

UNIVERSIDADE FEDERAL DE JUIZ DE FORA  
PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO

Wagner Silveira Rezende

**AVALIAÇÃO EM LARGA ESCALA DA ALFABETIZAÇÃO:** os casos de Paebes Alfa,  
Proalfa e Spaece Alfa

JUIZ DE FORA

2020

UNIVERSIDADE FEDERAL DE JUIZ DE FORA  
PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO

Wagner Silveira Rezende

**AVALIAÇÃO EM LARGA ESCALA DA ALFABETIZAÇÃO:** os casos de Paebes Alfa,  
Proalfa e Spaece Alfa

Tese apresentada como requisito parcial para a obtenção do título de Doutor em Educação, área de concentração Educação Brasileira: gestão e práticas pedagógicas, pelo Programa de Pós-Graduação em Educação da Universidade Federal de Juiz de Fora.

Orientadora: Hilda Aparecida Linhares da Silva Micarello.

JUIZ DE FORA

2020

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Rezende, Wagner Silveira.

Avaliação em larga escala da alfabetização: os casos de  
Paebs Alfa, Proalfa e Spaece Alfa / Wagner Silveira  
Rezende. -- 2020.

357 p.

Orientadora: Hilda Aparecida Linhares da Silva Micarello

Tese (doutorado) - Universidade Federal de Juiz de Fora,  
Faculdade de Educação. Programa de Pós-Graduação em  
Educação, 2020.

1. Avaliação da alfabetização. 2. Avaliação em larga escala. 3.  
Paebs Alfa. 4. Proalfa. 5. Spaece Alfa. I. Micarello, Hilda  
Aparecida Linhares da Silva, orient. II. Título.

**Wagner Silveira Rezende**

**Avaliação em larga escala da alfabetização:** os casos de Paebes Alfa, Proalfa e Spaece Alfa

Tese apresentada ao Programa de Pós-graduação em Educação da Faculdade de Educação da Universidade Federal de Juiz de Fora, como requisito parcial à obtenção do título de Doutor em Educação. Área de concentração: "Educação brasileira: gestão e práticas pedagógicas".

Aprovada em 25 de maio de 2020.

BANCA EXAMINADORA

Hilda A. L. da S. Micarello  
Dra. Hilda Aparecida Linhares da Silva Micarello – Orientadora  
Universidade Federal de Juiz de Fora

p/ Hilda A. L. da S. Micarello  
Dra. Beatriz de Basto Teixeira  
Universidade Federal de Juiz de Fora

p/ Hilda A. L. da S. Micarello  
Dr. Eduardo Magrone  
Universidade Federal de Juiz de Fora

p/ Hilda A. L. da S. Micarello  
Dra. Alicia Maria Catalano de Bonamino  
Pontifícia Universidade Católica do Rio de Janeiro

p/ Hilda A. L. da S. Micarello  
Dra. Delaine Cafiero Bicalho  
Universidade Federal de Minas Gerais

Para Heitor, Antonio e Roberta.

## AGRADECIMENTOS

A Roberta Palácios, por cuidar de mim.

A Armando, Edwiges, Diva, Nathália, Rosa, Manuel, Edna, Carlos, Mariana, Fernando e Gabi, por serem família e tomarem conta dos meus filhos quando foi necessário.

À professora Hilda Micarello, por aceitar, generosa e afetuosamente, a adversidade que lhe ofereci.

À banca, que gentilmente aceitou ler essas linhas: Beatriz de Basto Teixeira, Eduardo Magrone, Alicia Bonamino, Delaine Cafiero, Tufi Machado, Ilka Schaffer, Marcelo Burgos e Naira Muylaert.

A todos que ajudaram na construção deste trabalho, nominalmente: Aida Antunes, Bárbara Sousa, Carol Canedo, Carmilva Flores, Cristina Brandão, Débora Vieira, Francisca Sales, Higor Pifano, Irene Ribeiro, Josiane Silva, Leila Martins, Léo Campos, Márcia Mariana, Maíra Portela, Mayra Moreira, Michele, Rachel Finamore, Tuko, Wal Barbosa, Wellington Silva.

Aos personagens que, nas secretarias de educação, faziam a história da avaliação da alfabetização no Brasil: Alessio da Costa Lima, Ana Lúcia Gazzola, Andressa Rocha, Márcia Campos, Maria das Graças Pedrosa Bittencourt (Fite), Maurício Holanda, Raquel Elizabete de Souza Santos, Renata da Silva Marinho, Sueli Pires.

A Manuel Palácios e Lina Kátia Mesquita de Oliveira, por terem aberto as portas do CAEd. Sem o CAEd, não haveria o que escrever.

Ao Zé Francisco, que, junto com a Kátia, me deu Descoberto para escrever.

Ao Moacyr Oliveira, por saber esperar.

*Se podes olhar, vê. Se podes ver, repara* (SARAMAGO, 2006, p.10).

## RESUMO

Esta tese tem como objetivo analisar as principais características de três programas estaduais de avaliação da alfabetização no Brasil, no período entre 2009 e 2017: o Paebes Alfa, no Espírito Santo, o Proalfa, em Minas Gerais, e o Spaece Alfa, no Ceará. A análise busca identificar semelhanças e diferenças entre esses programas a partir de dez critérios considerados importantes para o estudo de políticas de avaliação na alfabetização. A proposição desses critérios para a análise dos programas é uma das contribuições da tese. Cada um dos critérios e os motivos de sua escolha são apresentados no capítulo 4. Parte da literatura sobre políticas públicas critica o excesso de trabalhos destinados à descrição de programas, sem estabelecer vínculos com o macrocontexto político, conforme discussão apresentada no capítulo 1. Contudo, no caso da avaliação da alfabetização, chama atenção a escassez de trabalhos dessa natureza. O capítulo 3 dedica-se a analisar a produção acadêmica sobre a avaliação da alfabetização em três periódicos relevantes na área de educação no Brasil (Estudos em Avaliação Educacional; Revista Brasileira de Educação e Revista Brasileira de Estudos Pedagógicos) e no banco de teses e dissertações da Capes, constatando o pouco interesse da pesquisa em educação sobre o tema. Conforme discussão apresentada no capítulo 2, as avaliações educacionais em larga escala passaram a ocupar um espaço cada vez mais central no contexto das políticas educacionais, ao mesmo tempo em que a alfabetização tornou-se prioritária na agenda dos governos. Apesar disso, a avaliação da alfabetização permanece sem um enfrentamento adequado pela pesquisa em educação. Com base em pesquisa documental e em entrevistas com atores que desempenharam papéis importantes nos três programas analisados, espera-se, com a presente tese, apontar caminhos para a pesquisa educacional preencher essa lacuna.

Palavras-chave: Avaliação da alfabetização. Avaliação em larga escala. Paebes Alfa. Proalfa. Spaece Alfa.



## ABSTRACT

This thesis aims to analyze the main characteristics of three literacy assessment programs of the states in Brazil, between 2009 and 2017: Paebes Alfa, in Espírito Santo; Proalfa, in Minas Gerais; and Spaece Alfa, in Ceará. The analysis seeks to identify similarities and differences between these programs based on ten important criteria for the research of literacy assessment policies. The proposition of these criteria for the analysis of the programs is one contribution of the thesis. Each criterion and the reasons for their choice are presented in chapter 4. Part of literature on public policies criticizes the excess of researches aimed at describing programs, without links with the political context (this is discussed in the chapter 1). However, in the case of literacy assessment, there is little research with such objectives. The chapter 3 is dedicated to analyzing the research on the literacy assessment in three relevant educational journals in Brazil (Estudos em Avaliação Educacional; Revista Brasileira de Educação e Revista Brasileira de Estudos Pedagógicos) and in the thesis and dissertation database of Capes, realizing that there is little interest of education research on the topic. The chapter 2 discusses how large-scale educational assessment have come to occupy a central place in the context of educational policies at the same time that literacy has become a priority on government's agenda. However, the literacy assessment remains without adequate confrontation by educational research. Based on documentary research and interviews with actors who played important roles in the three programs analyzed, we hope with this thesis to present ways for educational research to fill this gap.

Keywords: Literacy assessment. Educational assessment. Paebes Alfa. Proalfa. Spaece Alfa.

## RESUMÉN

Esta tesis tiene como objetivo analizar las principales características de programas de evaluación de alfabetización de tres estados en Brasil, en el período entre 2009 y 2017: Paebes Alfa, en Espírito Santo, Proalfa, en Minas Gerais, y Spaece Alfa, en Ceará. El análisis busca similitudes e diferencias entre estos programas en base a diez criterios considerados importantes para el estudio de las políticas de evaluación de la alfabetización. La propuesta de estos criterios para el análisis de los programas es una contribución de la tesis. Cada uno de los criterios y las razones de su elección se presentan en el capítulo 4. Parte de la literatura sobre políticas públicas critica el exceso de investigación dirigida a describir programas, sin establecer vínculos con el contexto político, como se discutió en el capítulo 1. Sin embargo, en el caso de la evaluación de la alfabetización, la escasez de investigación de esta naturaleza es notable. El capítulo 3 está dedicado a analizar la producción académica sobre la evaluación de la alfabetización en tres revistas relevantes en el campo de la educación en Brasil (Estudos em Avaliação Educacional; Revista Brasileira de Educação e Revista Brasileira de Estudos Pedagógicos) y en el banco de tesis de Capes, observando el poco interés en la investigación educativa sobre el tema. Como se discutió en el capítulo 2, las evaluaciones estandarizadas han llegado a ocupar un espacio cada vez más central en el contexto de las políticas educativas, al mismo tiempo que la alfabetización se ha convertido en una prioridad en la agenda de los gobiernos. A pesar de esto, la evaluación de la alfabetización permanece sin una confrontación adecuada por parte de la investigación en educación. Basado en investigaciones documentales y entrevistas con actores que jugaron papeles importantes en los tres programas analizados, se espera, con esta tesis, señalar formas para que la investigación educativa llene este vacío.

Palabras clave: Evaluación de la alfabetización. Evaluación estandarizada. Paebes Alfa. Proalfa. Spaece Alfa.

## LISTA DE FIGURAS

Figura 1- Mapeamento dos sistemas próprios de avaliação educacional no Brasil – 1990-2018 .....	107
---	-----

## LISTA DE GRÁFICOS

Gráfico 1- Número de alunos participantes do Paebes Alfa por rede – série histórica ....	173
Gráfico 2- Número de escolas participantes do Paebes Alfa por rede – série histórica ..	174
Gráfico 3- Número de redes municipais participantes do Paebes Alfa – série histórica .....	175
Gráfico 4- Número de alunos participantes do Proalfa por rede – série histórica .....	176
Gráfico 5- Número de escolas participantes do Proalfa por rede – série histórica (para a avaliação censitária, do 3º ano do ensino fundamental) .....	177
Gráfico 6- Número de alunos participantes do Spaece Alfa por rede – série histórica ...	179
Gráfico 7- Número de escolas participantes do Spaece Alfa por rede – série histórica (para a avaliação censitária, do 3º ano do ensino fundamental) .....	179
Gráfico 8- Proficiência média – Língua Portuguesa – 2º EF – Spaece Alfa –Linha histórica .....	256
Gráfico 9- Proficiência média – Língua Portuguesa – 3º EF – Proalfa –Linha histórica .....	257
Gráfico 10- Proficiência média – Leitura – 1º, 2º e 3º EF – Paebes Alfa – Rede estadual - Linha histórica .....	258
Gráfico 11- Proficiência média – Leitura – 1º, 2º e 3º EF – Paebes Alfa – Redes municipais - Linha histórica .....	258
Gráfico 12- Proficiência média – Escrita – 1º, 2º e 3º EF – Paebes Alfa – Rede estadual - Linha histórica .....	259
Gráfico 13- Proficiência média – Escrita – 1º, 2º e 3º EF – Paebes Alfa – Redes municipais - Linha histórica .....	259
Gráfico 14- Proficiência média – Matemática – 1º, 2º e 3º EF – Paebes Alfa – Rede estadual - Linha histórica .....	260
Gráfico 15 - Proficiência média – Matemática – 1º, 2º e 3º EF – Paebes Alfa – Redes municipais - Linha histórica .....	260

## LISTA DE QUADROS

Quadro 1- Proposta tipológica para formulação e implementação de políticas com base em Deubel .....	34
Quadro 2- Entrevistas para a pesquisa, atores e temáticas .....	168
Quadro 3- Critérios para a construção de uma tipologia de políticas de avaliação da alfabetização .....	170
Quadro 4- Formato da avaliação por edição do Proalfa – 2009 a 2017 .....	176
Quadro 5- Atores com acesso aos resultados por nível de agregação – Paebes Alfa, Proalfa e Spaece Alfa .....	183
Quadro 6- Tipos de avaliação, disciplinas e séries avaliadas – por edição e programa..	199
Quadro 7- Modelos e itens por caderno de teste – por edição e por programa .....	204
Quadro 8- O Paebes Alfa segundo os critérios selecionados para análise .....	246
Quadro 9- O Proalfa segundo os critérios selecionados para análise .....	247
Quadro 10- O Spaece Alfa segundo os critérios selecionados para análise .....	247
Quadro 11- Comparação entre os três programas segundo os critérios selecionados...	263

## LISTA DE TABELAS

Tabela 1- Mapeamento da EAE – 2004-2017 .....	125
Tabela 2- Mapeamento da RBE – 2004-2017 .....	130
Tabela 3- Mapeamento da RBPE – 2004-2017 .....	133
Tabela 4- Produção de teses e dissertações sobre o Proalfa .....	143
Tabela 5- Produção de teses e dissertações sobre o Spaece Alfa .....	152
Tabela 6- Gastos totais e com a política de avaliação da secretaria estadual de educação do Espírito Santo – 2009 a 2017 .....	241
Tabela 7- Gastos totais e com a política de avaliação da secretaria estadual de educação de Minas Gerais – 2009 a 2017 .....	243
Tabela 8- Montante e percentual dos gastos da secretaria de educação do estado do Ceará com a política de avaliação educacional – 2009 a 2017 .....	244

## LISTA DE SIGLAS E ABREVIações

ABAVE – Associação Brasileira de Avaliação Educacional.

Ana – Avaliação Nacional da Alfabetização

Aneb – Avaliação Nacional da Educação Básica

Anped – Associação Nacional de Pós-Graduação e Pesquisa em Educação

Anresc – Avaliação Nacional do Rendimento Escolar

BNCC – Base Nacional Curricular Comum

Capes – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

Ceale – Centro de Alfabetização, Leitura e Escrita

CAEd – Centro de Políticas Públicas e Avaliação da Educação

Crede – Coordenadoria Regional de Desenvolvimento da Educação

EAE – Estudos em Avaliação Educacional

FCC – Fundação Carlos Chagas

IBGE – Instituto Brasileiro de Geografia e Estatística

ICMS – Imposto sobre Circulação de Mercadorias e Serviços

Ideb – Índice de Desenvolvimento da Educação Básica

Idepe – Índice de Desenvolvimento da Educação Básica de Pernambuco

Inaf – Indicador Nacional de Alfabetismo Funcional

Inep – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

MEC – Ministério da Educação

OCDE – Organização para a Cooperação e Desenvolvimento Econômico

ONU – Organização das Nações Unidas

Paic – Programa Alfabetização na Idade Certa

Pnaic – Pacto Nacional pela Alfabetização na Idade Certa

Paebes – Programa de Avaliação da Educação Básica do Espírito Santo

Pde – Plano de Desenvolvimento da Educação

Pne – Plano Nacional de Educação

PIP/ATC – Programa de Intervenção Pedagógica/Alfabetização no Tempo Certo

Pisa – *Programme for International Student Assessment* (Programa Internacional de Avaliação de Estudantes)

Pnaic – Programa Nacional de Alfabetização na Idade Certa

PNLD – Programa Nacional do Livro Didático

PPGP/CAEd – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública do CAEd/UFJF

Proalfa – Programa de Avaliação da Alfabetização

Proeb – Programa de Avaliação da Rede Pública de Educação Básica

PSDB – Partido da Social Democracia Brasileira

PT – Partido dos Trabalhadores

PUC-MG – Pontifícia Universidade Católica de Minas Gerais

PUC-RJ – Pontifícia Universidade Católica do Rio de Janeiro

RBE – Revista Brasileira de Educação

RBPE – Revista Brasileira de Estudos Pedagógicos

Sadeam – Sistema de Avaliação do Desempenho Educacional do Amazonas

Saeb – Sistema de Avaliação da Educação Básica

Saego – Sistema de Avaliação Educacional do Estado de Goiás

Saems – Sistema de Avaliação da Educação Rede Pública de Mato Grosso do Sul

Saepe – Sistema de Avaliação do Estado de Pernambuco

Saers – Sistema de Avaliação do Rendimento Escolar do Rio Grande do Sul

Saethe – Sistema de Avaliação Educacional de Teresina

Saresp – Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo

Seape – Sistema de Avaliação do Estado do Acre

Seduc – Secretaria de Estado de Educação

Simave – Sistema Mineiro de Avaliação da Educação Pública



Spaace – Sistema Permanente de Avaliação da Educação Básica do Ceará

SRE – Superintendência Regional de Ensino

UCP – Universidade Católica de Petrópolis

UERJ – Universidade do Estado do Rio de Janeiro

UFC – Universidade Federal do Ceará

UFES – Universidade Federal do Espírito Santo

UFJF – Universidade Federal de Juiz de Fora

UFMG – Universidade Federal de Minas Gerais

UFPA – Universidade Federal do Pará

UFRGS – Universidade Federal do Rio Grande do Sul

UFSJ – Universidade Federal de São João Del Rei

UFU – Universidade Federal de Uberlândia

Undime - União Nacional dos Dirigentes Municipais de Educação

Unicef – Fundo da Nações Unidas para a Infância

TCT – Teoria Clássica dos Testes

TIC – Tecnologia da Informação e Comunicação

TRI – Teoria da Resposta ao Item

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	18
<b>2</b>	<b>POLÍTICAS PÚBLICAS COMO OBJETO DE ESTUDO: <i>POLICY SCIENCE</i> E SUA APLICAÇÃO PARA A ANÁLISE DE POLÍTICAS PÚBLICAS EDUCACIONAIS</b> .....	25
2.1	O CAMPO DE PESQUISA DAS POLÍTICAS PÚBLICAS.....	26
2.2	A ABORDAGEM DO CICLO DE POLÍTICAS.....	35
2.3	A AVALIAÇÃO DAS POLÍTICAS EM EVIDÊNCIA .....	50
<b>3</b>	<b>POLÍTICAS PÚBLICAS EDUCACIONAIS, DE ALFABETIZAÇÃO E DE AVALIAÇÃO</b> .....	61
3.1	POLÍTICAS PÚBLICAS EDUCACIONAIS NO BRASIL .....	61
3.2	POLÍTICAS DE ALFABETIZAÇÃO .....	70
3.3	A POLÍTICA DE AVALIAÇÃO EM LARGA ESCALA.....	82
<b>3.3.1</b>	<b>A avaliação educacional e a garantia de direitos</b> .....	82
3.3.1.1	<i>Fundamentos e características da avaliação educacional</i> .....	888
<b>3.3.2</b>	<b>Críticas e virtudes da avaliação</b> .....	94
<b>3.3.3</b>	<b>O sistema nacional e o surgimento dos sistemas próprios</b> .....	102
<b>3.3.4.</b>	<b>A importância dos resultados</b> .....	1111
3.3.4.1	<i>Os resultados como parte da implementação da política de avaliação</i> .....	113
<b>4</b>	<b>OS PROGRAMAS DE AVALIAÇÃO DA ALFABETIZAÇÃO NA LITERATURA EDUCACIONAL</b> .....	12020
4.1	ESTUDOS EM AVALIAÇÃO EDUCACIONAL.....	12424
4.2	REVISTA BRASILEIRA DE EDUCAÇÃO .....	12929
4.3	REVISTA BRASILEIRA DE ESTUDOS PEDAGÓGICOS .....	1322
4.4	BANCO DE TESES E DISSERTAÇÕES DA CAPES .....	1344
<b>4.4.1</b>	<b>A produção sobre o Paebes Alfa</b> .....	1366
<b>4.4.2</b>	<b>A produção sobre o Proalfa</b> .....	1388
<b>4.4.3</b>	<b>A produção sobre o Spaece Alfa</b> .....	1444
4.5	ALGUMAS OBSERVAÇÕES .....	1533
<b>4.5.1</b>	<b>Dos trabalhos de pós-graduação e da natureza de sua abordagem</b> .....	1533
<b>4.5.2</b>	<b>Dos desdobramentos possíveis para a pesquisa no banco da Capes</b> .....	1566
<b>4.5.3</b>	<b>Da invisibilidade da avaliação da alfabetização</b> .....	1577
<b>5</b>	<b>PROGRAMAS ESTADUAIS DE AVALIAÇÃO DA ALFABETIZAÇÃO: OS CASOS DE PAEBES ALFA, PROALFA E SPAECE ALFA</b> .....	1611
5.1	A ESCOLHA DOS PROGRAMAS E DO PERÍODO DE ANÁLISE.....	1611
5.2	METODOLOGIA .....	1655
5.3	OS CRITÉRIOS PARA ANÁLISE.....	1699

5.3.1	Alcance da política.....	1711
5.3.2	Público-alvo.....	18080
5.3.3	Contexto de criação (agenda) e modelo de implementação .....	1855
5.3.4	Desenho da avaliação .....	1944
5.3.4.1	Tipo de avaliação, séries e disciplinas avaliadas .....	1977
5.3.4.2	Matrizes de referência.....	1999
5.3.4.3	Escalas de proficiência, padrões de desempenho, cadernos de teste e aplicação .....	2022
5.3.4.4	A avaliação da escrita.....	2055
5.3.4.5	A matemática na avaliação da alfabetização.....	2111
5.3.5	Divulgação e uso dos resultados (natureza dos canais de divulgação e comunicação) .....	2155
5.3.6	Concepção de alfabetização .....	2199
5.3.7	Vínculo com programas relevantes e geração da avaliação ( <i>accountability</i> ).....	2211
5.3.8	Longevidade (duração e constância).....	2299
5.3.9	Participação dos municípios .....	2333
5.3.10	Volume e transparência dos recursos aplicados em avaliação .....	2388
5.4	A VISÃO DOS ATORES DAS SECRETARIAS DE EDUCAÇÃO SOBRE A IMPORTÂNCIA DOS PROGRAMAS PARA A GESTÃO DAS REDES.....	2488
6	CONSIDERAÇÕES FINAIS.....	2611
	REFERÊNCIAS .....	28080
	APÊNDICE A - Sobre três pontos relativos às pesquisas conduzidas no âmbito da <i>policy science</i> : conflito, descentralização e análise de programas específicos.....	2955
	APÊNDICE B - Uma reflexão sobre o conceito de políticas públicas .....	3033
	APÊNDICE C - Modelos de formulação e análise de políticas públicas .....	3144
	APÊNDICE D - Limites das tipologias de Lowi e Deubel para o estudo de programas de avaliação da alfabetização e proposta do critério como gerador do tipo .....	3244
	ANEXO A - Matrizes de Referência do Paebes Alfa – 2009 a 2017.....	33030
	ANEXO B - Matrizes de Referência do Proalfa – 2009 a 2017 .....	3366
	ANEXO C - Matrizes de Referência do Spaece Alfa – 2009 a 2017.....	3444
	ANEXO D – Roteiros das entrevistas semi-estruturadas .....	3488
	ANEXO E – Termo de cooperação com os municípios – Paebes Alfa e Paebes .....	3544

## 1 INTRODUÇÃO

A avaliação em larga escala é um dos temas mais espinhosos no cenário educacional brasileiro, fazendo divergir opiniões a seu respeito. Apesar disso, é reconhecida como uma das políticas públicas de maior duração nos últimos 30 anos, não apenas tendo sobrevivido a diferentes governos, mas experimentando desenvolvimentos ao longo do tempo, que a alçaram, para o bem ou para o mal, a um status de centralidade para os tomadores de decisão. A alfabetização também viu sua importância ser ampliada ao longo do tempo, em especial, a partir de fins do século XX e início do século XXI e para os chamados países em desenvolvimento. Alcançando a posição de tema mundial, ocupando a agenda internacional e a preocupação dos organismos internacionais e dos governos locais (fenômeno que carrega consigo muita controvérsia), a alfabetização, contudo, não experimentou a efetivação dos efeitos esperados dos esforços empreendidos para o combate ao analfabetismo, de modo que os problemas relacionados a essa etapa de escolaridade continuam vivos e emergentes. Os resultados da avaliação contribuem para o diagnóstico da crise na alfabetização, mas podemos afirmar que a relação entre avaliações em larga escala e essa etapa são, pelo menos no Brasil, recentes. A presente tese debruça-se sobre as relações entre esses dois temas, avaliação e alfabetização.

Seu objetivo central é realizar uma análise comparativa entre três diferentes sistemas estaduais de avaliação da alfabetização no Brasil: Paebes Alfa, Proalfa e Spaece Alfa. A escolha desses sistemas se deve a dois motivos, principalmente: i) são sistemas consolidados de avaliação da alfabetização, com um histórico temporal considerável; ii) são os sistemas de avaliação de três estados com papéis destacados no cenário da avaliação no Brasil (o Ceará, há alguns anos, é o estado que mais apresenta melhorias nos resultados da alfabetização; Minas Gerais tem tido, ao longo dos anos, os melhores resultados no ciclo de alfabetização nas avaliações externas; e o Espírito Santo, em especial nos últimos anos, é o estado que mais diversificou as ações de seu sistema de avaliação).

A análise comparativa será feita a partir de uma proposta tipológica original. Um conjunto de critérios será selecionado, e sua seleção devidamente justificada, para caracterizar cada um dos sistemas (entre os critérios: o tipo de federalismo em cena; a forma de divulgação e uso dos resultados; o vínculo e a articulação com outras políticas públicas; a geração de avaliação à qual pertencem; financiamento etc). As tipologias propostas não envolvem uma reunião de diferentes critérios para definir cada tipo, mas, antes, tratam cada critério como base para um tipo diferente de sistema de avaliação. Assim, os sistemas são diferentes ou

semelhantes tendo em vista cada critério adotado. Não se trata de um rol taxativo de critérios propostos. Em estudos derivados deste, eles podem ser selecionados ou não para a análise, de acordo com os interesses do pesquisador. Apesar disso, nosso objetivo é apresentar, da forma mais completa possível, os critérios que julgamos importantes para a análise. Dessa forma, acreditamos oferecer uma contribuição original para a pesquisa sobre o tema.

O pressuposto dessa discussão é um ponto que não nos parece trivial, qual seja, entender a avaliação educacional em larga escala como política pública. Metodologicamente, isso significa verificar a adequação de uma teoria das políticas públicas para sua análise, realizando ajustes na aplicação da *policy science* para lidar com as especificidades do objeto (as avaliações da alfabetização). Conceitualmente, essa afirmação implica o reconhecimento do caráter dúplice da avaliação em larga escala: avaliação de políticas públicas e política pública ela mesma. Para tanto, será necessário apresentar os fundamentos da avaliação educacional, buscando desfazer alguns mitos criados em torno dela.

A análise comparativa partirá da descrição dos sistemas de avaliação selecionados como objetos de pesquisa. A tarefa descritiva, longe de ser considerada como secundária ou menos importante, assumirá uma função documental. Esses sistemas de avaliação são analisados pela literatura de modo incompleto e fragmentário, isso quando despertam o interesse da pesquisa. Nossa proposta é descrever, de modo mais amplo e completo, como essas políticas foram e permanecem sendo implementadas, pontuando suas mudanças e desenvolvimentos ao longo do tempo. Assim, a tese pode ser usada como um registro de tais políticas, de seu funcionamento, dos problemas para sua implementação, de documentação de seus resultados, acompanhados, cada um desses pontos, de uma análise.

A pesquisa sobre políticas públicas tende a desmerecer esforços descritivos, concedendo maior importância a estudos que apresentem contribuições teóricas mais robustas, vinculadas à macropolítica e ao cenário internacional, e análises de políticas para fins pragmáticos, com suas conclusões servindo como base para propiciar ajustes nas políticas, orientando a ação dos governos. Contudo, pesquisas que analisem os impactos e os resultados das políticas exigem um aparato metodológico complexo, que, caso não seja observado, não permite a obtenção do tipo de conclusão esperada. Esse ponto parece ser olvidado por alguns teóricos das políticas públicas.

De todo modo, considerar as avaliações da alfabetização como políticas públicas nos faz recorrer aos achados e contribuições da *policy science*, campo de pesquisa com objetos, teorias e metodologias próprios. A um apanhado crítico desse campo se destina o segundo

capítulo da tese. Ele servirá como base para a análise dos sistemas de avaliação, seu objetivo geral. Os capítulos 3 e 4 destinam-se, respectivamente, a discutir as características das políticas públicas educacionais no Brasil (incluindo as políticas de alfabetização e de avaliação) e a analisar a produção da literatura nacional sobre a avaliação da alfabetização (uma contribuição que esperamos ser original para o campo). O quinto e último capítulo traz a apresentação de cada um dos critérios selecionados para compor a tipologia e a descrição de cada um dos sistemas de avaliação da alfabetização, a análise de suas características e de seus resultados, e a comparação através da proposta tipológica. Neste capítulo, integram o argumento as perspectivas de atores centrais para a formulação e a implementação dos sistemas de avaliação da alfabetização analisados, recolhidas através de entrevistas semi-estruturadas e de grupos focais.

Sobre o segundo capítulo. Ele não é dedicado, estritamente, a fazer uma espécie de estado da arte sobre a pesquisa de políticas públicas. Na literatura em língua portuguesa, já há textos que se dedicam de maneira bem-sucedida a essa tarefa, como se vê na produção de Celina Souza (2002, 2003 e 2006), Angela Maria Martins (2013) e Klaus Frey (2000). O intuito aqui é extrair da discussão aquilo que interessa diretamente à tese. Incontornavelmente, será feita menção às principais questões que envolvem o tema, como a dificuldade de uma conceituação estável, a ausência de uma metodologia clara e definida para análise, a preocupação e a dificuldade em criar tipologias de políticas públicas e a crítica destinada aos trabalhos que analisam programas, especificamente. Parte da literatura acredita que o campo da *policy science* ressent-se de abordagens mais gerais, que envolvam análises de macropolítica, em particular, as influências de organismos internacionais.

Diante disso, o segundo capítulo, dividido em três tópicos, destina-se a apresentar as características da *policy science* como campo de pesquisa, diferenciando-a de seu objeto, as políticas públicas, e reconhecendo a busca pela criação de tipologias como um de seus elementos constitutivos (tópico 2.1), algo que buscaremos fazer na tese. O estabelecimento de tipologias é um elemento importante para o pesquisador das políticas públicas, particularmente no que se refere a estudos comparativos, embora sua definição seja sempre arbitrária. O tipo depende dos critérios selecionados para tanto e a seleção dos critérios é influenciada pelos interesses de estudo do pesquisador. A arbitrariedade da definição tipológica, porém, não é tratada como um demérito ao exercício, não vazio, de construção dos tipos, mas, sim, como o reconhecimento de elementos circunstanciais para sua definição, significando que a tipologia construída, ao fim, poderia ser outra, caso critérios diferentes tivessem sido selecionados.

Em seguida, dentre os muitos modelos de formulação e análise de políticas públicas existentes na literatura, o modelo do ciclo de políticas, de Stephen Ball, é apresentado, bem como suas características e os motivos pelos quais foi selecionado para a análise (tópico 2.2), já que compreende a política pública como processo, permitindo a análise de suas diferentes etapas. De acordo com Mainardes (2006), o esquema analítico proposto pelo ciclo de políticas é pertinente para o contexto das políticas públicas no Brasil e deve ser utilizado para fins de interpretação de nossas políticas. É possível perceber a concentração do interesse das pesquisas na fase da definição da agenda, bem como nos conflitos que têm lugar quando dessa definição. Tal preferência pelo estudo da formulação da agenda revela mais a tomada de uma posição ideológica em relação ao que é mais importante para a compreensão das políticas públicas e também aponta para as fragilidades dos instrumentos de pesquisa direcionados ao entendimento das demais fases apresentadas pelo ciclo. Chamamos a atenção, ainda, para o enfoque dado à literatura aos modelos de formulação de políticas e discorremos sobre nossa proposta de apresentar uma análise de programas existentes, não orientações destinadas à sua formulação.

Por fim, apresentamos (tópico 2.3), com base em Faria (2005), uma discussão sobre a avaliação das políticas públicas, fase do ciclo de políticas que tem ganhado cada vez mais destaque e importância. Como uma das fases constituintes do ciclo de políticas, a avaliação goza de muito menos interesse, por parte dos analistas, do que a definição da agenda, embora ela tenha recebido maior atenção dos pesquisadores nos últimos tempos. Pesquisas sobre a avaliação de políticas públicas são os mais escassos dentro do campo de estudo ao qual elas pertencem. As dificuldades metodológicas apresentadas ao pesquisador para a realização de avaliações de políticas são somadas aos problemas políticos que uma avaliação de políticas pode gerar. Este é um dos pontos de maior interesse para a presente tese: as avaliações educacionais como avaliações de políticas públicas e como políticas públicas elas mesmas.

Em relação à discussão sobre as políticas públicas, levada a cabo pelo capítulo 2, a tese possui três apêndices, que aprofundam a discussão de alguns pontos importantes para a *policy science*, e que podem ser do interesse do leitor, mas que cabem melhor fora do corpo do texto. O Apêndice I discute três temas: i) a ênfase exagerada da literatura no conflito que caracteriza a formulação das políticas; ii) as críticas aos processos de descentralização das políticas, com a consequente fragilização do Estado; e iii) as críticas aos estudos sobre programas específicos, que, pretensamente, ignoram a macropolítica. O Apêndice II apresenta uma reflexão sobre o conceito de políticas públicas, abordando suas fragilidades na literatura e uma proposta de sua superação, ao passo que o Apêndice III discorre sobre os modelos de formulação de políticas

presentes na literatura, reforçando duas escolhas: i) pela adoção do modelo do ciclo de políticas para a análise empreendida na tese; e ii) pela proposição da perspectiva da análise, e não da formulação de políticas, como norte da tese.

O capítulo 3 também está organizado em três tópicos, destinados a discutir, respectivamente, no contexto nacional: i) as políticas públicas educacionais; ii) as políticas de alfabetização; e iii) as políticas de avaliação em larga escala. Assim como no capítulo 1, não se trata de um estado arte sobre esses temas, mas, antes, de uma análise seletiva sobre o que, em cada um desses temas, interessa à discussão levada a cabo na tese. Por conta disso, o tópico que trata das políticas públicas educacionais (tópico 3.1) dedica-se a discutir as influências dos organismos internacionais sobre as políticas locais (seu reconhecimento e as críticas a essa perspectiva), a continuidade das políticas e os processos de centralização que caracterizam o cenário nacional. Todas essas características das políticas educacionais aparecem, em maior ou menor medida, como críticas às políticas de avaliação em larga escala, por isso, foram selecionadas para análise no referido tópico. As influências dos organismos internacionais são reconhecidas, mas, ao mesmo tempo, seu exagero é colocado em xeque, em virtude das próprias características locais e dos obstáculos que elas colocam para uma simples cópia das políticas internacionais.

A internacionalização também é tematizada no âmbito específico das políticas de alfabetização (tópico 3.2), ligada à questão do conceito de alfabetização adotado pelas políticas nacionais, apresentando a discussão polarizada entre os defensores da importância do ensino de elementos fonéticos e os construtivistas. A influência do consenso do alfabetizar letrando é reconhecida no âmbito das mais diferentes políticas de alfabetização, da produção de materiais didáticos à formação continuada de professores, passando pela criação de uma avaliação externa para o ciclo. Sobre tal influência, são apresentados os argumentos contrários e favoráveis a essa perspectiva, bem como seus efeitos sobre a maneira de compreender a avaliação da alfabetização.

O tópico seguinte (3.3) destina-se a discutir a avaliação educacional em larga escala, reconhecendo seu caráter duplo (de política pública e de avaliação de políticas públicas), apresentando seus principais fundamentos (com destaque para sua natureza de direito), suas características (com ênfase na produção de informações a serem utilizadas como base para a tomada de decisões legítimas, em especial, para o desenho e a formulação de políticas públicas), as críticas que recebe e as virtudes que apresenta, o processo de criação do sistema nacional e



o desenvolvimento dos sistemas próprios de avaliação (os objetos de análise da tese) e a ênfase dada aos seus resultados.

O capítulo 4, por sua vez, apresenta uma análise do tema da avaliação da alfabetização na literatura educacional, concentrando seus esforços na observação de 4 conjuntos de material: i) a Revista Estudos em Avaliação Educacional (EAE); ii) a Revista Brasileira de Educação (RBE); iii) a Revista Brasileira de Estudos Pedagógicos; e iv) o Banco de Teses e Dissertações da Capes (com o complemento da análise do Banco de Dissertações do PPGP/UFJF). O objetivo do capítulo é analisar a atenção recebida pela avaliação da alfabetização por parte da pesquisa educacional, observado, ainda, em que se concentra essa atenção e como a avaliação para essa etapa é tratada na literatura. A seleção dos periódicos tem as seguintes justificativas: i) a EAE, publicada pela Fundação Carlos Chagas, é um periódico especializado em avaliação e uma das revistas mais importantes do país sobre o tema; ii) a RBE é publicada pela Anped e um dos mais importantes periódicos sobre educação do país; iii) a RBEP, também importante periódico educacional, é um dos mais antigos do país e é publicado pelo Inep, órgão responsável pelas avaliações nacionais. O Banco da Capes foi analisado a fim de verificar o interesse pelo tema avaliação da alfabetização nos programas de pós graduação. Cada um desses conjuntos de publicações ganhou um tópico próprio no capítulo, que conta ainda com um tópico final (4.5), trazendo observações consideradas pertinentes para a pesquisa realizada com esse material.

Finalmente, o capítulo 5 apresenta a análise dos três programas de avaliação da alfabetização selecionados. Sua escolha, bem como a seleção do período de análise, é justificada, assim como os critérios escolhidos para a análise dos programas. Cada critério ganhou um tópico próprio, com a apresentação do critério e a análise de cada um dos três programas no âmbito do próprio tópico. Ao fim do capítulo, um quadro geral de cada programa, com base nos critérios selecionados, é apresentado.

\*\*\*

Uma palavra sobre um aspecto importante para a escrita da tese. Entre 2011 e 2019, estive envolvido diretamente com uma série de sistemas de avaliação, estaduais e municipais, no Brasil, incluindo os programas de avaliação da alfabetização aqui analisados, de modo que minha atuação profissional no âmbito do CAEd/UFJF<sup>1</sup> não pode ser desprezada como elemento

---

<sup>1</sup> No referido período, atuei como analista de avaliação, coordenador de análise e publicação de resultados e supervisor de entrega de resultados e desenvolvimento profissional do CAEd/UFJF.

de influência para a produção deste texto. Aliás, o próprio interesse pelo estudo deste tema surgiu em função desse exercício profissional. No lugar de ocultá-lo, ele está aqui sendo reconhecido como determinante para a escolha dos argumentos, do tipo de análise e da condução da produção do texto.

## **2 POLÍTICAS PÚBLICAS COMO OBJETO DE ESTUDO: *POLICY SCIENCE* E SUA APLICAÇÃO PARA A ANÁLISE DE POLÍTICAS PÚBLICAS EDUCACIONAIS**

Enfrentar o tema da avaliação da alfabetização no Brasil a partir do viés das políticas públicas exige, primeiramente, uma apresentação da pesquisa sobre políticas públicas de modo geral. Esse é o ponto de partida para a tese: apresentar as características das pesquisas sobre políticas públicas que se adequam ao estudo da avaliação da alfabetização tomada como política pública educacional. O interesse deste primeiro capítulo, portanto, é discutir os elementos centrais da *policy science*, o ramo da ciência destinado à investigação das políticas públicas, selecionando, de maneira crítica, arcabouço teórico e instrumentação de pesquisa mais ajustados ao enfrentamento do tema.

A avaliação da alfabetização poderia ser estudada sob diferentes óticas: estatisticamente, através da análise da validade e adequação dos instrumentos avaliativos e das medidas produzidas; pedagogicamente, a partir da análise dos objetos selecionados para a avaliação, sua relação com o currículo e com ideologias dominantes durante a escolha do que deve ser avaliado; sociologicamente e antropologicamente, tendo em vista as consequências da avaliação para a cultura das escolas e redes de ensino; politicamente, quando pensamos no contexto político no qual a avaliação ganha espaço no cenário nacional etc. Contudo, analisar a avaliação da alfabetização como política pública, em maior ou menor medida, envolve todas essas óticas, acrescentando a elas uma percepção que, por mais que possa parecer, não é trivial. A avaliação da alfabetização pode ser tratada como um fenômeno correlato da avaliação de serviços públicos, característica atribuída ao que ficou conhecido como novo gerencialismo público, o que obscurece suas peculiaridades. Tratá-la como política pública dotada de características próprias parece não ser a regra. O estudo da avaliação da alfabetização, independentemente da ótica adotada, já tem seus méritos em virtude da parca produção acadêmica sobre o tema. A perspectiva das políticas públicas só qualifica esse esforço, pois coloca em cena um referencial de análise abrangente para o tratamento do assunto. Este será o esforço aqui empreendido.

O capítulo está organizado em cinco tópicos. O primeiro deles apresenta um panorama geral do campo de pesquisa de políticas públicas, com base nos esforços de autores que, no Brasil, se dedicaram a estabelecer um estado da arte sobre o tema, abordando as características e os pontos de interesse do campo desde suas origens até a atualidade. O segundo tópico se

dedica a discutir o conceito de políticas públicas, discorrendo sobre a pluralidade que caracteriza a expressão. O terceiro tópico faz um breve inventário dos modelos de formulação e análise de políticas públicas, enumerando suas virtudes e vicissitudes, e abrindo caminho para a escolha de um modelo que guiará nossa análise na escrita da tese, qual seja, a abordagem do ciclo de políticas, assunto que será tratado de forma mais minuciosa no quarto tópico. Finalmente, o quinto tópico concentra suas atenções em uma das fases do ciclo de políticas, a da avaliação, tema caro aos interesses deste trabalho.

## 2.1 O CAMPO DE PESQUISA DAS POLÍTICAS PÚBLICAS

As políticas públicas, como tema de interesse para o campo científico, têm sua emergência com data atribuída ao período posterior à II Guerra Mundial, já nos anos 1950, nos Estados Unidos. Em função da crise e das dificuldades enfrentadas pelos governos no pós-guerra, era necessário oferecer aos governantes instrumentos adequados para lidar com os problemas econômicos e sociais. Assim, as políticas públicas [a *policy science*<sup>2</sup>, conforme pontua Martins (2013, p.278)] surgem com um objetivo bem definido: fornecer ferramentas para que atores políticos pudessem tomar decisões acertadas no que dizia respeito à retomada do desenvolvimento econômico e social de sociedades afetadas pela guerra. A política pública a serviço da política. Sua maior contribuição, original, de acordo com a autora, seria a construção de um campo com enfoque nas relações entre Estado, sociedade e política<sup>3</sup>.

