 a 45 anos ( )
- 46 a 50 anos ( )
- 51 a 55 anos ( )
- Mais de 55 anos ( )

7. **TEMPO QUE ATUA NA REDE**

- 1 a 5 anos ( )
- 6 a 10 anos ( )
- 11 a 15 anos ( )
- 16 a 20 anos ( )
- 21 a 25 anos ( )
- 26 a 30 anos ( )
- Mais de 30 anos ( )

8. **TEMPO DE ATUAÇÃO NA FUNÇÃO ATUAL**

- 1 a 5 anos ( )
- 6 a 10 anos ( )
- 11 a 15 anos ( )
- 16 a 20 anos ( )
- 21 a 25 anos ( )
- 26 a 30 anos ( )
- Mais de 30 anos ( )

9. **EDIÇÕES DO PROGRAMA QUE JÁ PARTICIPOU NESTA FUNÇÃO:**

- Menos que 3 ( )
- Mais que 4 ( )
- Todas ( )

10. **CONSIDERE PARA RESPONDER AS QUESTÕES ABAIXO A SEGUINTE LEGENDA:**

0 – Não tenho como responder; 1 – Discordo/ não; 2 – Concordo / sim

Item	0	1	2
a. Conheço os objetivos do PAER			
b. Os diretores das escolas conhecem os objetivos do PAER			
c. Os professores conhecem os objetivos do PAER			
d. Os pais conhecem o objetivo do PAER			

e. Os alunos conhecem os objetivos do PAER			
f. Considero que o PAER alcançou seus objetivos			
g. Analisei os dados das avaliações do PAER dos últimos anos do meu campo			
h. Os dados do PAER são analisados pelas equipes administrativas das unidades escolares			
i. Os resultados do PAER são analisados pelos professores			
j. É possível perceber melhoria na qualidade do ensino após a introdução do PAER na rede			
k. Houve mudanças significativas na qualidade do ensino após a introdução do PAER na rede.			
l. Em meu planejamento de trabalho há uma meta relacionada ao PAER			

**11. EM SUA OPINIÃO, QUAIS OS GANHOS (ASPECTOS DESENVOLVIDOS) DO PAER?**

**12. QUAIS OS DESAFIOS OU DIFICULDADES (ASPECTOS A DESENVOLVER) DO PAER?**

**13. QUAIS AS IMPLICAÇÕES DO PAER NO PROCESSO EDUCATIVO?**

**14. O CENÁRIO DE 2020 COM O ENSINO REMOTO E A APLICAÇÃO DO PAER TOTALMENTE ON-LINE TROUXE ALGUMA IMPLICAÇÃO PARA A QUALIDADE DO PROCESSO?**

**15. QUAIS AS SUAS SUGESTÕES PARA O APERFEIÇOAMENTO DO PAER?**