Contudo, mais amplas do que um conjunto de ferramentas a serviço dos governos em contextos específicos, as políticas públicas alcançaram o status de campo científico, angariando espaço na academia. Conforme ressalta Martins (2013), o espaço adquirido como disciplina universitária foi impulsionado, também, pela demanda dos governos em relação a assessorias para a criação e análise de custos de programas (ação que recebeu e ainda recebe muitas críticas). Como campo disciplinar, as políticas públicas foram sensivelmente influenciadas pelos modelos da administração de empresas, o que lhes deu um caráter analítico quantitativo. Apesar disso, permaneceu sendo considerada uma espécie de área da ciência política. Frey (2000) ressalta que a ascensão desse campo de pesquisa na Europa, em particular, na Alemanha, data dos anos 1970, no contexto de desenvolvimento da social-democracia. As diferenças

---

<sup>2</sup> Conforme será exposto com maior clareza mais à frente, usaremos a expressão *policy science* como referência à ciência que se dedica ao estudo das políticas públicas, uma *ciência das políticas públicas*, e não como referência às políticas públicas em si.

<sup>3</sup> Na realidade, o estudo dessas relações é uma preocupação da teoria política clássica, remontando aos primórdios do contratualismo, e não com data no século XX.

iniciais entre as abordagens das políticas públicas como disciplina acadêmica nos Estados Unidos e na Europa se relacionam com o enfoque dado pelas pesquisas. O contexto europeu apresentava maior preocupação com uma teoria consolidada sobre o Estado, seus papéis e suas funções, ao passo que o contexto norte-americano se concentrava diretamente nas ações do governo. Mais precisamente, interessava produzir informações para planejar, com base na defesa do racionalismo, as ações governamentais, se valendo do conhecimento científico produzido por agentes independentes e não interessados (pelo menos, esse era o discurso oficial). O caráter pragmático é evidente. O estudo das políticas públicas tinha como objeto de interesse o entendimento do mundo empírico, de como os governos funcionavam na prática. Além de entender seu funcionamento, o horizonte apontava para informar suas ações, não apenas descrevê-las.

Isso fazia com que as decisões governamentais recebessem o verniz da legitimidade (característica herdada, como veremos, pelas avaliações educacionais, pelo menos, em termos de pretensão). A cientificidade e o racionalismo oferecidos pelos estudos sobre políticas públicas eram possíveis em contextos de democracias estáveis, como no caso dos Estados Unidos. Essa linha norte-americana marcou decisivamente o campo da pesquisa, com efeitos que se fazem sentir até hoje (em particular, no Brasil). A ideia mais difundida do que são as políticas públicas como disciplina acadêmica é tributária dessa concepção. Conforme define Martins, as políticas públicas se caracterizam por “estudos dedicados a examinar o mundo político, com foco na preocupação de procurar entender o que os governos fazem, porque e para quem eles fazem, ou ainda, em outras palavras, quem ganha e quem perde nas decisões de governo” (MARTINS, 2013, p.279). A origem do campo de pesquisa tem lugar em contextos de países desenvolvidos, o que leva ao questionamento sobre sua pertinência para o contexto brasileiro e dos demais países em desenvolvimento.

Como apontado anteriormente, a linha empírica e pragmática dos estudos de políticas públicas desenvolvida nos Estados Unidos se baseava no entendimento de como os governos tomavam suas decisões na prática. Como contexto para a análise, uma democracia estável, um Estado e uma sociedade com características próprias. Isso fez com que a *policy science* se ajustasse, como disciplina ancorada na análise da realidade (e produzindo teoria a partir dela), ao contexto dos países desenvolvidos. Por conta disso, duas questões se apresentam: i) o questionamento acerca da natureza dos achados científicos da *policy science*, de sua capacidade de generalização e de sua constituição como uma teoria de amplo alcance; e ii) o problema da

transposição para contextos diferentes daqueles onde uma disciplina tão ancorada na empiria se desenvolveu.

Em relação à primeira questão, cabe uma observação de caráter geral, que diz respeito às teorias no âmbito das ciências humanas e sociais, e não especificamente ao tipo de conhecimento produzido pela *policy science*. A heterogeneidade dos contextos onde políticas públicas se desenvolvem não pode fazer olvidar o que há de comum entre eles. As políticas pressupõem uma série de elementos sociais e políticos, tais como a existência de um Estado organizado, a identificação de problemas sociais selecionados para enfrentamento por parte do governo, um conjunto de ações sistematizadas, a necessidade (em maior ou menor grau) de legitimar as decisões diante da sociedade civil. Ao falarmos em países desenvolvidos e países em desenvolvimento, já pressupomos algum pano de fundo comum entre esses contextos, a ponto de nos permitirmos uma comparação (independentemente das críticas que podem ser direcionadas a essa categorização) entre o que é considerado desenvolvido e o que ainda está por se desenvolver. Assim, por mais que as características contextuais sejam muito significativas para a análise de uma política pública, há elementos gerais que permitem sua aplicação em outros contextos. Isso não significa a adoção completa, *ipsis litteris*, de uma teoria ou concepção de análise das políticas para contextos distintos. Contudo, definir algo como política pública já significa encontrar aspectos comuns em fenômenos diferentes, que fazem com que esse conceito seja utilizado para definir distintas ações em variados contextos. Isso significa que, mesmo com uma clara ancoragem em um contexto específico, a análise de políticas públicas produz conhecimento utilizável em contextos diferentes daqueles em que foi produzido.

Não é necessário, portanto, que uma teoria se encaixe perfeitamente em todos os contextos para que tenha validade e utilidade científica. Para tanto, é necessário realizar ajustes e adaptações. Essa é, por exemplo, a defesa de Frey (2000). Soma-se a isso outro ponto. Tratar os países desenvolvidos como um grupo homogêneo é negligenciar as diferenças que tais países apresentam em aspectos fundamentais para a análise social, incluindo a de políticas públicas. Países com democracias estáveis e com desenvolvimentos econômicos elevados não apresentam as mesmas características sociais e culturais em todos os aspectos. O mesmo pode ser dito para os países em desenvolvimento. Esse tema é sensível para as políticas de avaliação em larga escala. Como veremos, parte significativa da crítica a elas debruça-se sobre as influências internacionais.

Isso nos leva à segunda questão. Se, por um lado, as análises de políticas públicas não devem ser descartadas por sua matriz teórica ter sido gestada em países desenvolvidos, elas não devem, por outro, ser transpostas integralmente para outros contextos. Entra em cena o problema da transposição. De acordo com Martins, essa permanece sendo uma questão não resolvida pela pesquisa sobre políticas públicas<sup>4</sup>.

A defesa de uma adaptação da teoria e instrumentação da *policy science* para os países em desenvolvimento tem, a nosso ver, pelo menos duas implicações: i) a teoria termina servindo como medidor de nosso atraso em relação aos países em desenvolvimento, construindo um argumento do tipo ‘o instrumental teórico será adequado para nosso contexto social quando (ou se) nos tornarmos um país desenvolvido’; e ii) essa inadequação teórica de origem é um dos fatores que explicam os problemas da pesquisa de políticas públicas no país. Quanto ao primeiro ponto, ele se configura como um contrassenso: a adequação de uma teoria para analisar políticas públicas é usada como critério para definir o grau de desenvolvimento de um país. Se a teoria não é adequada para compreender nossa realidade, o problema é da teoria e não da realidade a ser investigada. Assim, deve-se ajustar a teoria, caso seja possível, ou descartá-la e lançar mão de um arcabouço teórico mais ajustado. Quanto ao segundo ponto, os problemas enfrentados pela pesquisa de políticas públicas no Brasil não se relacionam com o contexto de origem da teoria, mas sim com problemas de definição.

Se formos seguir à risca a proposta de análise das políticas públicas apresentada em linhas gerais por autores como Frey, Martins e Souza, não realizaríamos pesquisas sobre o tema no contexto nacional. A todo momento, a ressalva sobre a origem do campo é feita. São concepções pensadas com base no contexto das sociedades desenvolvidas, não se aplicando direta e adequadamente ao contexto brasileiro. No entanto, essa visão parece carregada de uma justificativa *post factum*: como não realizamos pesquisas da forma como as realizadas no contexto dos países desenvolvidos, os especialistas atribuem esse problema à inadequação da teoria ao nosso próprio contexto, reforçando a necessidade de adaptá-la. Os termos e medidas da adaptação, entretanto, não são definidos com clareza por esses autores.

A despeito das agruras que caracterizam qualquer campo científico, em particular, aqueles em vias de desenvolvimento e consolidação, a *policy science* é um campo de pesquisa em expansão, cujos limites e possibilidades ainda não foram plenamente mapeados. Mesmo

---

<sup>4</sup> “Assim, a transposição de uma determinada visão de Estado, de sociedade, de políticas públicas e de análise das políticas públicas – configuradas em determinado período histórico e em condições específicas – ainda constitui um dos principais desafios a serem superados nas pesquisas desenvolvidas em países periféricos” (MARTINS, 2013, p.280).

sendo uma ciência recente, engendrada e desenvolvida no século XX, autores como Souza (2002; 2006) falam em um ressurgimento atual da importância do campo de análise de políticas públicas, em particular, no que diz respeito ao estabelecimento de regras e modelos que regem sua decisão, sua implementação e sua avaliação.

De acordo com Souza (2002), tal ressurgimento está diretamente atrelado às políticas restritivas de gastos<sup>5</sup>, impostas aos governos nas últimas décadas, e aos fracassos das ações destinadas a resolver o problema da desigualdade social. No entanto, é preciso ressaltar que o interesse pelas políticas públicas em países como o Brasil não é tributário apenas da questão das desigualdades sociais, mas da análise das políticas em si mesmas.

Segundo a leitura de Martins (2013), as pesquisas sobre políticas públicas no Brasil são caracterizadas por preocupações mais localizadas e pontuais, sem interesses por temas mais gerais e com resultados de maior alcance explicativo. Com isso, as pesquisas brasileiras na área se dedicariam a temas pouco abrangentes e de pouco interesse para o desenvolvimento da *policy science* como campo do conhecimento<sup>6</sup>. Podemos afirmar que, de acordo com a leitura de Martins, os problemas das pesquisas sobre políticas públicas no Brasil são de duas naturezas: de um lado, o conteúdo, centrado nos processos de negociação de políticas setoriais, com ênfase em seus efeitos; de outro, as metodologias, dotadas de pouca complexidade e muita fragilidade analítica (MARTINS, 2013, p.283-284). Há, ainda, um terceiro problema, que motiva a crítica direcionada à *policy science* como um todo: a alegada falta de embasamento teórico das pesquisas com políticas públicas. Apesar de, em casos específicos, essa crítica ser pertinente, é preciso destacar que a própria história de desenvolvimento do campo, em particular no contexto norte-americano, fez com que a empiria fosse a grande fonte de inspiração para a análise de políticas, sem contudo desprezar a teoria (utilizada quando era capaz de auxiliar nessa compreensão).

Conforme narra Souza (2002), a *policy science* surge como parte da ciência política, sua terceira geração<sup>7</sup>, interessada em orientar as decisões do governo a partir da compreensão de

---

<sup>5</sup> A nova visão sobre o papel dos governos teve como consequência a substituição de políticas keynesianas por políticas de ajuste fiscal, fazendo com o que o tema da restrição de gastos tenha dominado a agenda nos últimos anos, com a consequente defesa da adoção de orçamentos equilibrados e de restrições à intervenção do Estado na economia e nas políticas sociais

<sup>6</sup> “A autora (Souza, 2006) assinala que investigações preocupadas em abranger temas mais amplos e/ou articulados não são comumente encontradas nesse campo, tais como: o papel das burocracias responsáveis pela implementação das políticas públicas; ausência de modelos fundamentados na teoria das elites ou das redes sociais; estudos comparativos entre políticas implementadas por esferas estaduais, tendo em vista que as preocupações se voltam aos governos e à esfera federal” (MARTINS, 2013, p.283; acréscimo nosso, em itálico).

<sup>7</sup> A primeira geração se dedicava ao estudo das instituições, visto que, sob a influência da Madison, adotava certo ceticismo em relação à natureza humana, e, portanto às suas ações; as instituições eram vistas como elementos necessários para limitar os excessos e paixões característicos das ações humanas. A segunda geração, por sua vez, dedicava interesse à compreensão das organizações locais, fonte de onde nasciam as



seus motivos. Historicamente, houve uma valorização da tecnocracia e do racionalismo para a tomada de decisões, com a aplicação de métodos científicos para guiar a tomada de decisão dos governos em relação a problemas públicos específicos, com a posterior crítica a leituras puramente racionalistas, como foco nas ações dos decisores. Conforme mapeamento de Souza (2002; 2003), nos anos 1930, Laswell se preocupava em conciliar o conhecimento produzido pela ciência com as decisões dos governos (seu agir prático). Na década de 1950, a preocupação de Simon se relacionava com a denúncia de uma racionalidade limitada por parte dos decisores, já que as informações recebidas por eles nem sempre eram suficientes para dar suporte a uma decisão ou eram imprecisas<sup>8</sup>. Na esteira de Simon, Lindblom questionou a ênfase dada ao racionalismo, defendendo a necessidade de incorporar outras variáveis para a análise de políticas públicas, como as relações de poder entre os diferentes grupos de interesse, o papel das eleições e a existência de diferentes fases que caracterizam as políticas. Posteriormente, já na década de 1960, Easton defendeu que as políticas públicas deveriam ser entendidas como um sistema, baseado nas relações estabelecidas entre sua formulação, seus resultados e o ambiente circundante (SOUZA, 2002, p.3). As propostas destes autores históricos lançaram as bases para a *policy science* como campo científico, sendo importantes até hoje como linhas de orientação para a pesquisa na área, e demonstram uma preocupação do campo em compreender as ações do governo, explicitando seu enfoque empírico e pragmático.

Os estudos de políticas públicas educacionais beberam da mesma fonte da *policy science* em geral, qual seja, a preocupação com os serviços prestados no pós-guerra. A crença era fundamentalmente ancorada no racionalismo: a ciência seria capaz de direcionar os governos para o melhor caminho, para as melhores decisões, para as políticas mais adequadas em cada caso. Essa visão inicial é tratada como ingênua e utópica, pois se baseia em uma transposição perfeita entre o que é planejado e o que é executado como política pública. As ações do governo são pensadas como a encarnação, sem desvios, de suas decisões. Essa visão alimentava a ideia de progresso, da qual a educação era parte central (MARTINS, 2013). O enfraquecimento dessa perspectiva, estimulada por análises mais realistas sobre as políticas públicas (e pelos limites do racionalismo científico), levou a uma espécie de pessimismo, nas décadas seguintes aos anos 1950, no que tange às possibilidades de construção de uma sociedade mais igualitária e justa pelas mãos da educação.

---

virtudes cívicas da sociedade norte-americana para a constituição de um bom governo (nesse caso, a influência é de Tocqueville) (SOUZA, 2002, p.2).

<sup>8</sup> Além disso, ele chamou a atenção para a influência dos interesses próprios dos decisores na definição de políticas públicas, o que exigiria um sistema bem definido de regras e incentivos capaz de limitar a arbitrariedade desses interesses particulares diante dos interesses públicos.

A produção do campo da análise das políticas públicas educacionais, a partir dos anos 1990, conforme ressalta Martins (2013), se concentra em esforços de construção de estados da arte e registros da literatura produzida sobre o tema<sup>9</sup>. Grande parte desse mapeamento dirige críticas aos estudos sobre políticas públicas no que tange à metodologia utilizada, que, em regra, não é clara e não goza do respaldo dado por outras pesquisas. Os estudos de caso, em detrimento de esforços de pesquisa cujas conclusões possuem maior alcance, são igualmente criticados. Além disso, a crítica se direciona, também, à falta de suporte teórico substantivo.

Outro ponto de destaque no que tange às pesquisas sobre políticas públicas educacionais se relaciona com a percepção de que a agenda local é influenciada pela agenda internacional. Os estudos sobre redes internacionais e sobre os efeitos da globalização, como o de Castells (2016), chamam a atenção para a dissolução de fronteiras nacionais no que tange a aspectos valorativos, culturais e econômicos. Nessa linha, os estados nacionais são se apresentam como territórios isolados e idiossincráticos, mas, antes, como compartilhando um pano de fundo valorativo e, do ponto de vista dos indivíduos, comportamental. Aplicando essa discussão para a *policy science*, isso significa que as condições e questões sociais que, ocupando a agenda política, exigem enfrentamento através de políticas públicas, podem, em maior ou menor medida, ser semelhantes em diferentes contextos. Assim, demandas que se transformam em questões públicas em um contexto podem ser elevadas ao grau de problema público em contextos diferentes de sua origem.

O entendimento acerca dessa questão, em regra, adota um tom crítico. Globalização, redes e agendas comuns são encaradas negativamente, como um problema a ser enfrentado, e não apenas como uma característica da sociedade contemporânea. Na visão mais radical, os países ricos impõem sua agenda aos países mais pobres. De modo mais realista, contudo, é possível perceber que esse movimento é mais complexo do que a mera imposição. Sem negar a influência de uma agenda mundial, cabe reconhecer que as políticas são implementadas de modo adaptado a cada contexto, que oferece obstáculos a essa implantação. Além disso, decisores, implementadores e pesquisadores de políticas públicas podem aprender com outros contextos, adaptando, quando necessário, conhecimentos e experiências desenvolvidos em outros cenários<sup>10</sup>. Voltaremos a discutir esse ponto com mais deter no capítulo seguinte, em tópico dedicado às políticas educacionais.

---

<sup>9</sup> É o caso do texto da própria autora, destinado a mapear o campo de pesquisas na área.

<sup>10</sup> Se não fosse assim, haveria uma *policy science* para cada contexto, bem como teorias sociológicas de alcance apenas local. A existência de uma ciência das políticas públicas aplicada ao âmbito da educação implica exatamente isso: o conhecimento desenvolvido nessa área é suficientemente generalizável para se constituir como uma ciência, sem que isso signifique que todos os instrumentos dessa ciência (conceitos, conhecimento, experiência, artefatos metodológicos, conclusões) sejam aplicáveis, da mesma forma, em todas as circunstâncias.

Um dos instrumentos que exemplificam o argumento de que as políticas públicas apresentam certo grau de generalidade é a construção de tipologias de políticas (a generalidade está presente na própria ideia de tipologia). O objetivo da criação de tipologias no campo se relaciona com o propósito atribuído originalmente a ele, qual seja, auxiliar os formuladores, implementadores e avaliadores na seleção da política mais ajustada para a resolução de um problema tratado como estratégico. No caso da presente tese, o que interessa é a tipologia como instrumento analítico, como auxílio à descrição e compreensão de como a política se constitui. Diante de estudos comparados, a tipologia pode ser um instrumento particularmente interessante e útil, orientando a comparação em cada caso.

Maia & Grando (2016), por sua vez, afirmam a pertinência das tipologias com vistas à formulação de políticas públicas, diferente da proposta desta tese, que baseia-se no uso da tipologia para fins de análise de políticas já existentes<sup>11</sup>. Os autores veem dificuldades na aplicação da conhecida tipologia de Lowi<sup>12</sup>, desenvolvida nos anos 1970.

Por serem demasiadamente genéricas, torna-se difícil enquadrar as políticas em apenas um desses grupos. Se a tipologia não é capaz de operacionalizar uma distinção razoavelmente clara entre políticas diferentes, ela se torna pouco útil para o analista. Além disso, a própria política pública, multifacetada como pode ser, oferece dificuldades para ser enquadrada em tipologias como a de Lowi. Quanto mais diversificada é uma política, mais difícil defini-la como pertencente a um tipo específico<sup>13</sup>.

---

<sup>11</sup> Não nos interessa diretamente, conforme a proposta de Maia & Grando, propor um modelo de formulação de políticas públicas, mas sim de análise. O objetivo da presente pesquisa não é deontológico, oferecendo aos tomadores de decisão e formuladores mecanismos sobre como as políticas devem ser definidas, implementadas e avaliadas, de acordo com os achados da *policy science*. É discutível que o campo das políticas públicas deva produzir informações do tipo dever-ser. Guiar a ação política não é o mesmo que compreendê-la e analisá-la, embora seja possível que essas atividades caminhem juntas, visto que não são excludentes. Contudo, é preciso deixar claro que esse não é o interesse do presente esforço de pesquisa. O que interessa ao trabalho aqui desenvolvido é uma análise comparativa de políticas de avaliação da alfabetização, sem que uma proposta de como tais políticas devam ser conduzidas emergja ao fim do processo de pesquisa. Quanto mais reconhecemos que as políticas públicas se caracterizam pela influência do contexto na qual se inserem, por sua flexibilidade e adaptação a diferentes cenários econômicos, políticos e sociais, e pela interferência de uma série de outros fatores (como o papel das instituições e dos grupos de interesse), menos propícia se torna uma proposta deontológica para sua elaboração e implementação.

<sup>12</sup> Baseando-se na ideia de que a política pública faz a política, e não o contrário, Lowi propõe uma divisão das políticas em quatro tipos distintos: i) *políticas distributivas* – desconsiderando a limitação dos recursos, são políticas com impactos mais individuais do que gerais, privilegiando certos grupos sociais sem acarretar prejuízos para outros; apresentam baixo grau de conflito n que diz respeito aos beneficiários; ii) *políticas redistributivas* – atingindo um maior número de beneficiários do que as políticas distributivas, essas políticas levam a perdas de curto prazo para certos grupos sociais, oferecendo ganhos futuros, porém incertos, para outros grupos; iii) *políticas regulatórias* – são políticas que envolvem a máquina burocrática pra a emissão de ordens, proibições, decretos etc; elas são mais visíveis ao público e os efeitos de custos e benefícios não são determináveis de antemão; iv) *políticas constitutivas* – políticas que lidam com procedimentos, determinando as regras do jogo para as demais políticas; em geral, apresentam pouco impacto na esfera pública.

<sup>13</sup> Torna-se difícil, por exemplo, enquadrar as políticas de avaliação educacional em larga escala em um dos tipos construídos por Lowi.

Diante disso, Maia & Grando encontram em Deubel (2009) uma proposta de tipologia mais ajustada a seu propósito de orientação para a formulação e implementação de políticas, conforme exposta no quadro a seguir. Na abordagem dos autores, não há menção à importância da tipologia como instrumento analítico.

Quadro 1 - Proposta tipológica para formulação e implementação de políticas com base em Deubel

<b>Tipo de política</b>	<b>Características de formulação</b>	<b>Enfoque de implementação</b>	<b>Critério de avaliação</b>	<b>Custos de implementação</b>	<b>Variáveis críticas de implementação</b>
<b>Tipo 1</b>	Objetivos claros e procedimentos flexíveis	<i>Top-down</i> clássico, comando e controle	Realização dos objetivos	Custos de negociação baixos/custos de informação baixos	Troca de circunstâncias externas ou em torno de uma política
<b>Tipo 2</b>	Objetivos abertos e procedimentos flexíveis	<i>Bottom up</i> , por retrocesso	Criatividade e liderança administrativa e capacidade de inovação	Custos de negociação altos/custos de informações baixos	Liderança administrativa e respaldo de políticos
<b>Tipo 3</b>	Objetivos claros e procedimentos precisos	Experimentos e reformulações	Aprendizagem	Custos de negociação baixos/custos de informações altos	Capacidade administrativa
<b>Tipo 4</b>	Objetivos claros e procedimentos flexíveis	Negociações	Acordos e respaldos	Custos de negociação altos/custos de informações altos	Relacionamento com os atores beneficiados

Fonte: Maia & Grando (2016), adaptado de Deubel (2009).

A tipologia apresentada pretende reunir, em quatro tipos, todas as políticas públicas. Cada tipo é definido por uma reunião de critérios. Embora alguns elementos da proposta possam nos ser úteis como critérios da nossa própria tipologia, é preciso ressaltar que essa proposta não é suficiente para servir como base de análise para políticas de avaliação da alfabetização. Em primeiro lugar, por não abarcar critérios importantes para lidar com a especificidade das políticas de avaliação, tais como o tipo de divulgação dos resultados, o modelo de federalismo que lhe dá base e o tipo de *accountability* atrelado a ela. Trata-se de uma tipologia demasiadamente genérica, assim como a de Lowi. Além disso, apresenta algumas falhas em termos de coerência. No que diz respeito às características de formulação, por exemplo, nem todas as combinações possíveis para formar um tipo são utilizadas.

O que nos interessa aqui é o reconhecimento de que as tipologias são elementos importantes no âmbito da pesquisa com políticas públicas. Somam-se a isso: i) o reconhecimento da necessidade de construir critérios próprios para uma tipologia mais adequada para as políticas

de avaliação da alfabetização (por isso, no capítulo 4 da tese, são apresentados critérios específicos); e ii) a defesa da ideia de que as tipologias servem para fins de análise e não apenas para a formulação de políticas, como em Maia & Grando (2016) e Deubel (2009).

Antes que passemos à discussão sobre o conceito de políticas públicas, vale chamar atenção para três pontos relativos às pesquisas conduzidas no âmbito da *policy science*: o excessivo enfoque no conflito para explicar, especialmente, a formação da agenda; a descentralização da execução de políticas para entidades privadas; e a crítica aos estudos sobre programas específicos. Em relação ao conflito, cumpre ressaltar que seu reconhecimento como elemento integrante das políticas públicas, em todas as suas fases, fornece uma visão mais realista do que a leitura proposta pelo racionalismo técnico, mas, em excesso, representa o perigo de transformar o conflito na explicação para todo e qualquer aspecto da política, sendo alçado à condição de premissa, sem necessidade de explicação. Quanto à descentralização da execução de políticas para entidades privadas, se esse é um fenômeno real, que pode ter consequências negativas para uma visão que coloca o Estado como o único agente condutor das políticas públicas, ao mesmo tempo, é preciso reconhecer que a influência de grupos não governamentais nas políticas públicas não implica, necessariamente, uma incapacidade do Estado em executá-las. Além disso, reconhecer as complexas relações entre o Estado e essas entidades não significa defender o capitalismo contemporâneo e ocultar suas mazelas. Por fim, no que tange às críticas destinadas ao estudo de programas específicos<sup>14</sup>, cabe ressaltar que a defesa de uma visão geral das políticas públicas, envolvendo a análise do contexto internacional e de sua influência sobre as decisões dos governos locais, não é antagônica à análise de programas específicos, abordagem muito criticada no âmbito da pesquisa em políticas públicas. Mais sobre esses três pontos é apresentado no Apêndice I da tese.

## 2.2 A ABORDAGEM DO CICLO DE POLÍTICAS

A definição de um conceito geral de políticas públicas é uma tarefa não plenamente resolvida pela literatura (SOUZA, 2002), algo que ocorre com termos e expressões polissêmicas, como é o caso. Uma discussão mais aprofundada sobre a conceituação da expressão e os problemas que algumas definições apresentam na literatura encontra-se no Apêndice II da tese. Para fins pragmáticos, no âmbito deste trabalho, adotamos uma das três

---

<sup>14</sup> Apesar da crítica (ou como base para sua existência), há muitos estudos sobre programas específicos, não no campo da educação, em especial, mas no que diz respeito às políticas públicas em outras áreas (saúde, economia, segurança pública etc), como pode ser visto em Costa (2017); Coelho, Costa e Schalch (2017), Filho et al (2017), Eiró (2017).

dimensões das políticas públicas abordadas pela literatura (DEUBEL, 2009; SOUZA, 2002; FREY, 2000) como sua definição, a *policy*, significando os programas e ações desenvolvidos pelos governos e autoridades políticas, representando o conteúdo concreto das políticas. A atuação das instituições políticas (*polity*) e as atividades de organização do poder, incluindo o processo político (*politics*), não nos interessam diretamente na tese, apenas como parte da contextualização e análise dos programas e ações em si. Dessa forma, diferenciamos o campo de pesquisa (*policy science*) de seu objeto (*policy*)<sup>15</sup>.

Já afirmamos que o interesse deste trabalho não é propor como a política deve ser formulada (dever-ser), mas, antes, analisar políticas que já existem. Isso significa que os modelos de formulação de políticas só nos interessam se puderem ser utilizados como modelos de análise. Souza (2002), por exemplo, propõe um conjunto de critérios para um modelo de formulação de políticas que nos interessam como critérios para um modelo de análise (objetivos, temporalidade prevista para o alcance dos resultados, custos, sistema de incentivos etc). A literatura sobre modelo de formulação de políticas apresenta um rol considerável de exemplos. Eles são apresentados e discutidos no Apêndice III da tese. Aqui, interessa apresentar o modelo selecionado para nossa análise, o ciclo de políticas. Os modelos apresentam características distintas, com virtudes e limitações, de modo que não há um único modelo completo e adequado para todos os casos. Modelos diferentes podem ser mais ou menos úteis de acordo com o tipo de pesquisa que se quer realizar. Para nossos propósitos, o modelo do ciclo de políticas se apresenta como o mais ajustado. Ele reconhece o caráter processual das políticas, apresentando uma vantagem analítica: a separação da política em diferentes etapas (implementação, execução e avaliação). De todos os modelos, ele é o que apresenta essa proposta de maneira clara.

A abordagem do ciclo de políticas (*policy cycle approach*) foi desenvolvida por Stephen Ball e colaboradores na década de 1990 e vem sendo utilizada para analisar políticas públicas educacionais. Conforme destaca Mainardes (2006), o campo de pesquisa em políticas educacionais no Brasil é relativamente novo, sem que esteja consolidado e suas referências analíticas e teóricas ainda não são consistentes. Há poucos autores trabalhando com referenciais analíticos mais específicos, ajustados à pesquisa de políticas públicas educacionais. Como vimos, muitos autores afirmam a necessidade de adaptação de modelos teóricos oriundos de outros contextos (dos países desenvolvidos) para a compreensão de nosso próprio contexto. De todo modo, a abordagem do ciclo de políticas nos parece um referencial teórico particularmente

---

<sup>15</sup> Em português, ambos são denominados de políticas públicas.

útil para analisar políticas públicas educacionais, em nosso caso, políticas de avaliação da alfabetização.

Segundo Mainardes, a abordagem do ciclo de políticas destaca “a natureza complexa e controversa da política educacional” (2006, p.49), dando ênfase às ações dos atores envolvidos com a política pública (de sua decisão à sua implementação) e relacionando o microcontexto ao macrocontexto. A ênfase no macrocontexto é uma característica dos principais referenciais teóricos sobre políticas públicas e não apenas do ciclo de políticas. Para Frey (2000), trata-se de um modelo explicativo da vida de uma política, ressaltando o caráter dinâmico e a complexidade temporal das políticas públicas. Segundo Maia & Grando (2016), a abordagem do ciclo capta bem a natureza processual complexa das políticas públicas, que passam por contínuas mudanças, em virtude das negociações entre os diferentes grupos e atores interessados. O processo de decisão é interativo e as prioridades dos formuladores são definidas pelo poder político que diferentes atores sociais possuem e por sua interação com as instituições políticas.

Conforme aponta Mainardes, a ideia de ciclo se baseia na continuidade. Inicialmente, reconhecia-se três arenas da política pública: a política proposta, a política de fato e a política em uso. A política proposta dizia respeito à política oficial, às intenções manifestadas pelo governo, mas, no caso da educação, também àquelas dos departamentos educacionais e das escolas. A política de fato era encarnada, por sua vez, pelos textos políticos e legislativos, as bases para que a política pública pudesse ganhar vida. A política em uso, por fim, se referia aos discursos e práticas que ocorriam no momento da implementação da política, ou seja, às ações dos atores educacionais na prática (MAINARDES, 2006, p.49).

Essa primeira proposta de abordagem para o ciclo de políticas ignorava uma série de elementos importantes e se apresentava, ainda, de forma excessivamente engessada para a análise das políticas públicas. Com isso, ela deu origem a uma nova proposta para a abordagem do ciclo, mais ajustada às interações que de fato acontecem entre as fases do ciclo de vida de uma política. Um dos principais pontos de ajuste em relação à primeira proposta foi o reconhecimento da importância de que a formulação e a implementação de uma política, para efeitos de análise, não poderiam ser tratadas como uma única etapa, em especial pelo fato de ser necessário reconhecer que a formulação da política é uma etapa extremamente importante e complexa, permeada por inúmeras disputas e conflitos. Além disso, conforme leitura de Mainardes (2006) sobre a proposta de Ball & Bowe, seria preciso reconhecer a participação, na formulação e implementação de uma política, de atores que se encontram no nível da prática,

como os profissionais que atuam nas escolas (professores, diretores, coordenadores etc). Se tais profissionais não estão excluídos dessas fases do ciclo de políticas, sua participação, nem sempre, se dá da mesma forma. Ela pode ser mais ou menos ativa, dependendo do tipo de texto que a política produz. Assim, se o texto é demasiadamente prescritivo, definindo os sentidos que serão ou não aceitos, a participação do leitor (no caso, dos profissionais das escolas) é limitada e ele apenas recebe do texto os sentidos já definidos (Roland Barthes chama esse tipo de texto de *readerly*). Por outro lado, se o texto abre espaço para que o leitor seja uma espécie de coautor, contribuindo para construir seus sentidos, que, de todo modo, não estão plenamente fechados, a participação é expandida e o leitor troca a posição inerte por uma posição ativa (Barthes chama esse tipo de texto de *writerly*; MAINARDES, 2006). Em um mesmo texto político, esses dois tipos podem ser utilizados.

Alterando a proposta inicial de três arenas da política, a abordagem do ciclo de políticas passou a propor três contextos: o contexto da influência, o contexto da produção do texto e o contexto da prática. Vale notar que cada um desses contextos é permeado por disputadas contínuas. Além disso, é importante reconhecer que eles não estão organizados de forma linear no tempo, pois o ciclo das políticas é recheado de idas e vindas. Na prática, os contextos estão tão inter-relacionados que não é tarefa fácil definir com clareza onde um contexto termina e onde começa o outro. De todo modo, para fins analíticos, trata-se de uma distinção importante.

O contexto da influência marca o momento em que as políticas públicas são iniciadas, no qual os discursos políticos são organizados. Aqui, a agenda é construída, influenciada por um conjunto de interesses diversos, que, em última instância, disputam a legitimidade de dizer a agenda, de definir o que deve ou não ser tratado como um problema público, a ser enfrentado através de uma política, e o tipo de política adotada. No âmbito da educação, isso significa definir quais são os principais problemas educacionais que exigem uma atuação mais sólida do governo, estabelecendo, também, quais são os objetivos da educação e como eles devem ser alcançados. É nesse contexto que conceitos adquirem sentidos mais estabilizados e passam a gozar de legitimidade. As influências para a formação de um discurso são inúmeras: os partidos políticos, grupos de apoio aos governantes, grupos de interesse em torno de determinada temática, mídia especializada, sindicatos etc. No caso da educação, pensemos em empresas, em organizações não-governamentais destinadas à educação, em universidades, sindicato de professores, editoras etc.

Na esteira de Ball, Mainardes (2006) chama a atenção, também, para as influências internacionais no âmbito local. Seguindo seu argumento, cada vez mais, no contexto da



expansão das redes internacionais, políticas públicas adotadas em países mais desenvolvidos têm influenciado (o argumento mais radical dirá determinado) políticas locais, trazendo uma lógica nem sempre (ou nunca) adequada para lidar com a realidade específica dos países em desenvolvimento.

A influência internacional sobre os contextos locais ocorre a partir de, pelo menos, dois caminhos. O primeiro deles é o percorrido através do próprio fluxo de ideias, que circulam por meio de redes constituídas ao redor do mundo em torno de temas de interesse das políticas. Essa circulação de ideias, por sua vez, pode ter lugar em virtude da circulação através de redes sociais e políticas, através da produção científica e acadêmica sobre temas de interesse para as políticas (sem excluir o caráter mercadológico que esse tipo de fluxo de ideias pode assumir, a partir da produção e venda de livros, periódicos, conferências, palestras, cursos etc, formando um grande mercado intelectual de circulação de ideias) e através do empréstimo de políticas.

O segundo caminho é aquele que se dá diante da atuação de agências multilaterais, não-governamentais, mas com grande influência em determinados cenários e contextos. É o caso, por exemplo, de agências internacionais como o Banco Mundial, a OCDE (Organização para Cooperação e Desenvolvimento Econômico), o FMI (Fundo Monetário Internacional) e a UNESCO. De acordo com Mainardes (2006), tais agências patrocinam, quando não impõem, determinadas políticas internacionais para cenários nacionais. A avaliação educacional em larga escala é tratada, pelos críticos, como um exemplo de política pública que, no Brasil, só ganhou espaço em virtude do patrocínio das agências multilaterais. Em geral, essa influência se dá através de apoio técnico-científico e, principalmente, através de transferência de recursos financeiros. As críticas a esse tipo de influência sobre os contextos nacionais pode gerar uma imagem de palidez dos governos nacionais, plácidos e passivos diante das imposições das agências internacionais.

Contudo, uma visão mais acurada da realidade negará que tais influências sejam simplesmente imposições irrecusáveis, sem deixar, evidentemente, de reconhecer seus efeitos. As agências internacionais possuem demandas próprias e tentam convencer os Estados nacionais a encampem-nas. Apesar disso, esse movimento não é apenas unilateral. A tentativa de oferta ou imposição de determinadas políticas públicas não se dá sem que haja algum nível, mesmo que tímido, de adaptação ao contexto nacional. Mainardes (2006) vê esse reconhecimento em Ball, ao afirmar que toda política internacional é reinterpretada, relida, adaptada, de uma maneira ou de outra, ao contexto nacional. Essa mesma interpretação pode ser aplicada a um dos caminhos do fluxo internacional de ideias, qual seja, o empréstimo de

políticas. Por mais que os empréstimos sejam feitos e possam ser criticados (na medida em que ocorrem sem um processo reflexivo que dê a eles sustentação), a transposição é sempre permeada por algum nível de reinterpretação. Esse processo interpretativo contínuo, vale notar, não afasta a crítica a determinadas políticas desajustadas ao contexto em que se inserem, mas busca afastar o entendimento radical de que as políticas são simplesmente transpostas sem qualquer tipo de adaptação.

Se o contexto da influência é pautado por uma ideologia mais clara e declarada por parte dos governantes e atores que dele participam, o contexto da produção do texto assume um discurso mais articulado com o interesse público em geral. O momento de produção dos textos exige um tratamento mais direcionado da política pública, buscando também suporte na legislação. Como no caso do contexto da influência, o contexto de produção é atravessado por conflitos e pela busca da legitimidade de dizer o texto da política. Esse texto nunca é final, visto que permanece sujeito a reinterpretações contínuas, mesmo depois de oficializado, e é o produto de acordos e disputas por parte dos grupos de interesse. Isso significa que o texto da política não é a tradução literal das decisões tomadas no contexto de influência. É preciso, ainda, contextualizar essa produção, visto que ela sempre tem lugar em um tempo e em um cenário específicos. Por fim, é forçoso reconhecer que o texto produzido, ao ser influenciado por diferentes atores, não se apresenta de modo coerente e coeso. Suas diferentes partes, fragmentos que representam diversos acordos e ajustes, não necessariamente formarão um todo dotado de uma única lógica.

O contexto da prática, por fim, responde pela implementação da política. Mais uma vez, é necessário reconhecer que a passagem do texto produzido para a implementação da política não ocorre de maneira direta ou literal. O texto produzido e a política implementada, em regra, não são idênticos. O texto é a base para a política ser implementada, mas sofre ajustes e interpretações quando ganha a realidade. Portanto, essa é mais uma fase caracterizada por reinterpretações contínuas, a ponto de Mainardes (2006) falar em recriação da política. É nesse contexto que interessa o estudo das consequências e efeitos das políticas públicas, reconhecendo as transformações operadas na proposta da política original.

Há, pelo menos, três questões importantes a serem observadas no contexto da prática: i) a questão linguística da abertura dos textos; ii) a questão sociológica da participação dos atores; e iii) a questão metodológica da investigação dos resultados e dos efeitos da política. Em relação à primeira questão, cumpre ressaltar que, de sua formulação inicial até sua implementação, a política passa por uma série de transformações. Mais do que considerar esse processo de

mudança como algo pejorativo, que aponta para a fragilidade das políticas públicas, a questão envolve o reconhecimento linguístico de que os sentidos de um texto não podem ser absolutamente controlados (BAKHTIN, 2011). As palavras, em si mesmas, carregam diversas possibilidades semânticas e conceitos muito complexos, tais como sociedade, escola, educação etc, são polissêmicos. Eles carregam uma série de significados, pela sua própria natureza complexa, que se somam aos significados que receberam historicamente através de seu uso no tempo e no espaço. Além disso, os conceitos estão sempre abertos a novas interpretações, que podem ou não ser consagradas pelo uso. Isso significa que um texto sempre pode ser interpretado. Uma nova circunstância, um novo cenário, novos atores, qualquer novo elemento pode levar a uma nova interpretação. Assim, as políticas não são formuladas e implementadas de maneira direta. A cada fase de seu ciclo, elas passam por transformações e ajustes interpretativos.

Uma linha de pesquisa interessante para a *policy science* é investigar as mudanças semânticas experimentadas pelos textos das políticas em cada uma de suas fases e ao longo do tempo. Soma-se a ela outro ponto de interesse. Os textos são produtos de diferentes influências. Portanto, são compostos por diferentes fragmentos de significados. No entanto, os grupos interessados na política não se apresentam em posição de igualdade no que diz respeito à sua capacidade de influência. Alguns grupos possuem mais capacidade de influência do que outros e os textos refletirão essas diferenças. Há vozes mais e menos ouvidas: “a política como discurso estabelece limites sobre o que é permitido pensar e tem o efeito de distribuir ‘vozes’, uma vez que somente algumas vozes serão ouvidas como legítimas e investidas de autoridade (...). Na prática, os atores estão imersos numa variedade de discursos, mas alguns discursos serão mais dominantes do que outros” (MAINARDES, 2006, p.54).

A segunda questão, da participação dos atores, coloca em xeque a leitura determinista que pode ser feita em relação às políticas públicas. Como vimos, o campo de pesquisa tende a dar muita ênfase aos conflitos e às disputas de poder, mas sem conceder tanta atenção ao papel dos atores que implementarão as políticas. Quando o fazem, as pesquisas tendem a ver esses atores como caracterizados por uma postura passiva, sem exercer grandes influências sobre as políticas. Isso é particularmente marcante quando o que está sob investigação são os contextos de influência (definição da agenda, formação dos discursos que darão início à política) e de produção do texto. É como se os atores do nível prático não tivessem nenhuma influência nesses dois contextos. Reconhecer sua participação no contexto da prática é perceber que a política pública, por mais que tenha sido formulada com base no alijamento de vários atores do

processo, não pode ser tratada como uma mera imposição. Essa questão, de caráter sociológico, se relaciona com a questão linguística, exposta anteriormente. Como os textos são abertos e sua interpretação está em disputa, os atores do nível prático (professores, diretores, alunos, coordenadores pedagógicos, funcionários etc) exercem um papel ativo na implementação da política. Eles reinterpretam os textos em seu fazer profissional.

A questão da análise dos resultados envolve um desenvolvimento posterior proposto por Ball para o ciclo das políticas (MAINARDES, 2006). Dois outros contextos foram acrescentados aos três originais: o contexto dos resultados ou efeitos e o contexto da estratégia política. A análise dos resultados da política é um tema central para a *policy science*, como veremos em outro tópico. Mainardes (idem) afirma que falar em efeitos é mais adequado do que em resultados. A preocupação, segundo ele, gira em torno dos impactos que as políticas têm no que tange à redução das desigualdades. Assim, efeitos gerais e específicos devem ser analisados. No âmbito da educação, o que interessaria são os efeitos da política na educação como um todo (na organização do ensino, no currículo, na atuação dos profissionais etc) e seus efeitos específicos (digamos, o impacto da formação continuada de professores sobre a aprendizagem dos alunos). Os efeitos chamados de primeira ordem seriam aqueles mais evidentes, observados no sistema educacional, ao passo que os efeitos de segunda ordem seriam aqueles observados no âmbito mais amplo da sociedade, como equidade e justiça social (MAINARDES, 2006, p.55). Além disso, é importante que se observe a interação dos efeitos da política com as demais políticas educacionais. Esse ponto levanta um problema para a análise dos resultados de uma política.

A partir da concepção de Ball exposta por Mainardes (2006), é possível observar uma prevalência dos interesses da pesquisa sobre os efeitos da política no contexto macrossocial em prejuízo do interesse pelos efeitos específicos da política, no âmbito educacional. Sem desconsiderar a importância dos efeitos no âmbito macrossocial, é necessário conceder importância aos efeitos que a política exerce em plano micro. Esses dois tipos de análise não se excluem, antes, se complementam. Nem toda política educacional é destinada e será capaz de alterar a estrutura enraizada das desigualdades que caracterizam a sociedade. Isso não significa que ela não seja importante ou que, por isso, contribui para o adensamento dessas desigualdades. Políticas menos ambiciosas do ponto de vista social podem exercer importantes efeitos educacionais que, em alguma medida, podem contribuir para o cenário social mais amplo mediatamente. Evitando leituras radicais sobre o tema, que selecionam alguns pontos de

interesse para a pesquisa negando importância a outros, é preciso reconhecer o valor de políticas de menor alcance e seu interesse para a *policy science*.

A análise dos resultados de uma política envolve uma importante discussão metodológica. A avaliação da política, com foco nos resultados, só pode fornecer informações contundentes caso lance mão de instrumentos adequados a obter conclusões seguras. O grande desafio para a análise dos resultados da política é decidir como eles serão avaliados. Voltaremos a tratar do assunto em linhas vindouras. Por ora, cumpre ressaltar que os resultados da política precisam lidar com o tema da causalidade e a avaliação deve responder se os resultados obtidos são, de fato, produto daquela política em questão e não de outros fatores não devidamente monitorados. O difícil tema da causalidade envolve a escolha do aparato metodológico mais adequado para permitir conclusões sobre a relação entre a política e a observância dos efeitos<sup>16</sup>.

O quinto contexto do ciclo, acrescentado posteriormente por Ball, é o contexto da estratégia política. Posterior ao contexto dos resultados, por dele se alimentar, a estratégia política se refere ao enfrentamento de desigualdades e efeitos colaterais (previstos ou não previstos) de determinada política, com base em atividades e ações destinadas a reduzir esses efeitos.

Autores diferentes se apropriaram da abordagem do ciclo de políticas, propondo fases diferentes daquelas originalmente propostas por Ball ou renomeando as fases originais. Para Frey (2000, p.26) por exemplo, as fases do ciclo são as seguintes: percepção e definição de problemas, *agenda setting*, elaboração de programas e decisão, implementação de políticas, e avaliação das políticas e eventuais correções. Já Saravia (2006) define sete etapas do processo<sup>17</sup>, incluindo também a avaliação. De acordo com Maia & Grando (2016), a avaliação foi a etapa

---

<sup>16</sup> A título de exemplo, pensemos nas análises de efeitos combinados entre políticas. As políticas não possuem efeitos que se somam apenas, de modo direto e linear. Os efeitos conjuntos podem ser potencializados, obstaculizados, sobrepostos ou simplesmente não convergirem. Os impactos e efeitos das políticas são multifacetados e sua análise exige seleção. Não é a regra da pesquisa em políticas públicas educacionais que essas questões sejam tematizadas e apareçam com destaque no debate.

<sup>17</sup> São elas: i) definição da agenda, momento no qual determinadas questões são categorizadas como problemas públicos, sendo tratadas como prioritárias e clamando pela intervenção do governo; ii) elaboração da política, momento em que o problema é delimitado e ações são desenhadas para enfrentá-lo de acordo com o contexto e os possíveis beneficiários; iii) formulação da política, etapa que define e seleciona a alternativa mais adequada entre as apresentadas para o enfrentamento do problema (a adequação envolve elementos jurídicos, políticos e financeiros – o respeito à legislação, a conveniência política e os custos de execução); iv) implementação da política, fase em que há a especificação do planejamento e a organização dos recursos; v) execução da política, etapa na qual são efetivadas as ações previstas para o alcance dos objetivos traçados; vi) acompanhamento (monitoramento da execução), em que as ações são supervisionadas de acordo com o que foi previsto, permitindo correções e adequações ao longo desse processo; e vii) avaliação da política, fase final na qual os resultados são analisados, bem como os efeitos produzidos e o alcance das metas projetadas.

que mais se desenvolveu. Faria (2005) também chega a essa conclusão, apontando que a principal causa para tanto foi o desenvolvimento e a expansão o novo gerencialismo público<sup>18</sup>.

Mainardes (2006) aponta uma série de críticas dirigidas à abordagem do ciclo de políticas. A principal delas diz respeito à ausência de uma teoria do Estado mais robusta que lhe dê sustentação. Tal crítica é oriunda, principalmente, de abordagens marxistas, que se ressentem da ausência, na pesquisa sobre políticas públicas educacionais, de uma relação clara (e causal) entre as políticas educacionais e econômicas. Essa crítica é complementada pela ausência de reconhecimento do papel do Estado como o grande controlador dos resultados das políticas. Além disso, outras críticas são dirigidas à ausência de uma posição do ciclo que traga à cena desigualdades raciais e de gênero no âmbito das políticas públicas.

No que tange ao papel do Estado, a abordagem do ciclo não exclui seu enfoque, bem como não o prepondera. A proposta do ciclo serve tanto aos defensores de uma abordagem de pesquisa centrada no papel do Estado, como o grande ator responsável pelas políticas, como àqueles que advogam uma análise conjunta dos fatores macrossociais (como o papel desempenhado pelo Estado) e microssociais (como o papel desempenhado por atores posicionados no nível prático, de implementação da política, por exemplo). A rigor, o foco na ação do Estado e em suas instituições ou em atores posicionados no nível microssociológico da ação é uma decisão do analista de políticas públicas. Trata-se de uma escolha da pesquisa, que, em regra, precisa selecionar os elementos que deseja investigar. Tal escolha depende de uma série de fatores, incluindo o tipo de política a ser analisado.

Uma teoria mais robusta do Estado será importante para o analista que, se valendo da abordagem do ciclo, decida dar ênfase ao papel do Estado para a execução de determinada política, mas, sua ausência, a princípio, não é um problema da abordagem, que serve a outros propósitos. O mesmo pode ser dito para as demais críticas listadas por Mainardes (2006). A ausência de estudos que envolvam questões raciais e de gênero não é um problema intrínseco da abordagem do ciclo, que, por certo, não se destina ao conteúdo das políticas. Trata-se de uma instrumentação analítica para o estudo das políticas, independentemente de seu conteúdo e de seus efeitos. A ausência de estudos sobre esses temas está relacionada às escolhas dos analistas e não ao modelo de análise.

Chama a atenção, na abordagem do ciclo de políticas, o tema da implementação de políticas, fase que, embora fundamental para o próprio conceito de políticas públicas, não recebe a mesma atenção, por parte dos pesquisadores, que a definição da agenda. Souza (2002),

---

<sup>18</sup> Para mais sobre o novo gerencialismo público, ver Apêndice C.

por exemplo, como defensora da perspectiva da *policy science* como orientadora das ações governamentais, vê os estudos sobre implementação como importantes para detectar, a tempo de corrigi-los, problemas no desenho da política. Sua defesa é a da realização de estudos de caso para identificar a miríade de variáveis que podem afetar a implementação de uma política. Esse é um tema sensível, já que há inúmeros fatores que podem influenciar a forma como uma política é implementada. Para realizar uma análise desse tipo, parte-se dos achados da literatura especializada, mas, na análise empírica, uma variável interveniente não prevista pode aparecer diante do analista. Apesar da ampla gama de variáveis influenciadoras, Souza (2002) dá destaque ao papel desempenhado pela burocracia que implementa a política (a chamada burocracia de linha), ou seja, aos atores que tomam decisões rotineiras para a execução contínua de uma política pública.

No que diz respeito a esse ponto, os desenvolvimentos da teoria da escolha racional podem ser úteis para a análise da implementação de políticas públicas, em particular, para a compreensão das relações entre os principais (aqueles que decidem, os tomadores de decisão, os políticos eleitos) e os agentes (aqueles que executam o que foi decidido, a burocracia). Conforme ressaltado por Souza (2002), o controle do principal sobre o agente é complexo, difícil de ser executado e dispendioso, pois há conflitos de interesses entre eles. Para lidar com esses conflitos, a proposta é criar um sistema de interesses, de modo que o agente tenha a intenção de cumprir o que foi previsto pelo principal<sup>19</sup>.

Um dos pontos mais sensíveis do ciclo de uma política pública é justamente sua implementação, a passagem do âmbito das intenções e documentos para a realização prática. Para que tal passagem ocorra, alguns autores destacam a necessidade de mapear alguns elementos centrais para a implementação, tais como: i) o conhecimento e a capacidade das organizações e dos implementadores sobre a política; ii) os processos que compõem a implementação, incluindo pontos de conflito; iii) o papel dos agentes na implementação; iv) o grau de discricionariedade de burocracia ao implementar a política (SOUZA, 2002, p.21). Outros, por sua vez, afirmam a necessidade de uma intervenção do Estado, de modo a garantir a implementação da política através de instrumentos específicos<sup>20</sup> (MAIA; GRANDO, 2016).

---

<sup>19</sup> De acordo com Souza (2002, p.20), há, pelo menos, quatro mecanismos de incentivo para que o agente aja de acordo com aquilo que foi previsto pelo principal: i) o contrato firmado com o agente preveria mecanismos de compensação, estimulando a competição entre os agente; ii) mecanismos de seleção eficientes para a contratação meritocrática dos agentes; iii) monitoramento interno das ações dos agentes; e iv) controles institucionais das ações dos agentes, por grupos externos e contrários à política.

<sup>20</sup> Citando Deubel, eles elencam um rol de instrumentos dos quais o Estado pode se valer para garantir a implementação: a) instrumentos de incentivos, que podem ser baseadas em uma sanção (em regra, jurídica) ou em uma recompensa; b) instrumentos prescritivos, que buscam, através da obrigatoriedade, determinar que

Seguindo, Deubel, os referidos autores apontam para a existência de dois modelos de implementação das políticas: i) o modelo *top-down* é uma forma mais tradicional de desenvolvimento da política, com os decisores agindo a partir de sua autoridade, sem contar com participação efetiva dos implementadores e beneficiários da política; assim, a política fica sob o controle completo da autoridade e os problemas identificados em sua implementação tendem a ser tratados como pertencentes à ordem do controle e da coordenação dos programas através dos quais a política é efetivada; ii) o modelo *bottom-up* é o oposto do modelo anterior, visto que os problemas identificados são resolvidos em cada nível onde aparecem, com a participação e o envolvimento dos implementadores e dos beneficiários; a política, nesse caso, não é construída de cima para baixo, mas em todos os seus níveis de execução.

Cada um desses modelos tende a ser vinculado a contextos políticos específicos. Assim, pela definição, o modelo *top-down* poderia ser associado a cenário políticos com democracia menos consolidada, mais propícios a autoritarismos. Contando com pouca ou nenhuma participação de atores fora dos círculos decisórios, o modelo poderia ser taxado como centralizador. Essa centralização, por sua vez, levaria a uma espécie de miopia para a identificação de problemas que afetam a política em qualquer uma de suas fases. Problemas que aparecem nas franjas da política, em sua execução microssociológica, no plano dos implementadores e beneficiários, podem não ser identificados com clareza por um controle centralizado<sup>21</sup>. O modelo *bottom-up*, por sua vez, encarna as características de cenários sociais e políticos caracterizados pela maturidade democrática e por uma sociedade participativa, sem a presença de um Estado centralizador. Assim, as decisões são tomadas em conjunto, contando com a participação de setores não governamentais. Mais condizente com o contexto de países desenvolvidos, esse modelo tende a ser defendido como o mais adequado para o desenvolvimento das políticas públicas.

Contudo, se, por definição, parece haver um modelo mais adequado do que o outro, é importante ressaltar que não há evidências de que um modelo deve ser adotado em prejuízo do

---

uma ação seja realizada; c) instrumentos de organização, que definem, através de procedimentos específicos, como uma política pública deve ser implementada; d) instrumentos de coordenação, desenvolvidos pelo Estado para lidar com a multiplicidade de políticas; e) instrumentos materiais, que regulam a distribuição de bens e serviços; e f) instrumentos de delegação, através dos quais o Estado estabelece parcerias de colaboração com instituições privadas e não-governamentais para fins de execução das políticas (citado por MAIA; GRANDO, 2016, p.167).

<sup>21</sup> No Brasil, historicamente, o Estado seria o grande ator responsável pela condução das políticas públicas, desde sua implementação, sendo taxado como centralizador e pouco afeito à participação de outros setores sociais. O modelo *top-down* para o desenvolvimento de políticas públicas, dessa forma, seria alvo de um conjunto de críticas dirigido à sua própria concepção. Característica de países com desenvolvimento tardio (seja pelo modelo prussiano ou colonial), a forte presença do Estado em todos os setores sociais é tratada como um sinal de atraso e de pouca maturidade democrática.



outro. De um lado, é forçoso reconhecer que tais modelos servem mais como instrumentos para o analista, com aplicação *post factum*, do que como linhas mestras para o planejamento e implementação da política. Diante disso, o modelo é sempre uma análise posterior sobre a política, não uma forma necessariamente consciente de concebê-la e orientá-la. De outro, pelo fato dos contextos sociais onde as políticas públicas se desenvolvem serem mais complexos do que o modelo pode prever. Assim, o modelo nunca se adequa perfeitamente à realidade<sup>22</sup>.

Além disso, é preciso questionar os pressupostos que dão base à qualificação dos dois modelos. Não resta dúvida de que a participação de atores, além dos decisores, é importante e condizente com os pressupostos de condução de um Estado democrático. A participação, a princípio, traz ao palco da discussão visões heterogêneas, fundamentais para a elaboração e implementação de uma política, que nem sempre são percebidas pelos decisores. Ainda, é importante que aqueles que implementarão a política ou receberão seus benefícios sejam ouvidos, não apenas para aprimorar as ações da política, mas para torná-la mais legítima. Contudo, é preciso notar, nem toda participação é, necessariamente, benéfica. Isso não quer dizer que a participação não deva existir. Ela é defensável como princípio (democrático). Porém, não é garantia de sucesso da política. Uma política não terá bons resultados garantidos só por haver abertura para participação. Em outras palavras, o modelo *bottom-up* não é garantia de que a política cumprirá seus objetivos. De outro lado, a participação ativa do Estado para a definição e implementação de uma política não é sinal de que ela estará fadada ao fracasso e à centralização absoluta<sup>23</sup>.

Ainda no que tange à implementação, Maia & Grando (2016) chamam a atenção para os seus custos, que podem representar, a princípio, dois cenários. Por um lado, os custos de implementação podem ser considerados baixos, situação em que existe um consenso diante de

---

<sup>22</sup> A dificuldade de aplicação completa desses modelos de implementação das políticas públicas é afirmada por Maia & Grando: “Apesar de todos os estudos realizados e modelos propostos para se analisar a implementação de políticas públicas, não existe um consenso sobre qual pode ser considerado mais adequado, ou utilizado como referência para uma aplicação prática, devido à dificuldade de se identificar quais são os elementos fundamentais de definição e implementação e as circunstâncias em que as ações são realizadas, ou seja, onde começa e onde termina uma implementação, bem como quais elementos são relevantes para a avaliação dessa implementação” (2016, p.170).

<sup>23</sup> Aliás, é necessário colocar em dúvida também esse pressuposto. Nem toda centralização é condenável de saída. Em cada caso, é preciso observar como o Estado se comporta, bem como os demais atores, para analisar a contribuição de suas ações para o sucesso de uma política. Com isso, esses modelos não podem ser julgados de antemão. Eles precisam ser analisados no caso concreto, observando suas vantagens e desvantagens de acordo com o contexto. Soma-se a isso outro aspecto. Um modelo não exclui o outro, na medida em que seus elementos característicos podem ser utilizados em fases distintas da mesma política. Como exemplo, num primeiro momento, a implementação da política pode seguir um caminho mais centralizado e, num momento seguinte, abrir maior espaço para a descentralização.

uma situação em particular para o enfrentamento de um problema. Nesse caso, os interesses em jogo estão mapeados, são claros e bem definidos. Por outro lado, os custos da implementação podem ser considerados altos, situação na qual o dissenso impera e interesses muito divergentes, e nem sempre claros, fazem com que as negociações sejam mais demoradas e complexas. Além disso, os custos para a implementação de uma política podem ser altos ou baixos em função do volume e qualidade das informações disponíveis. Informações têm custos. Para que uma política seja definida e implementada com segurança, as informações disponíveis sobre cada uma das alternativas propostas para o enfrentamento do problema em questão e sobre o próprio problema são decisivas para a escolha de um caminho no lugar de outro.

Até aqui, é possível perceber que a abordagem do ciclo tem propósitos heurísticos. Ela tem a vantagem de propor uma organização didática para o estudo das políticas públicas, incluindo as educacionais. Se, na prática, os contextos do ciclo não podem ser absolutamente separáveis, para a análise, tal separação é perfeitamente cabível e útil, segundo os interesses de pesquisa de cada analista. Porém, como outros modelos de análise de políticas públicas, essa abordagem apresenta algumas obstáculos.

Uma das dificuldades do ciclo é a questão temporal. Como analisar, por exemplo, o contexto da influência, da definição da agenda, para públicas com longa duração no tempo? Os discursos e influências iniciais podem nem ter sido registrados. Isso não diminui a pertinência do estudo, mas coloca limitações a ele. Esse é um problema com o qual teremos que lidar na tese: os programas de avaliação da alfabetização tiveram início em outro momento, apresentando considerável duração temporal (um dos critérios de seleção dos programas para análise foi justamente sua longevidade). Soma-se a isso dificuldades metodológicas para investigar cada uma das fases.

No contexto da influência, dentre uma enorme constelação de problemas, é preciso explicar como alguns são selecionados como parte da agenda, recebendo atenção através de políticas públicas próprias, ao passo que outros ficam de fora. A questão central é saber como um problema se torna um problema público, como a agenda é definida. Segundo Mainardes (2006), o contexto da influência pode ser investigado através de pesquisa bibliográfica, entrevista com os formuladores e com os demais profissionais envolvidos. Porém, essa proposta apresenta uma série de limitações. Em geral, para explicar a definição da agenda, o enfoque recairá sobre relações de poder. Em concepções influenciadas pelo marxismo, os grupos econômicos mais poderosos serão tratados como os grandes influenciadores da agenda (numa postura mais radical, como determinando a agenda). Mesmo as concepções mais pluralistas e

menos deterministas verão a construção da agenda como pautada por relações de conflito e poder<sup>24</sup>.

O mesmo tipo de problematização metodológica pode ser levantado em relação aos outros contextos e fases do ciclo de políticas. O estudo do contexto de produção do texto exige a análise de documentos e, seguindo a proposta de Mainardes (2006), também entrevistas com os autores da política. Em relação aos autores, eles sequer podem ocupar a mesma posição anterior, não integrando mais a equipe de governo (se a política se estende ao longo do tempo e a pesquisa é feita em momento distante do início da política, corre-se o risco da ausência de acesso aos autores da política). Além disso, as políticas, como tem sido defendido pelos modelos mais recentes de análise, não são produto de indivíduos isolados, mas oriundas da contribuição de diferentes vozes. Nesse caso, para uma pesquisa que abra espaço para todas elas, seria preciso identificar cada um dos atores envolvidos. Em relação ao contexto da prática, a proposta de Mainardes (*idem*) é pela pesquisa etnográfica, com a inserção no ambiente no qual a política é desenvolvida. Nesse ponto, a observação recairia sobre como a política é reinterpretada pelos atores. Em contextos educacionais, isso significa selecionar os atores, dando voz à maior representatividade possível, e os momentos em que as observações serão realizadas. O contexto dos resultados por sua vez, se baseia na comparação entre os resultados obtidos, efetivamente, pelas políticas e os resultados pretendidos originalmente por ela. A interpretação de Mainardes (2006, p. 60), por exemplo, em relação aos resultados dá ênfase à redução das desigualdades sociais.

Entretanto, é importante ressaltar que nem toda política pública educacional se destina, como objetivo direto e imediato, à redução de desigualdades sociais. Esse é um ponto polêmico, que exigiria maior fôlego para um enfrentamento adequado. Porém, por ora é necessário pontuar que nem todos os problemas educacionais enfrentados através de políticas públicas se resumem a problemas sociais imediatos. Evidentemente, há uma relação entre problemas sociais e problemas escolares. Contudo, eles não podem ser confundidos. Uma política de

---

<sup>24</sup> Apesar deste tipo de enfoque, chama a atenção a ausência de uma teoria do poder como suporte para a *policy science* no que diz respeito ao contexto da influência. Assim, há a afirmação de que a agenda é definida com base em relações de poder, mas tal afirmação é tratada como premissa, sem que seja explicada. Afinal, que tipo de relações de poder estão em jogo? Quais são os mecanismos utilizados? Que recursos e posições são responsáveis por conceder mais capacidade de influência? Eles operam em todo e qualquer cenário? Em nenhuma das concepções aqui apresentadas, uma teoria do poder tem lugar. O mesmo pode ser dito em relação ao tema do conflito. As relações são taxadas como conflituosas, mas as bases do conflito não são apresentadas. Para um campo de disputas para dizer a agenda, chama a atenção a ausência de uma teoria sociológica que explique como as políticas públicas são, de fato, selecionadas e alvo de disputas pela legitimidade de dizer a política, com conversões do capital acumulado em um campo determinado para o campo das políticas públicas [Bourdieu (2003; 2011) seria o principal representante de uma teoria desse tipo, seja para servir como base à teoria ou como contraponto a ela].

formação continuada de professores para a alfabetização, por exemplo, pode ser bem sucedida com base em objetivos pedagógicos próprios, sem que tenha alcance para a solução de problemas sociais imediatamente. Atribuir à educação a resolução de problemas sociais nos parece ser atitude, no mínimo, temerária, que exige uma reflexão mais profunda. Apesar disso, o contexto dos resultados coloca em evidência a fase final do ciclo de políticas, a avaliação, tema que tem recebido notoriedade por parte dos especialistas da área. A ele, é dedicada a seção seguinte.

### 2.3 A AVALIAÇÃO DAS POLÍTICAS EM EVIDÊNCIA

A avaliação de políticas públicas é um dos temas mais importantes para a presente tese. Como fase final do ciclo de políticas, ela tem recebido, cada vez mais, a atenção dos analistas de políticas públicas. A avaliação desempenha o importante papel de encerrar o ciclo de uma política, funcionando como engrenagem para que um novo ciclo se inicie. Seu papel primordial é fornecer informações qualificadas que orientem as ações dos formuladores de políticas e tomadores de decisão: a política será mantida, ampliada, reduzida ou encerrada? Em contextos de valorização da qualidade dos serviços oferecidos pelo Estado e de eficácia de suas ações, a avaliação tem se tornado um elemento central para a análise de políticas públicas. Em última instância, a avaliação deseja responder ao questionamento acerca da eficácia da política e do alcance de seus objetivos. O próximo capítulo tratará do tema da avaliação educacional com mais detalhe. O presente tópico é destinado a apresentar a avaliação como uma importante fase da abordagem do ciclo de políticas, descrevendo seu contexto de desenvolvimento e problematizando algumas análises sobre a forma como ela tem sido percebida.

Faria (2005) vê os anos 1990 como um marco normativo para a avaliação de políticas públicas, para o que chamou de “função avaliativa na gestão governamental” (p. 97). É desse período que datam sistemas de avaliação de políticas públicas na América Latina. O principal motivo para isso é o contexto de busca por uma legitimação e um aumento da eficácia da ação governamental, sob o discurso da modernização da gestão pública.

Como vimos, considerando a abordagem do ciclo de políticas, a avaliação se apresenta como a última etapa. Através dela, são analisados os resultados de determinado ciclo da política, atividade que dá suporte para decisões relacionadas à manutenção da política tal qual ela se apresenta, à promoção de ajustes na política ou à sua suspensão. Soma-se a isso um importante

aspecto da avaliação, qual seja, a prestação de contas do governo perante a sociedade civil, elemento fundamental para o sistema de *accountability*.

De acordo com Faria (2005), nesse cenário de desenvolvimento da avaliação, justificado a partir da necessidade de modernização e prestação de contas do Estado, o discurso que se constrói em torno dela assume um caráter técnico e normativo, despindo, dessa forma, o viés político que a avaliação pode assumir. Ela é apresentada como um instrumento gerencial para os governos e não como uma opção política dos mesmos. Nesse ponto, vale notar, há uma inadequação na leitura de Faria (idem) se a aplicarmos à avaliação no âmbito educacional, em particular, à avaliação em larga escala. Uma das principais críticas destinadas à avaliação educacional é justamente sua instrumentalização pelos governos, para fins propagandísticos, de modo que os resultados, manipulados ao sabor das necessidades governamentais, mostram uma realidade distorcida, que beneficiaria a imagem dos governos e das políticas que eles adotam. Nessa perspectiva, os elementos técnicos da avaliação não são ressaltados e sequer são mencionados. Na verdade, poderíamos afirmar que há uma inversão desse discurso predominante. O que Faria (2005) vê como o discurso oficial da avaliação como instrumento gerencial ganhou menos espaço, na academia, do que o discurso que a vê como um instrumento a serviço da propaganda de governo, pelo menos, no âmbito educacional. A crítica se destina, justamente, ao uso político (e pejorativo) da avaliação.

A partir da leitura de Guba & Lincoln, Faria (2005) aponta para a existência de 4 gerações de pesquisas em avaliação no contexto norteamericano: i) no período do New Deal, a avaliação assumiu um caráter mais técnico; ii) entre o fim da II Guerra Mundial e os anos 1960, a avaliação era caracterizada por ser descritiva; iii) entre 1960 e 1970, a avaliação ficou caracterizada por fornecer julgamentos abalizados para a tomada de decisões; iv) por fim, sem entrar em muitos detalhes, Faria define a quarta geração, pertencente aos anos Reagan, como reativa<sup>25</sup>.

A década de 1960 foi o período de maior expansão e desenvolvimento da pesquisa em avaliação nos Estados Unidos, quando era entendida como uma ferramenta fundamental para o planejamento de políticas públicas. Seus principais destinatários eram os tomadores de decisão, o alto escalão do governo, e seu principal objetivo era dar legitimidade às decisões governamentais, concedendo a elas um caráter racional, principalmente no que dissesse respeito ao aspecto econômico. Nesse cenário, o modelo de avaliação era *top-down* (FARIA, 2005, p.98).

---

<sup>25</sup> Exceção feita à quarta, que não recebe nenhuma descrição no texto de Faria, as diferentes gerações de avaliação que caracterizam o contexto norteamericano, principal fonte de influência para a construção de nossos próprios modelos de avaliação, em especial, os educacionais, não são excludentes. Em períodos diferentes, o enfoque pode recair em uma de suas características, mas, a rigor, a avaliação pode, simultaneamente, ser técnica (primeira geração), descritiva (segunda geração) e destinada a dar base para decisões (terceira geração).

Esse modelo<sup>26</sup> foi sendo progressivamente substituído pelo modelo *bottom-up*, que pressupunha uma incorporação dos resultados da avaliação na prática profissional, portanto, quando da rotina prática da política, e não apenas em seu momento de formulação e decisão sobre sua continuidade. Já no contexto de expansão do novo gerencialismo público (*New Public Management*), o enfoque passou a ser na mensuração dos resultados da política, para fins de legitimação das decisões governamentais. Mais uma vez, cumpre ressaltar que a distinção das três funções básicas atribuídas à avaliação não leva à conclusão de que elas são excludentes. Inicialmente, a função de destaque era a produção de informações, de modo a fornecer conhecimento sobre a realidade. Essa função é inarredável de qualquer processo avaliativo, não importa que outras funções possa vir a desempenhar. A função de realocação de recursos se baseia na primeira. Sem a produção de informações seguras sobre a política, seria impossível reajustar a lógica de alocação de recursos. Essa função de realocação já se baseia no uso da avaliação como instrumento gerencial. A avaliação é um instrumento de auxílio para a tomada de decisões. Além disso, ela traz consigo uma questão: a própria avaliação exige recursos. Realocar recursos pressupõe contar com recursos para a avaliação. Por fim, e também articulada às demais, a função de legitimação. Decisões tomadas com base em informações abalizadas possuem mais legitimidade do que aquelas que não se baseiam em nenhuma evidência. Como as decisões pressupõem escolhas, optar por uma política pública significa deixar de implementar outra, levando a desgastes diante de diferentes grupos de interesses. Escolher tem custos políticos. A avaliação, ao fornecer informações que ajudam na tomada de decisões, contribui para reduzir o desgaste diante dos grupos cujos interesses não foram atendidos.

A questão da alocação de recursos com base nos achados da avaliação leva a outro ponto, ressaltado pelos críticos, que tem sido tratado, no Brasil, como uma ‘epidemia’ de avaliação (VIEIRA, 1997). A alocação de recursos baseada em informações produzidas pelas avaliações pode levar a decisões de descentralização do Estado, seguindo o caminho de privatizações, com base no argumento de que o Estado gastaria menos transferindo a execução de determinado serviço para a iniciativa privada. A privatização é vista, pelos críticos, como o principal caminho para a desresponsabilização do Estado. As críticas dirigidas à privatização são de mesma natureza daquelas endereçadas ao novo gerencialismo público e à avaliação como instrumento gerencial.

---

<sup>26</sup> Esses modelos, como vimos, não dizem respeito apenas à avaliação, mas ao ciclo das políticas como um todo, em especial, no que tange à implementação.

Estamos diante de um ponto importante para a presente tese. A reforma do Estado baseada no paradigma do novo gerencialismo público, a defesa de um Estado mais eficaz e controlado em relação a seus gastos, mais flexível, pautado pela transparência de suas ações e pela responsabilização de seus agentes, cenário em que a avaliação de políticas públicas encontrou terreno para seu desenvolvimento no Brasil, não podem obscurecer o fato de que a avaliação não é, por essência, pertencente a um tipo específico de Estado. Não há nada, em suas características, que a vinculem a um modelo social e político específico. De fato, o que diferirá um modelo de Estado do outro, uma concepção de governo da outra, no que se refere à avaliação, é o uso que se faz de seus resultados, não sua existência e sua importância. No contexto brasileiro, por exemplo, a avaliação educacional em larga escala não pertence a uma concepção específica de governo. Como veremos no próximo capítulo, a avaliação educacional se desenvolveu no cenário nacional ao longo do tempo em governos dotados de aparatos ideológicos muito distintos, muitas vezes, antagônicos.

A crítica de Faria (2005) ao modelo gerencialista, que deu, e ainda dá, base à reforma do Estado na América Latina, a despeito de sua pertinência, não pode ser confundida com uma crítica que retira da avaliação de políticas toda sua legitimidade. Instrumento e uso do instrumento não podem ser aqui confundidos. Se Faria (2005) identifica uma hegemonia do gerencialismo (o que ele critica), isso não significa que a avaliação de políticas, que, de fato, se desenvolveu sob a égide desse modelo, serve apenas aos propósitos de tal modelo (em geral, tratados como escusos).

Como vimos, o que Faria (2005) critica é o discurso técnico e gerencial da avaliação de políticas, que olvida seu caráter político. Contrapondo essa visão de Faria, afirmamos que o discurso mais difundido sobre avaliação é justamente o contrário: omite o elemento técnico para dar voz apenas ao aspecto político (no caso, tratado como pejorativo). Estamos diante de mais um caso onde uma escolha polarizada não precisa ser feita. Reconhecer o uso político da avaliação não pressupõe negar seu caráter técnico. O contrário também é verdadeiro.

O cerne dessa discussão se relaciona com o uso feito da avaliação. Em geral, ela é tipologizada em função de seus objetivos [Faria fala em *timing* – antes (diagnóstica), durante (formativa) ou ao fim (somativa) da execução de determinado programa ou ciclo do programa], da posição do avaliador (externa ou interna) e do tipo de objeto avaliado (avaliação de resultados, de processos, de contexto etc) (FARIA, 2005, p.101). É possível, igualmente, construir uma tipologia de uso dos resultados da avaliação. Inicialmente, quando da expansão da avaliação, a expectativa era de que os resultados funcionassem como subsídios para a tomada

de decisões (ver BOUDETT et ali, 2013). Decisões qualificadas por informações qualificadas, produzidas segundo critérios rigorosos e confiáveis.

Não se pode dizer que essa expectativa é parte do passado, do momento inicial em que as avaliações ganharam o cenário das políticas públicas. Da mesma forma, a crítica que avaliação pode (e deve, já que, como todo instrumento de mensuração de um fenômeno, está sujeita a aprimoramentos contínuos) receber não deve obliterar seu objetivo de dar suporte a decisões mais qualificadas. A busca desse objetivo permanece viva e não deve ser caracterizada, cinicamente, como mera ingenuidade. A maior dificuldade enfrentada pela avaliação não é a ausência de clareza quanto ao tipo de uso que justifica sua existência (o suporte para decisões mais qualificadas), mas, antes, a incorporação de seus resultados na rotina de trabalho dos profissionais. Ao discutir os fatores que interferem em seu uso, o próprio Faria fala em propósito original da avaliação, qual seja, “melhorar a qualidade das decisões e garantir a maximização da consecução dos objetivos definidos pelas políticas e programas” (FARIA, 2005, p.102).

No que diz respeito aos problemas enfrentados para que os resultados de uma avaliação sejam utilizados, Faria, citando Carol Weiss, apresenta um conjunto deles:

(...) existência de crenças e interesses conflitantes na organização que gerencia o programa; ocorrência de conflitos de interesses entre as distintas unidades do programa; mudanças no pessoal encarregado, quando, por exemplo, os novatos têm prioridades diferentes daquelas vigentes na época do início da avaliação; eventual inflexibilidade das regras e dos padrões operacionais da organização encarregada, que pode impedir a adoção das recomendações feitas quando da avaliação; e mudanças nas condições externas, tais como cortes orçamentários e alterações no ambiente político, que podem tornar impossível para a organização responder à necessidade de mudança revelada pelos avaliadores (FARIA, 2005, p.102).

Em relação aos resultados da avaliação educacional, as dificuldades que se levantam para seu uso apresentam características próprias. Os conflitos de interesses entre as diferentes unidades do programa e na própria gerência do programa, bem como a troca de pessoal e as mudanças no cenário político, são fatores que podem obstaculizar os usos dos resultados da avaliação. No entanto, como veremos, a avaliação educacional em larga escala enfrenta problemas característicos de sua história no cenário brasileiro, especialmente, em virtude da forma como foi implantada entre nós.



Faria (2005) prevê uma tipologia para o uso dos resultados da avaliação. Como toda tipologia, ela se vale de critérios arbitrariamente selecionados, de acordo com o juízo de valor que guia o pesquisador, e não se apresenta como um rol taxativo, antes, exemplificativo. O autor fala em quatro tipos de uso da avaliação. O primeiro deles é o instrumental<sup>27</sup>, que se apresenta como o mais comum, no argumento do autor, diante de quatro cenários: i) quando os resultados da avaliação apresentam questões com implicações não muito controvertidas; ii) quando as mudanças que podem ser propostas a partir dos resultados já estavam previstas pelo programa ou, mesmo não previstas, não possuem custos adicionais elevados; iii) quando o programa é estável (ambiente, lideranças, orçamento, beneficiários etc); iv) quando o programa passa por uma crise a ponto de não se ter clareza sobre o que deve ser feito com ele (2005, p.103).

O segundo tipo de uso é chamado de conceitual, significando a utilização dos resultados feita exclusivamente por técnicos do programa e tomadores de decisão. Os resultados levariam, nessas circunstâncias, a uma mudança desses profissionais em relação à forma como entendem e conduzem determinado programa, mudando sua concepção, características, modo de implementação etc. Trata-se um uso voltado para a aprendizagem daqueles que dão vida ao programa, acentuando-se que, quanto mais tais atores participam da avaliação do programa, mais esse processo de aprendizagem ganha espaço.

O terceiro tipo é o uso da avaliação como instrumento de persuasão. Aqui, o uso se destina a obter maior legitimidade diante de uma tomada de decisão ou de uma posição assumida diante do enfrentamento de um problema através de uma política pública. A utilização dos resultados, nesse caso, busca reforçar o apoio já dado por determinados grupos e angariar novos apoiadores. É importante, em relação a esse tipo de uso, estabelecer uma diferença. De um lado, a avaliação como instrumento de persuasão pode ser entendida de forma pejorativa, como elemento propagandístico e enganoso, destinado apenas a envernizar determinada tomada

---

<sup>27</sup> Vale notar que o termo instrumental pode não ser o mais adequado para servir como um dos elementos de uma tipologia. Na medida em que serve de apoio à tomada de decisões qualificadas, tal como definidos em seu sentido original (ponto reconhecido pelo próprio Faria), os resultados da avaliação servem sempre a um propósito maior do que eles mesmos, a algo externo à própria avaliação (no caso, à tomada de decisões sobre o programa ou política). Nesse sentido, se todo resultado serve a algum propósito externo, ele é sempre instrumental (usado como instrumento para alguma outra coisa). A instrumentalidade seria, assim, uma característica da avaliação, e não um critério para definir tipos. Como critério, ele não se presta a realizar uma distinção. É importante ressaltar, ainda, que Faria (2005) não apresenta uma definição do uso instrumental da avaliação, se limitando a citar os cenários em que tal uso encerra sentido. Soma-se à ausência de especificidade desse tipo de uso o fato de que o autor chama atenção para a necessidade de que a avaliação tenha qualidade, com resultados divulgados de forma adequada e inteligível, apresentando recomendações factíveis de mudança. Essas características não são particulares do uso instrumental da avaliação, sendo sua observância necessária para qualquer uso que se faça dos resultados de uma avaliação.

de decisão com base no caráter de evidência que seus resultados assumem. A avaliação, dessa forma, preencheria uma lacuna da política, cobrindo um vazio real com resultados fictícios e manipulados. No caso das avaliações educacionais em larga escala no Brasil, por exemplo, essa interpretação é oriunda de sua própria história de implementação, além de usos equivocados do instrumento por parte de governantes. Há, em complemento, uma questão linguística. A própria ideia de persuasão é contaminada pelas críticas que a retórica recebeu ao longo da história do pensamento ocidental. A persuasão é vista, equivocadamente, como a arte do engodo. De outro lado, superando a contaminação semântica e prestando a devida atenção às mudanças experimentadas pelo uso dos resultados ao longo do tempo, o uso persuasivo da avaliação é oriundo da proposta original de uso dos resultados. Qualificar decisões é uma forma de legitimá-las em contextos de escolhas difíceis. As escolhas precisam ser justificadas. Legitimá-las não é apenas um ato propagandístico. É uma maneira de justificar uma decisão em um cenário onde outras seriam possíveis. Isso parte do modelo democrático de organização social. Nesse sentido, o uso persuasivo dos resultados da avaliação não é um problema, mas uma necessidade. Os resultados da avaliação ajudam a tomar decisões, tornando-as mais legítimas aos olhos da sociedade como um todo. Esse é um dos princípios que justificam a existência das avaliações.

O quarto uso da avaliação diz respeito ao esclarecimento. As avaliações produzem informações sobre os fenômenos sobre os quais se debruçam. Essa é sua função imediata e direta. Com base nessas informações, a princípio, qualificadas, já que seguem uma metodologia própria de produção que garante sua validade, as outras funções tornam-se possíveis. Podemos dizer que esse é o uso primário da avaliação. Faria (2005) destaca que o uso para esclarecimento faz com que a avaliação seja capaz de alterar crenças e hábitos dos profissionais e instituições aos quais seus resultados são direcionados. É como instrumento de conhecimento que a avaliação torna-se capaz de dar suporte a decisões mais qualificadas, retirando disso sua legitimidade. Conforme já salientado, o grande desafio da avaliação educacional no Brasil atualmente é fazer com que os resultados da avaliação sejam, de fato, incorporados às rotinas dos profissionais.

A despeito do tipo de uso que é dado para a avaliação, Faria (2005), seguindo, mais uma vez, Carol Weiss, destaca que o principal elemento da avaliação utilizado é seu resultado, tratado como descoberta, evidência. Voltaremos a tratar mais detidamente sobre esse ponto em tópico próprio, mas, por agora, vale apresentar os outros elementos da avaliação utilizados. O primeiro deles se refere às ideias e generalizações que podem ser extraídas da avaliação, ou

seja, mesmo que os resultados específicos não sejam direta e imediatamente utilizados, a avaliação pode ser o ponto de partida para ideias mais gerais acerca de um programa ou política. O segundo elemento diz respeito à própria existência da avaliação. Na descrição de Faria, o fato da avaliação existir já seria uma demonstração de racionalidade e compromisso do governo com suas políticas e programas, mostrando sua abertura para o aprimoramento contínuo e para a responsabilização de seus agentes. O terceiro ponto é controverso e, no âmbito educacional, foi alvo de muitas críticas. Trata-se da concentração dada ao que é avaliado. No caso da avaliação educacional, significa o enfoque dado pelas escolas naquilo que é avaliado. Como a avaliação não abarca todo o currículo, a crítica aponta que o instrumento avaliativo é reducionista e leva, necessariamente, a uma desconsideração daquilo que não é objeto de avaliação. Retomaremos esse ponto mais à frente. Por ora, é preciso notar que dar foco a determinados aspectos não é algo necessariamente ruim. O último elemento apresentado é o desenho da pesquisa. A metodologia, a forma de coleta e análise dos dados, bem como a maneira de torná-los públicos, podem ser utilizados como exemplos para o governo, levando à criação de determinados aprendizados e hábitos no que tange às políticas públicas.

Em relação a esse último ponto, qual seja, o da publicização dos dados, é necessário chamar atenção para os destinatários das informações produzidas pelas avaliações. Tradicionalmente, os destinatários dos resultados das avaliações eram os tomadores de decisão, os governantes e os planejadores das políticas e programas, posicionados em um nível mais elevado da hierarquia governamental. As informações das avaliações eram pensadas como mais adequadas aos atores do alto escalão do governo. O modelo *top-down* é ajustado a essa concepção. À medida que esse modelo sofreu críticas e abriu espaço para a participação fora do comando de governo, os resultados passaram a ser destinados a outros atores, como aqueles que, de fato, implementariam determinado programa na prática (a chamada *street level bureaucracy*). O passo seguinte foi levar em consideração os beneficiários da política, divulgando os resultados de modo ajustado a seus interesses e necessidades, tornando-os mais responsivos aos participantes dos programas. Evidentemente, outros usuários podem ser pensados, como os pesquisadores, estudiosos, acadêmicos, instituições interessadas e a sociedade civil como um todo.

Em relação aos resultados, vale chamar a atenção para uma observação feita por Faria, que assume o tom de crítica. De acordo com ele (2005, p.103), o que chama mais atenção nas avaliações são os seus resultados finais, ou suas descobertas. De fato, eis aqui um problema que não pode deixar de ser mencionado. No que se refere à avaliação educacional, muitas vezes,

seu uso faz com que seus resultados pareçam verdades absolutas, incontornáveis, como se qualquer outra informação produzida sobre a mesma realidade, mas sem passar pelo crivo do instrumento avaliativo, não tivesse valor. Os críticos verão nisso uma espécie de tirania da avaliação, principalmente como critério para definir a qualidade da educação. É importante, nesse ponto, a distinção entre uso e instrumento, tema que nos ocupará em tópico próprio. Os resultados da avaliação educacional não encerram uma verdade absoluta. Eles podem ser contestados. Aliás, é exatamente isso que os coloca no plano científico e não dogmático. A informação produzida é qualificada em razão da confiança que se tem na metodologia adotada, mas não é a única possível para analisar o fenômeno educacional. A força da avaliação não está, portanto, em seu dogmatismo, mas no fato de seguir determinados procedimentos para que a informação seja produzida.

No texto de Faria (2005), tem destaque o vínculo entre o desenvolvimento do que ele chamou de uma política de avaliação das políticas e a expansão da perspectiva do novo gerencialismo público. Esse vínculo é visto pelo autor como um problema e a expansão da avaliação, nesses termos, é compreendida sobre o viés da desconfiança, como necessariamente vinculada a organismos financeiros, colocando muita ênfase na relação entre custo e benefício. Além disso, outros dois pontos apontados por Faria direcionam críticas à avaliação: i) as avaliações pressupõem atribuição de valor e definição dos critérios a serem avaliados, atividades discricionárias; e ii) o *timing* das avaliações é o ciclo eleitoral, ou seja, as avaliações se prestam a fins eleitorais, sem que se espere a maturação dos resultados de uma política, quando, na verdade, a formulação inicial da política, com a definição de seus objetivos, é que deveria ser o guia para sua avaliação.

Em relação ao primeiro ponto, a discricionariedade para a escolha do que será avaliado, não é, por si só, um problema da avaliação. A escolha se dá por uma série de motivos, entre eles, os de caráter técnico: só se avalia o que pode ser avaliado em termos metodológicos. A metodologia da avaliação tem suas limitações, embora tenha evoluído cada vez mais ao longo dos anos. Atribuir valor é uma tarefa exigida por todo processo avaliativo, incluindo as avaliações no âmbito da escola, e não uma característica da avaliação de políticas. O problema não está em atribuir valor a determinado resultado, mas, sim, em não ter critérios bem definidos para que a avaliação ocorra. O segundo ponto deixa claro o problema gerado pela confusão entre instrumento e uso do instrumento. O uso eleitoral das avaliações não é um problema intrínseco ao instrumento, antes, diz respeito ao uso dado a ele.

Seguindo outros autores que chamaram a atenção para os mecanismos de poder que caracterizam todas as fases de definição das políticas públicas, Faria (2005, p.102) entende a etapa da avaliação como sujeita a todos as disputas de poder que compõem esse cenário. A política da avaliação de políticas públicas, portanto, em sua leitura, é mais um componente das disputas eleitorais. O problema desse tipo de perspectiva é que ele pode levar a interpretações polarizadas e maniqueístas sobre o fenômeno da avaliação. Interessante notar que o desenvolvimento da avaliação nunca é visto sob o prisma de sua função como base para a tomada de decisão, por sua utilidade técnica e por sua capacidade de auxiliar o aprimoramento de políticas públicas. Ou seja, a avaliação não costuma ser vista como dotada de mérito próprio, apenas como joguete nas mãos de atores disputando a legitimidade do campo. Essa é uma forma de reduzir a realidade, empobrecendo-a. Se os mecanismos de poder, de fato, existem, a avaliação, por outro lado, apresentou grande capacidade de se adaptar a diferentes orientações ideológicas, permanecendo como uma das políticas públicas mais duradouras no cenário educacional brasileiro.

A avaliação de políticas tem como objetivo primordial produzir informações para corrigir o curso da ação. A avaliação tem um olhar para o passado, em termos de prestação de contas, e um olhar para o futuro, de correção da ação. Na abordagem do ciclo de políticas, ela encerra um ciclo e funciona como gatilho para o início de outro. Tanto a avaliação de processos<sup>28</sup> quanto a avaliação de impacto<sup>29</sup> deveriam ser tratadas como instrumentos para a aprendizagem política (como potencializar os efeitos de uma política, que ações repetir, que ações evitar etc). Contudo, no contexto das políticas públicas, há pouco espaço para essa aprendizagem, impedida de acontecer em virtude das disputas de poder.

No entanto, acreditamos que os obstáculos enfrentados para que a avaliação se constitua como um instrumento de aprendizagem política vão além das disputas que permeiam o campo. Antes de tudo, é preciso reconhecer que ela possui altos custos políticos, em virtude da exposição do fracasso diante de volumosos gastos (por isso, muitas vezes, os resultados das avaliações são desqualificados como propagandísticos a favor da imagem do governo). Além

---

<sup>28</sup> A avaliação de processos busca aferir a eficácia de uma política, ou seja, se ela está sendo implementada de acordo com o que ficou estabelecido para sua implementação. É uma avaliação das ações de execução.

<sup>29</sup> A avaliação de impacto busca aferir os efeitos previstos pela política, ou seja, se ela produz alterações no contexto social no qual tem lugar. Aqui, a análise é de causalidade, e essa é uma grande dificuldade. O que se busca é reconhecer o vínculo causal entre a existência da política e a observância dos efeitos. Para que essa avaliação possa ser realizada de maneira adequada, é necessário que os objetivos e metas da política sejam claros, que o critérios de sucesso da política sejam especificados e que a metodologia escolhida seja adequada para reconhecer tal vínculo. Voltaremos a tratar desse assunto em capítulo próprio.

disso, há dificuldades metodológicas para sua realização (os instrumentos avaliativos não são triviais, bem como os resultados que produzem; a exigência de uma qualificação técnica é fator preponderante para que uma avaliação seja realizada, inclusive, por entidades externas ao governo, o que, a princípio, seria um fator de credibilidade para seus resultados). Por fim, como consequência desses dois fatores, a avaliação ainda se ressentir da falta de uma cultura e de uma tradição estabelecidas para sua realização.

### **3 POLÍTICAS PÚBLICAS EDUCACIONAIS, DE ALFABETIZAÇÃO E DE AVALIAÇÃO**

O objetivo deste capítulo não é realizar um inventário das políticas públicas educacionais no Brasil, algo que exigiria um esforço diferente e uma tese com outros propósitos. O amplo espectro de políticas educacionais, num esforço dessa natureza, exigiria que se levasse em conta as ações em cada um dos entes federados, tendo em vista a estrutura organizacional da educação no país, com estados e municípios assumindo parte significativa das políticas educacionais em nível local. Soma-se a isso a relação entre essas iniciativas regionais e aquelas de alcance nacional. Da mesma forma, seu objetivo também não é selecionar determinadas políticas educacionais nacionais para uma análise mais profunda de sua elaboração, implementação e efeitos. O que se busca com esse capítulo é i) apresentar algumas características gerais das políticas educacionais no Brasil, ii) apresentar os principais aspectos enfrentados pelas políticas de alfabetização e iii) discutir, com mais deter, elementos centrais da política de avaliação educacional. O capítulo está organizado com essa linha de raciocínio: primeiro, uma breve análise das políticas educacionais no Brasil, de modo a extrair dela elementos pertinentes para uma compreensão mais contextualizada das políticas de avaliação, quais sejam, a influência da agenda internacional, a descentralização/centralização, a descontinuidade/continuidade e a influência do gerencialismo e da racionalidade técnica; em seguida, a exposição dos principais eixos de enfrentamento das políticas de alfabetização, no intuito de analisar a posição ocupada pela avaliação neste cenário; e, por fim, objetivo principal do capítulo, apresentar uma análise das políticas de avaliação no Brasil, apontando suas principais características e fins, seus fundamentos e as críticas que recebe, vinculando-os com as características gerais reconhecidas nas políticas educacionais.

#### **3.1 POLÍTICAS PÚBLICAS EDUCACIONAIS NO BRASIL**

Conforme apresentado anteriormente, este tópico busca discutir, de modo breve e com o objetivo de estabelecer relações com as políticas de avaliação, algumas das principais características atribuídas às políticas públicas educacionais no Brasil. A primeira delas diz respeito à influência da agenda mundial na área de educação.

O desenvolvimento de uma sociedade globalizada, propiciada pelo desenvolvimento tecnológico, em particular, no que diz respeito à comunicação, pode ser avaliado em diferentes

sentidos, entre a quebra de valores tradicionais, aumento da liberdade e busca pela identidade com base na diferenças e o avanço, concomitante, de uma forma única de pensar e agir, dominada pelo mercado mundial e pelo capitalismo predatório, conduzido, em sua dimensão ideológica, pelo neoliberalismo. A sociedade em rede de Castells (2016) compartilha mais do que informações. O diagnóstico é de um avanço da lógica capitalista, e, portanto, dos valores que a acompanham, por todo o cenário mundial.

A educação, como produto social, não escapa a essa lógica. O discurso é de uma espécie de colonização das políticas educacionais pelo modelo neoliberal, pernicioso e nocivo à própria ideia de escola, transformada em instituição reprodutora do cenário de desigualdades já estabelecidas (BOURDIEU; PASSERON, 1992). Nesse sentido, é impossível interpretar a escola e as políticas educacionais fora do contexto do capitalismo. Entender as políticas em educação passa pela compreensão do contexto social, econômico e cultural que o permeia.

A internacionalização das políticas educacionais é parte do fenômeno de expansão do capitalismo, acompanhando as políticas econômicas (o fortalecimento do livre mercado e o enfraquecimento do Estado de Bem-Estar Social) e a agenda da democracia, ambas também internacionalizadas. Essa relação, entre políticas educacionais e capitalismo, se dá em duas direções: i) de um lado, implica o reconhecimento de que as políticas educacionais sofrem os efeitos das mudanças no mercado e na vida política, sendo influenciadas pela lógica do capitalismo; ii) de outro, e simultaneamente, as políticas em educação servem ao propósito de alavancar e consolidar a lógica capitalista vigente, dando base para sua expansão. Oliveira (2000) afirma que, no processo de globalização do início do século XXI, as reformas educacionais desempenharam um papel crucial para a reestruturação do capital. Em outro trabalho, a mesma Oliveira (2013) aponta que as reformas educacionais contemporâneas têm como objetivo primordial a formação de “força de trabalho adequada às exigências últimas do capitalismo” (idem, p. 18), servindo também ao propósito de disciplinar a pobreza. No lugar de buscar a justiça social em uma sociedade desigual, as políticas em educação servem à reprodução da desigualdade. Frigotto (2013) segue a mesma linha de Oliveira (2013), identificando a “hegemonia de concepções neoconservadoras e mercantis na sociedade e nos processos educativos” (FRIGOTTO, 2013, p.65). A essa hegemonia, ele chama de “regressão no plano da teoria social e educacional” (FRIGOTTO, 2013, p.67), que ocorre tanto no plano nacional quanto no mundial.

Dourado (2007) também defende que as políticas educacionais dos últimos anos só podem ser entendidas no contexto de reforma do Estado, que serve, justamente, aos ditames de



uma agenda que não é nacional, mas definida mundialmente. Esse movimento é encabeçado por “agências e/ou organismos multilaterais e fortemente assimilados e/ou naturalizados pelos gestores de políticas públicas” (DOURADO, 2007, p.922). Frigotto (2013) também ressalta o papel desempenhado pelos organismos internacionais na definição das agendas nacionais para a educação:

Nesse contexto, as políticas e a gestão educacional tendem, especialmente em países de capitalismo dependente, a se pautar pelas determinações dos organismos internacionais – legítimos representantes do capital mundial. Esses organismos têm forçado políticas focais fragmentadas e a gestão educacional centrada na perspectiva da mercantilização. (FRIGOTTO, 2013, p.69).

No mesmo sentido, Oliveira (2013):

Na América Latina e no Caribe viveu-se uma onda de reformas educativas, a partir dos anos 1990, em que os organismos internacionais vinculados à ONU tiveram forte influência na determinação das políticas nacionais. Tal influência foi exercida por meio de assistência técnica prestada aos ministérios ou na forma de empréstimos que terminaram por definir o tipo de empreendimento educativo que era orientado pelas linhas de financiamento disponíveis (OLIVEIRA 2013, p.25).

Outros exemplos de argumentos nessa linha podem ser encontrados na literatura<sup>30</sup>: em Feldfeber (2013), ao tratar da internacionalização da educação na América Latina, questiona se esse movimento aponta para uma convergência e um intercâmbio positivo de ideias ou se representa apenas uma mercantilização da educação (deixando claro que estamos diante do segundo movimento – ela fala em educação vítima do livre comércio – FELDFEBER, 2013, p.177); em Lima e Afonso (2002), identificando as influências do neoliberalismo e dos organismos internacionais, como a Unesco, na transformação da escola em um modelo empresarial; em Krawczyk (2000), reconhecendo o papel dos bancos na definição de políticas educacionais na América Latina; em Castro (2002), questionando a possibilidade e capacidade de bancos, mais especificamente, o Banco Mundial, para a condução de reformas educacionais.

Não se trata de negar a influência dos organismos internacionais na agenda educacional dos países em desenvolvimento, nem mesmo os efeitos perniciosos que tal influência pode gerar, visto que pautada por interesses específicos desses organismos, nem sempre casados com

---

<sup>30</sup> Diferentemente do que é elencado a seguir, há quem busque pesquisar a influência dos organismos internacionais como facilitadores no processo de construção de uma padrão mundial de educação baseada na equidade, como, por exemplo, McNelly (1995).

as necessidades das políticas locais. Com apoio técnico, mas, principalmente, com financiamento, a ação desses organismos pode ser abusiva e desequilibrada. É o que grande parte da literatura sobre tema apresenta. Em que pese este aspecto, absolutamente real e digno de crítica, parece que, muitas vezes, escapam aos críticos alguns elementos importantes para uma compreensão mais acurada da internacionalização das políticas educacionais.

Conforme vimos nos autores supracitados, o fenômeno da internacionalização ocorre de maneira unilateral, uma imposição dos países desenvolvidos aos países em desenvolvimento. Essa imposição é dada pela ideologia, convencendo-nos de que o melhor para nós é que aceitemos as políticas mundiais, e/ou pela dependência técnica e financeira, na medida em que aceitaríamos a imposição por precisarmos de transferência de conhecimentos e de valores. Parece-me uma visão não totalmente adequada para o entendimento dessa relação, por dois motivos, em especial: i) as políticas mundiais jamais são integralmente impostas, copiadas tal qual ocorrem nos países desenvolvidos ou em outros cenários; e ii) os países em desenvolvimento são formados apenas por tomadores de decisão absolutamente inocentes, cegos e suscetíveis de enganação ideológica, a ponto de serem vítimas de organismos internacionais. Nos dois casos, a pesquisa em educação iria se beneficiar com estudos mais acurados no sentido de captar, *in loco*, como ideias e políticas internacionais são apropriadas, realizadas e reajustadas pelos atores.

Nesses dois casos, o que está em jogo é reconhecer que, por mais que o capitalismo tenha se alastrado mundialmente, arrastando consigo não apenas práticas e instituições, mas também ideias e visões de mundo, esse movimento não é liso, fluido, despido de incongruências. Os contextos locais resistem, consciente e inconscientemente, e, muitas vezes, simplesmente não são ajustados para uma cópia irrefletida da política. Os atores educacionais estão sujeitos a influências ideológicas, evidentemente, mas não é possível presumir que são apenas marionetes nas mãos de organismos internacionais. O procedimento de implementação de uma política educacional, por mais que seja influenciado por elementos do capitalismo mundial, não ocorre de forma tão previsível e estanque. Os atores refletem, tomam decisões, com níveis diferentes de consciência e informação, é verdade, mas não podemos eliminá-los da equação. Uma política educacional internacional, ao se defrontar com um contexto local, não é simplesmente imposta, mas ajustada, rearranjada, com retirada de aspectos não ajustáveis à realidade local e contando com a inclusão de outros. O capitalismo se expande, inclusive na esfera educacional, mas essa história nunca é linear como parece nas palavras de parte significativa da crítica.

A agenda mundial e a influência internacional nas políticas educacionais apresentam desdobramentos em outras características dessas políticas. É o caso da expansão da racionalidade técnica. As reformas educacionais dos últimos 30 anos encontram na Conferência de Jomtien, em 1990, uma espécie de data de nascimento (BROOKE, 2012; MACHADO; ALAVARSE; ARCAS, 2015). A partir dali, com a educação sendo tratada como elemento central para o desenvolvimento econômico, os vínculos entre mercado, neoliberalismo, gerencialismo e educação passaram a ser denunciados como parte do movimento global de produção de políticas educacionais, cujas críticas foram apresentadas em linhas antecedentes.

Ainda falta uma pesquisa mais profunda sobre o gerencialismo em educação, sobre suas características e seus efeitos. As críticas a esse modelo são conhecidas, mas suspeito que haja mais por trás desse processo. De todo modo, em geral, podemos afirmar que fazem parte das características do gerencialismo em educação uma preocupação com a eficácia dos processos educacionais, uma maior preocupação com o papel da gestão (o que leva a críticas sobre a transformação da escola em um modelo de empresa, como em Dourado, 2007, por exemplo) e da formação do gestor, uma educação baseada no desenvolvimento de competências, entre outras. A preocupação com elementos técnicos levou a críticas sobre a ausência de preocupações políticas e sobre a invasão, na escola, de uma lógica mercadológica, que enaltece o mérito ao passo em que aprofunda desigualdades sociais. O mérito está na própria origem da escola republicana francesa, modelo herdado pelo Ocidente (mais sobre isso será dito em tópico vindouro), de modo que não pode ser atribuído ao avanço do neoliberalismo. Além disso, não parece convincente acreditar que a preocupação com o elemento técnico é antagônica ao elemento político. Eles podem conviver. No caso da gestão escolar, em específico, espera-se que eles convivam. É o caso de Polon (2009), ao buscar identificar perfis diferentes para a gestão (organizacional, pedagógica e relacional), e de Lück (2009), ao discutir padrões de competência para o exercício da função escolar.

Para a crítica ao neoliberalismo e a seus efeitos na educação, a racionalidade técnica característica do gerencialismo é uma ameaça ao ideal da escola como uma instituição, numa sociedade desigual como a nossa, destinada a promover justiça social. Ancorado cegamente no mérito, o gerencialismo valorizaria o indivíduo abstrato, sem observar as desigualdades que os caracterizam, fenômeno típico do ideário de uma educação voltada para competências e a serviço da economia de mercado, que submete a escola à sua lógica:

Trata-se de reduzir, no plano educacional, um ideário em que não há lugar para todos, e o problema não é coletivo, mas individual. As competências que serão desenvolvidas e que garantem empregabilidade são as que o mercado

reconhece como as que tornam cada trabalhador produtivo ao máximo. (FRIGOTTO, 2013, p.68).

O resultado disso só pode ser o aprofundamento do quadro de injustiça e o diagnóstico da escola como instituição a serviço do reforço desse cenário. No âmbito da gestão, o avanço do gerencialismo implicaria a despolitização da função, transformando o gestor em um mero administrador, alienado em relação aos problemas de natureza política e social que afetam a escola. A mesma preocupação é direcionada para os professores, que seriam reduzidos a disseminadores de técnicas e metodologias de aprendizagem, mas não seriam capazes de compreender a natureza política do processo de ensino e aprendizagem (DOURADO, 2007).

Essas críticas atingem em cheio à avaliação em larga escala, vista como um instrumento típico dessa onda neoliberal que invade a educação (FRIGOTTO, 2013). A avaliação representa, aos olhos da crítica, uma ode ao mérito individual<sup>31</sup> e à eficácia, critérios do mercado, mas não da escola. A avaliação é a própria racionalidade técnica encarnada, motor do movimento de despolitização da escola e de seus atores, portanto, num salto lógico, a conclusão é que ela se apresenta como um instrumento antidemocrático. A oposição entre gerencialismo, técnica e avaliação, de um lado, e democracia e elementos políticos, do outro, pode ter origem na própria forma de organização das políticas públicas educacionais no Brasil a partir da metade do século XX, de modo que as características de um lado e de outro foram mais ou menos ressaltadas de acordo com o cenário vivido em determinado período de tempo.

Ao analisarem diferentes gerações de políticas educacionais para fins de melhoria da qualidade da educação, desde a década de 1970, Franco, Alves e Bonamino (2007) observam seu movimento em torno da racionalidade que as caracteriza. De modo geral, conforme argumento dos autores, a década de 1970 foi caracterizada pela valorização na racionalidade técnica, baseada em uma legislação que afirmava a necessidade de profissionais da educação, com formação adequada para tanto, para o exercício de funções de administração, supervisão e orientação educacional, incluindo uma formação de professores nos moldes do comportamentalismo, baseada em iniciativas em tecnologia educacional (FRANCO; ALVES; BONAMINO, 2007, p.993-994). A crítica dos professores ao comportamentalismo adotado como modelo de formação e sua resistência ao poder dada pela legislação aos especialistas oriundos dos cursos de Pedagogia foram, segundo os autores, elementos centrais para o

---

<sup>31</sup> A dimensão heurística da avaliação, que encontra nas pesquisas em eficácia escolar sua principal encarnação, não é de todo conhecida, inclusive pelos críticos, e lida, justamente, com a consideração do contexto, individual e escolar, e não com uma leitura abstrata dos atores e instituições. Mais sobre isso será dito em tópico à frente.

desenvolvimento de uma forte resistência à racionalidade técnica, principalmente, quando consideramos o contexto político da época, marcado pela autoritarismo da ditadura militar.

A geração de políticas educacionais seguintes a essa, já na década de 1980, buscou solapar as características progressas, levantando a bandeira da democratização e da renovação do elemento político como identitário para a escola e para a educação de maneira geral. Entre as principais características do período estavam a defesa da eleição de diretores, como instrumento a serviço da democratização e crítica ao exercício predefinido de cargos pelas especialistas da Pedagogia, e a autonomia dos professores e das escolas, contra a centralização marcada do período anterior (FRANCO, ALVES, BONAMINO, 2007, p.996). A ênfase no compromisso político dos atores escolares levou à transformação de uma circunstância em uma regra: a superação ao período anterior significava negar suas características, entre elas, a racionalidade técnica, criando um antagonismo inarredável entre envolvimento político e capacidade técnica, algo que caracteriza a literatura educacional majoritária até hoje.

As duas gerações de políticas educacionais seguintes, compreendendo o período dos governos de Fernando Henrique Cardoso e Lula (o primeiro mandato), reavivaram a racionalidade técnica como um elemento estrutural para a definição e implementação de políticas (idem, 2007). Em especial, importante para o argumento desta tese, é o reconhecimento da valorização da política baseada em evidências, própria da racionalidade técnica e base para a avaliação educacional em larga escala (a educação baseada na prova, questionada por SAUSSEZ; LESSARD, 2013). Como veremos em tópico à frente, a avaliação tem como objetivo principal dar suporte à tomada de decisões, fornecendo informação de qualidade para tanto (evidências). Nesse sentido, ela é contundente como exemplo de racionalidade técnica. A consolidação da avaliação nacional no período e seu desenvolvimento no governo Lula mostram o papel central assumido por ela para o quadro de políticas educacionais que buscam a melhoria da qualidade da educação. O tema da qualidade como o centro das preocupações educacionais não pode ser dissociado, entre nós, do papel desempenhado pela política de avaliação. Por isso, a avaliação em larga escala é tomada como uma das características mais marcantes das reformas educacionais: “as reformas educativas implantada nas últimas décadas caracterizam-se, entre outros traços, pela utilização de avaliações em larga escala como instrumento de gestão de redes de ensino e de responsabilização de profissionais da educação” (BAUER; ALAVARSE; OLIVEIRA, 2015, p.1367).

Ainda dois temas característicos das políticas educacionais, a continuidade/descontinuidade e a centralização/descentralização. Dourado (2007) aponta a descontinuidade como uma marca indelével da trajetória histórica das políticas educacionais no Brasil:

**(as políticas educacionais)** têm sido marcadas hegemonicamente pela lógica da descontinuidade, por carência de planejamento de longo prazo que evidenciasse políticas de Estado em detrimento de políticas conjunturais de governo. Tal dinâmica tem favorecido ações sem a devida articulação com os sistemas de ensino, destacando-se, particularmente, gestão e organização, formação inicial e continuada, estrutura curricular, processos de participação. (DOURADO, 2007, p.925-926) (acréscimo nosso, em negrito).

O mesmo argumento, das políticas educacionais sendo caracterizadas pela descontinuidade, em virtude, precipuamente, das trocas de governo, pode ser encontrado em Couto (2017). Spink (1987) vê a descontinuidade como um problema que ultrapassa as políticas educacionais, sendo característico das organizações públicas. Em sua visão, esse é um paradoxo da democracia: a rotatividade do governo leva a trocas constantes no âmbito das políticas e da própria forma de organização do poder. Diante disso, as políticas não têm tempo suficiente de maturação para exercer efeitos significativos, tendo em vista que, em regra, principalmente em educação, os efeitos das políticas tendem a ser mediatos.

Embora a descontinuidade seja tratada como uma característica de nossas políticas educacionais, Franco, Alves e Bonamino (2007) apresentam uma leitura que encontra, em período recente do país, uma continuidade nas políticas educacionais, principalmente em torno do que já mencionamos anteriormente, o consenso acerca da racionalidade técnica como guia para sua definição e implementação. Uma dessas políticas é a avaliação em larga escala, que, além de ter resistido a governos de diferentes orientações ideológicas (PSDB e PT), pelo menos no discurso oficial, foi capaz de se expandir e se consolidar em um sistema robusto (CASTRO, 2009b). A política de avaliação é uma exceção ao diagnóstico de alguns autores, como os supracitados, de que a descontinuidade é uma característica marcante de nossas políticas. Alguns aspectos que ajudam a compreender essa continuidade serão discutidos em tópico próprio, à frente.

Ainda um tema antes de passarmos à discussão sobre as políticas de alfabetização. A centralização é outro aspecto ressaltado na análise das políticas educacionais no Brasil. Mello (1991) já chamava atenção para o debate entre centralização e descentralização na vira dos anos 1990, quando a Constituição Federal de 1988 ainda era recente. A centralização era vista, e

permanece sendo, em grande medida, como um atavismo do período militar (o que pode ajudar a explicar seu vínculo com a racionalidade técnica, percebido, especialmente, no âmbito da gestão escolar, como em PARO, 2015; 2017), portanto, contrário ao ideal democrático. No caso de Paro (2015), por exemplo, a própria ideia de liderança do gestor é uma forma de romper com os valores democráticos, na medida em que supõe uma ausência de divisão de protagonismo no exercício dos papéis na escola, concentrando todo o poder nas mãos de uma única figura. Como o gestor e a liderança na escola são elementos importantes para a racionalidade técnica defendida pelo gerencialismo, o vínculo está dado: racionalidade técnica, centralização, gerencialismo.

Essa lógica estende-se ao plano macroeducacional, com a imposição de políticas e programas de cima para baixo (FRIGOTTO, 2013). A centralização do governo federal no âmbito das políticas é visto como uma forma de rompimento com o pacto federativo que, na Carta de 1988, estabelecida a divisão de responsabilidades educacionais entre os entes federados, prevendo a cooperação entre eles como um imperativo político-jurídico. A descentralização proposta pelo pacto federativo é correlata à autonomia do professor e da escola (MELLO, 1991), elemento chave para as políticas da década de 1980, conforme vimos, e a centralização é vista com desconfiança.

Num modelo prototípico, a política educacional a ser criticada é, simultaneamente, determinada pela agenda mundial, pautada pela racionalidade técnica, descontínua e centralizada. Uma política desse tipo encarna o que há de pior: i) é colonizada por mecanismos internacionais, de modo que não apresenta autonomia para o atendimento das necessidades locais; ii) encarna a racionalidade técnica que despolutiza a educação, valorizando o mérito individual e reproduzindo desigualdades; iii) por ser política de governo, não dura o suficiente para exercer efeitos, sofrendo oscilações como aquelas que caracterizam o mercado; e iv) é centralizada, portanto, contrária aos princípios democráticos que deveriam basilar a escola, dando força ao autoritarismo. Mais uma vez, trata-se de uma análise baseada em tipos ideais. Na prática, as diferentes políticas podem encarnar, em diferentes momentos, uma combinação dessas características.

A avaliação em larga escala é vista, muitas vezes, como uma política com essas características. Exceção feita à descontinuidade, visto que é uma política contínua no país desde a institucionalização do Saeb, no início dos anos 1990, ela é criticada como parte da agenda mundial para a educação, portanto, determinada pela força de organismos internacionais a serviço do mercado, portadora de uma racionalidade técnica despolutizada e utilizada como

parte de ações centralizadas e impostas de cima pra baixo. Em tópico a seguir, analisaremos cada um desses pontos em relação à política de avaliação.

Uma palavra final sobre a análise desse tópico. Trata-se de um exercício difícil de ser realizado. A leitura aqui apresentada é dominante na literatura educacional e tem sua razão de ser. Afinal de contas, as políticas educacionais sofrem mesmo com a influência de organismos internacionais, com as vicissitudes da racionalidade técnica, com problemas de descontinuidade e com a centralização das decisões. O problema não está em reconhecer essas críticas, mas, antes, em *perceber seus limites*. Se a agenda internacional influencia, ela, contudo, não determina absolutamente (no caso da avaliação, por exemplo, o sistema nacional e os sistemas próprios não são uma cópia do Pisa<sup>32</sup>); a racionalidade técnica, radicalizada, pode levar aos problemas apontados pela crítica, como o aguçamento das desigualdades, mas ela não é, *per si*, antagônica ao envolvimento político (ambas as dimensões são importantes); reconhecer a importância do papel do gestor na escola, algo particularmente importante para a apropriação dos resultados da avaliação educacional, não significa defender a centralização como princípio, mesmo porque liderança do gestor e centralização não são sinônimos. Esses temas são delicados e sensíveis. Uma má interpretação sobre o que está sendo dito pode levar a equívocos acerca do argumento apresentado, extrapolando o terreno científico e adentrando no moral. Tenho consciência dos riscos de ser mal interpretado, por isso, a escrita deste parágrafo, mas considere maior risco não abordar o tema, que me parece urgente e fundamental.

### 3.2 POLÍTICAS DE ALFABETIZAÇÃO

Tratada como direito humano<sup>33</sup>, a alfabetização vê, simultaneamente, crescer sua importância nas políticas educacionais, em âmbito mundial, inclusive, e fracassar os esforços para combater o analfabetismo no mundo, problema que ainda não foi contornado, tendo em vista que cerca de 20% da população adulta não sabem ler e escrever (GONTIJO, 2014). Sobre a influência da agenda internacional nas políticas locais, já tratamos no tópico anterior. Contudo, no caso da alfabetização, tal influência apresenta matizes um pouco diferentes.

Em obra especificamente destinada à análise das políticas de alfabetização no Brasil e da influência que recebe das políticas mundiais, Gontijo (2014) reconhece a “centralidade da alfabetização nas orientações dos órgãos internacionais” (idem p. 5), analisando se o conceito

---

<sup>32</sup> Para uma discussão sobre os impactos do Pisa na literatura educacional no Brasil, ver Dickel, 2010.

<sup>33</sup> Declaração Universal dos Direitos Humanos, art. XXVI.



de alfabetização proposto por eles teve ou não impacto nas políticas de alfabetização no país (entre 2003 e 2014, período analisado pela autora, que coincide, em grande parte, com a seleção do período de análise desta tese). A referida obra de Gontijo será a base das análises aqui empreendidas, por três motivos: i) trata-se de uma obra específica sobre o tema, destinada a analisar em detalhe as políticas de alfabetização; ii) seu período de análise coincide com o período selecionado para a tese; e iii) seu pensamento é representativo das críticas às avaliações da alfabetização, conforme relata Silva (2013), tema importante para o argumento aqui construído.

De acordo com a autora (*idem*), o principal motivo que levou à centralidade da alfabetização nas políticas educacionais do século XXI foi o descompasso entre o galopante avanço tecnológico e a incapacidade de resolver os problemas relacionados ao fracasso escolar, encarnados, entre outras coisas, nos altos índices de analfabetismo (UNESCO; MINISTÉRIO DA EDUCAÇÃO, 2003). Conforme definição de Cook-Gumperz (2008), podemos compreender a taxa de alfabetização como uma espécie de medidor do clima social e da saúde de uma sociedade. A alfabetização é um dos principais indicadores de desenvolvimento de um país (o aumento de suas taxas representa, em regra, redução das desigualdades sociais e dos índices de pobreza). Tal centralidade pôde ser percebida na atenção que a alfabetização recebeu nas metas estipuladas no âmbito do Fórum Mundial de Educação de 2000, que levou à publicação do documento Educação para todos: o compromisso de Dakar<sup>34</sup>, em 2001, e na declaração da ONU, através de resolução própria, da década das Nações Unidas para a Alfabetização, com início em 2001. Apesar dos esforços, os avanços reais foram insuficientes.

A questão central analisada por Gontijo (2014) é o conceito de alfabetização que informa as políticas educacionais. O Plano de Ação Internacional da Unesco, voltado para o combate ao analfabetismo, estabelece, na visão da autora (*idem*), uma ampliação fundamental do conceito de alfabetização, de forma que esse novo conceito deve ser a base para todas as políticas de alfabetização no mundo (GONTIJO, 2014, p.12). A Unesco passou a incluir, na definição de alfabetização, mais do que o domínio da leitura, da escrita e da matemática (aritmética), abrindo espaço para o pensamento crítico e para os valores associados à cidadania. De inspiração freireana (FREIRE, 2009), a alfabetização como prática destinada à libertação do homem revela uma nova fase do conceito para os organismos internacionais, em especial, a

---

<sup>34</sup> Das seis metas definidas, duas delas tratam diretamente da alfabetização: a meta IV, com o objetivo de melhorar em 50% os níveis de alfabetização de adultos até o ano de 2015; e a meta VI, com o objetivo de melhorar a qualidade da educação para todos, garantindo resultados mensuráveis e reconhecíveis, em especial, em alfabetização e matemática.

Unesco, superando concepções diversas<sup>35</sup> (e insuficientes) de alfabetização, a principal delas, a que a entende como um conjunto autônomo de competências. De acordo com Gontijo (2014, p.15), esse é o entendimento mais comum do termo e aquele que deu base, ao longo da história brasileira, à produção de materiais e aos métodos de ensino. A divergência entre os adeptos dessa concepção girava em torno dos modos de aquisição de tais competências, dando mais ênfase ao aspecto fonético (linguístico) ou ao semântico, com foco na construção do sentido e na compreensão global do significado. Aqui entra em cena um debate que permanece vivo e polêmico, lançando luz sobre a relação saussuriana entre significante (a atenção à relação entre fonemas e grafemas) e significado (a atenção ao sentido), transformada já em oposição (vale lembrar que Saussure entendia as duas dimensões como partes essenciais do signo, SAUSSURE, 2006)<sup>36</sup>.

Essa oposição entre o ensino com atenção ao fonético e o ensino que privilegia o significado global é mais do que pedagógica e epistemológica. Ela revela uma oposição política e moral também. Os defensores do ensino das unidades da língua são vistos por seus opositores como tecnicistas, reducionistas, excessivamente abstratos e descontextualizados, oferecendo aos alunos apenas o domínio de competências técnicas, sem vínculo com a realidade à qual pertencem. Por outro lado, os defensores da construção do sentido defendem que o significado é o que importa no aprendizado da língua, pois ele permite a compreensão contextualizada do que está sendo dito, lido e escrito, colocando em cena uma dimensão política negligenciada por seus antagonistas. De um lado, a decodificação, tratada como operação mecânica e desprovida de sentido. De outro, a compreensão contextualizada e consciente, revelando o que é, de fato, o processo de alfabetização (no Brasil, o termo letramento<sup>37</sup> é o utilizado para definir os usos sociais da alfabetização, a leitura com significado e contextualizada).

Gontijo (2014) vê como um problema o avanço do enfoque fonético em alfabetização no Brasil, em particular, a partir de 2006, com a criação, pela Câmara dos Deputados, de um Grupo de Trabalho destinado a elaborar um relatório sobre a alfabetização infantil (idem). O grupo, tratado como antidemocrático pela autora, por não ter sido constituído com base na divergência de ideias sobre um tema plural e não consensual, como é o caso do conceito de alfabetização,

---

<sup>35</sup> Além do conjunto de competências e do enfoque libertador, outras concepções de alfabetização utilizadas pela Unesco ao longo dos anos foram: a alfabetização como processo de aprendizagem, de inspiração piagetiana (para ver sobre sua influência nas políticas nacionais, ver Ferreiro e Teberosky, 2007); e a alfabetização como texto, observando, precipuamente, seu conteúdo, a partir da influência de Foucault.

<sup>36</sup> Defendendo Ferreiro e Teberosky (2007), Gontijo acredita que a ênfase no significante, promovida pela ênfase na decodificação, leva à destruição do signo (2014, p.75).

<sup>37</sup> Para mais sobre os conceitos de alfabetização como o aprendizado da escrita como código e de letramento como a compreensão da leitura e da escrita como práticas sociais, ver Soares (2007).

coloca em destaque a decodificação como a competência central do processo de alfabetização, reavivando seu conceito como conjunto de competências autônomas e, por sua vez, a ênfase no ensino das unidades da língua (fonético) como o caminho mais adequado para esse processo. Para Gontijo (2014), essa abordagem causa um empobrecimento do conceito de alfabetização, reduzindo o ato de ler ao de decodificar e o de escrever ao de codificar (idem, p. 31).

É interessante notar, neste ponto, uma diferença de interpretação em relação às influências das políticas educacionais internacionais sobre o contexto local. Para Gontijo (2014), no caso da alfabetização, as políticas brasileiras negligenciam o conceito de alfabetização adotado em escala global e propagado por organismos internacionais como a Unesco. Conforme vimos no tópico anterior, a influência internacional sobre políticas nacionais é vista, em grande medida, como um problema, uma imposição do capitalismo global. Aqui, Gontijo inverte a perspectiva: as políticas de alfabetização no Brasil operam com um conceito reduzido de alfabetização (como conjunto de habilidades autônomas), quando, na verdade, deveria observar o conceito proposto por organismos internacionais<sup>38</sup>. Estes são entendidos, nesse caso, como uma referência positiva, visto que defendem o conceito também defendido pela autora, e não como instrumentos a serviço do mercado mundial. O que o GT sobre Alfabetização Infantil fez foi ignorar o modelo de alfabetização proposto pelos organismos internacionais, algo que é tratado como um problema para Gontijo (2014). A influência das políticas internacionais sobre as nacionais é defendida quando convém às perspectivas dos autores que a estão analisando.

O referido GT, ao alçar a decodificação como o elemento central para o processo de alfabetização, contrapôs-se ao construtivismo de Ferreiro e Teberosky (2007). Emilia Ferreiro foi particularmente influente nas políticas de alfabetização no Brasil na década de 1990, com uma postura claramente ancorada no construtivismo de Piaget (ver Ferreiro, 1993; 2011). Todas as políticas de alfabetização desenvolvidas no Brasil desde então são marcadas por um conceito restrito de alfabetização, alavancado pelo consenso (entre construtivistas e defensores da importância do ensino dos elementos fonéticos, presene na ideia do alfabetizar letrando; na leitura de Gontijo, um falso consenso, 2014, p. 76) em torno da ideia de letramento. Parte do sucesso desse consenso é tributário à influência das ideias de Magda Soares. Defendendo que o aspecto fonético é fundamental para o processo de alfabetização (seu abandono leva a uma espécie de

---

<sup>38</sup> “É estranho que o seminário não tenha abordado as principais contribuições de Paulo Freire, da perspectiva histórico-cultural no campo da psicologia, e da abordagem discursiva para a alfabetização, quando se sabe que até os organismos e as comunidades de pesquisadores internacionais têm reconhecido essas contribuições, principalmente as de Paulo Freire.” (GONTIJO, 2014, p.67). Gontijo refere-se, aqui, ao seminário organizado pela Secretaria de Educação Básica – SEB, do MEC, com o título Alfabetização e Letramento em Debate.

“desinvenção” da alfabetização, SOARES, 2004), e distinguindo-o (mas tratando-o como complementar) do letramento, entendido como os usos sociais do aprender a ler e escrever, Soares procura superar os problemas do construtivismo e da discussão sobre o melhor método para alfabetizar com a proposta de que o pior é alfabetizar sem método (SOARES, 2016).

O trabalho do GT e as políticas de alfabetização que ali encontraram sua fonte, de acordo com a crítica de Gontijo (2014), além de ignorarem as orientações da Unesco, olvidaram, também, a produção nacional sobre o assunto, baseando-se em resultados de avaliação em larga escala e em pesquisas internacionais pontuais, elogiosas do enfoque no aspecto fonêmico (aqui, as referências internacionais voltam a ser um problema). No fundo, o GT e sua posição enviesada sobre a importância do elemento fonético para o processo de alfabetização foram os grandes responsáveis pelas políticas de alfabetização no país desde 2006, todas elas perpassadas pelo destaque dado à decodificação: formação de professores, produção de material didático, a ampliação do ensino fundamental para nove anos e a política de avaliação da alfabetização.

Em relação à formação de professores, o principal programa do governo federal nos anos 2000 foi o Pró-Letramento<sup>39</sup>. Criado em 2005 e entendendo a formação inicial e a continuada como processos complementares, o programa é visto por Gontijo (2014) como aligeirado, baseado em um sistema de monitoria ineficaz. A participação por adesão só reforça seu caráter não-essencial (idem, p.70). Os materiais instrucionais são uniformes, de rápida difusão e distribuídos em todo o país com o selo do discurso oficial (do MEC) sobre alfabetização. E sua base conceitual é exatamente aquela criticada por Gontijo como limitada, tecnicista e reducionista. A decodificação é o cerne do curso de formação, fio condutor do material de formação, destinado a capacitar os professores a alfabetizar letrando. O nome do programa já informa sobre sua posição em relação ao conceito de alfabetização que adota. Por esse motivo, Gontijo vê o programa como um mecanismo ideológico, a serviço do poder oficial, para difundir as ideias de alfabetização e letramento, na esteira de Soares (2004), aos professores. O diagnóstico para a construção do curso, novamente, recai sobre os resultados da avaliação em larga escala: as escolas são incapazes de alfabetizar e a formação dos professores incide diretamente no desempenho dos estudantes.

Em relação aos livros didáticos, a análise do PNLD (Programa Nacional do Livro Didático)<sup>40</sup> segue na mesma direção. A partir de 2007, sob a influência do Ceale/UFMG, a avaliação e a seleção dos livros didáticos passaram a usar como critério fundamental a

---

<sup>39</sup> Para uma análise mais detida do programa, ver Salomão, 2014.

<sup>40</sup> Para um estudo sobre como os professores alfabetizadores usam os livros do programa em sala de aula, ver Albuquerque e Ferreira, 2019.

perspectiva do letramento, ou seja, a que traz consigo a valorização do ensino das habilidades relacionadas à decodificação<sup>41</sup> (GONTIJO, 2014, p.83). Somente materiais que seguem esse entendimento passam a ser aprovados. Os livros sofrem grande influência dos exames padronizados (e da ideia de se ensinar para o exame, de acordo com o que já foi mencionado anteriormente), substituindo a noção de aprendizagem pela noção de aquisição. A influência dos exames é vista também no repertório limitado de textos, que privilegiam aliterações e repetições de palavras. Isso provoca o “retorno de métodos e materiais de ensino que já demonstraram a sua insuficiência na prática” (GONTIJO, 2014, p.89).

O ensino fundamental de nove anos<sup>42</sup> também é uma política influenciada pela perspectiva do letramento. O ensino baseado na decodificação, desde o início do primeiro ano do ensino fundamental, não faria a transição adequada entre o universo da educação infantil, orientado, precipuamente, para o brincar, e o ensino fundamental, nem respeitaria a idade da criança, com 6 anos, e suas características psicológicas. A ampliação é bem vista por Gontijo, mas não o tempo e a organização curricular (2014, p.93). Os guias orientadores, produzidos pelo MEC, para a inserção de crianças de seis anos no ensino fundamental trazem o letramento e a ideia de consciência fonológica como conceitos basilares e tratam a alfabetização, meramente, como tecnologia da escrita, seguindo a concepção de Soares (2004).

Em 2007, foi criada a Provinha Brasil, aplicada pela primeira vez em 2008, com o intuito de avaliar crianças de 6 a 8 anos de idade. Para Saviani (2007), a Provinha Brasil foi o instrumento pelo qual o governo buscou monitorar, através do PDE, a meta 2 do movimento Todos pela Educação, que estabelecia a alfabetização de todas as crianças até os 8 anos de idade. Os documentos orientadores dessa avaliação, produzidos pelo MEC, definem o diagnóstico das possíveis insuficiências das habilidades de leitura e escrita como seu maior objetivo, dando subsídios para o aprimoramento de práticas de alfabetização (característica primordial de uma política de avaliação, como veremos). A Provinha Brasil foi pensada para ser aplicada duas vezes ao ano, no início e ao fim do ano letivo, de modo que alterações no desempenho dos alunos pudessem ser observadas.

O fato de ser aplicada e corrigida pelo próprio professor da turma, somado ao uso da TCT como metodologia para a produção dos resultados, dá à Provinha Brasil características distintas daquelas dos sistemas próprios de avaliação em alfabetização, como os que aqui serão

---

<sup>41</sup> Desde o uso adequado dos instrumentos de escrita e da escrita e leitura na direção correta da página até o conhecimento das relações entre fonemas e grafemas.

<sup>42</sup> Para uma discussão sobre a ampliação do ensino fundamental para nove anos e a defesa do direito à educação, ver ARELARO; JACOMINI; KLEIN, 2011.

analisados, e da Ana (Avaliação Nacional da Alfabetização), da qual falaremos em seguida. O tipo de aplicação (e quem é o aplicador) é um dos elementos que influenciam os resultados. Ainda faltam pesquisas mais aprofundadas sobre isso, mas uma hipótese que pode ser levantada é a de que o professor da turma pode ajudar, mesmo que de maneira inconsciente, os alunos a resolverem as questões propostas. Mesmo uma entonação diferente ao ler as alternativas ou o comando do item pode provocar tal influência. Isso sem falar na ajuda deliberada, tendo o professor tomado essa posição em virtude de temer a forma como os resultados poderão ser utilizados pelos tomadores de decisão da política de avaliação.

Além disso, o uso da TCT não permite a comparação dos resultados. Os resultados de um ano para o outro, por exemplo, não podem ser usados para estabelecer análises no sentido de dizer se houve ou não melhoria dos resultados ao longo do tempo. As diferenças entre os testes, mesmo dentro do ano letivo, desautorizam qualquer comparação mais acurada dos resultados. Isso se deve à ausência de itens comuns entre os testes, com capacidade de discriminação e nível de dificuldade calculados, algo propiciado pela TRI.

A Provinha Brasil não foi pensada para ser uma avaliação de caráter classificatório, mas, antes, para a produção de um diagnóstico que servisse diretamente aos professores em sala de aula. Isso não significa que ela não tenha tido importância, afinal: i) ela inaugura, em nível nacional, uma avaliação da alfabetização, algo inexistente até então; ii) estabelece a avaliação como um processo de aprendizagem e capacitação para os próprios professores aplicadores, algo que não é trivial; e iii) sua criação ajudou, em maior ou menor medida, dependendo do contexto, a mitigar a resistência dos atores escolares com a ideia de uma avaliação em larga escala (aspecto da avaliação que será tratado no tópico seguinte). Pensá-la como processo de aprendizagem para os professores coloca em cena o reconhecimento de que avaliar os alunos no ciclo de alfabetização pode ser um processo bastante complexo, nem sempre bem compreendido pelos professores alfabetizadores, ainda mais se levarmos a sério as críticas destinadas à formação desses profissionais, principalmente, no que diz respeito ao tema da avaliação. Ainda, a Provinha Brasil abriu espaço para a criação de um programa nacional de avaliação da alfabetização em outros moldes, a Ana.

Seguindo sua análise da influência do conceito restrito de alfabetização, proposto pelo GT da Câmara dos Deputados, sobre as políticas nacionais, Gontijo (2014) vê a Provinha Brasil como mais um exemplo de programa que coloca em cena o empobrecimento da ideia de alfabetização. De acordo com ela (idem, p.112), a despeito do caráter não-classificatório da Provinha, um sistema de avaliação é sempre um instrumento de controle e produz efeitos

“singulares nos sujeitos envolvidos” (idem). Gontijo refere-se aqui ao treinamento para o teste, em sua análise, uma consequência necessária (e nefasta) dos sistemas de avaliação. Além disso, ao analisar os eixos que compõem a Matriz de Referência da avaliação, ela chama atenção para a ênfase dada aos descritores relacionados à apropriação do sistema de escrita (reconhecer letras, reconhecer sílabas, estabelecer relações entre unidades sonoras e suas representações gráficas). Este último, em particular, é a própria encarnação da defesa de que os elementos fonéticos são importantes para o processo de alfabetização.

O eixo relacionado à leitura, além de receber menor atenção na Provinha, propõe uma gradação das unidades: palavras, frases e pequenos textos. Para Gontijo (2014, p.116), as palavras são usadas como contextos para a identificação das unidades linguísticas (letras e sílabas) e o mesmo pode ser dito para as frases e pequenos textos, que não se prestam à construção de sentidos, mas à aquisição de habilidades relacionadas à decodificação. O eixo relacionado à escrita, fundamental para o processo de alfabetização, teve de ser retirado da matriz, conforme proposta inicial.

A análise de Gontijo (2014) sobre a Provinha Brasil chega à conclusão de que a leitura não é tratada como um processo de produção de sentidos, mas, antes, como mera “identificação, reconhecimento do que está expresso na imagem” (idem, p.123), encarnando o conceito de alfabetização como conjunto de competências autônomas. A Matriz de Referência nada mais é do que a reunião dessas competências. O que sobra é o texto como objeto para ser decifrado, como reconhecimento do código escrito.

Silva (2013), em análise da Provinha Brasil, afirma que a maior parte dos trabalhos sobre o programa possui um olhar crítico, no sentido de desqualificar o exame, em especial, por acreditar que a língua é tratada como código (dando passagem a uma visão estruturalista da língua) e por entender que a Provinha apresenta uma visão segmentada de alfabetização, servindo, além disso, para controlar a infância. Esse é o caso de Gontijo, que encarna, em sua visão, exatamente essa perspectiva, e de uma série de outros autores. Goulart (2014) critica a perspectiva do alfabetizar letrando, que transforma as crianças em “cabeças que aprendem aspectos do sistema alfabético” (idem, p. 330), e acusa a avaliação da alfabetização de inserir na escola uma lógica de mercado, deixando de fora aspectos importantes como as emoções e a criatividade. Chama a atenção, ainda, para a importação de modelos internacionais de avaliação, dos quais ela discorda. Mortatti (2013) é outra autora a vincular os sistemas de avaliação a um conceito restrito e rudimentar de alfabetização, cuja base é a lógica do neoliberalismo. Na mesma linha, Esteban (2012) vê a Provinha Brasil como uma redução do elemento educacional

ao campo meramente técnico e pedagógico, desconsiderando aspectos culturais e sociais que fazem parte da vida escolar.

Mas há quem defenda uma posição diferente em relação à avaliação da alfabetização. Morais (2014, p.282), por exemplo, defende a necessidade de políticas de avaliação no ciclo de alfabetização para ajudar a “democratizar, de fato, a vivência de um ensino de alfabetização de qualidade e a indicar fatores que, e, nosso entender, têm dificultado que avancemos no cumprimento dessa republicana tarefa”<sup>43</sup>. Em outro trabalho, o autor defende a avaliação da alfabetização como forma de monitorar as aprendizagens dos alunos de modo a orientar os processos de ensino e de formação continuada dos professores (MORAIS; LEAL; ALBUQUERQUE, 2009) e, em um terceiro, apresenta evidências de usos dos resultados da Provinha Brasil para a reformulação de práticas pedagógicas (MORAIS, 2012). Ao observar as defesas e as críticas à Provinha, é possível perceber a polarização do debate, entre aqueles que defendem a aprendizagem fonética como elemento importante para o processo de alfabetização e aqueles que a criticam. Morais (2014), seguindo Soares (2004), é um defensor da perspectiva de que o aprendizado da leitura e da escrita não ocorre somente através da participação espontânea em atividades dessa natureza.

Em 2012, foi instituído o Pacto Nacional pela Alfabetização na Idade Certa – Pnaic, um compromisso assumido pelo governo federal em parceria com os demais entes federados, com o intuito de assegurar a alfabetização das crianças até o final do ciclo (3º ano do ensino fundamental), em atendimento à meta 5 do PNE. O Pnaic é um política abrangente, envolvendo diversa gama de ações, embora poucos trabalhos tenham se dedicado ao seu estudo como política pública (ALFERES; MAINARDES, 2019). Trata-se de um esforço que envolve formação continuada de professores, com produção, entrega e distribuição de material didático, e avaliação, no caso, a Ana. A Avaliação Nacional da Alfabetização foi criada em 2013, com aplicações em 2015 e 2016. Depois disso, foi descontinuada até o ano de 2020. A Ana, ao contrário da Provinha Brasil, utiliza a TRI como metodologia, permitindo a produção de medidas de proficiência, a criação de escalas e a comparação entre os resultados das escolas e ao longo dos anos.

Avaliando os alunos do 3º ano do ensino fundamental, a Ana tem como objetivo produzir um diagnóstico de alfabetização e letramento em Língua Portuguesa e de alfabetização em Matemática. Conforme seu documento básico (BRASIL, 2013), o marco teórico da Ana envolve a perspectiva da alfabetização e do letramento, tratados como processos distintos, mas

---

<sup>43</sup> A tarefa de alfabetizar.



complementares e importantes (idem, p. 9), em sintonia com Soares (2014). Gontijo (2014) não analisou a criação da Ana em seu livro, mas, seguindo os argumentos que apresentou para as demais políticas de alfabetização, podemos afirmar que entenderia a Ana como uma política que coloca em cena um conceito restrito de alfabetização. As críticas que ela direcionou à Provinha Brasil, contudo, não podem ser aplicadas, todas e diretamente, à Ana. Esta última avalia a escrita (3 itens no teste). Algumas habilidades previstas na Matriz de Referência de Língua Portuguesa, como a leitura de palavras com estrutura silábica canônica e a leitura daquelas com sílabas não-canônicas, permanecem no registro daquilo que Gontijo (2014) tratou como a redução do conceito de alfabetização à decodificação. Além disso, no que diz respeito à escrita, duas das três habilidades previstas pela matriz dizem respeito a conhecimentos técnicos sobre a língua, envolvendo o valor sonoro das letras. Zanchetta Junior (2017), por sua vez, critica a abordagem do Pnaic em virtude da proposta de um ensino de literatura escolarizado, distante da estética que os textos podem propiciar.

A Ana integra o Saeb, completando o sistema nacional de avaliação, que não contava com instrumentos destinados ao ciclo de alfabetização. Vincula-se também ao Pnaic, sendo um dos eixos do programa o trabalho com os resultados da avaliação para a produção de diagnósticos e o suporte para a tomada de decisões<sup>44</sup>. Embora não possua o mesmo tipo de consequência que a Prova Brasil, por exemplo, visto que esta última integra o Ideb, a Ana, de acordo com o entendimento de Dickel (2016), possui consequências que vão além de um mero diagnóstico da qualidade da alfabetização:

(...) a ANA traz um elemento novo que a distingue das demais modalidades de avaliação integradas ao SAEB e corrobora o caráter indutor de currículo assumido nesse cenário: ela vem acompanhada de um projeto de formação de professores e, portanto, de um projeto curricular explícito (DICKEL, 2016, p.198).

Nem a Ana nem a Provinha Brasil, seja por quais motivos forem, não foram capazes, ainda, de se estabelecer como políticas nacionais de alfabetização. As rupturas no campo, em especial no que diz respeito àquilo que Gontijo (2014) chamou de falso consenso em torno da perspectiva do letramento, podem ter prejudicado essa consolidação. Soma-se a isso a disputa

---

<sup>44</sup> Em relação a isso, possuo uma experiência pessoal. Fui coordenador de avaliação do Pnaic – polo UFJF, entre os anos de 2015 e 2017. O trabalho, ao longo das capacitações, com o tema da avaliação foi caracterizado tanto pela resistência inicial de boa parte dos participantes, quanto pela surpresa dos mesmos em relação ao tema e a suas possibilidades. De maneira geral, é possível afirmar que eles não conheciam em detalhe a Ana, ou qualquer avaliação em larga escala, mas já possuíam uma certa resistência, da qual eles mesmos não sabiam explicar bem o motivo, em relação à ideia de uma avaliação externa.

sobre a idade certa (termo que dá nome ao Pnaic) para que uma criança esteja alfabetizada: ao fim do ciclo de alfabetização, conforme previsto pela Ana, ou ao fim do segundo ano do ensino fundamental, conforme previsto pela BNCC. Seja como for, a demora para o estabelecimento de uma política nacional de alfabetização (quando comparada à avaliação de outras etapas) e essas divergências, serviram como combustível para a criação e desenvolvimento de sistemas próprios de avaliação da alfabetização.

\*\*\*

O fato da presença do elemento fonético nas políticas de alfabetização ter tido como uma de suas bases os resultados das avaliações em larga escala pode ter contribuído para a desconfiança e para a resistência em relação à avaliação no âmbito da alfabetização. Uma das críticas do GT às políticas de alfabetização no Brasil é que elas desconsideravam as evidências fornecidas pelas avaliações em larga escala. Por extensão, uma recusa da autoridade do GT, como vemos em Gontijo (2014), pode significar uma recusa da própria ideia de avaliação. Soma-se a isso o vínculo da avaliação com o conceito de competências, visto como extremamente técnico e abstrato, muitas vezes, como a antítese do que é político. A ideia de competência vincula-se, na crítica, ao conceito de alfabetização como conjunto de competências autônomas, como vimos, conceito que é tratado como restrito e empobrecido (como em GONTIJO, 2014; ESTEBAN, 2012; MORTATTI, 2013; GOULART, 2014). A avaliação é contaminada pela crítica feita à forma como as políticas de alfabetização foram pensadas no Brasil na última década e pelo conceito de alfabetização que adotaram.

A avaliação em larga escala é vista como o espelho do conceito inadequado, restrito e meramente técnico-linguístico de alfabetização. Os testes padronizados, nessa visão crítica, nada mais fazem do que avaliar as limitadas habilidades de decodificação e codificação, assumindo o conceito de alfabetização como o domínio de um conjunto de técnicas autônomas e descontextualizadas, sem que os usos sociais do aprender a ler e a escrever, objetivo legítimo do processo de alfabetização, sejam avaliados. A avaliação, assim, reduz a alfabetização à sua dimensão linguística, excluindo a social:

Tendo em vista a criação da avaliação da alfabetização, é necessário construir objetivos de aprendizagem que possam ser medidos por meio de testes. Sendo assim, a necessidade de mensuração é um dos fatores que produz, no atual contexto, o empobrecimento da alfabetização e de suas dimensões, ou seja, o empobrecimento dos conhecimentos ligados à leitura e à escrita (GONTIJO, 2014, p.89).

A crítica já direcionada à avaliação em larga escala, conforme veremos no tópico seguinte, ganha especificidade na alfabetização: a avaliação padronizada tem como uma de suas consequências a redução do currículo, visto que força professores a adotarem as matrizes dos testes como currículo, ensinando os alunos apenas para o teste, deixando de fora, de modo sistemático e proposital, outras tantas habilidades importantes para o desenvolvimento dos alunos; na alfabetização, isso significa ensinar, como precedência e destaque, as habilidades relacionadas aos processos de decodificação e codificação, as mais cobradas nos testes.

Para a formação de professores, a avaliação aparece como diagnóstico de sua importância para a aprendizagem dos alunos e, ao mesmo tempo, como denúncia da incapacidade da escola, pois os professores são mal formados. Esse é uma forma de distorcer o aspecto heurístico da avaliação que, funciona, entre outras coisas, como denúncia de desigualdades no âmbito escolar. Aqui, a avaliação aparece como uma informação que deturpa a realidade, levando a conclusões descontextualizadas. Diante desse cenário, não é difícil compreender a resistência em relação a um sistema nacional de avaliação da alfabetização.

Uma palavra antes de encerrar o tópico. O objetivo desta tese não é discutir a oposição entre construtivistas e defensores da aprendizagem fonética como elemento importante para o processo de alfabetização, de modo a apontar qual dos dois caminhos é melhor para a alfabetizar. A perspectiva do alfabetizar letrando inclui o ensino dos elementos fonéticos da língua como parte do processo de alfabetização. Merece atenção a forma como os críticos dessa perspectiva constroem seus argumentos. O aprendizado dos aspectos fonéticos da língua é tratado como um empecilho para o domínio social da leitura e da escrita, de modo que as duas posições só podem ser percebidas como contraditórias. Contudo, não há evidências de que essa relação seja antagônica. Aprender sobre a relação entre fonemas e grafemas não leva, necessariamente, a uma incompreensão do uso social do ler e do escrever, ou seja, não resulta em alienação política. Uma discussão técnica e pedagógica (epistemológica, sobre os caminhos para a alfabetizar alguém) é transformada em uma discussão moral. As analogias são imediatas: se você defende o aprendizado dos fonemas, logo, está do lado da lógica dominante e excludente do capitalismo, é um serviçal de ideias estrangeiras e é a favor da alienação dos alfabetizandos, reduzidos, aqui, a meros reprodutores a serviço do sistema social que os condena. De outro lado, se você critica o aprendizado do aspecto fonético, passa a ocupar a posição de crítico do sistema capitalista, defensor da alfabetização como liberdade e adepto das ideias de Paulo Freire. São muitos os problemas de uma sequência de correlações como essa, a começar pela completa falta de evidência que liga todos esses pontos. A crítica ao capitalismo

e a seus efeitos não é exclusiva dos críticos do ensino da relação entre fonemas e grafemas, assim como acreditar que o aprendizado de aspectos técnicos da língua não é contraditório à busca por uma escola que seja capaz de formar pessoas críticas e reflexivas. Aprender a língua em uma dimensão mais técnica não retira a criticidade de ninguém. Até que surjam pesquisas que evidenciem essa relação de modo inequívoco, ela é apenas uma conjectura sem lastro empírico, revelando uma posição moral.

### 3.3 A POLÍTICA DE AVALIAÇÃO EM LARGA ESCALA

Neste tópico, a política de avaliação em larga escala é discutida a partir de quatro dimensões: i) como instrumento de garantia de direitos educacionais; ii) a partir das críticas que recebe e de suas virtudes como política educacional; iii) no que diz respeito ao sistema nacional e à criação de sistemas estaduais e municipais; e iv) no que tange à importância dos resultados para a política.

#### 3.3.1 A avaliação educacional e a garantia de direitos

A afirmação de que a avaliação é um instrumento de garantia do direito à educação (SOARES, 2012) pode levar, equivocadamente, à ideia de que, se a serviço da efetivação de um direito, ela seria aceita consensualmente como um elemento importante para o cenário educacional. Como veremos a seguir, a avaliação educacional em larga escala está longe de ser um consenso e uma série de críticas tem sido direcionada a ela desde sua implementação no Brasil, no início dos anos 1990.

A despeito disso, para aqueles que acreditam que a avaliação pode ter usos positivos no âmbito educacional, seu vínculo com a defesa e a garantia do direito à educação é uma de suas principais justificativas. Se a previsão jurídica de garantia ao direito não é suficiente para sua efetivação, ela é, na esteira de Cury (2002), um passo fundamental para tanto. Conforme ressalta o autor (idem), a desconfiança em relação à lei (e ao direito de maneira geral, em especial, de acordo com a tradição marxista) foi sendo superada paulatinamente entre os educadores, de modo que o reconhecimento de um direito passou a ser postulado como um aspecto importante na luta pela igualdade (um dos princípios estruturadores do direito à educação):

Hoje cresceu, enfim, a importância reconhecida da lei entre os educadores, porque, como cidadãos, eles se deram conta de que, apesar de tudo, ela é um instrumento viável de luta porque com ela podem-se criar condições mais

propícias não só para a democratização da educação, mas também para a socialização de gerações mais iguais e menos injustas. (CURY, 2002, p.247)

Mas o que significa, de fato, defender a avaliação como um instrumento a serviço da garantia do direito à educação? Há duas formas, não excludentes, de compreender essa afirmação. De um lado, é preciso especificar de que direito à educação falamos, afinal, *não existe um direito à educação*, mas, antes, um conjunto de direitos educacionais. Seguindo lição de Boto (2005)<sup>45</sup>, é possível reconhecer diferentes gerações de direitos educacionais, que, ao longo do tempo, foram sendo garantidos, à maneira como ocorreu com as gerações de direitos civis, políticos e sociais, conforme argumento de Thomas Marshall (1967). Essa garantia, contudo, nunca é imediata e livre de disputas e conflitos. Conforme ressalta Bobbio (1992), os direitos nascem e experimentam desenvolvimentos não por nossa pura vontade, mas, antes, pela conjuntura de forças históricas da sociedade. Boto (2005) propõe, dessa forma, três gerações de direitos educacionais: i) o acesso à educação; ii) a qualidade do ensino ofertado; e iii) a garantia da diversidade<sup>46</sup>.

O acesso à educação como direito é encarnado em sua transformação em direito público, a ser garantido pelo Estado. Cury (2002) descreve como esse processo ocorreu a partir do modelo liberal burguês e da necessidade de criar uma sociedade letrada para fins de cidadania, democracia e participação no mercado. O acesso à educação, segundo Boto (2005, p. 779) só é efetivado quando a escola pública é universalmente garantida, fenômeno que, entre nós, ocorreu formalmente apenas no início dos anos 1990 (para o ensino fundamental). Embora seja parte dessa primeira geração de direitos, o acesso à educação tende a ser confundido como o próprio direito à educação. Há, pelo menos, dois problemas advindos dessa confusão. Primeiramente, ela oblitera a existência de outros direitos educacionais, reduzindo direitos a um único direito. Em segundo lugar, como consequência disso, o acesso à educação é tratado como suficiente para o reconhecimento de que um direito foi garantido. Assim, ter acesso à educação torna-se sinônimo de garantia do direito à educação. Neste ponto, surge uma bifurcação: o acesso à educação pode ser entendido de modo formal, a partir da oferta suficiente de matrículas tendo em vista a população em idade escolar, e de modo substancial, com a garantia de requisitos mínimo para que, de fato, essa população tenha acesso à escola (oferta e garantia de transporte escolar, merenda etc). De um modo ou de outro, a confusão entre acesso e direito à educação

---

<sup>45</sup> Passa despercebido, pela própria autora, a ideia de direitos, no plural, visto que fala em educação escolar como direito humano, no singular.

<sup>46</sup> Essa nomenclatura é dada pela minha interpretação do texto de Boto, que usa descrições para definir as gerações, mas não uma nomenclatura específica.

tem efeitos restritivos, não dando conta de abarcar uma ideia plural de direitos à educação. Apenas para fins de registro, seria possível argumentar que o próprio acesso é constituído por um conjunto de direitos (como, por exemplo, defendendo a ideia de existência de um direito ao transporte escolar), não se apresentando como um direito único.

A segunda geração de direitos, ainda seguindo a proposta de Boto (2005), caracteriza-se por um salto qualitativo em relação ao direito ao acesso, geração precedente. Trata-se da garantia de padrões mínimos de qualidade do ensino ofertado, estabelecidos através de uma proposta curricular que avança além dos conhecimentos puramente acadêmicos (Boto fala também do reconhecimento das ideias democráticas internas à vida escolar, 2005, p.779). O direito à qualidade à educação coloca em cena um elemento qualitativo necessário para a efetivação da ideia de um direito à educação, algo que não pode ser alcançado pelo simples acesso. Os objetivos educacionais não são cumpridos quando crianças e jovens ingressam na escola. Sem a garantia de uma qualidade mínima, o ideal de um direito à educação é vazio. Como veremos adiante, definir o que é qualidade da educação é uma tarefa nada trivial e traz uma série de consequências para o tema da avaliação educacional.

É necessário ressaltar aqui a tensão existente entre as duas gerações, entre o acesso e a qualidade. De acordo com Boto (2005, p.779), esses dois direitos são colocados como opostos, como ideias antagônicas. Assim, a ampliação do acesso viria necessariamente acompanhada de uma queda nos padrões de qualidade:

(...) geralmente, quando se irradia a instrução pública – com o fito de progressivamente universalizar seu acesso –, serão incorporadas populações anteriormente excluídas do acesso à escola. Consequentemente, há quem denuncie que isso ocasiona um declínio de supostos padrões de qualidade, os quais assegurariam – por definição – alguma excelência do modelo anterior, aquele que contemplava apenas uma parcela diminuta dos que, então, passariam a compor a população escolar.

Contudo, a autora deixa claro a seguinte prerrogativa: não há motivos para se falar em qualidade da educação se o acesso não é universal, se toda a população não faz jus a ele. Diante disso, se é verdade que o acesso sem qualidade não configura a garantia plena dos direitos à educação, não é menos verdade que qualidade sem acesso também não prefigura tal garantia. A qualidade precisa ser buscada, sob pena de um vazio, mas pressupõe, de modo inarredável, o acesso universal.

A terceira geração de direitos educacionais diz respeito à inclusão, à garantia da educação para além do acesso e da qualidade, reconhecendo que determinados segmentos

sociais possuem maior dificuldade em ver garantida sua participação nesse conjunto de direitos. Essa geração é característica de uma sociedade cada vez mais plural em termos de valores, que vê na construção de identidades próprias, singulares, o cerne da reconstrução do contrato social (STOER; MAGALHÃES, 2003). Da mesma maneira, Cury (2002) reconhece a importância do direito à diferença como princípio estruturador, ao lado do direito à igualdade, do direito à educação. Em minha leitura, essa terceira geração de direitos, que se preocupa com a diferença e com a inclusão, está diretamente relacionada a uma dimensão fundamental da educação, qual seja, a permanência. Problemas históricos enfrentados pelo Brasil, como as altas taxas de reprovação (RIBEIRO, 1991) e os fenômenos de abandono e evasão, estão relacionados com questões ligadas à não observância e proteção da diferença e da inclusão. Dubet (2004), ao analisar o sistema educativo francês, chama atenção para as formas com as quais os “vencidos” de uma escola meritocrática reagem a seu fracasso. A principal delas é desistir da escola. Assim, um tratamento a respeito da diferença (entendida aqui em sentido amplo, sexual, racial, étnica, religiosa, estética etc) passa a figurar como um direito educacional necessário para que o ciclo iniciado com o acesso possa resultar em uma garantia efetiva de direitos.

Mas como a avaliação educacional relaciona-se com essas gerações de direitos educacionais? Uma resposta mais imediata seria identificá-la com a segunda geração, com a qualidade da educação. De fato, a avaliação educacional encontra na qualidade da educação seu vínculo mais óbvio. Nessa esteira, ela seria um instrumento a serviço da qualidade, na medida em que os diagnósticos que produz versam sobre a qualidade do ensino ofertado. Veremos mais sobre esse aspecto a seguir. Por ora, vale acentuar que, além da dimensão da qualidade, é possível perceber o vínculo da avaliação com as demais gerações de direitos apresentadas anteriormente.

Em relação ao acesso à educação, sua relação com a avaliação em larga escala, de modo mais imediato e claro, diz respeito ao monitoramento periódico feito pelos instrumentos avaliativos em relação aos alunos matriculados nas redes públicas. Em especial nos sistemas próprios de avaliação, no âmbito dos quais esse controle é feito anualmente, os alunos são avaliados de modo censitário (em regra). Se a perspectiva apresentada anteriormente, a respeito do acesso à educação como caracterizado por um conjunto de direitos, fica de pé, a avaliação é capaz de mostrar a relação entre a garantia de um acesso substancial (observância das condições mínimas, que extrapolam apenas a matrícula) e a aprendizagem dos alunos.

No que diz respeito à defesa da diversidade e à garantia da diferença como princípio, a avaliação é um instrumento que, a partir de sua dimensão heurística, é capaz de identificar os

principais fatores que se associam ao desempenho, denunciando a influência de categorias como classe social, raça e gênero, de um lado, e, de outro, reconhecendo os efeitos de enturmação, distribuição dos alunos por turno, características dos professores, dos gestores e do próprio processo de ensino sobre a aprendizagem dos estudantes. Essa é uma linha de investigação que se relaciona diretamente com a ideia de diferença e de diversidade, já que permite identificar quais características dos alunos são mais ou menos potencializadas pelas escolas como fundamentos para uma desigualdade de aprendizagem. Assim, a avaliação pode, e tem sido, utilizada como um instrumento a serviço da identificação e da denúncia de desigualdades e injustiças escolares, alinhada, portanto, com a garantia dos direitos educacionais de terceira geração.

De maneira geral, apresentados esses argumentos, não seria despropositado afirmar que a avaliação educacional pode ser utilizada como instrumento a serviço da garantia dos direitos educacionais das três gerações e não apenas daqueles relacionados mais diretamente à qualidade. O modal da possibilidade é aqui utilizado pela dependência da observância da forma como a avaliação é utilizada. Esse tema será retomado adiante.

Antes de passarmos ao próximo ponto, é necessário fazer uma observação. Mais do que garantia de direitos, é possível argumentar que a avaliação, ela própria, é um direito do aluno, no plano individual, e da escola, no plano da estrutura. Essa afirmação, para ter poder de persuasão, requer, evidentemente, a construção de uma argumentação que seja capaz de superar a imagem histórica que nós mesmos temos do que significa avaliar.

Em uma breve organização dos argumentos, poderíamos dizer que a avaliação foi, e ainda é, usada, em muitos casos, como mecanismo punitivo. Em uma leitura que deve grande parte de sua força a Foucault (em particular, em *Vigiar e Punir*, 1987), a avaliação seria um dos muitos instrumentos que a escola (assim como outras instituições disciplinares) usa como técnicas de punição, para fins de domesticação e adestramento. A preocupação com a avaliação punitiva extrapola os círculos mais próximos à filosofia da educação, sendo tema de enfrentamento para aqueles mais preocupados com sua utilização para fins mercadológicos, condizentes com o regime capitalista (por exemplo, Luckesi, 2010, e Libâneo, 1994). De acordo com Luckesi, na sociedade capitalista atual, a avaliação tem sido necessariamente utilizada de modo autoritário, pois seu objetivo primordial é reproduzir e conservar a sociedade da maneira como está organizada (LUCKESI, 2010). Libâneo, por sua vez, critica o uso da avaliação a partir do registro meritocrático: para uns, prêmio, para outros, instrumento de intimidação e avaliação (LIBÂNEO, 1994).



Uma pesquisa mais densa, ainda por ser feita, poderia enfrentar o tema da leitura pejorativa e da desconfiança em relação à avaliação educacional (e não apenas da avaliação externa, fenômeno recente nesse processo) como uma espécie de característica do próprio desenvolvimento da escola moderna. O vínculo imediato entre avaliação e julgamento promove uma espécie de ojeriza imediata a qualquer processo avaliativo (uma contribuição da psicanálise nesse processo poderia ser muito produtiva). A princípio, ninguém gosta de ser julgado e é isso o que uma avaliação oferece de saída. A dimensão do julgamento promovido por uma avaliação obscurece o que vem a seguir: o julgamento é premissa para um processo de compreensão, seguido, por sua vez, de uma ação. Toda avaliação tem em seu horizonte o agir. Esse aspecto será retomado adiante. De todo modo, o que interessa aqui é o reconhecimento de que a avaliação é um tema espinhoso na educação e isso não está ligado apenas à avaliação em larga escala. Qualquer tipo de avaliação padece dessa mesma leitura.

Se esse diagnóstico está correto, torna-se mais simples compreender o quão difícil é produzir o convencimento de que a avaliação é um direito, em particular, no que tange aos alunos. Dificilmente, um aluno seria capaz de, sem uma explicação convincente, compreender que ele tem o direito de ser avaliado. A ele, a imagem que parece mais condizente com essa tarefa é a do dever. Afinal, se pudesse escolher, certamente, não desejaria ser avaliado. De uma maneira ou de outra, os professores (e a escola como um todo), reproduz essa mesma ideia, já que todos foram alunos. É por termos passado pela escola como alunos que construímos uma imagem de resistência em relação à avaliação.

A despeito disso, a avaliação é um direito. O direito de ser avaliado significa receber informações qualificadas sobre o desenvolvimento de sua aprendizagem, dever, por excelência, da escola. A partir da avaliação, o aluno deveria ser capaz de aprender a gerenciar sua própria aprendizagem, reconhecendo problemas e virtudes nesse processo, amparado pelos professores e demais profissionais da escola. A avaliação, nessa leitura, é um instrumento do e para o aluno. Ela monitora para o aluno aquilo que lhe interessa diretamente (pelo menos, em tese), a sua aprendizagem. Assim, numa inversão completa de imagens, a avaliação passaria de um obstáculo para o aluno a um instrumento exigido pelo aluno. O teor utópico dessa leitura, em especial, quando comparada com o que a escola tem demonstrado na realidade quanto a esse ponto, não deve ser demérito para o raciocínio apresentado. Apesar de não vermos isso em realidade, não é absurdo construir um argumento nesse sentido. Para uma mudança, evidentemente, muito mais do que uma argumentação dessa natureza é necessária, tendo em

vista o grau de enraizamento da visão pejorativa da avaliação na escola brasileira. Contudo, ela é um ponto de partida.

### *3.3.1.1 Fundamentos e características da avaliação educacional*

A perspectiva de que a avaliação é um instrumento a serviço da garantia do direito à educação (e a perspectiva de que ela própria é um direito) é um de seus fundamentos jurídicos. Uma política tão questionada como a avaliação educacional precisou, e ainda precisa, em muitos contextos, de fundamentos para justificar e legitimar sua existência. A avaliação não foi produto de um governo específico e não é datada. Ela permanece em expansão e não pode ser entendida apenas como uma política unilateral e autoritariamente imposta, a despeito do histórico brasileiro com a implementação de um sistema nacional de avaliação (que, ao longo do processo, deu pouco espaço para a voz dos atores educacionais implicados, em especial, os escolares).

Assim, mesmo sem uma efetiva participação dos atores na implementação dos sistemas, a avaliação em larga escala não é despida de fundamentos, ilegítima. Sem eles, pela pura força de imposição, a avaliação não teria podido se sustentar como política pública e, mais do que isso, ter sobrevivido e mesmo se expandido, através de governos distintos, com orientações ideológicas diversas, por vezes, antagônicas. A expansão dos sistemas de avaliação pode ser melhor compreendido quando identificamos os fundamentos responsáveis por sua força. Aliás, tal identificação é crucial para a superação de boa parte das críticas que são dirigidas à avaliação, como veremos em tópico adiante.

A seguir, são apresentados e discutidos três fundamentos da avaliação educacional em larga escala: jurídico, político e científico-pedagógico. Não se trata, vale ressaltar, de um rol taxativo, mas, antes, de um exemplificativo. Outros fundamentos podem ser encontrados para a justificação da existência das avaliações em larga escala. Os três aqui apresentados, contudo, são suficientes para dar uma sólida base à avaliação, contornando boa parte das críticas que a ela atribuem um caráter transitório e volitivo.

Do ponto de vista jurídico, algo já foi dito anteriormente (a avaliação como garantia dos direitos à educação e a avaliação como direito). Resta agora apontar para a previsão jurídica dos sistemas de avaliação, sua ancoragem no ordenamento. Esse ponto é particularmente importante, tendo em vista o contexto de um Estado Democrático de Direito. A previsão legislativa serve aos propósitos de recusar a crítica de que a avaliação é uma mera imposição

governamental. A passagem pelo legislativo, em tese, garante sua legitimidade democrática (produzida no âmbito da colisão de forças antagônicas representadas no parlamento), recusando sua imposição unilateral e autoritária. Além disso, ela implica o reconhecimento da existência de um direito subjetivo (que pode ser exigido pelos indivíduos, ideia que reforça o que foi apresentado no tópico anterior). A previsão legislativa da avaliação é o reconhecimento de que ela foi produzida pela vontade democrática.

A Constituição Federal de 1988 e a Lei de Diretrizes e Bases da Educação (LDB – lei 9394/96) são os documentos normativos nos quais a fundamentação jurídica da avaliação pode ser encontrada. Na Constituição Federal, a avaliação encontra amparo no artigo 205<sup>47</sup>. Ao garantir a educação como um direito de todos e um dever do Estado, o legislador não precisou adjetivar o tipo de educação que está sendo garantida constitucionalmente. O acréscimo da locução adjetiva *de qualidade* ao substantivo educação só expressaria o que a intenção, por motivos óbvios, já havia garantido. Não faz sentido garantir uma educação qualquer aos indivíduos. A garantia do direito à educação só pode se referir a uma educação de qualidade. Se a educação de qualidade é um direito de todos e um dever do Estado, cabe a este último garantir a qualidade da oferta educacional. Caso tal qualidade não seja garantida, a sociedade está habilitada a exigir o cumprimento desse dever estatal, afinal, ela é o destinatário desse direito. Como garantir qualidade sem estabelecer instrumentos apropriados para aferir se ela está sendo atendida ou não? A avaliação é a consequência lógica da previsão constitucional do artigo 205. Nele, a avaliação já encontra seu fundamento jurídico, assim como nele está garantida a ideia de avaliação como direito: a sociedade tem direito a uma educação de qualidade, ofertada pelo Estado; para saber se o direito está sendo garantido, é preciso monitorar se a qualidade da educação está sendo atendida; para monitorar a qualidade, é necessário que existam instrumentos aptos a fazer isso; o que garante que a sociedade tem direito de ver avaliada a educação oferecida pelo Estado<sup>48</sup>.

Somam-se à previsão constitucional as normas previstas pela LDB/96, que tratam mais especificamente da avaliação educacional. O artigo 9º da LDB<sup>49</sup>, em seus incisos V e VI,

---

<sup>47</sup> Artigo 205 (que abre a Seção I do Capítulo III, Título VIII): A educação, direito de todos e dever do Estado e da família, será promovida e incentivada com a colaboração da sociedade, visando ao pleno desenvolvimento da pessoa, seu preparo para o exercício da cidadania e sua qualificação para o trabalho.

<sup>48</sup> Pela própria natureza jurídica da Constituição Federal, ela não é o documento responsável pela definição do polêmico tema sobre o que é qualidade da educação, algo a ser definido por leis infraconstitucionais e específicas. De todo modo, fica a claro, em meu entendimento, que o fundamento da avaliação é, antes de tudo, constitucional.

<sup>49</sup> O artigo 9º, integrante do Título III, “Do Direito à Educação e do Dever de Educar”, prevê: A união incumbir-se-á de: (...) V - coletar, analisar e disseminar informações sobre a educação; VI - assegurar processo nacional

explicitam o que a Constituição de 1988 já havia previsto. O inciso V estabelece como dever da União a coleta, análise e disseminação de informações sobre a educação. Ele já é suficiente para reconhecer o fundamento legal da avaliação em norma infraconstitucional e específica. A avaliação, do ponto de vista técnico, é justamente a coleta e análise de informações (evidentemente, não é uma característica exclusiva das avaliações). A disseminação prevista em lei é a garantia do princípio da publicidade que rege as avaliações e a base da etapa de divulgação de resultados (como veremos uma parte fundamental da política de avaliação). O inciso VI especifica ainda mais a previsão, dizendo respeito explicitamente a um processo nacional de avaliação do rendimento escolar, que deve ser assegurado pela União (em todas as etapas de escolaridade). Este é o fundamento direto para a criação do SAEB, por exemplo. O objetivo de um sistema dessa natureza também é dado pela lei: definir prioridades e possibilitar a melhoria da qualidade do ensino (como vimos, um direito constitucional).

A fundamentação jurídica da avaliação é importante, mas pode parecer insuficiente em um contexto em que direitos são previstos, mas não efetivados. Do ponto de vista político, a avaliação encontra abrigo em, pelo menos, duas ideias: i) a de que sua própria finalidade é política, na medida em que a avaliação existe, primordialmente, para informar decisões; e ii) a de que a avaliação é uma forma do Estado de prestar contas à sociedade em relação aos serviços que deve oferecer, nesse caso específico, os educacionais, prezando pela qualidade dos mesmos, sob pena de ser pressionado para tanto, política e juridicamente. Em relação ao primeiro ponto, mais será dito a seguir. Por ora, falemos da segunda ideia.

Em uma democracia, baseada na divisão dos poderes, o Estado não está acima do direito ou da sociedade. Do ponto de vista do modelo adotado, em teoria, o funcionamento da democracia pressupõe a participação da sociedade como premissa para a legitimidade do Estado. O princípio da soberania popular pressupõe que o povo soberano, num modelo representativo, elege seus representantes na arena política, responsáveis por decidir acerca de políticas públicas (os fins do Estado) e de direitos que as garantam. O Estado perde legitimidade se age fora dessa lógica. As prestações estatais em relação à realização dos fins previstos e ao cumprimento dos direitos estabelecidos passam pela prestação de contas de suas ações diante da sociedade. A sociedade é o critério último da legitimidade do Estado (HABERMAS, 2012). Nesse sentido, as avaliações educacionais, principalmente com base no princípio da publicidade (a divulgação e a disseminação dos resultados), são instrumentos necessários para que o Estado

---

de avaliação do rendimento escolar no ensino fundamental, médio e superior, em colaboração com os sistemas de ensino, objetivando a definição de prioridades e a melhoria da qualidade do ensino.

informe a quem lhe sustenta (política, jurídica, simbólica e economicamente) sobre a qualidade dos serviços prestados (o que está sendo feito para atingir os fins definidos politicamente, para assegurar os direitos previstos e com o dinheiro público). As avaliações, do ponto de vista dos direitos à educação e das políticas educacionais, são instrumentos a serviço da democracia, permitindo a sociedade pressionar o Estado por mudanças.

Ainda há o fundamento técnico-científico. Todo processo educacional baseia-se no desenvolvimento de um conjunto de conhecimentos, habilidades, competências, comportamentos, epistemologias. A aprendizagem é um traço latente, uma característica que não pode ser medida diretamente. Apesar disso, é o grande horizonte do processo educacional e não pode não ser mensurada. Através de sua mensuração, sabe-se se o ensino está alcançando ou não seus objetivos. Isso serve para qualquer tipo de avaliação, não apenas para a avaliação em larga escala. Produzir informações sobre a aprendizagem é indispensável e requer instrumentos especializados e adequados a cada tipo e etapa de aprendizagem. A avaliação, portanto, é um desdobramento, uma consequência do ensino sistematizado pela escola. O que um sistema de avaliação faz é produzir essas informações em larga escala, para toda uma rede de ensino.

Além disso, as informações produzidas pela avaliação experimentam um duplo caráter: ao passo em que i) servem como base e fundamento para a tomada de decisões, nos mais diferentes níveis da estrutura educacional, prestam-se também ii) a promover uma espécie de autoanálise do sistema educacional e, em última instância, de uma meta-avaliação. A primeira faceta será abordada adiante. Quanto à segunda, é preciso reconhecer que a avaliação oferece às escolas e às redes informações que revelam, em boa medida, seus principais pontos de fracasso e força, podendo ser o ponto de partida para uma reflexão mais aprofundada. Em acréscimo, a avaliação é composta por um conjunto significativo de conhecimentos que, ao longo do processo de sua aplicação, permitem aos atores aumentar o conhecimento sobre a própria maneira de avaliar. Nessa perspectiva, a avaliação é uma fonte de conhecimento, um processo de pesquisa e investigação que, com observância de uma metodologia rigorosa, oferecerá informações de qualidade a seus operadores. Somente por essa perspectiva, a avaliação educacional encontraria fundamento e legitimidade, mesmo que os fundamentos jurídicos e políticos não estivessem presentes.

Em relação a esse último fundamento, a avaliação em larga escala costuma ser definida. Qual o objetivo primordial da avaliação educacional? Quais são suas principais características? A resposta mais direta dada a essa pergunta, e, possivelmente, a mais comum, pode ser

encontrada em Klein e Fontanive, já em 1995, nos primórdios do sistema nacional de avaliação no Brasil:

A avaliação educacional é um sistema de informações que tem como objetivos fornecer diagnóstico e subsídios para a implementação ou manutenção de políticas educacionais. Ela deve ser concebida também para prover um contínuo monitoramento do sistema educacional com vistas a detectar os efeitos positivos ou negativos de políticas adotadas (KLEIN; FONTANIVE, 1995, p.29).

A definição dos autores é direta e contém os elementos que, até hoje, permanecem centrais para a definição da política de avaliação. Como sistema de informações (fundamento técnico-científico) a avaliação primordialmente tem como objetivo produzir diagnósticos e funcionar como subsídio para a implementação e manutenção de políticas públicas. Essa perspectiva está presente também em Silva e Carvalho (2014), Bauer, Alavarse e Oliveira (2015)<sup>50</sup>, Franco, Alves e Bonamino (2007) e Bonamino e Sousa (2012). Outra característica importante já está presente na definição dos referidos autores: a ideia de que a avaliação é contínua. Seu caráter de continuidade é um importante aspecto de sua duração como política pública e explica em parte, como veremos, o desenvolvimento de sistemas próprias de avaliação. Resta dizer que as informações precisam ser qualificadas (Klein e Fontanive falam em consistência, 1995, p.30), algo garantido pelo uso de uma metodologia adequada para realizar a mensuração proposta da maneira mais rigorosa possível<sup>51</sup>.

Importante, a partir dessa definição, esmiuçar suas características. A principal delas é seu vínculo com as demais políticas públicas. Nesse ponto, entra em cena o caráter dúplice da avaliação educacional: i) é política pública e, ao mesmo tempo, ii) avaliação das demais políticas públicas. A idiossincrasia da avaliação como política pública concentra-se, justamente, em seu objetivo maior: produzir informações de suporte para a implementação de outras políticas públicas. Uma política pública como base, prerrogativa, para outras políticas públicas. Nesse sentido, é possível reconhecer o caráter instrumental da avaliação, e ele pode ser

---

<sup>50</sup> Os autores reconhecem “a contribuição das avaliações em larga escala e das medidas educacionais como norteadoras de políticas públicas e programas educacionais (...)” (BAUER; ALAVARSE; OLIVEIRA, 2015, p. 1367).

<sup>51</sup> Um dos elementos metodológicos responsáveis pela garantia dessa consistência e desse rigor é a Teoria da Resposta ao Item – TRI, que passou a fazer parte das avaliações nacionais a partir de meados dos anos 1990. Conforme pontuam Klein e Fontanive, a “TRI é um conjunto de modelos onde a probabilidade de resposta a um item é modelada como função da proficiência do aluno (variável não observável) e de parâmetros (que expressam certas propriedades) do item. Quanto maior a proficiência, maior a probabilidade de o aluno acertar o item” (1995, p.31). Ela é responsável também por garantir a comparabilidade entre os desempenhos dos alunos. Para mais sobre a TRI, ver Pasquali e Primi, 2003.

entendido, mais uma vez duplamente: i) a avaliação é um conjunto de instrumentos, de ferramentas, a serviço da produção de informações; ii) ela não possui existência descolada de outras políticas públicas, sendo necessária como critério qualificado de informação para potencializar as chances de outras políticas públicas educacionais terem sucesso.

Ainda um ponto merece ser considerado. Além de dar base para a tomada de decisões e para a implementação de outras políticas, a avaliação é utilizada como critério para avaliar o sucesso (no sentido de alcance dos objetivos propostos) das políticas educacionais como um todo. Duas questões estão envolvidas nessa perspectiva. A primeira delas parte do princípio de que toda política educacional tem como objetivo último, mediato, a melhoria da qualidade da educação. Isso diz muito sobre a qualidade no âmbito educacional: ela é o objetivo por excelência de qualquer ação no campo, mesmo que não seja de modo imediato. Sendo assim, se toda política educacional busca, em maior ou menor medida, a qualidade da educação, a avaliação torna-se o critério mais adequado para confirmar (ou não) o alcance dos objetivos dessa política (se houve melhora da qualidade, atestada pela avaliação, é sinal de que a política avaliada surtiu efeitos). A segunda questão relaciona-se com o fato de que, a despeito desse uso das avaliações educacionais, elas, entre nós, não têm sido desenhadas especificamente para avaliar o impacto de políticas educacionais. Seu objetivo direto, permitido pelo desenho que assume, é monitorar o desempenho dos alunos, escolas e redes, não avaliar, especificamente, o efeito de cada uma das políticas educacionais atuantes no período e no contexto em que as avaliações são aplicadas. Avaliar impactos de políticas públicas é uma tarefa mais complexa e envolve, em geral, instrumentos com características específicas (JANUZZI, 2016). De todo modo, mesmo com a ausência de um desenho específico para tanto, os resultados das avaliações têm se beneficiado de técnicas que buscam isolar os efeitos de outras variáveis com o intuito de avaliar o impacto de políticas educacionais específicas. As consequências das políticas educacionais podem ser melhor compreendidas a partir dos resultados das avaliações:

Ainda que a avaliação nacional tenha importantes limitações para a investigação de efeitos causais (Franco, 2001), é inegável que os dados da avaliação em larga escala oferecem oportunidade ainda ímpar para que se investigue empiricamente as consequências de políticas e práticas educacionais (FRANCO; ALVES; BONAMINO, 2007).

O que fica da definição de Klein e Fontanive (1995) é a natureza da avaliação como base para a tomada de decisões. Embora os autores afirmem que as avaliações não têm como objetivo produzir informações sobre alunos e escolas individuais (p. 30), é preciso reconhecer que, no momento em que escreviam, a avaliação ainda não havia experimentado o

desenvolvimento e a expansão que só se apresentariam alguns anos depois. Naquele momento (1995), a avaliação educacional no Brasil ainda vivia sua primeira geração (BONAMINO; SOUSA, 2012), tema que será retomado em tópico adiante. Contudo, hoje, podemos afirmar como característica da avaliação educacional sua vocação como suporte para a tomada de decisões nos mais diferentes níveis da estrutura educacional, das redes de ensino às turmas na escola. Isso significa que as avaliações em larga escala não são instrumentos privilegiados para uso das secretarias de educação, se apresentando como importantes também para as escolas<sup>52</sup>. Os resultados da avaliação, sua divulgação e sua apropriação pelos agentes escolares desempenham um papel particularmente importante em relação a isso. Esse tema será retomado a seguir.

A existência de um tópico sobre os fundamentos e características da avaliação, mais do que apenas descrever a política em seus principais elementos, justifica-se pelas polêmicas que envolveram e continuam envolvendo a avaliação educacional em larga escala. Um dos desafios permanentes da avaliação no contexto brasileiro é legitimar-se. O tópico a seguir apresenta as principais críticas que avaliação recebe e busca desmistificá-las, na medida em que não encontrem força argumentativa e justificativa.

### **3.3.2 Críticas e virtudes da avaliação**

Em trabalho anterior (REZENDE, 2014), apresentei um conjunto de argumentos em relação às críticas que a avaliação recebe, apontando que a maior parte delas tem origem em um duplo movimento: i) o desconhecimento acerca dos limites e virtudes da política de avaliação, incluindo suas etapas e os próprios instrumentos (testes e questionários); e ii) o papel desempenhado por uma ideologia fortemente disseminada no âmbito educacional, que vincula a avaliação educacional, de maneira irrestrita, com o neoliberalismo e com o avanço da lógica do mercado sobre a escola. Esses movimentos, quando somados, produzem uma forte resistência à ideia de avaliação em larga escala, reforçando a desconfiança que a avaliação já recebe em virtude de toda a história de seu uso no ambiente escolar, conforme apresentado em tópico anterior.

O desconhecimento em relação à avaliação pode ser atribuído, em parte, à forma como os sistemas de avaliação foram implementados no Brasil (tanto o nacional, quando os próprios,

---

<sup>52</sup> O significativo número de trabalhos de pós-graduação que versam sobre o tema de apropriação de resultados pelas escolas é um exemplo disso. Esse ponto será retomado de modo mais detido no capítulo 3 desta tese.



estaduais e municipais), sem que houvesse um efetivo diálogo e envolvimento dos atores escolares. Isso alimenta, de modo justo, a crítica de que a avaliação, em um primeiro momento, veio de modo imposto e pouco debatido. Os atores educacionais precisavam lidar com o instrumento que eles não conheciam. Contudo, o desconhecimento, mesmo passado esse momento inicial, foi alimentado pela resistência em relação à política (não apenas não conheciam, mas não queriam conhecer). Desconhecimento e ideologia se retroalimentam, sem que seja possível identificar uma relação de causa e efeito entre eles.

Por outro lado, isso provoca reações naqueles que veem valor na avaliação, que buscam justificar a política diante das numerosas críticas recebidas. Um perigo que se corre é a polarização excessiva do debate. Assim, duas posições arquetípicas se apresentam (como tipos ideais weberianos, WEBER, 1999). Os defensores da avaliação correm o risco de, no intuito de defendê-la, atribuir a ela um caráter soteriológico, vendo na política uma espécie de solução para todos os problemas educacionais, vocação que, evidentemente, ela não possui. Com isso, tendem a ver apenas os aspectos positivos, sem, contudo, identificar os limites da avaliação e os problemas da política. Os detratores, por sua vez, em regra, caracterizados pelo desconhecimento (as críticas mais ferozes são oriundas daqueles que nunca trabalharam efetivamente com a avaliação em larga escala), tendem a ver a avaliação como um inimigo público da educação, atribuindo a ela os principais males do sistema educacional contemporâneo (REZENDE, 2014). Como estamos diante de tipos ideais, esses polos servem a propósitos analíticos, não existindo, em todas as suas características, na prática. Empiricamente, os críticos e defensores da avaliação equilibram-se entre cada um desses polos radicais, ora aproximando-se mais da crítica demonizada, ora mostrando mais afinidades com a defesa deificada. De um modo geral, as críticas que parecem mais acertadas e justificadas em relação à avaliação vêm de autores<sup>53</sup> que conhecem bem a política e o instrumento, identificando suas virtudes e seus limites.

O processo de deificação e de demonização da avaliação promovido por essas polarizações tem a mesma causa: a avaliação é personalizada, sendo atribuído a ela um conjunto de características que pessoalizam o instrumento. É como se a avaliação tivesse vontade própria, caráter e pudesse, ela própria, tomar decisões<sup>54</sup>.

---

<sup>53</sup> Ver, por exemplo, BAUER; ALAVARSE; OLIVEIRA, 2015, e BONAMINO; SOUSA, 2012.

<sup>54</sup> “A crítica, nesse caso, parte de um processo de personalização da avaliação, que passa a ser tratada como uma espécie de pessoa, dotada de caráter duvidoso. Daí a observância de expressões tais como ‘a avaliação é responsável pela situação crítica das escolas’, ‘a avaliação faz ranqueamento de alunos’, ‘a avaliação penaliza os professores e diretores’, e outras do gênero” (REZENDE, 2014, p.157).

Esse processo de personalização da política de avaliação, e mesmo do instrumento, somado ao desconhecimento acerca da política, tem como consequência a significativa confusão entre o instrumento e seu uso. As críticas feitas ao uso da avaliação (em especial, o que é feito com seus resultados), tais como os ranqueamentos e as políticas de bonificação são estendidas ao instrumento como um todo, à política inteira, de modo que a maneira como a avaliação é usada contamina a própria ideia de avaliação. Nesse contexto, as virtudes da política são obliteradas, sem que se possa ver nela qualquer aspecto positivo para a educação. As críticas ao uso são importantes (e elas têm sua razão de ser, conforme veremos), mas não podem obscurecer o reconhecimento da importância que a avaliação em larga escala pode ter para a garantia da qualidade da educação. Contudo, é preciso “estabelecer com clareza a que a crítica se destina. Ao criticar todo o instrumento a partir apenas de seu uso, o resultado é a desconsideração daquilo que a avaliação pode, efetivamente, produzir em benefício da educação pública” (REZENDE, 2014, p.159). Em relação ao uso, é preciso reconhecer que ele pertence ao âmbito da decisão e da ação humana. Atores com capacidade de decisão, ocupando posições específicas para tanto, decidem usar os resultados da avaliação de uma maneira ou de outro. Como instrumento, a avaliação pode ser usada para diferentes fins. Seus resultados podem assumir um caráter punitivo (como, por exemplo, no uso para promover punições a diretores de escolas com baixo desempenho) ou reflexivo (nos casos em que as escolas são estimuladas – e capacitadas para tanto - a utilizar os resultados como base do planejamento). Depende muito mais dos atores do que do instrumento e da política em si.

Da mesma forma, em termos técnicos, há problemas a serem solucionados. Parte significativa da pesquisa em avaliação dedica-se ao tema da validade da medida produzida (a importância da TRI se faz notar ainda mais forte nesse quesito). No entanto, os críticos raramente estabelecem objeções a questões específicas de natureza técnica<sup>55</sup>, já que, de saída, tendem a rechaçar a política como um todo.

De modo mais realista, fugindo da dicotomia polarizada apresentada, a avaliação é uma política que possui limites e potencialidades (REZENDE, 2014). Reconhecê-los é o melhor que se pode fazer para a eficácia e aprimoramento dessa política. Tanto no que diz respeito a seu uso quanto no que tange à sua natureza técnica como instrumento, a avaliação apresenta problemas e forças. Seu uso pode levar a cenários indesejáveis, mas pode conduzir a um processo de autorreflexão da escola, com consequências muito importantes para a melhoria do

---

<sup>55</sup> A relação entre desconhecimento e crítica fica aqui mais patente, visto que as questões de natureza técnica tendem a ser desconhecidas por críticos mais contumazes e radicais.

ensino e potencialização da aprendizagem. Sua técnica é limitada (a metodologia é probabilística e opera, todo o tempo, com o cálculo e o controle do erro na produção da medida), contudo, é o mais apurado que temos no que diz respeito à mensuração de um traço latente (PASQUALI; PRIMI, 2003), como é o caso da aprendizagem. Portanto, para que a crítica seja bem direcionada e produtiva, é necessário que se estabeleça, com clareza, o que pertence ao âmbito do uso e o que é próprio do instrumento e inerente à política.

Essa distinção, que, como argumentado, é tributária do desconhecimento acerca da política e do instrumento, para que possa ser superada, tem contado com a presença quase obrigatória de programas de capacitação para interpretação e uso dos resultados nos sistemas de avaliação (em particular nos sistemas próprios estaduais). Ou seja, mais do que tornar o resultado público, é necessário capacitar os atores educacionais para que sejam capazes de interpretá-los corretamente, orientado quanto ao seu uso. O tema será retomado em tópico adiante.

Como os resultados, conforme veremos, ocupam um papel de destaque na política de avaliação, a consequência é que as críticas destinam-se, em sua maior parte, ao uso que se faz deles. Bauer, Alavarse e Oliveira (2015) apontam para as principais críticas na literatura sobre o uso dos resultados. Elas se direcionam majoritariamente às políticas de responsabilização (*accountability*) associadas aos resultados. Mesmo entre autores que buscam estabelecer um ajuste entre a importância da avaliação e as críticas destinadas a aprimorar a política, reconhecem o quão polêmica é a prática da responsabilização a partir de resultados educacionais, como Brooke, Alves e Oliveira (2015). Freitas (2013) critica a atribuição de bônus e premiação a professores e escolas, devido aos altos impactos que uma política como essa pode causar. Em outro trabalho, Freitas (2013b) aponta para dois problemas das políticas de responsabilização: a falta de evidência e os problemas éticos que ela gera. Por trás dela, um grande mercado, rico e volumoso, o que explicaria sua duração e expansão (os Estados Unidos são tomados como o grande exemplo novamente, FREITAS, 2013b, p.348-349). O mesmo Freitas (2012) vincula a responsabilização à meritocracia e à privatização, como se todas fizessem parte do mesmo e necessário arranjo. Desta forma, a própria avaliação é acrescentada à equação, contaminando a política e o instrumento pelo uso que se faz dela (de acordo com o que já discutimos anteriormente). Segundo ele, essa combinação é típica de uma postura adotada pelos “reformadores empresariais da educação no Brasil” (FREITAS, 2012, p.380), que fazem com que as ideias da produção americana sobre o tema colonizem as secretarias estaduais e municipais de educação no país. A crítica de Freitas é um exemplo do vínculo

pretensamente necessário entre avaliação, responsabilização, meritocracia e mundo empresarial. Dois problemas principais aparecem nessa crítica: i) a ideia de que a meritocracia é uma característica trazida à escola pela avaliação em larga escala; e ii) a ideia de que importamos irrefletidamente políticas de outros países, em especial, dos Estados Unidos. Quando ao segundo ponto, algo já foi dito anteriormente (as políticas não são importadas sem algum nível de ajuste à realidade nacional e esse é o caso do Brasil – as políticas de responsabilização não são iguais aqui e lá; o próprio Freitas reconhece isso, 2012). Em relação à meritocracia, ela é uma característica da escola como a conhecemos, o modelo adotado pelo Brasil e por boa parte do mundo moderno. A meritocracia é a característica por excelência do modelo escolar republicano francês<sup>56</sup> e não uma idiossincrasia do mercado trazida à luz pela avaliação em larga escala.

Outras críticas são direcionadas a outros aspectos. Minhoto (2012) chama a atenção para o engessamento curricular que as avaliações podem promover, dando extrema importância ao conteúdo das provas e negligenciando outras funções do desenvolvimento humano que deveriam ser promovidas pela escola. Oliveira (2011), em pesquisa de dissertação de mestrado sobre os usos dos resultados da Prova Brasil, ouviu professores<sup>57</sup> que afirmaram redefinir o conteúdo programático para adequação ao que era cobrado na avaliação nacional, crítica frequentemente dirigida à avaliação educacional, qual seja, a redução do currículo. Analisando os efeitos dos resultados do Saesp em uma escola da rede estadual, Freire (2008), encontrou práticas destinadas ao treinamento dos alunos para a realização dos testes, mais um desdobramento do empobrecimento do currículo, também muito criticada, assim como ações de outra natureza, tais como a atribuição de notas aos alunos para a participação no dia dos testes e a utilização dos resultados dos testes em larga escala como parte das notas bimestrais dos alunos. Em relação a este último ponto, o vínculo entre avaliação externa e interna está longe de ser óbvio, seja para defender seu antagonismo, seja para propor sua complementaridade. Mesmo autoras dispostas a reconhecer que as avaliações possuem virtudes chamam a atenção para os riscos de avaliações baseadas em políticas de responsabilização, como é o caso de Bonamino e Sousa (2012). Sobre o estudo acerca das três gerações de avaliação (o qual retomaremos a seguir), elas afirmam:

---

<sup>56</sup> Para mais sobre esse tema, ver François Dubet, 2004, *O que é uma escola justa?* O objetivo não é fazer uma defesa da meritocracia, mas reconhecer que seu vínculo com a escola não está atrelado ao surgimento da avaliação em larga escala, fenômeno muito mais recente.

<sup>57</sup> De duas escolas do Distrito Federal, acerca do uso de resultados da Prova Brasil edição 2007.

Apontou (**o artigo**), por um lado, os riscos de as avaliações relativas a políticas de responsabilização exacerbarem a preocupação de diretores e professores em preparar seus alunos para os testes, levando a um estreitamento do currículo escolar. Indicou, ainda, as implicações para a avaliação da aprendizagem quando as escolas passam a organizá-la tomando como referência o tipo de teste utilizado pela avaliação em larga escala (BONAMINO; SOUSA, 2012, p.386) (acréscimo nosso, em negrito).

Um ponto crucial do debate gira em torno da própria ideia de qualidade. A crítica é dirigida, então, à redução do conceito de qualidade ao que pode ser medido pelos testes dos sistemas de avaliação em larga escala (MACHADO; ALAVARSE, 2014), limitação essa que se estende ao Ideb (OLIVEIRA, 2011b). Entendendo que a avaliação é limitada, estabelecer como qualidade aquilo que os testes padronizados são capazes de medir é negligenciar outros aspectos fundamentais, que não são mensurados pelos instrumentos, tais como componentes atitudinais e comportamentais, relacionados à vida cívica e a uma postura ligada à cidadania. De fato, os testes são limitados, não apenas nesse sentido, mas diante da própria metodologia adotada, de modo que nem todo o conteúdo curricular consegue ser plenamente avaliado. Em relação aos aspectos atitudinais e não conteudísticos, as chamadas competências socioemocionais vêm sendo paulatinamente inseridas no debate acerca da avaliação e sendo avaliadas em contextos pontuais (para mais sobre vínculo com a aprendizagem, ver SANTOS; PRIMI, 2014). No que diz respeito às limitações dos instrumentos, plenamente reconhecidas, a pergunta que cabe aqui é outra: o que conseguimos avaliar em termos curriculares, mesmo com as limitações inerentes à metodologia e ao tipo de instrumento utilizado, é suficiente para produzir diagnósticos seguros acerca da qualidade da educação no Brasil? Mesmo sabendo que a qualidade da educação é um fenômeno mais complexo do que aquilo que pode ser mensurado pelos testes em seu estágio atual de desenvolvimento, os diagnósticos obtidos não se prestam a fornecer uma visão segura sobre a qualidade do ensino oferecido? Não parece ser o caso, sobretudo quando observamos os resultados nacionais. O restrito que pode ser mensurado pelos testes já nos mostra os problemas que a educação brasileira enfrenta em relação ao aprendizado dos alunos.

Somam-se a isso os gastos exigidos pela política de avaliação e seu retorno quanto à melhoria da educação. O questionamentos aqui giram em torno de seu custo-benefício. Em um país que apresenta problemas profundos em relação à sua infraestrutura educacional (incluindo aqui os recursos humanos), por que destinar vultuosas quantias para avaliar se essa prática não tem mostrado nenhum efeito sobre a melhoria da qualidade educacional que ela pretende avaliar? O mesmo pode ser expandido para a realidade latino-americana, como faz Tedesco (2003). A pesquisa

em avaliação, neste ponto, ainda se ressentem da ausência de informações mais precisas e aprofundadas sobre o efeito das próprias avaliações sobre o desempenho dos alunos. A existência de um sistema de avaliação pode, isolando os efeitos de demais ações e políticas, impactar na aprendizagem dos alunos? Ou a avaliação não se presta a esse fim, limitando-se a avaliar o efeito de outras políticas (o que não é pouca coisa)? Apesar de se considerar que a própria avaliação possa ter efeitos sobre a aprendizagem, seu intuito fundamental é produzir informações sobre os efeitos de outras políticas. Os gastos com esse tipo de política são justificados pelos tomadores de decisão justamente por aquilo que a avaliação pode lhes oferecer: suporte para a tomada de decisões, inclusive as de caráter financeiro. Além disso, não há na literatura pesquisas mais aprofundadas sobre o aspecto financeiro ligado à avaliação. Um eventual cancelamento dos sistemas de avaliação geraria alívio financeiro para os entes, a ponto de áreas mais críticas (como o pagamento dos profissionais da educação) resolverem seus problemas? Ou a avaliação destina-se, em parte, justamente a auxiliar na escolha de zonas prioritárias nas quais os investimentos devem ser maiores?

Uma crítica fundamental à avaliação diz respeito às influências externas. A política de avaliação é criticada, muitas vezes, por ser uma importação de políticas desenvolvidas em contextos nacionais distintos do nosso, em especial, nos Estados Unidos. Boa parte da lógica economicista e meritocrática que a caracterizaria, segundo os críticos, é tributária dessa origem. Produto da expansão do neoliberalismo, a avaliação em larga escala seria um instrumento tipicamente a serviço do mercado, levando à escola instrumentos, pressupostos e mesmo uma forma de pensar que é contrária ao que é (ou deveria ser) defendido no ambiente escolar. A avaliação externa é vista como competitiva (pelo menos, como um instrumento que estimula a competição) e excludente, características do mercado neoliberal. Importada, ela padece com todos os problemas que esse fenômeno de importação apresenta: desconsideração do contexto local, levando a uma espécie de artificialismo da política; esse aspecto artificial leva, necessariamente, à ineficácia da política, que não produz os efeitos pretendidos; criada e disseminada em um ambiente externo (no caso, neoliberal), a política não é capaz de se livrar de suas características de origem (sendo também neoliberal). Hagopian (2014), por exemplo, nos Estados Unidos, critica o vínculo entre avaliação e neoliberalismo, atribuindo a ela a função de aprofundar as desigualdades educacionais já existentes. Soma-se a isso a ideia de que somos uma espécie de etapa do que os países desenvolvidos já viveram: aconteceu primeiro com eles e, trágica e inevitavelmente, acontecerá conosco. Ou seja, basta olharmos para a avaliação nos Estados Unidos de modo a vermos o que nos espera. Eles são o espelho do nosso futuro (FREITAS, 2013b).

Diante disso, a avaliação é vista, de modo reducionista, como uma mera cópia dos programas de avaliação que acontecem em cenários mais desenvolvidos economicamente do que

o nosso, levando à negligência de aspectos fundamentais de nossos sistemas educacionais, em particular, a desigualdade gritante que os define. Outro problema é achar que a avaliação, em seus contextos de origem, é livre de crítica. Eis um ponto de semelhança entre o Brasil e os Estados Unidos, por exemplo: como aqui, a avaliação não é consensual. Há críticos<sup>58</sup> e defensores.

É interessante notar, contudo, conforme lição de Bauer, Alavarse e Oliveira (2015, p.1369), que a “contribuição internacional do debate a respeito de avaliações em larga escala é praticamente ausente no debate brasileiro acerca do tema”. A influência internacional é muito menor do que os críticos tendem a afirmar, inclusive, pelo próprio papel de resistência da crítica em relação a essas contribuições. Sem negar a influência internacional das políticas educacionais internacionais em nosso contexto (elas estão na agenda mundial, como apontado por Bonamino e Sousa, 2012), particularmente no que diz respeito à sua difusão e percepção de necessidade, é preciso reconhecer que elas não são cópias puras do que acontece em outros países. As características de cada cenário educacional não são tão passivas de modificação pela mera chegada de uma política. Esse processo é sempre bem menos fluido do que sugere a ideia da simples importação. Ao serem implementadas aqui, as políticas passam por ajustes, resistências e adaptações, realizados de modo consciente e planejado ou inconsciente (a partir da resistência oferecida pela própria estrutura de organização das redes e do ensino). A expansão dos sistemas próprios de avaliação no Brasil, com o protagonismo sendo assumido pelos estados (em especial) e municípios é um exemplo disso. Trata-se de um fenômeno desenvolvido entre nós, sem que seja cópia de outros países.

No que diz respeito à responsabilização, não apenas críticas são encontradas. Literatura internacional tem mostrado os efeitos de programas de accountability, como é o caso de Cullen e Reback (2006) e Hanushek e Raymond (2005) e Jacob (2005), no contexto norteamericano. Sem rechaça-lo ou aceita-lo absolutamente e de modo irrefletido, o fenômeno da responsabilização parece ser mais bem enfrentado quando segue a lição de Afonso (2012): uma vez aceita a responsabilização como parte da cobrança legítima da sociedade acerca da oferta educacional estatal, cumpre procurar os meios de realizá-la de maneira tal que seus objetivos sejam os mais democráticos possíveis.

O mesmo acontece em relação aos demais aspectos alvos da crítica. Embora reconheçam os problemas e os riscos que a avaliação pode gerar, dependendo do uso que dela seja feito, Bonamino e Sousa (2012, p.386-387) reconhecem a importância que a política de avaliação pode

---

<sup>58</sup> Para mais detalhes sobre as críticas à avaliação no contexto norteamericano, ver Hagopian, 2014, em obra que reúne uma série de contribuições de críticos à disseminação dos testes entre eles.

assumir para gerar debates mais aprofundados e informados acerca do currículo escolar, em especial, sobre as habilidades básicas ainda não desenvolvidas pelos alunos da educação básica. Em seu mapeamento acerca da pesquisa sobre os usos das avaliações, Bauer et al (2015) encontraram na literatura outros argumentos favoráveis à sua adoção: i) desenvolvimento de uma cultura de transparência, com a publicização dos resultados; ii) que levaria, por sua vez, ao apoio, aos pais e responsáveis pelos estudantes, para a tomada de decisão no que tange à escolha das escolas; iii) o sentimento de responsabilidade, por parte da escola e dos professores, em relação aos resultados alcançados (gerando a ideia de que algo deve ser feito para a melhoria do ensino ofertado); iv) redução da subjetividade que afeta os resultados, visto que lança mão de uma metodologia mais objetiva (a TRI como elemento primordial); v) a possibilidade de acompanhar a evolução da educação, mais um virtude atribuída à técnica utilizada; vi) são capazes de impulsionar mudanças em currículos desajustados (conforme Bonamino e Sousa, 2012, já haviam pontuado).

A defesa da avaliação como política base para outras políticas gira em torno não de uma visão cega sobre suas virtudes. Ao mesmo tempo, exige um enfrentamento menos enviesado e apriorístico acerca de seus problemas. Como instrumento, ela está sujeita a problemas e é passível de ajustes. A avaliação não é apenas um conjunto de técnicas estatísticas, portanto, na visão de muitos, fria e distante do que acontece na escola, com linguagem, em regra, cifrada e inacessível para a maior parte dos atores escolares. Seguindo definição de Vianna (2003), a avaliação deve se preocupar com sua validade consequencial, não apenas no que diz respeito à fidelidade de seus resultados, mas no que tange à forma como impacta seus usuários, o comportamento e as atitudes daqueles para os quais ela é destinada, ou seja, a diferença que é capaz de produzir na educação. Essa visão parece fazer mais justiça à avaliação, escapando de sua personificação polarizada, conforme argumentado apresentado no início do tópico.

### **3.3.3 O sistema nacional e o surgimento dos sistemas próprios**

Desde o início dos anos 1990, a avaliação educacional em larga escala não apenas resistiu às diferentes trocas de governo no cenário federal como se consolidou (CASTRO, 2009a) e se desenvolveu. Embora criticada em diversos aspectos, a política de avaliação mostrou um grau de continuidade excepcional em relação a outras políticas educacionais. Parte dessa durabilidade pode ser explicada pela própria natureza da avaliação educacional: produtora de informações qualificadas sobre o efeito de outras políticas, a avaliação não se apresenta como uma iniciativa pontual, que esgota seus efeitos uma vez aplicada. Com o intuito de ser base de um



monitoramento da evolução da educação, a avaliação requer tempo e continuidade. Não há sentido em um monitoramento pontual. Seus efeitos podem ser sentidos ao longo do tempo.

Outro fator que explica sua continuidade relaciona-se com o objetivo último da avaliação (produzir informações qualificadas para a tomada de decisão) e a forma como tal objetivo foi entendido e apropriado pelos atores educacionais. As decisões sobre políticas educacionais passaram a depender, sobremaneira, dos diagnósticos oferecidos pelas avaliações, de modo que, mesmo em governos com orientações ideológicas discursivamente distintas, a importância da avaliação educacional não arrefeceu. De acordo com o que vimos anteriormente, para Franco, Alves e Bonamino (2007), a centralidade da avaliação educacional pode ser creditada ao resgate da importância da racionalidade técnica no âmbito das políticas educacionais brasileiras. Para alguns autores (GATTI; BARRETO; ANDRÉ, 2011; AFONSO, 2000), a transformação das avaliações em instrumento central das políticas educacionais é consequência de dois movimentos: i) da eficácia das pressões de organismos internacionais sobre o contexto educacional nacional (em especial, por força do papel desempenhado por instituições como a OCDE e o Banco Mundial, que condicionavam investimentos à implementação de sistemas de avaliação); e, ii) como consequência dessa influência internacional, a adoção do modelo gerencialista como padrão a ser seguido no âmbito das reformas educacionais. Sobre a influência de políticas educacionais internacionais em nosso contexto, algo já foi dito em tópicos anteriores. Em relação ao segundo ponto, resta uma pesquisa mais profunda sobre o que significa de fato, o gerencialismo em educação, seus efeitos sobre as políticas e práticas educacionais, seus limites, seus problemas e, também, suas virtudes<sup>59</sup>. De todo modo, um efeito do gerencialismo seria, de acordo com os autores supracitados, o foco das políticas educacionais apenas nos resultados de desempenho, entendidos de modo instrumental, e levando ao já referido reducionismo curricular<sup>60</sup>.

O ano de 1990 ficou estabelecido como o início, entre nós, de um sistema nacional de avaliação. A criação do Saeb, contudo, foi antecedida, conforme pontua Gatti (2013), por um conjunto de ações e iniciativas do governo federal, em parceria com alguns estados e fundações, para a realização de avaliações de rendimento (como exemplo, as avaliações realizadas pela secretaria estadual do Paraná, em 1988, Gatti, *idem*). Segundo a autora, os baixos resultados dessas avaliações incipientes (ainda não constituindo um sistema) surpreenderam os agentes do MEC e das secretarias de educação envolvidas, funcionando como um impulso importante para

---

<sup>59</sup> Um esforço nesse sentido encontra-se em andamento, em formato de livro, conduzido pelo professor Nigel Brooke e por mim.

<sup>60</sup> Os questionamentos sobre a exclusividade da avaliação de conteúdos de Matemática e Língua Portuguesa encontra abrigo nessa crítica.

a criação do Saeb em seguida. Em 1993<sup>61</sup>, o Saeb realizou uma avaliação com uma amostra nacional de alunos, do ensino fundamental e médio, já contando com a aplicação de questionários contextuais. Sua regulamentação, através de Portaria Ministerial, ocorreu em 1994, e uma nova edição do Saeb ocorreu em 1995. Desta vez, ao contrário das iniciativas anteriores, a avaliação contou com uma significativa mudança metodológica. A Teoria da Resposta ao Item - TRI substituiu<sup>62</sup> a Teoria Clássica dos Testes – TCT<sup>63</sup>, permitindo a comparação dos resultados (conforme já vimos em Klein e Fontanive, 1995). A partir da edição de 1995, também, o Saeb tornou-se bienal, avaliando 4ª e 8ª séries (5º e 9º anos, respectivamente) do ensino fundamental e a 3ª série do ensino médio, em Língua Portuguesa e Matemática.

Nova grande mudança acontece na edição de 2005. Entre 1995 e 2002, o Saeb seguiu com o mesmo padrão de avaliação, avançando em técnicas de elaboração de itens e no ajuste das matrizes de referência que dão suporte aos testes. No entanto, de acordo com leitura de Freitas (2013c), o formato do Saeb até então tinha se mostrado insuficiente para promover ações no âmbito das escolas. A explicação para isso estava no desenho da avaliação, que era amostral. Por conta disso, em 2005, foi criada a Anresc, conhecida como Prova Brasil (o formato amostral, que permaneceu sendo chamado de Saeb, foi mantido, recebendo o nome de Aneb). Avaliando as mesmas séries/anos e as mesmas disciplinas, a grande inovação da Prova Brasil foi a mudança do desenho da avaliação, de amostral para censitário. Desta forma, os resultados poderiam ser produzidos num nível mais rizomático do que aquele oferecido pelo Saeb amostral, chegando a escolas e turmas.

Com formato censitário, a Prova Brasil abriu espaço para a criação do Ideb – Índice de Desenvolvimento da Educação Básica, em 2007, combinando indicadores de fluxo escolar e de desempenho nas avaliações externas nacionais (FERNANDES, 2007). O Plano de Metas Compromisso Todos Pela Educação (Decreto n. 6.094/2007) estabeleceu o Ideb como base para a estipulação de metas bianuais, para escolas e redes de ensino, até o ano de 2021. Gatti (2013) e Freitas (2013c) apontam que o Ideb e a centralidade que assumiu na política educacional brasileira terminaram por fortalecer o modelo gerencial na educação e por contribuir para a percepção restritiva do que é qualidade da educação, na medida em que o próprio índice foi confundido com a ideia de qualidade.

É preciso reconhecer que a Prova Brasil e o Ideb deram início a uma nova etapa no cenário da avaliação educacional no Brasil. A possibilidade de produzir resultados no nível das

---

<sup>61</sup> Detalhes sobre a edição de 1993 do Saeb podem ser encontrados em Pilati (1994).

<sup>62</sup> Sobre a adoção da TRI no âmbito do Saeb, ver Klein (2003).

<sup>63</sup> Para mais sobre a TCT, seus limites e capacidades, ver Sartes e Souza-Formigoni (2013) e Primi (2012).

escolas colocou em cena um novo desenho para a avaliação, no que diz respeito ao alcance de seus efeitos. A Prova Brasil é considerada uma avaliação de segunda geração, na já clássica distinção estabelecida por Bonamino e Sousa (2012), que será retomada com mais detalhe no tópico seguinte. Por segunda geração de avaliação, entende-se aquele conjunto de avaliações que, além da divulgação pública dos resultados, característica marcante da primeira geração de avaliação, na qual o Saeb amostral se encaixa, entrega os resultados para cada uma das escolas, estabelecendo consequências simbólicas, mas não materiais (idem, 2012, p. 375).

A primeira geração de avaliações, como o Saeb, caracteriza-se, primordialmente, por seu objetivo de promover um diagnóstico e propiciar um monitoramento da qualidade da educação básica. O fato de ser amostral é fundamental para compreender a categorização do Saeb como avaliação de primeira geração e para reconhecer suas lacunas. Produzindo resultados agregados, sem chegar no nível das escolas, o Saeb apresenta baixo nível de interferência no dia a dia das escolas, impedindo seu uso para o estabelecimento de políticas de responsabilização e, ao mesmo tempo, criando certo distanciamento entre a avaliação em larga escala e as escolas.

De acordo com Bonamino e Sousa (2012), estados e municípios começaram a sentir necessidade de implantar avaliações que chegassem a todas as escolas. Esse é um ponto particularmente importante para a presente tese. O surgimento de sistemas próprios de avaliação, estaduais e municipais, respondeu, inicialmente, a uma demanda dos entes federados por sistemas de avaliações que se comunicassem mais diretamente com as escolas, permitindo um vínculo maior entre as avaliações e a promoção de ações e políticas educacionais com efeitos de mais amplo alcance. As autoras (idem, 2012, p. 377) apontam que as demandas regionais dos estados em relação a isso estão ligadas ao surgimento dos primeiros sistemas de avaliação, ainda no início dos anos 1990 (como os casos de Minas Gerais e Ceará). Freitas (2013c) também reconhece que os limites do desenho amostral do Saeb, que não dialogava diretamente com as escolas, foi um dos motivos pelos quais os sistemas próprios de avaliação começaram a se desenvolver e a se expandir no Brasil. A Prova Brasil, segundo ela, foi uma espécie de reação do governo federal a essa perda de espaço no âmbito da avaliação.

No caso da segunda geração, a responsabilização passa a ser atrelada ao diagnóstico, que deixa de ser o objetivo último da avaliação, passando a se configurar como uma etapa para o objetivo real. O pressuposto dessa geração é o de que a avaliação pode funcionar como um elemento de pressão, para os pais, responsáveis e pela sociedade em geral, pela melhoria da qualidade da educação. Tal pressuposto remonta ao fundamento político da avaliação em larga

escala, conforme apresentado em tópico antecedente. Por ser censitária<sup>64</sup>, a Prova Brasil permitiu a devolução dos resultados por escola, levando a uma responsabilização branda (*low-stakes*): publicização dos resultados por escola, estratégias de divulgação na mídia (gerando ranqueamentos, alvos dos críticos das avaliações) e distribuição dos materiais para as escolas (incluindo o conteúdo das matrizes de referência e exemplos de itens que fizeram parte dos testes) (BONAMINO; SOUSA, 2012, p. 380). Como já afirmado, a Prova Brasil levou à criação do Ideb e ao estabelecimento de metas educacionais, vinculando as escolas a seu cumprimento.

Se o Brasil já possuía um sistema nacional de avaliação desde o início dos anos 1990, aprimorado, em 2005, a partir das necessidades dos entes federados, como explicar o surgimento de inúmeros sistemas próprios de avaliação a partir de 2007, em especial? Machado, Alavarse e Arcas (2015, p. 669) afirmam que os estados brasileiros foram implantando seus sistemas de avaliação gradativamente, mas é possível notar uma explosão desse movimento depois da criação da Prova Brasil. Até 2005, o número de sistemas próprios de avaliação era modesto. A partir de então, o número cresceu consideravelmente. Como explicar esse movimento?

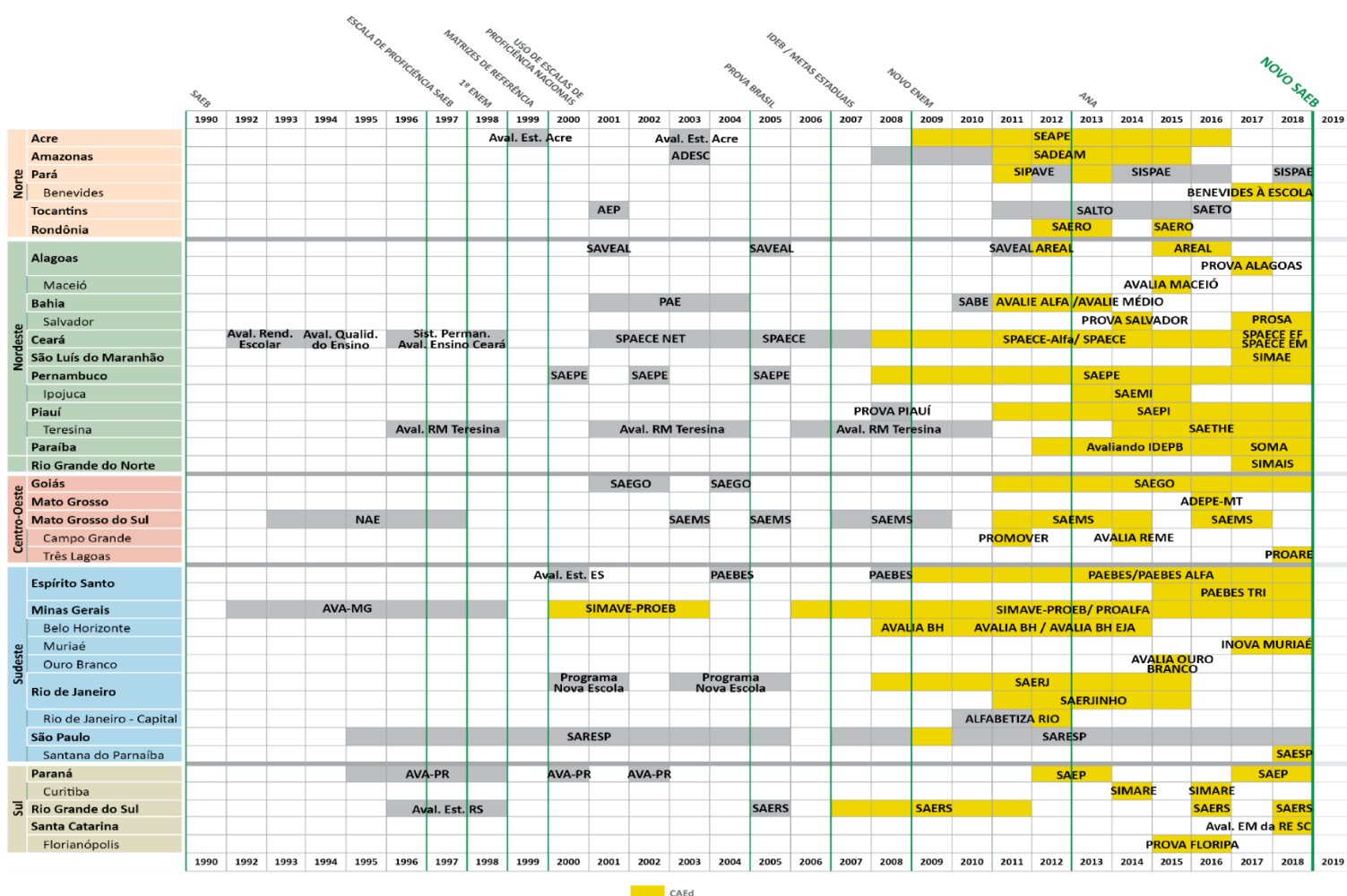
A figura 1 mostra, de modo sintético, os sistemas próprios de avaliação da educação desenvolvidos no Brasil desde a criação do Saeb, em 1990. Até 2005, é possível perceber que os sistemas próprios de avaliação já existiam, mas muitos já haviam sido encerrados, como é o caso do Saego, do Saems e do próprio Paebes, objeto de estudo deste trabalho. O Simave-Proeb, em Minas Gerais, também havia sido descontinuado. Depois da Prova Brasil, os sistemas estaduais se proliferaram, criando uma identidade, até o momento em que esta análise está sendo feita, baseada na continuidade.

Esse movimento nos leva a pensar sobre a capacidade da Prova Brasil de atender às demandas dos entes federados no que diz respeito à avaliação. Mesmo estabelecendo um formato censitário para avaliação nacional, os estados continuaram desenvolvendo sistemas próprios. Como hipótese, dois elementos podem nos ajudar a compreender esse movimento de expansão: i) a periodicidade e ii) a especificidade.

---

<sup>64</sup> O desenho da Prova Brasil: caráter censitário, aplicada aos alunos de 5º e 9º anos do ensino fundamental, a cada dois anos, para avaliar habilidades de Língua Portuguesa e Matemática, em escolas rurais e urbanas com, pelo menos, 20 alunos por turma avaliada.

Figura 1 - Mapeamento dos sistemas próprios de avaliação educacional no Brasil – 1990-2018.



Fonte: CAEd/UFJF, 2018.

A Prova Brasil é aplicada a cada dois anos, de modo que, considerando o tamanho de sua base de dados para o recolhimento do material e seu processamento, a produção da medida e a divulgação dos resultados, eles só são divulgados quase ao fim do ano seguinte ao da aplicação. Recebendo os resultados no fim do ano, como utilizá-los de modo a produzir alguma forma de intervenção pedagógica nas turmas a partir desse diagnóstico? Essa distância entre a aplicação e o resultado compromete uma das principais características da avaliação, qual seja, seu uso como suporte para a tomada de decisões no âmbito do planejamento pedagógico. A defesa do uso dos resultados como um elemento central para a própria implementação da política de avaliação e para a produção dos efeitos esperados passa pela chegada das informações aos seus destinatários finais, o atores escolares. Com resultados a cada dois anos, mesmo divulgados no nível da escola, a avaliação nesse modelo apresenta-se muito mais como

um instrumento destinado aos atores que ocupam posições nas secretarias de educação, portanto, em um nível macroeducacional, do que aos atores que estão na escola.

Os sistemas estaduais de avaliação, em regra, possuem um desenho censitário, como no caso da Prova Brasil, mas sua periodicidade é anual. Dessa forma, os resultados também são divulgados anualmente, em geral, no primeiro trimestre do ano. Isso altera perspectiva de uso dos resultados, visto que eles chegam às escolas e aos atores escolares a tempo de permitirem uma análise para fins de planejamento do ano escolar que ainda se inicia. Os diagnósticos do fim do ano anterior passam a poder dar base para ações do ano seguinte, fazendo valer a premissa de que o uso central dos resultados das avaliações está ligado justamente a esse suporte para a tomada de decisões no âmbito do planejamento escolar. Para os tomadores de decisão, tanto no âmbito das secretarias quanto no âmbito das escolas, essa é uma grande vantagem em relação ao sistema nacional. Aqui, há uma aproximação entre avaliação em larga escala e escola, no sentido de reconhecer que aquela pode ser um instrumento útil para esta.

Além disso, a Prova Brasil avalia anos (5º e 9º anos do ensino fundamental) e disciplinas específicas (Língua Portuguesa e Matemática), aplicando questionários contextuais para o levantamento de informações que possam ajudar a explicar os resultados. Os estados apresentam uma preocupação, de base constitucional, com o ensino médio, etapa pela qual estão responsáveis com exclusividade. Os municípios, por sua vez, concentram seus esforços nos iniciais, com especial ênfase na alfabetização. Contudo, nem o ensino médio nem o ciclo de alfabetização<sup>65</sup> são avaliados pelo sistema nacional, que não atende, portanto, às áreas prioritárias dos entes federados. Esse é um ponto importante para explicar a criação de sistemas próprios de avaliação, que, em geral, avaliam o ciclo de alfabetização e o ensino médio. A entrada do ensino médio no desenho destina-se a lidar com as preocupações dos estados em relação à qualidade da educação da etapa. Em relação à alfabetização, os estados avaliam o ciclo em virtude do pacto federativo com os municípios: os alunos das redes municipais serão, em algum momento, alunos das escolas estaduais e interessa aos estados acompanhar como esses alunos chegam às suas redes. É interessante notar aqui uma característica da política de avaliação no Brasil, diferente de outros contextos, como o norteamericano: a renovação do pacto federativo via avaliação educacional ocorre, precipuamente, no que diz respeito ao ciclo de alfabetização, responsabilidade dos municípios, mas avaliado por políticas estaduais. Isso

---

<sup>65</sup> Conforme vimos em tópico anterior, a avaliação nacional não se debruçava sobre o ciclo de alfabetização. A primeira iniciativa nacional neste sentido foi a Provinha Brasil, em 2008. Contudo, suas limitações como política de avaliação eram claras. Apenas em 2013, foi criada a ANA, com última edição em 2016.

mostra que a política de avaliação não é mera cópia irrefletida de parâmetros internacionais, mas ajusta-se e acomoda-se de acordo com as características locais.

Assim, desejando avaliar etapas distintas daquelas contempladas pelo sistema nacional, os entes federados, principalmente os estados, passaram a desenvolver sistemas próprios de avaliação capazes de atender às suas demandas em relação a etapas prioritárias. O mesmo pode ser dito em relação às disciplinas avaliadas. A concentração da avaliação em Língua Portuguesa e Matemática permaneceu como característica de boa parte dos sistemas próprios, mas alguns passaram a avaliar outras áreas do conhecimento, como Ciências Humanas e Ciências da Natureza, no ensino médio. Amazonas, Bahia, Ceará e Espírito Santo são exemplos de estados com sistemas que expandiram a avaliação para outras disciplinas. Ainda, os questionários contextuais aplicados pelos sistemas próprios permitem investigar ações, políticas e aspectos locais, pontuais, bem como sua associação com o desempenho dos alunos, algo que não é contemplado pelos questionários aplicados pelo sistema nacional. A esse conjunto de elementos (etapas, disciplinas e atenção a ações locais) estamos dando o nome de especificidade.

Os sistemas próprios trouxeram outra inovação: a inauguração, no cenário nacional, de políticas de avaliação de terceira geração, retomando a distinção proposta por Bonamino e Sousa (2012). As autoras tomam o Saresp (de São Paulo) e o Saepe (de Pernambuco) como exemplos, mas outros estados também caracterizam-se, ou já puderam ser caracterizados, como detentores de políticas de terceira geração). As políticas avaliativas de terceira geração atrelam aos resultados da avaliação políticas de responsabilização forte (*high-stakes*<sup>66</sup>). As consequências, neste caso, são mais sérias do que aquelas observadas nas políticas de segunda geração, ultrapassando o plano do simbólico e estabelecendo sanções ou recompensas a partir dos resultados (idem, 2012, p. 375). Em Pernambuco, por exemplo, foi criado o Idepe – Índice de Desenvolvimento da Educação Básica de Pernambuco, nos moldes do Ideb, estabelecendo metas por escola e premiando, através de um bônus (BDE), as escolas que alcançaram as metas.

A distinção das autoras (idem, 2012) no que tange às gerações de avaliação serve a propósitos analíticos e não obedece a uma espécie de evolucionismo, nem mesmo temporal. As gerações não são necessariamente cumulativas, de modo que uma política de terceira geração tenha que ter passado pelas duas gerações anteriores. Nesse sentido, a política de avaliação já pode ter nascido de terceira geração (ou de segunda). Da mesma forma, a evolução não é certa: uma política de primeira geração não precisa se transformar em uma política de segunda ou terceira. Além disso, mudanças de cenário podem fazer com que o modelo de política mude:

---

<sup>66</sup> Para mais sobre a distinção entre políticas de responsabilização *low-stakes* e *high-stakes*, ver Brooke, 2008.

uma política de terceira geração pode tornar-se de segunda geração, na medida em que retire de cena consequências fortes em relação aos resultados da avaliação. De todo modo, essa proposta de categorização é muito útil para distinguir as políticas de avaliação de acordo com as consequências de seus resultados e será base para uma das classificações propostas para as políticas de avaliação da alfabetização analisada no capítulo 4 desta tese.

Antes de passar à análise dos resultados e do papel que desempenham na política de avaliação, uma palavra sobre os sistemas próprios no contexto brasileiro. É possível argumentar que o desenvolvimento e a consolidação de sistemas estaduais entre nós não obedece, pura e simplesmente, a uma lógica imposta por organismos internacionais ou mesmo por um governo federal com vocação centralizadora. Como reconhecem muitos autores (entre eles, MACHADO; ALAVARSE; ARCAS, 2015; BONAMINO; BESSA; FRANCO, 2004; FREITAS, 2013c), a avaliação em larga escala pertence a uma agenda educacional mundial, com caráter universalista, e o Saeb, no caso brasileiro, representa uma referência incontornável para os sistemas estaduais (SOUSA; OLIVEIRA, 2010) (não seria exagero dizer que muitos estados mantiveram e mantêm sistemas próprios como uma espécie de termômetro para o resultados do Ideb), que se manifesta no desenho da avaliação, nos elementos comuns das matrizes, no uso da mesma metodologia, na existência de itens comuns etc. Contudo, não se pode negar as especificidades dos sistemas próprios, que surgem como uma forma de lidar com as vicissitudes identificadas pelos entes federados no sistema nacional. Esse movimento levou a própria União a se movimentar, de modo a aprimorar o seu sistema (a criação da Prova Brasil é um bom exemplo disso). Assim, os entes federados não podem ser pensados como meros reprodutores de políticas nacionais e mesmo internacionais no âmbito da avaliação educacional. A relação é muito menos unilateral do que esse tipo de perspectiva tende a apresentar. Da mesma forma em que são influenciados pelas agendas nacional e internacional, os estados influenciam o sistema nacional. É interessante observar que os dois planos permanecem existindo simultaneamente: a existência de um sistema nacional não impediu a criação e o desenvolvimento de sistemas próprios, ao passo que surgimento destes não levou à destruição daqueles. Isso mostra que as relações entre os dois níveis dos sistemas de avaliação não são antagônicas, mas, em grande medida, complementares. O papel dos estados, portanto, não pode ser reduzido ao de mero reprodutor. No âmbito da avaliação da alfabetização, por exemplo, eles reativam o pacto federativo avaliando redes municipais através de um programa estadual de avaliação, com caráter nitidamente colaborativo. Essa é uma idiosincrasia do cenário da



avaliação no Brasil, que está longe de ser mero refém, irreflexivo e passivo, de políticas internacionais.

#### **3.3.4. A importância dos resultados**

A expansão e a consolidação da avaliação educacional em larga escala como política pública, conforme vimos anteriormente, fizeram com que não apenas essa política fosse lançada no centro do debate educacional, assumindo protagonismo em governos com orientações ideológicas distintas, em especial, devido à sua função de subsídio à tomada qualificada de decisões, mas também colocaram em evidência uma etapa específica da avaliação educacional, a partir da qual a política passou a ser conhecida, qual seja, os seus resultados, particularmente, no que diz respeito a seus usos pelos atores educacionais.

A avaliação educacional é política pública formada por um conjunto complexo e diversificado de etapas, envolvendo profissionais de diferentes áreas do conhecimento e, exigindo, de acordo com o que já foi ressaltado em tópico anterior, a participação de atores situados em diferentes posições da organização educacional. De modo simplificado, podemos enumerar essas etapas: i) a elaboração da política, envolvendo, em regra, as secretarias de educação (e suas equipes internas, cujo número e complexidade variam de acordo com a complexidade da própria estrutura da secretaria) e as instituições parceiras, responsáveis pelo aparato técnico da avaliação; nesta etapa, uma série de decisões importantes é tomada, da definição das séries a serem avaliadas à definição do tipo de instrumento e das habilidades e competências; ii) a elaboração das matrizes de referência (momento crucial para a definição do que, de fato, será avaliado), dos itens e dos testes); iii) a logística de aplicação, envolvendo impressão (com os devidos cuidados em relação ao sigilo), transporte, treinamento de aplicadores e a aplicação propriamente dita; iv) o processamento dos dados; v) a produção da medida; vi) a divulgação dos resultados; e vii) a apropriação dos resultados pelos atores educacionais.

Cada uma dessas etapas, aqui expostas de modo sintético, poderia ser dividida num amplo conjunto de subetapas. De maneira geral, poderia se pensar que a divulgação dos resultados, obedecendo ao princípio da publicidade que rege a política de avaliação (pelo menos, na maioria dos casos<sup>67</sup>), é a última etapa da avaliação, quando seu ciclo como política

---

<sup>67</sup> A maior parte dos sistemas de avaliação, incluindo o sistema federal, opera com base em uma divulgação aberta dos resultados da rede, para consulta pública. No que diz respeito aos resultados das escolas, tomadas individualmente, e das turmas e alunos, os resultados, em regra, não são abertos ao público (nos sistemas

se encerra. De fato, é a última etapa da política que permanece sobre o controle estrito de seus idealizadores e executores. Por essa perspectiva, o uso dos resultados divulgados pertenceria a um momento *pós-política*, em que as consequências da política escapam ao controle de seus idealizadores, ficando a cargo da escola, exercendo o princípio de sua autonomia, o uso ou não desses resultados e a maneira como serão usados. Entretanto, seguindo argumento já apresentando anteriormente, o uso dos resultados pelas escolas é fator decisivo para a política, não apenas para seu sucesso, mas para a própria ideia de uma política de avaliação. Poderíamos dizer que o uso dos resultados faz parte da *natureza* da política de avaliação. Sem isso, a avaliação careceria de sentido como política pública. E o principal uso do resultado é servir como informação de base para a tomada de decisões, em qualquer âmbito educacional, conferindo às decisões, desta forma, uma legitimidade da qual elas costumam carecer.

Os resultados, além de essenciais para a política, são também os mais visíveis. Não seria exagero dizer que os resultados são a *cara da política de avaliação*. Em muitos casos, eles não são apenas a etapa mais visível, mas a única. Isso leva a duas consequências, pelo menos: i) de um lado, coloca em cena a importância dos resultados e de seu uso, que, conforme já mencionado, é um aspecto fundamental da política; ii) de outro, costuma obscurecer a complexidade da política de avaliação, frequentemente reduzida a seus resultados, sem a observância da importância das demais etapas. Reconhecer a importância dos resultados não significa obliterar as demais etapas da política de avaliação. Para o pesquisador da área, a compreensão mais profunda dessa política passa pela análise de todas as suas etapas, a despeito da centralidade dos resultados. Esse é o argumento que dá base à seleção dos aspectos a serem analisados em relação aos três sistemas de avaliação objetos de estudo desta tese, como ficará claro no capítulo 4.

O uso dos resultados é tão central para a política de avaliação que tornou-se campo de pesquisa. No âmbito da pesquisa em avaliação, ele tem se tornado, conforme argumento apresentado por Calderón (2017), um objeto privilegiado de estudo. A transformação do uso dos resultados da avaliação em tema de pesquisa vem a cabo da transformação da própria avaliação em objeto de estudo. Como será visto no capítulo 3, a seguir, os programas de pós-graduação em educação, em especial aqueles que concedem à avaliação um espaço privilegiado, têm observado o aumento do número de trabalhos dedicados ao tema do uso dos resultados (o

---

próprios, os resultados das escolas também são, em geral, abertos): apenas os atores escolares pertencentes a cada uma das unidades têm acesso a esses resultados. De todo modo, os resultados são sempre divulgados. A escola sempre recebe os resultados de suas turmas e alunos.

PPGP CAEd/UFJF é um exemplo claro desse movimento). Em relação a isso, Silva e Carvalho reconhecem:

(...) a relevância dessa agenda de pesquisa que focaliza a influência e a apropriação dos resultados desse tipo de avaliação em escolas e sistemas de ensino público brasileiros como objeto de pesquisa que merece atenção de pesquisadores e profissionais da área com diferentes abordagens de investigação e possibilidades de aportar conhecimento sobre o tema. (SILVA; CARVALHO, 2014, p.12)

Conforme já vimos em tópico precedente, Bauer, Alavarse e Oliveira (2015) fazem um mapeamento profundo das tensões que caracterizam a literatura sobre o uso dos resultados da avaliação educacional. A centralidade do uso dos resultados para a política de avaliação fica em evidência também na tipologia proposta por Bonamino e Sousa (2012), a qual já nos referimos anteriormente. As gerações de avaliação são definidas justamente com base no uso dos resultados e na forma como os mesmos impactam na realidade das escolas, em especial, nos currículos. Uma das maiores polêmicas acerca da avaliação gira em torno das políticas de responsabilização (BROOKE, 2008), que é uma dimensão do uso dos resultados.

#### *3.3.4.1 Os resultados como parte da implementação da política de avaliação*

Como vimos no capítulo 1 desta tese, a implementação de uma política pública pelos atores responsáveis por essa ação é parte integrante e fundamental do ciclo de políticas, da qual depende a efetividade dos efeitos pretendidos pela política. No caso da avaliação educacional em larga escala, os atores responsáveis pela implementação são muitos e diversos, ocupando posições distintas na estrutura organizacional da educação pública: do secretário de educação ao professor na escola, passando pelos técnicos da secretaria, pelos gestores escolares, funcionários da escola e entidades de apoio à política (elaboradores de itens e testes, aplicadores, atores envolvidos com a logística de recolhimento, estatísticos, analistas de dados, comunicadores etc) , todos estão envolvidos no processo de implementação e nas chances de sucesso das ações previstas. Os próprios alunos (os destinatários finais dos resultados da política de avaliação) e os responsáveis por eles fazem parte desse processo. Mesmo não sendo implementadores no sentido estrito, sua ação interfere, em maior ou menor grau, dependendo do contexto, nos resultados da política.

Especificamente no âmbito da escola, professores e gestores são atores particularmente importantes para o processo de implementação de uma política de avaliação. A ideia de um

sistema de avaliação externa não pode ser confundida com a ausência de participação e protagonismo desses atores. A externalidade da avaliação, vinculada à sua amplitude (em larga escala), caracteriza-se pela busca de uma avaliação sistêmica, macroeducacional, tendo em vista a produção de informações para toda a rede de ensino (e não apenas para os alunos, tomados individualmente), sendo elaborada por agentes externos para fins de garantir o mínimo de interferência nos resultados e a possibilidade de generalização desses resultados e de suas análises (cada professor elaborando avaliações para sua realidade responde bem à análise dessa realidade, mas não do todo). Isso não significa que os resultados não possam (e devam) servir como instrumento diagnóstico para ações pedagógicas e de gestão no âmbito da escola. O antagonismo entre avaliação externa e avaliação interna não é necessário, mas fruto da forma como essas avaliações são usadas. A complementaridade entre esses instrumentos pode ser o caminho trilhado. O fato é que não estamos diante de instrumentos vinculados a realidades absolutamente distintas, mas que estabelecem relações entre si (SOUZA, 2013). Decisões consequentes no âmbito da rede passam pelo reconhecimento de que as secretarias de educação, isoladamente, não são capazes de alterar o cenário educacional de problemas estruturais e gerais de aprendizagem, conforme diagnóstico das avaliações no Brasil ao longo dos anos, sem a participação das escolas, conhecedoras potenciais, em profundidade, de sua própria realidade. Do mesmo modo, os problemas de cada unidade escolar não podem ser totalmente equacionados pela própria escola, tendo em vista que muitos deles extrapolam seu raio de ação, dizendo respeito à rede como um todo.

Isso quer dizer que a avaliação educacional em larga escala não é uma política pública com implementação apenas por parte do governo. Se, historicamente, a secretaria de educação foi, por excelência, o ponto de partida da política, esta última não pode realizar completamente a implementação sem a participação dos atores escolares. A avaliação educacional é uma política particularmente complexa, conforme já ressaltado anteriormente. Sua natureza etapista, multifacetada e multidisciplinar exige a participação de muitos atores.

No que diz respeito aos professores e gestores escolares, particularmente, sua ação diante das avaliações é o que permite que essa política seja implementada e possa ver seus resultados aparecerem. Para tanto, é preciso reconhecer que: i) a política de avaliação não encerra sua ação com a entrega e a divulgação dos resultados; e ii) a atuação dos atores escolares é o que permitirá à política atingir seus objetivos. A entrega dos resultados não é o fim da política, mas o início de seu maior desafio. A partir da divulgação dos resultados para a escola, toda uma engrenagem precisará funcionar para que a avaliação exerça, de fato, sua vocação

para a mudança na escola. A chegada da informação, pura e simplesmente, é apenas uma etapa, instrumental, de seu verdadeiro objetivo: fazer com que essas informações, qualificadas por métodos mais acurados de mensuração da aprendizagem, sejam o ponto de partida para um processo de reflexão da escola sobre suas próprias dificuldades, seus desafios e suas virtudes, servindo como base para a alteração ou o reforço das práticas pedagógicas e de gestão existentes. Como afirma Souza:

Hoje, e sempre, a grande preocupação quando se fala em avaliação de escola/sistema/redes de ensino é não somente fazer chegar até a escola os resultados em termos estatísticos-quantitativos, mas como fazer para que toda a escola entenda esses resultados e possa fazer uso deles, transformando-os em informações importantes para melhorar a qualidade do seu trabalho (SOUZA, 2013, p.163).

A avaliação externa encarna, dessa forma, um objetivo de todo processo avaliativo: ser a chave de um processo de autoavaliação. Neste ponto, a externalidade da avaliação aparece somente como uma característica metodológica, de forma, mas não substancial. Em substância, como objetivo, a avaliação em larga é também autoavaliação. Transformar a avaliação externa em instrumento de planejamento da escola, a partir de sua própria reflexão, faz com que as fronteiras entre o externo e o interno fiquem embaçadas.

Sendo assim, fazer uso dos resultados da avaliação é parte da política pública, etapa de sua implementação. A ideia é a de que a implementação não se encerra com a aplicação dos testes e com a divulgação dos resultados. Isso seria reduzir em demasia o processo avaliativo, o que, em última instância, colabora com as críticas que a avaliação tem recebido (incluindo a de que a avaliação externa e em larga escala é antagônica à avaliação interna). Se a apropriação dos resultados é parte da implementação da política, é necessário que os atores escolares tenham consciência disso, além de compreenderem seu papel no processo, as características e limitações dos instrumentos e da política e a forma como podem contribuir para impulsioná-la.

Contudo, parte considerável desses atores desconhece a política e, conseqüentemente, o papel que devem nela desempenhar. Isso se aplica, particularmente, aos resultados da avaliação. De maneira geral, podemos dizer que há, pelo menos, três grandes obstáculos técnico-interpretativos para o uso dos resultados da avaliação: i) conhecer os resultados, no sentido de sua publicidade, de ter acesso a eles, de saber que determinada informação sobre a escola existe; ii) conhecendo os resultados, saber interpretá-los, de modo a evitar equívocos sobre a amplitude e natureza da informação, o que pode levar a conclusões não autorizadas pelos resultados; e iii) mesmo conhecendo e compreendendo o que os resultados querem dizer, há ainda o desafio de

seu uso, de colocar em prática, através de ações refletida e planejadas, as consequências autorizadas pelos resultados.

A superação desses obstáculos pode ser pensada de forma etapista e uma analogia com a ideia de gerações não está descartada (não perfeitamente correspondentes às gerações de Bonamino e Sousa, 2012, mas em relação à ideia de geração, de fases distintas). Inicialmente, o maior desafio para a avaliação, no que diz respeito aos resultados, era fazer conhecê-los. Ainda hoje, esse desafio não está completamente superado. As redes de ensino são renovadas constantemente com a entrada de novos profissionais, oriundos dos cursos de Pedagogia e das Licenciaturas, que não tiveram, em sua formação inicial, acesso adequado às discussões que envolvem mais profundamente a avaliação educacional em larga escala. Secretarias de educação e o próprio MEC, através do Inep, permanecem enfrentando o desconhecimento dos atores escolares em relação à existência de dados que dizem respeito à escola e podem ser úteis a ela (não somente os produzidos pelas avaliações educacionais, mas oriundos de outras fontes, como o Censo Escolar, por exemplo). Hoje, as redes de ensino parecem estar mais informadas sobre a existência de dados que a ela dizem respeito do que no início dos anos 1990, o que se vê pelo uso mais constante desses dados para a tomada de decisão e pela existência, em boa parte das secretarias, de equipes encarregadas especificamente da avaliação educacional. Contudo, uma boa parte do vasto manancial de dados educacionais permanece desconhecida, inclusive pelas equipes das secretarias. No caso das escolas, essa ausência torna-se ainda mais evidente.

A interpretação dos dados apresenta-se como um problema de outra natureza, mais qualitativa. Os dados educacionais, incluindo os da avaliação em larga escala não são autoevidentes e óbvios. Isso significa que eles precisam ser compreendidos, em seus limites e em sua capacidade de explicação. Como qualquer informação científica, os resultados da avaliação são limitados e passíveis de diferentes interpretações. Eles não são dogmáticos. Uma das grandes dificuldades desse processo de interpretação por parte dos atores educacionais está depositada nas conclusões autorizadas pelos dados. Não é raro encontrar casos em que atores chegam a conclusões que extrapolam aquilo que os dados oferecerem. Outro risco é interpretar menos do que os dados dizem, não levando a interpretação até seu esgotamento. Tudo isso, evidentemente, prejudica essa etapa crucial da política de avaliação, levando a usos equivocados dos resultados.

Em relação ao uso, dois problemas principais: i) o uso equivocado; e ii) o não uso. O uso equivocado ocorre, principalmente, de uma interpretação equivocada dos dados, conforme dito no parágrafo anterior. Um exemplo disso é a não compreensão em torno da natureza da

medida produzida pela TRI, comparando-a com uma nota de natureza linear. Assim, olhando para uma escala de 0 a 1000, o intérprete pode acreditar que uma proficiência acima de 500 é um bom indício, por estar acima da média. Problema mais grave, e mais comum, é o do não uso. Mesmo quando a interpretação dos dados é correta, resta a questão, nem um pouco trivial, acerca de como esses dados podem ser usados pela escola, de forma a transformá-los de informação estatística em ações práticas. Somente com o uso dos resultados é que a política de avaliação cumpre seu ciclo de implementação, podendo gerar os efeitos pretendidos sobre a mudança no planejamento e sobre a aprendizagem dos alunos.

Para lidar com essa tríade de questões, os sistemas de avaliação, em especial os próprios, têm utilizado, como etapa importante para a efetivação da política, capacitações voltadas aos atores escolares e educacionais (os técnicos das secretarias) justamente para lidar com problemas de conhecimento, interpretação e uso dos dados educacionais. Essa é uma etapa particularmente importante da devolutiva dos resultados da avaliação, sem a qual, é provável, a chance da escola usá-los como parte do planejamento seria pequena. No contexto norteamericano, por exemplo, o uso de protocolos para estimular e orientar as escolas em como utilizar os dados educacionais tem sido cada vez mais disseminado. Uma obra importante nesse sentido foi escrita por Boudett, City e Murnane (2013), um guia de ações, descritas de forma minuciosa, para orientar diretores e professores em relação ao uso de resultados de avaliação com o fim de aprimorar a gestão e o ensino. As capacitações no âmbito das devolutivas dos resultados das avaliações serão um dos critérios observados na análise dos sistemas próprios de avaliação da alfabetização, no capítulo 4 da tese.

Ainda uma observação sobre os resultados. Podemos afirmar que eles apresentam uma dupla dimensão: i) descritiva e ii) heurística. Do ponto de vista descritivo, o resultados da avaliação oferecem informações que apresentam uma espécie de fotografia das escolas e das redes de ensino, em relação ao desenvolvimento de habilidades em diferentes disciplinas, previstas pelo currículo. O resultados dos testes cognitivos (as médias de proficiência e a distribuição dos alunos pelos padrões de desempenho, em especial) possuem essa função: apresentar um estado de desenvolvimento das habilidades curriculares. Contudo, é preciso reconhecer que essa dimensão descritiva dos resultados, embora tenha a capacidade de colocar em cena os problemas de aprendizagem dos alunos, não é capaz, em si, de *explicar o motivo pelo qual os resultados são aqueles*<sup>68</sup>. Eles entregam um diagnóstico, mas não explicam suas causas. Essa tarefa, heurística, cabe a outro tipo de investigação, conduzida, com especial

---

<sup>68</sup> Sobre os limites da dimensão descritiva do Ideb, por exemplo, ver Alves e Soares (2013).

relevância, através das informações obtidas pelos questionários contextuais. A vasta pesquisa em eficácia escolar (BROOKE; SOARES, 2008) se debruça sobre essa perspectiva heurística. Quais são os fatores associados ao desempenho escolar? Quais deles estão associadas à escola (fatores internos) e quais não estão (fatores externos)? A produção de efeitos educacionais pela escola (ANDRADE; SOARES, 2008) implica o reconhecimento e a superação dos efeitos sociais, econômicos e culturais sobre a aprendizagem dos alunos, e o diagnóstico de desigualdades educacionais. Nesse sentido, há um conjunto de pesquisas: Soares e Alves (2003), sobre a desigualdade racial nas escolas; Basso et al (2012), sobre o mesmo tema; Rezende (2016), sobre a influência do clima escolar, em sua dimensão normativa, sobre o desempenho dos alunos. A heurística é uma dimensão complementar à descritiva e muito relevante da avaliação educacional, embora seja ainda pouco conhecida e explorada pelas redes e escolas.

\*\*\*

A avaliação em larga escala como política pública implica o reconhecimento de que ela pode ser entendida como formada por ciclos ou etapas, tais como a proposta do ciclo de políticas de Stephen Ball. Uma das etapas mais importantes da política de avaliação são seus resultados. Mais do que a divulgação, espera-se que os resultados sejam utilizados pelas redes de ensino e pelas escolas, de modo que a política possa produzir efeitos. Desta forma, podemos argumentar que a política de avaliação não encerra seu ciclo com a divulgação dos resultados, visto que sua implementação é estendida ao uso dos resultados, através da capacitação dos atores educacionais para tanto. Implementar uma política de avaliação implica em capacitar os atores para que os resultados sejam conhecidos, compreendidos e utilizados como parte das ações de escolas e secretarias. Isso concede aos resultados um papel de proeminência no âmbito da política. As principais críticas às avaliações, como aquelas direcionadas às ações de accountability, por exemplo, têm sua razão de ser no uso dos resultados, o que faz com que haja uma confusão entre o uso e o instrumento, levando, muitas vezes, a uma espécie de personalização da avaliação (seja para defendê-la, seja para atacá-la). De fato, os resultados são fundamentais para a política, mas não podemos olvidar que as avaliações são formadas por outras etapas, também importantes para uma compreensão mais profunda de sua natureza, de seus limites e de suas virtudes. Mesmo com a ênfase dada aos resultados, as demais etapas são negligenciadas, conforme será percebido pela análise que nos propomos a fazer no capítulo 5.

Como política, a avaliação é *sui generis*: política de avaliação de políticas. Espera-se que ela seja capaz de dizer o que está dando efeito nas políticas educacionais. Sua centralidade



fica evidenciada pelo seu objetivo: dar suporte a decisões em diferentes níveis. Do sistema nacional ao desenvolvimento dos sistemas próprios, a avaliação mostrou-se contínua, a despeito da instabilidade de outras políticas educacionais. Os sistemas estaduais desempenham um importante papel entre nós, colocando em cena um reforço do pacto federativo em que municípios, estados e União colaboram para fins de implementação da política. No capítulo 4, a seguir, veremos uma parte específica dessa história: como a literatura educacional tem enfrentado o tema dos sistemas estaduais de avaliação da alfabetização, que assumem protagonismo em relação ao sistema nacional, já que possuem uma história mais antiga e sólida do que a proposta federal para avaliar a alfabetização.

#### 4 OS PROGRAMAS DE AVALIAÇÃO DA ALFABETIZAÇÃO NA LITERATURA EDUCACIONAL

Em uma obra de fôlego, Maria do Rosário Longo Mortatti (2011), certamente, um importante nome da pesquisa em alfabetização no Brasil (uma “líder no campo de investigação sobre o tema”, conforme adjetivação de Carlota Boto, na apresentação da referida obra), organizou um substantivo número de trabalhos destinados a contar, conforme indica o próprio título do livro, a história da alfabetização no país. São 16 textos, distribuídos em 3 partes, sobre a história da alfabetização no Brasil, contada a partir da perspectiva selecionada por Mortatti (como ela mesma propõe no título, trata-se de uma história de sua história). Os temas navegam entre homenagens à obra de autores consagrados (como Magda Soares e Rogério Fernandes), a história do ensino de leitura em diferentes cenários, a questão do uso das cartilhas, a constituição dos principais grupos de pesquisa e discussão sobre o tema. Apesar da obra datar de 2011 e termos, desde o ano de 2008, um avaliação nacional da alfabetização, não há nenhum texto, ou mesmo parte de texto, dedicado à avaliação da alfabetização.

Evidentemente, um livro tem suas limitações e a história da alfabetização poderia ser contada sob diferentes perspectivas. Seria uma tarefa hercúlea reunir, em um único material, todas as facetas que o processo de alfabetização pode assumir. De todo modo, a organização de um livro é feita de escolhas, e, mesmo diante do reconhecimento dessas limitações, cabe ainda indagar sobre os motivos pelos quais a avaliação da alfabetização não faz parte de uma obra que conta sua história. Longe de transformar essa indagação em uma crítica à qualidade do material selecionado por Mortatti, o que está em cena é o reconhecimento de que a avaliação é uma dimensão negligenciada no âmbito da alfabetização, em particular, quando estamos falando de avaliação externa. A Provinha Brasil é mencionada em um dos capítulos<sup>69</sup> que compõem o livro, escrito por Iole Maria Faviero Trindade (2011). O capítulo tem como foco a narrativa acerca da constituição e do desenvolvimento das ações de um grupo de pesquisa sobre alfabetização na UFRGS, o Alfa NECCSO (Núcleo de Estudos sobre Currículo, Cultura e Sociedade), e a Provinha Brasil apenas é mencionada em uma passagem do texto, quando a autora cita o trabalho de Darlize Teixeira de Mello, uma tese de doutorado, defendida na UFRGS, em 2009, sobre como a Provinha Brasil foi recebida pela rede municipal de ensino de Porto Alegre.

---

<sup>69</sup> Intitulado Caminhos e descaminhos na área da alfabetização.

O livro de Mortatti é de 2011. À época, além da Provinha Brasil, alguns sistemas estaduais de avaliação da alfabetização já contavam com várias edições (incluindo Proalfa, Paebes Alfa e Spaece Alfa). O caso do Ceará, em específico, com o acréscimo do Paic, era conhecido em todo o Brasil, em virtude das melhorias nos resultados que o estado vinha apresentando no ciclo de alfabetização. Além disso, os sistemas de avaliação estavam funcionando a pleno vapor, sendo disseminados por todo o Brasil. O sistema nacional havia se consolidado, chamando atenção, entre outras coisas, pela ausência de uma avaliação para a alfabetização, o que já daria um capítulo para o livro. Certamente, tudo isso não passou despercebido pelos estudiosos da alfabetização e, em especial, da história da alfabetização. Sendo assim, a que pode ser atribuída essa patente ausência?

A crescente expansão e o desenvolvimento dos sistemas de avaliação educacional no Brasil, caracterizados, particularmente, por sua continuidade, na contramão de boa parte das políticas educacionais, não foram acompanhados, na mesma proporção, pelo interesse da pesquisa em educação por esse tema. Críticos e defensores, embora diverjam em relação à postura diante da avaliação, à sua importância e a seus impactos, concordam em reconhecer que, para o bem ou para o mal, a avaliação passou a ocupar um lugar central no cenário educacional brasileiro, em especial, por seu uso como suporte para a tomada de decisões. Se, de um lado, há quem veja nisso a real missão da avaliação (informar e diagnosticar com qualidade para permitir o agir consciente, a tomada de decisão), de outro, há quem entenda a avaliação como um problema para a educação, por colocar em cena uma lógica perversa e contrária àquela que deveria reger a escola (a imposição da meritocracia via instrumento avaliativo). De um modo ou de outro, não escapa aos olhos de qualquer observador a centralidade que a avaliação assumiu no Brasil nos últimos 30 anos.

Se tem sido parte inarredável da política educacional, como vimos em tópico anterior, tanto no nível macroeducacional, das secretarias de educação e do próprio MEC, quanto no nível microeducacional, das escolas, era de se esperar que a pesquisa em educação passasse a se interessar, de forma crescente, por esse fenômeno, seja para tentar compreender suas potencialidades, seja para criticar suas características e seu uso.

A alfabetização, como fenômeno, interessa a muitas áreas do conhecimento, da pedagogia à psicologia, passando pela linguística, pelas políticas públicas e pela matemática. Sua importância, ao contrário do processo experimentado pela avaliação educacional, em particular, do tipo em larga escala, não é algo recente. Contudo, é possível afirmar que, também nos últimos 20 anos, a alfabetização foi assumindo papel cada vez mais relevante, a ponto de

se tornar um dos principais pilares da política educacional nacional recente, a despeito dos acertos ou equívocos com que essa política é conduzida.

Esse ganho de relevância e, mais do que isso, essa emergência experimentada pela alfabetização podem ser lidos através do desenvolvimento e consolidação da avaliação educacional em larga escala. Não se trata de atribuir à avaliação a responsabilidade por ter alçado a alfabetização ao posto de prioridade da educação pública brasileira, visto que esse fenômeno é multicausal e pode ser explicado com base na observância de muitos fatores, mas de reconhecer as relações entre a expansão da avaliação em larga escala e a emergência do tema da alfabetização. Os resultados das avaliações educacionais (em especial os do Saeb), diagnosticando problemas de aprendizagem nos alunos das redes públicas, dispararam um alerta para o MEC e as redes de ensino no Brasil. Por meio de uma análise pedagógica mais atenta dos resultados, foi possível inferir que alunos nos anos finais do ensino fundamental, e mesmo no ensino médio, apresentavam problemas de aprendizagem oriundos do ciclo de alfabetização. Habilidades relacionadas ao domínio e à utilização da língua portuguesa e relacionadas às operações matemáticas básicas, previstas para o ciclo de alfabetização, não haviam sido plenamente desenvolvidas pelos alunos do ensino médio, por exemplo, prejudicando, e mesmo impedindo, o desenvolvimento das habilidades previstas para esta etapa de escolaridade.

Nesse sentido, é possível afirmar que os resultados dos sistemas de avaliação lançaram luz sobre os problemas de aprendizagem experimentados no âmbito das redes públicas de ensino, evidenciando, em escala nacional, um cenário já conhecido pelos professores e diretores em âmbitos locais. Tais problemas de aprendizagem puderam ser observados ao longo de todas as etapas de escolaridade, mas a ausência de desenvolvimento de habilidades relacionadas à alfabetização em alunos dos anos finais do ensino fundamental e do ensino médio encarnaram a maximização do problema. A preocupação com o ciclo de alfabetização foi intensificada desde então (referência temporal entre meados e fins da década de 2000, anos de 2006, 2007).

Os diagnósticos da avaliação e a preocupação com a alfabetização influenciaram não apenas políticas públicas destinadas ao ciclo, mas também sistemas próprios destinados à alfabetização. No que diz respeito à influência dos resultados da avaliação impulsionando o surgimento de novas políticas públicas para a alfabetização, incluindo um sistema de avaliação próprio para o ciclo, o caso do Ceará, embora não seja o único, pode ser considerado o mais emblemático. De acordo com o entendimento de Lima e colaboradores (2008), a não alfabetização das crianças no início da escolarização ajuda a explicar o fracasso escolar exposto

pelos resultados da avaliação nacional (e os do próprio estado, através do Spaece). Esse foi o principal motor para a criação de políticas voltadas à alfabetização no estado cearense, em especial, o Paic e o Spaece Alfa.

Outros sistemas de avaliação da alfabetização já haviam surgido antes do Spaece Alfa, como é o caso do Proalfa, e outros vieram em sua esteira, como o Paebes Alfa. A combinação entre avaliação e alfabetização, num contexto em que esses dois temas tonaram-se essenciais para o debate sobre educação no Brasil, deveria ter produzido um conjunto bastante significativo de trabalhos e pesquisas. Os sistemas próprios de avaliação destinados à alfabetização, conforme apresentado em tópico anterior, produzem informações anuais sobre as redes públicas estaduais e municipais no Brasil e deram base a inúmeras políticas públicas. São fontes fundamentais de informação para a tomada de decisão no âmbito dos entes federados.

Contudo, o que é possível notar a partir da análise de algumas das mais importantes publicações e canais de divulgação da pesquisa acadêmica no Brasil é a parca produção sobre a avaliação da alfabetização. Se, por um lado, os programas próprios de avaliação da alfabetização desempenham um papel central na condução das políticas educacionais no âmbito de estados e municípios, por outro, eles permanecem, de certa forma, invisíveis para a pesquisa em educação.

Com o intuito de investigar a produção acadêmica especializada sobre o tema da avaliação da alfabetização, foram analisadas as publicações de três importantes revistas educacionais brasileiras (Estudos em Avaliação Educacional, Revista Brasileira de Educação e Revista Brasileira de Estudos Pedagógicos), além do banco de teses e dissertações da Capes. A cobertura da análise das revistas foi de 2004 a 2017, período superior àquele selecionado para a análise, no âmbito desta tese, dos três programas de avaliação da alfabetização (2009 a 2017). A escolha por iniciar a análise em 2004 deveu-se, exclusivamente, à realização da primeira edição do Proalfa, em 2005. Se alguma discussão que precedia o Proalfa existia em nível significativo, a esperança era captá-la na literatura, pelo menos, um ano antes do início do programa. Em relação ao banco da Capes, as pesquisas realizadas não estabeleceram nenhum filtro temporal.

A seguir, são apresentadas, em tópicos distintos, as análises da produção sobre avaliação da alfabetização em cada um desses periódicos e no banco da Capes. Antes, contudo, é necessário fazer uma observação que diz respeito aos critérios para a realização da análise e para a contagem dos trabalhos (em relação aos três periódicos). A palavra *artigo*, nas tabelas,

possui, na verdade, um significado mais amplo. Ela faz referência não apenas a artigos no sentido convencional do termo, mas envolve produções como resenhas, discursos publicados, ensaios, entrevistas, resumos, relatos de pesquisa, relatos de experiência, notas de leitura, depoimentos e publicação de documentos. Esses tipos de publicação faziam parte dos periódicos e foram considerados para a análise. Apesar da variedade de tipos de publicação, a maior parte dos trabalhos veiculados pelos três periódicos, para o período analisado, foi mesmo de artigos. Para a contabilização do trabalho como parte do grupo “artigos sobre avaliação da alfabetização”, conforme mostra a última coluna das tabelas oriundas da análise dos periódicos, ele precisaria ter: i) *avaliação e/ou alfabetização* como palavras componentes do título (ou palavras que dão nomes a programas de avaliação, tais como Provinha Brasil, ANA, Spaece Alfa, Proalfa, Paebes Alfa etc); ou ii) *avaliação e/ou alfabetização* como palavras-chave (o raciocínio quanto aos programas de avaliação aplica-se aqui também, da mesma forma como no título); ou iii) no resumo, um conteúdo diretamente vinculado ao tema da avaliação da alfabetização. Essa mesma lógica foi aplicada para o grupo “artigos sobre alfabetização” (contudo, apenas com a palavra *alfabetização* como referência), conforme mostra a penúltima coluna das tabelas. Os artigos sobre alfabetização só eram contabilizados quando tratavam da alfabetização sem que tratassem, especificamente, da avaliação da alfabetização. Assim, trata-se de um grupo que funciona como gênero, envolvendo a espécie “artigos sobre avaliação da alfabetização”. A contabilização de trabalhos sobre alfabetização (como gênero) foi realizada apenas para efeitos de parâmetro e comparação com a espécie (trabalhos sobre avaliação da alfabetização).

#### 4.1 ESTUDOS EM AVALIAÇÃO EDUCACIONAL

A revista Estudos em Avaliação Educacional (EAE) é um periódico científico da Fundação Carlos Chagas (FCC), com publicação quadrimestral. Fundada em 1990 (substituindo a revista Educação e Seleção, que existiu entre 1980 e 1989), a EAE é um periódico especializado em temas relacionados à avaliação educacional, sendo um dos mais importantes do país nessa temática (A2 em Educação no Qualis Capes), o que justifica sua inclusão como foco da análise da presente tese. A tabela a seguir apresenta o resultado da análise.

Tabela 1 - Mapeamento da EAE – 2004-2017

Estudos em Avaliação Educacional - EAE						
Nº	Vol.	Ano	Meses	Artigos da revista	Artigos sobre alfabetização	Artigos sobre avaliação da alfabetização
29	-	2004	jan-jun	10	0	0
30	15	2004	jul - dez	9	0	0
31	16	2005	jan-jun	8	0	0
32	16	2005	jul - dez	7	0	0
33	17	2006	jan - abr	9	0	0
34	17	2006	mai - ago	9	0	0
35	17	2006	set - dez	8	0	0
<b>36</b>	<b>18</b>	<b>2007</b>	<b>jan - abr</b>	<b>10</b>	<b>0</b>	<b>1</b>
37	18	2007	mai - ago	7	0	0
38	18	2007	set - dez	8	0	0
39	19	2008	jan - abr	8	0	0
40	19	2008	mai - ago	6	0	0
<b>41</b>	<b>19</b>	<b>2008</b>	<b>set - dez</b>	<b>9</b>	<b>0</b>	<b>3</b>
<b>42</b>	<b>20</b>	<b>2009</b>	<b>jan - abr</b>	<b>8</b>	<b>2</b>	<b>0</b>
<b>43</b>	<b>20</b>	<b>2009</b>	<b>mai - ago</b>	<b>11</b>	<b>0</b>	<b>1</b>
<b>44</b>	<b>20</b>	<b>2009</b>	<b>set - dez</b>	<b>7</b>	<b>0</b>	<b>1</b>
<b>45</b>	<b>21</b>	<b>2010</b>	<b>jan - abr</b>	<b>10</b>	<b>0</b>	<b>1</b>
46	21	2010	mai - ago	10	0	0
47	21	2010	set - dez	10	0	0
48	22	2011	jan - abr	9	0	0
49	22	2011	mai - ago	9	0	0
50	22	2011	set - dez	10	0	0
<b>51</b>	<b>23</b>	<b>2012</b>	<b>jan - abr</b>	<b>13</b>	<b>0</b>	<b>1</b>
52	23	2012	mai - ago	14	0	0
<b>53</b>	<b>23</b>	<b>2012</b>	<b>set - dez</b>	<b>12</b>	<b>0</b>	<b>2</b>
<b>54</b>	<b>24</b>	<b>2013</b>	<b>jan - abr</b>	<b>12</b>	<b>0</b>	<b>1</b>
<b>55</b>	<b>24</b>	<b>2013</b>	<b>mai - ago</b>	<b>12</b>	<b>1</b>	<b>0</b>
56	24	2013	set - dez	10	0	0
57	25	2014	jan - abr	11	0	0
58	25	2014	mai - ago	11	0	0
59	25	2014	set - dez	11	0	0
60	25	2014	dez (esp.)	14	0	0
61	26	2015	jan - abr	9	0	0
62	26	2015	mai - ago	10	0	0
63	26	2015	set - dez	10	0	0
64	27	2016	jan - abr	9	0	0
65	27	2016	mai - ago	10	0	0
66	27	2016	set - dez	12	0	0
<b>67</b>	<b>28</b>	<b>2017</b>	<b>jan - abr</b>	<b>10</b>	<b>0</b>	<b>1</b>
68	28	2017	mai - ago	11	0	0
<b>69</b>	<b>28</b>	<b>2017</b>	<b>set - dez</b>	<b>11</b>	<b>1</b>	<b>0</b>
<b>Total</b>				<b>404</b>	<b>4</b>	<b>12</b>

Fonte: Estudos em Avaliação Educacional, 2004-2017. Tabulação própria.

Entre 2004 e 2017, a EAE publicou 41 números, distribuídos em 14 volumes diferentes. No total, foram 404 trabalhos publicados no período, sendo 4 deles destinados a temas diretamente relacionados à alfabetização (o que responde por 0,99% do total) e 12 dedicados à avaliação da alfabetização (2,9% do total). Como é possível notar, a produção sobre a avaliação

da alfabetização é muito pequena quando comparamos com a produção sobre os demais temas relacionados à avaliação da educação.

O artigo de Benigna Maria de Freitas Villas Boas (2007) destina-se a apresentar os resultados de uma pesquisa realizada no Distrito Federal com o intuito de analisar as práticas avaliativas dos professores do ciclo de alfabetização em escolas de Ceilândia. O trabalho versa sobre as avaliações internas, realizadas pelos professores em sala de aula, sem referência à avaliação externa. Como conclusão, Villas Boas (2007) aponta para a necessidade de se pensar melhor as práticas avaliativas diante da implementação do BIA – Bloco Inicial de Alfabetização no Distrito Federal.

A edição da EAE de setembro-dezembro de 2008 (volume 19, número 41) publicou três artigos relacionados ao tema da avaliação da alfabetização (no total, a referida edição publicou 9 artigos). O artigo de Lima et al (2008) é um dos poucos artigos e trabalhos aqui analisados que têm um programa estadual de avaliação da alfabetização como objeto de estudo. Sua publicação relaciona-se, em grande medida, ao frescor da descrição a que se propunha a fazer então: o Spaece Alfa havia sido criado em 2007 e os autores se propuseram a analisar sua primeira edição. Mesmo diante da incipiência do programa e da proposta de apresentação feita pelo artigo, o grande destaque é dado para os resultados dessa primeira experiência do Spaece Alfa, o que mostra como os resultados são, em regra, o elemento da avaliação educacional que mais recebe a atenção dos estudiosos e interessados pelo tema. Somam-se à apresentação dos resultados, as informações de caráter metodológico que deram base à avaliação, bem como as estratégias de divulgação dos resultados utilizadas.

O artigo de Vera Masagão Ribeiro e Tufi Machado Soares (2008) tem como objetivo descrever a construção de uma escala de proficiência para mensurar o analfabetismo de jovens e adultos, com a combinação de habilidades relacionadas à matemática e à leitura em língua portuguesa. A escala foi construída com base nos dados obtidos pelo Inaf (Indicador Nacional de Alfabetismo Funcional) e seu enfoque é técnico e estatístico. O artigo tem como objetivo central apresentar a construção de uma escala sob um ponto de vista metodológico, não dizendo respeito a nenhum programa de avaliação da alfabetização em específico. Na verdade, a escala construída faz referência à alfabetização em contextos não escolares. É de Vera Masagão Ribeiro também, em parceria com Maria da Conceição Ferreira Reis da Fonseca (2010), artigo sobre o desdobramento da escala criada para a Inaf. As autoras apresentam uma matriz de referência para o indicador de alfabetismo funcional, trazendo como novidade o mesmo



elemento que caracterizava a escala, apresentada em trabalho anterior, qual seja, a integração de dois domínios do alfabetismo: o letramento e o numeramento.

O Paic é o objeto de pesquisa do artigo de Cláudio de Albuquerque Marques, Ana Paula de Medeiros Ribeiro e Maira Isabel Filgueiras Lima Ciasca (2008). Os autores apresentam as principais características do programa do Ceará, dando destaque para seu eixo de avaliação e o objetivo de auxiliar os municípios participantes com o desenvolvimento de sistemas municipais de avaliação (o ano de análise foi o de 2007, quando o Spaece Alfa foi criado). A ênfase é dada à autonomia dos municípios no que tange à criação de seus sistemas avaliativos. O artigo não faz referência ao Spaece Alfa e o Spaece é apenas mencionado. O foco é mesmo a autonomia dada aos municípios para a criação de seus sistemas.

Cláudio Albuquerque Marques (2009), junto de dois outros autores (Rui Rodrigues Aguiar e Márcia Oliveira Cavalcante Campos), também é autor de outro artigo sobre o Paic. Dessa vez, o esforço foi direcionado a explicar o modelo de avaliação adotado pelo programa para o ciclo de alfabetização para o primeiro semestre de 2006, apresentando sua metodologia e seus resultados. A avaliação foi realizada em 55 municípios do Ceará com crianças matriculadas no 2º ano do ensino fundamental, antes da criação do Spaece Alfa, que só surgiria no ano seguinte, em 2007.

A avaliação de um programa destinado ao ensino de leitura e escrita é o tema do artigo de Thaize de Souza Reis, Deisy das Graças de Souza e Júlio César de Rose (2009). A avaliação de programas é um tema ainda pouco abordado pela literatura educacional e, quando soma-se ao tema da avaliação da alfabetização, torna-se ainda mais incomum. Os autores dedicaram-se a avaliar a eficácia de um programa informatizado de ensino de leitura e escrita, concluindo pelo reconhecimento da efetividade do programa como atividade suplementar ao currículo escolar.

Gladys Rocha e Raquel Fontes Martins (2012) publicaram artigo, resultado de uma pesquisa mais densa, com a proposição de um corpus de escrita infantil, com base na avaliação do Paebes Alfa (usando os resultados das edições de 2010 e 2011). O objetivo das autoras, com a criação do referido corpus, foi investigar elementos ligados à aquisição da escrita, como os caminhos trilhados pelos alunos para a escrita de palavras e o desenvolvimento de habilidades ligadas ao uso da página. Exceção feita ao resumo, onde o Paebes Alfa é brevemente descrito, o texto não apresenta uma descrição mais profunda do programa, apenas fazendo uso dos dados das já citadas edições.

O artigo de Juliana de Lucena Ruas Riani, Vânia Cândida da Silva e Tufi Machado Soares (2012) teve com objeto de estudo as desigualdades regionais entre escolas e alunos da rede pública de Minas Gerais. Para avaliar a evolução da desigualdade no estado, os autores analisaram dados do Simave (com foco no Proalfa) das edições entre 2006 e 2010. A conclusão do trabalho indicou melhorias nos resultados de alfabetização dos alunos, tendo em vista a proficiência média, e a diminuição das desigualdades entre os alunos, mas o aumento das diferenças entre as escolas.

Alessio Costa Lima, em artigo de 2012, buscou analisar a política de avaliação da educação do Ceará no período entre 2007 e 2010, que ganhou uma avaliação própria para a alfabetização com o Spaece Alfa, em 2007. O autor descreve as principais ações desenvolvidas no período, bem como apresenta os resultados obtidos e a forma como foram comunicados. Ele conclui afirmando que os resultados mostravam, à época, uma melhoria do desempenho dos alunos em todas as séries e disciplinas avaliadas, com maior intensidade nos anos iniciais do ensino fundamental. Ao longo do texto, Lima (2012) apresenta o Spaece Alfa e o descreve com mais atenção do que a maioria dos trabalhos que, quando abordam o programa, apenas o citam (como veremos em tópicos a seguir).

O artigo de Marcelo Câmara (2013) tem como objeto de estudo a Provinha Brasil, mais especificamente, uma análise do desempenho dos alunos do 2º ano do ensino fundamental no pré-teste da avaliação, com ênfase na resolução de problemas envolvendo adição. O autor conclui que elementos componentes dos problemas apresentados, tais como o contexto, a presença ou não de imagens (como suporte dos itens), a localização dos dados e a magnitude dos números envolvidos, são variáveis que influenciam os resultados dos alunos.

Por fim, o artigo de Karla Oliveira Franco e Adolfo Ignacio Calderón (2017) faz uma análise do Simave, o sistema mineiro de avaliação, com base na proposta de Bonamino e Souza (2012) acerca das três gerações de avaliação da educação básica. O artigo aborda o Simave desde sua implementação, em 2000, até a edição de 2013, e conclui que, ao longo do período analisado, o programa de avaliação em Minas Gerais passou pelas três gerações de avaliação educacional. Trata-se de um dos poucos trabalhos em que o programa de avaliação é o objeto de estudo e está no centro da análise, como protagonista, e não como elemento de apoio a análises de outras naturezas (como, por exemplo, no caso em que o elemento que interessa no programa são os seus resultados como instrumento para outro tipo de análise).

A partir da análise de cada um dos artigos classificados no grupo dos que tratam da avaliação da alfabetização, é possível observar que sete deles fazem referência a algum sistema

estadual de avaliação. Dois deles (Lima, 2008, e Franco e Calderón, 2017) possuem programas de avaliação da alfabetização como objeto de estudo (Spaace Alfa e Proalfa, respectivamente). Estamos diante de casos de exceção (2 entre os 404 artigos publicados pela revista no período analisado). Os outros cinco fazem referência ao programa de avaliação, mas sem que o mesmo seja objeto central do estudo (os resultados são utilizados como apoio para os temas centrais que ocupam a preocupação dos autores). Em uma revista especializada em avaliação, a produção sobre a avaliação da alfabetização é muito tímida.

#### 4.2 REVISTA BRASILEIRA DE EDUCAÇÃO

A Revista Brasileira de Educação – RBE é uma publicação da Anped, de fluxo contínuo, que existe desde 1995. Os trabalhos publicados pela revista são dos mais variados temas na área de educação, sem que haja uma temática exclusiva e pré-definida. A importância da Anped no cenário nacional, em especial, no que diz respeito à disseminação e ao incentivo à produção de conhecimento na área educacional, e a qualidade da revista (A1 no Qualis Capes) motivaram sua inclusão como objeto de análise desta tese. A tabela a seguir apresenta os resultados da análise da RBE.

De acordo com o que pode ser extraído da leitura da tabela a seguir, a RBE, entre 2004 e 2017, publicou 47 números, organizados em 12 volumes. Ao todo, no período, foram 668 trabalhos publicados, sendo 22 deles sobre temas relacionados à alfabetização (o que corresponde a 3,29% do total) e 4 deles sobre a avaliação da alfabetização, em específico (o que corresponde a 0,59% do total).

O artigo de Tufi Machado Soares (2005) destina-se a explicar a proficiência dos alunos da 4ª série do ensino fundamental, em língua portuguesa, na edição do Proeb de 2002. O estudo é técnico, com enfoque na apresentação de modelos de regressão em três níveis hierárquicos (alunos, turmas e escolas), assumindo um caráter mais estatístico. Isso explica a preocupação do autor com o elemento metodológico e a breve apresentação do programa (Proeb) e de suas características. Não é um trabalho sobre um sistema de avaliação, mas, antes, um trabalho que lança mão dos dados de um sistema de avaliação para um estudo sobre fatores associados ao desempenho, com forte preocupação estatística.

Tabela 2 - Mapeamento da RBE – 2004-2017

Revista Brasileira de Educação - RBE						
Nº	Vol.	Ano	Meses	Artigos da revista	Artigos sobre alfabetização	Artigos sobre avaliação da alfabetização
25	-	2004	jan-abr	14	3	0
26	-	2004	mai-ago	17	2	0
27	-	2004	set-dez	15	0	0
28	-	2005	jan-abr	17	0	0
<b>29</b>	-	<b>2005</b>	<b>mai-ago</b>	<b>14</b>	<b>0</b>	<b>1</b>
30	-	2005	set-dez	16	0	0
31	11	2006	jan-abr	15	0	0
32	11	2006	mai-ago	14	0	0
33	11	2006	set-dez	15	1	0
34	12	2007	jan-abr	15	0	0
35	12	2007	mai-ago	13	0	0
36	12	2007	set-dez	14	2	0
37	13	2008	jan-abr	17	0	0
38	13	2008	mai-ago	18	1	0
39	13	2008	set-dez	16	0	0
40	14	2009	jan-abr	14	0	0
41	14	2009	mai-ago	14	2	0
42	14	2009	set-dez	13	0	0
43	15	2010	jan-abr	15	1	0
44	15	2010	mai-ago	16	3	0
45	15	2010	set-dez	15	0	0
46	16	2011	jan-abr	13	0	0
47	16	2011	mai-ago	12	0	0
48	16	2011	set-dez	12	0	0
49	17	2012	jan-abr	13	0	0
50	17	2012	mai-ago	11	0	0
<b>51</b>	<b>17</b>	<b>2012</b>	<b>set-dez</b>	<b>14</b>	<b>0</b>	<b>2</b>
52	18	2013	jan-mar	13	0	0
53	18	2013	abr-jun	13	0	0
54	18	2013	jul-set	14	1	0
55	18	2013	out-dez	12	0	0
56	19	2014	jan-mar	12	0	0
57	19	2014	abr-jun	12	1	0
58	19	2014	jul-set	13	1	0
59	19	2014	out-dez	13	0	0
60	20	2015	jan-mar	12	1	0
61	20	2015	abr-jun	12	0	0
62	20	2015	jul-set	11	0	0
63	20	2015	out-dez	13	0	0
64	21	2016	jan-mar	13	0	0
65	21	2016	abr-jun	12	0	0
66	21	2016	jul-set	12	0	0
67	21	2016	out-dez	13	0	0
<b>68</b>	<b>22</b>	<b>2017</b>	<b>jan-mar</b>	<b>13</b>	<b>3</b>	<b>1</b>
69	22	2017	abr-jun	13	0	0
70	22	2017	jul-set	13	0	0
71	22	2017	out-dez	37	0	0
<b>Total</b>				<b>668</b>	<b>22</b>	<b>4</b>

Fonte: Revista Brasileira de Educação, 2004-2017. Tabulação própria.

O artigo de Maria Teresa Esteban (2012) destina-se ao estudo das políticas de avaliação da alfabetização, dando ênfase à Provinha Brasil. Seu objetivo era discutir os sentidos que uma

avaliação dessa natureza poderia ter no contexto das políticas públicas e também das ações escolares. Seu tom é crítico, em particular no que diz respeito a como as avaliações em larga escala podem afetar o cotidiano escolar a partir da maneira como têm sido usadas. Não há referência aos sistemas próprios de avaliação.

A Provinha Brasil é também o objeto de estudo do artigo de Artur Gomes de Moraes (2012). O autor apresenta os resultados de uma pesquisa que acompanhou, ao longo de todo um ano letivo, 12 turmas de 2º ano em três redes públicas no Brasil, de modo a estabelecer relações entre as características do ensino recebido pelos estudantes e seus resultados na Provinha Brasil. Como conclusão, Moraes (2012) observou que o ensino de compreensão da leitura não era muito praticado pelos professores das escolas analisadas e, quando ocorria, havia diferenças entre as habilidades trabalhadas e o que era cobrado pela avaliação. Ele propõe, ao fim do artigo, que sejam discutidos os currículos para a alfabetização no Brasil, bem com a matriz de referência da Provinha Brasil e a necessidade de envolvimento dos professores com a formulação e o uso dos resultados dessa avaliação.

O tema mobilizado por Moraes (2012) no artigo ao qual nos referimos no parágrafo anterior é retomado por ele e três outras autoras em outro trabalho (LEAL et al. 2017). Em novo artigo, mais uma vez, a Provinha Brasil é o objeto da análise. O foco do trabalho continua sendo as relações entre práticas pedagógicas em alfabetização, mais especificamente, no que tange ao ensino da compreensão em leitura, e os resultados dos alunos na avaliação em larga escala. Novamente, a conclusão caminha no sentido de reconhecer que os docentes nessa etapa de escolaridade dedicam-se pouco ao ensino de leitura de textos e que as questões da Provinha Brasil não exigiam dos estudantes habilidades mais sofisticadas de compreensão.

Na RBE, os poucos trabalhos que mobilizam (em maior ou menor grau) o tema da avaliação da alfabetização fazem referência à Provinha Brasil (3 dos 4). O outro trabalho usa os dados de um sistema de avaliação (o Simave, em Minas Gerais) para conduzir uma discussão mais técnica e estatística sobre a avaliação (o que está em jogo é a validade dos modelos de regressão construídos para a análise dos dados, que, inclusive, não diziam respeito ao ciclo de alfabetização). Diante disso, não há nenhum trabalho sequer que tenha algum programa de avaliação como objeto de estudo e nenhum trabalho que aborde a avaliação da alfabetização no âmbito dos sistemas estaduais. Não estamos diante, aqui, da ínfima produção sobre o tema

apresentada pela EAE, como vimos no tópico anterior, mas da ausência completa de publicações de trabalhos relacionados à avaliação da alfabetização<sup>70</sup>.

#### 4.3 REVISTA BRASILEIRA DE ESTUDOS PEDAGÓGICOS

A Revista Brasileira de Estudos Pedagógicos é um periódico dos mais antigos no Brasil. Criado em 1944, é quadrimestral, publicado em formato impresso e eletrônico. Na área de educação, não apresenta enfoques temáticos específicos, publicando trabalhos sobre a educação em geral, sendo um dos periódicos mais importantes em educação no país (A1 em Ensino e A2 em Educação no Qualis Capes). Publicada pelo Inep, sua história, qualificação e importância justificam sua inclusão na lista de periódicos aqui analisados. O Inep é o órgão responsável pelas avaliações nacionais, o que aumenta o interesse pelo periódico em minha pesquisa.

Conforme podemos observar a partir das informações da tabela a seguir, entre 2004 e 2017, a RBPE publicou 39 números, em 14 volumes diferentes. No total, foram 473 trabalhos publicados ao longo do período. Destes, 12 versavam sobre a alfabetização (correspondendo a 2,53% do total) e apenas 1 tratava do tema da avaliação da alfabetização (correspondendo a 0,21% do total). Os trabalhos sobre alfabetização giraram em torno de temas variados: a formação do professor alfabetizador (2), a distinção entre alfabetização e letramento (1), a alfabetização matemática (com foco na atuação do docente em turmas multisseriadas) (1), as contribuições da neurociência para a alfabetização (1), a apropriação do processo de escrita (1), a análise do material didático de um programa destinado ao 1º ano do ensino fundamental (1), as estratégias dos alunos do ciclo de alfabetização para a solução de problemas envolvendo números decimais (1), a notação da língua escrita ao fim da educação infantil (traz *alfabetização* como palavra-chave) (1), a representação docente nos cadernos de formação do Pnaic (1), a alfabetização científica (1), um mapeamento dos trabalhos sobre alfabetização publicados pela própria revista (a partir de 1998) (1).

---

<sup>70</sup> Numa revista que reúne a produção dos Programas de Pós-Graduação em Educação nacionais. Fica a dúvida se esse tema não é uma demanda dos autores do campo ou se textos que poderiam tratar da temática não são aprovados quando da seleção dos textos que serão apresentados a cada reunião anual pelo comitê científico do evento.

Tabela 3 - Mapeamento da RBPE – 2004-2017

Revista Brasileira de Estudos Pedagógicos - RBEP						
Nº	Vol.	Ano	Meses	Artigos da revista	Artigos sobre alfabetização	Artigos sobre avaliação da alfabetização
209-211	85	2004	-	11	0	0
212	86	2005	-	9	0	0
213-214	86	2005	-	11	0	0
215	87	2006	-	14	0	0
216	87	2006	-	13	0	0
217	87	2006	-	10	0	0
218	88	2007	-	10	0	0
219	88	2007	-	10	0	0
220	88	2007	-	11	0	0
221	89	2008	-	9	0	0
222	89	2008	-	12	0	0
223	89	2008	-	11	1	0
224	90	2009	-	12	0	0
225	90	2009	-	13	0	0
226	90	2009	-	10	0	0
227	91	2010	-	11	0	0
228	91	2010	-	9	0	0
229	91	2010	-	11	0	0
230	92	2011	-	11	1	0
231	92	2011	-	11	0	0
232	92	2011	-	17	0	0
233	93	2012	-	13	0	0
234	93	2012	-	13	1	0
235	93	2012	-	16	0	0
236	94	2013	-	15	0	0
<b>237</b>	<b>94</b>	<b>2013</b>	-	<b>14</b>	<b>0</b>	<b>1</b>
238	94	2013	-	11	0	0
239	95	2014	-	11	1	0
240	95	2014	-	10	0	0
241	95	2014	-	13	0	0
242	96	2015	-	14	3	0
243	96	2015	-	14	0	0
244	96	2015	-	11	1	0
245	97	2016	-	12	1	0
246	97	2016	-	13	0	0
247	97	2016	-	12	1	0
248	98	2017	-	14	0	0
249	98	2017	-	15	2	0
250	98	2017	-	16	0	0
<b>Total</b>				<b>473</b>	<b>12</b>	<b>1</b>

Fonte: Revista Brasileira de Estudos Pedagógicos, 2004-2017. Tabulação própria.

O único trabalho que diz respeito à avaliação da alfabetização tem como objeto a Provinha Brasil (nº 237, vol. 94, de 2013). Freitag e colaboradoras (2013) analisam essa avaliação a partir de seus limites e incompletudes, concentrando seus esforços em três dimensões para fins de aprimoramento do instrumento: a avaliação da competência narrativa dos alunos, o tratamento dado pelo instrumento à variação linguística e a formação do professor alfabetizador.

Assim, não há um único trabalho sequer referente a um dos programas de avaliação da alfabetização aqui pesquisados. Mais, não há um único trabalho sobre qualquer sistema próprio de avaliação da alfabetização. O único artigo acerca do tema destina-se à Provinha Brasil. Nem mesmo a ANA mereceu a atenção dos estudiosos e pesquisadores que publicaram na RPBE no período analisado. Por se tratar de uma revista cuja organização e publicação são de responsabilidade do Inep, essa ausência chama ainda mais a atenção.

#### 4.4 BANCO DE TESES E DISSERTAÇÕES DA CAPES

Conforme visto anteriormente, três importantes periódicos da área, veículos centrais para a divulgação de trabalhos acadêmicos sobre educação no Brasil, apresentam um quadro que ressalta a parca (quase inexistente) produção sobre a avaliação da alfabetização entre nós, em um período que envolve o surgimento e o desenvolvimento dos principais programas de avaliação da alfabetização, que permanecem em ação até hoje (foram 14 anos analisados). No total, foram publicados 17 trabalhos sobre o tema, entre 1545 artigos (o que corresponde a ínfimos 1,1% do total). A princípio, poderíamos argumentar que pouco foi publicado, mas não necessariamente pouco foi produzido sobre esse tema. As revistas, que não versam apenas sobre alfabetização, tampouco sobre avaliação da alfabetização, publicam sobre os mais variados temas em educação, uma área de pesquisa reconhecidamente caracterizada, entre outras coisas, pela amplitude de seus pontos de interesse e dos tipos de pesquisa que abraça. Assim, seria possível pensar que, a despeito das publicações nesses três periódicos, a avaliação da alfabetização se apresentaria como objeto de interesse para trabalhos acadêmicos não necessariamente publicados, em especial, nos programas de pós-graduação brasileiros (as dissertações e teses). Por isso, foi tomada como necessária a consulta a essa produção acadêmica, de modo a averiguar se a escassa publicação nos periódicos era ou não acompanhada de baixa produção sobre o tema também no âmbito da pós-graduação.

A consulta ao banco da Capes<sup>71</sup> foi feita para identificar o número de teses e dissertações dedicadas ao estudo de, pelo menos, um dos programas de avaliação da alfabetização objetos da presente pesquisa. No caso dos três periódicos pesquisados, de acordo com o que foi apresentado em tópicos anteriores, a pesquisa foi mais ampla, buscando trabalhos sobre a avaliação da alfabetização como um todo. No que diz respeito ao banco da Capes, pelas próprias

---

<sup>71</sup> Consulta realizada entre 10 de junho e 15 de julho de 2019.



limitações da ferramenta de busca, a pesquisa dedicou-se aos três programas de alfabetização objetos desta tese (Proalfa, Paebes Alfa e Spaece Alfa).

Para o Proalfa, a palavra de busca foi a própria abreviação do programa de avaliação, ao passo que, nos outros dois casos, foi necessário pesquisar por Spaece e Paebes (não acrescidos da palavra *alfa*). A palavra *alfa*, componente das abreviaturas desses dois programas, gerava problemas para a busca, visto que, sendo separada, levava a um enorme retorno de trabalhos que usavam a palavra alfabetização em seus títulos ou palavras-chaves. Assim, um grande contingente de trabalhos não relacionados a esses programas aparecia na busca. Por isso, a opção por pesquisar sem a inclusão dessa palavra nesses dois casos.

A busca inicial não foi realizada através dos refinamentos permitidos pela ferramenta do banco de teses e dissertações. O objetivo inicial era verificar o número de trabalhos em que esses programas apareciam, independentemente da área de conhecimento na qual o trabalho foi categorizado. Da mesma forma, nenhum filtro acerca de ano, autor ou instituição foi feito. Nessa primeira busca, o retorno foi de 16 trabalhos para o Paebes, 35 para o Proalfa e 71 para o Spaece. Vale lembrar que, até esse momento, nos casos dos programas de avaliação do Espírito Santo e do Ceará, não havia diferenciação entre o Paebes e o Spaece e o Paebes Alfa e o Spaece Alfa, respectivamente.

A partir disso, para uma análise mais fina e detida dos trabalhos, uma segunda busca foi realizada, dessa vez, usando o refinamento de pesquisa que a ferramenta da Capes permite. A busca foi feita com a seleção de algumas áreas do conhecimento: *educação*, como área central, e áreas conexas, como *ensino* (e derivados – *ensino de matemática*, por exemplo), *políticas públicas*, *planejamento educacional* e *multidisciplinar*. A inclusão de *multidisciplinar* como área para a busca deve-se à existência de programas associados à gestão, por exemplo, que, apesar de não estarem incluídos como parte da área de conhecimento *educação*, apresentam particular interesse no tema da avaliação educacional. Contudo, é importante ressaltar, como a análise foi feita caso a caso, foi possível analisar a natureza de cada programa multidisciplinar, de modo que foram selecionados apenas aqueles relacionados diretamente à educação.

Mais uma vez, não houve refinamento da busca por tipo (mestrado ou doutorado), ano, autor ou instituição. O objetivo foi verificar a produção sobre esses três programas de avaliação da alfabetização em nível de pós-graduação ao longo do tempo e em todos os programas de pós-graduação brasileiros. Como resultado dessa busca mais específica, temos: 10 trabalhos para o Paebes, 32 trabalhos para o Proalfa e 41 trabalhos para o Spaece.

A partir dos resultados dessa segunda busca, cada trabalho passou a ser analisado caso a caso, com o objetivo principal de identificar os trabalhos que, de fato, dedicaram-se à pesquisa sobre um ou mais desses programas. Vale lembrar, de acordo com o que já foi explicitado anteriormente, que os termos de busca para os programas de Espírito Santo e Ceará não incluíram a palavra *alfa*, o que levou a resultados para o Paebes e para o Spaece, e não especificamente para Paebes Alfa e Spaece Alfa. Além disso, o intuito foi o de encontrar trabalhos que trataram esses programas como objeto central do estudo e não como temas incidentais (por exemplo, o estudo da desigualdade escolar usando dados do Paebes Alfa, como veremos a seguir).

#### 4.4.1 A produção sobre o Paebes Alfa

Dos 10 trabalhos relativos ao Paebes, 4 deles estão relacionados, direta ou indiretamente, ao Paebes Alfa (todos posteriores à Plataforma Sucupira). É uma produção extremamente tímida para um programa que existe há uma década. Dos 4 trabalhos, 2 deles versam especificamente sobre o Paebes Alfa (Côco, 2014, e Silva, 2013), 1 deles analisa o Paebes Alfa em conjunto com outros programas, ao passo que o outro estuda um tema específico a partir dos dados do Paebes Alfa, sem que o programa em si seja seu objeto principal de estudo.

A tese de doutorado de Naira da Costa Muylaert Lima<sup>72</sup>, defendida em 2016, no Programa de Pós-Graduação em Educação da PUC-RJ, tem como tema central as desigualdades educacionais, mais especificamente, a desigualdade de conhecimento entre os alunos. A autora lança mão dos dados do Paebes Alfa para propor um debate sobre as desigualdades educacionais no cenário nacional. Assim, o Paebes Alfa apresenta-se, no contexto desse trabalho, mais como fonte de dados do que como objeto central do estudo<sup>73</sup>.

A tese de doutorado de Dilza Côco<sup>74</sup>, defendida em 2014, no Programa de Pós-Graduação em Educação da UFES, tem como objeto de estudo o Paebes Alfa. A tese dedica-se à análise do programa, particularmente, no que diz respeito à produção de sentidos para essa avaliação e sua relação com as práticas de alfabetização nas turmas de 1º e 2º anos na rede municipal de Serra/ES. Trata-se de um estudo sobre uma faceta importante do programa de

<sup>72</sup> Intitulada *Educação e desigualdades sociais: suas relações no Estado de Espírito Santo*.

<sup>73</sup> O próprio resumo da tese atesta que: “Este trabalho busca contribuir para o debate acerca do tema das desigualdades educacionais, por meio dos dados do PAEBES-Alfa – avaliação longitudinal da alfabetização do estado do Espírito Santo” (LIMA, 2016) (destaque meu, em negrito).

<sup>74</sup> Intitulada *Avaliação externa da aprendizagem: o Paebes-Alfa no Espírito Santo*.

avaliação (a forma como produz efeitos nas práticas de alfabetização em sala de aula) e, embora não seja um trabalho sobre o programa como um todo, é uma exceção na parca produção bibliográfica sobre a avaliação externa da alfabetização, ainda mais quando se trata de programas estaduais de avaliação. Mais sobre a tese de Côco será dito no capítulo seguinte, destinado a análise de cada um dos programas de avaliação da alfabetização selecionados para a pesquisa.

A tese de doutorado de Josiane Toledo Ferreira Silva<sup>75</sup> analisa os resultados de escrita ao final do 2º ano do ensino fundamental, tendo como referência o Paebes Alfa. O trabalho dedica-se a uma análise técnica da matriz de referência e das chaves de correção da avaliação de escrita, culminando com a proposição de uma nova chave de correção para os itens. Já a dissertação de Maíra Miranda Portela<sup>76</sup> dedica-se à análise das matrizes de referência de três sistemas de avaliação da alfabetização, incluindo o Paebes Alfa (os outros dois são Saepe e Seape). A proposta da autora é comparar as matrizes de referências dos três programas com a matriz da ANA, propondo, por fim, uma matriz unificada para as avaliações de matemática no 3º ano do ensino fundamental. Trata-se, também, de um trabalho técnico, que analisa uma dimensão importante do Paebes Alfa (em comparação com outros programas). Esses dois trabalhos, o de Silva (2013b) e o de Portela (2016) são exceções dentro da exceção: abordam dois temas invisíveis dentro da invisibilidade já diagnosticada para a avaliação em larga escala da alfabetização, quais sejam, a avaliação da escrita e a avaliação da alfabetização matemática, respectivamente.

A tese de Selma Lúcia de Assis Pereira<sup>77</sup> analisa as estratégias de divulgação dos resultados do Paebes em Língua Portuguesa, tendo como recorte a rede estadual do município de Cariacica/ES. Seu foco é o ensino fundamental (a partir do 5º ano), não tendo o Paebes Alfa como objeto. Da mesma forma, a dissertação de Dayane Louzada de Oliveira<sup>78</sup>, que analisa os resultados do Paebes em Ciências da Natureza, com foco no 9º ano do ensino fundamental da rede municipal de Marataízes/ES, não aborda o Paebes Alfa. O Paebes (juntamente com

---

<sup>75</sup> Defendida em 2013, no Programa de Pós-Graduação em Educação da PUC-RJ, com o título *A escrita na avaliação da alfabetização em larga escala*.

<sup>76</sup> Com o título *Alfabetização matemática: um paralelo entre a avaliação nacional e cenários estaduais*, foi defendida em 2016, no âmbito do Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, da Universidade Federal de Juiz de Fora.

<sup>77</sup> Intitulada *PAEBES: modos, formas e diálogos a partir dos usos dos resultados em língua portuguesa da avaliação externa estadual no município de Cariacica-ES*. Defendida no Programa de Pós-Graduação em Educação da UFES, em 2015.

<sup>78</sup> Intitulada *Análise dos resultados do Paebes de Ciências da Natureza do município de Marataízes-ES e seu reflexo na prática docente*. Defendida em 2018, no programa de pós-graduação em Ensino, Educação Básica e Formação de Professores da UFES, em 2018.

Sadeam e Avalia Bahia) também é um dos programas analisados no âmbito da dissertação de Patrícia de Lima Paula<sup>79</sup>, com o objetivo de compreender os fatores que afetam a qualidade dos itens de biologia no âmbito desses programas. Com foco no ensino médio, não há menção ao Paebes Alfa nesse trabalho. Ainda, o Paebes é mencionado, embora não analisado com deter, como parte da dissertação de mestrado de Ruy Barboza de Oliveira Júnior<sup>80</sup>, na qual o autor se propõe a analisar as políticas educacionais do segundo governo de Paulo Hartung (entre 2006 e 2010). No trabalho, o Paebes é encarado sob o prisma do monitoramento que promove e das políticas de bonificação a ele associadas. Não há qualquer menção ao Paebes Alfa.

Por fim, a tese de Tarliz Liao<sup>81</sup> está nomeada de forma equivocada no banco da Capes (aparece com o mesmo título da tese de Dilza Côco). Aparece como resultado para a busca por *Paebes*, mas não trata do programa, nem sequer do tema da avaliação educacional. Defendida em 2014, no departamento de educação da UFES, a tese aborda o tema do currículo. Dessa forma, não pode ser considerada como resultado para o Paebes Alfa. Da mesma forma, a dissertação de Leonardo Polese Alves<sup>82</sup>, que apresenta uma proposta de integração das disciplinas ciências e matemática com disciplinas técnicas no curso de geoprocessamento, não possui nenhuma referência ao Paebes, Paebes Alfa ou à avaliação educacional.

#### 4.4.2 A produção sobre o Proalfa

Há trabalhos em que o Proalfa aparece como parte do estudo, tem seus resultados analisados ou é mencionado ao longo do texto. A tese de doutorado de Edna Martins Borges<sup>83</sup> teve como objetivo analisar as percepções dos diretores das escolas estaduais de Minas Gerais sobre os efeitos do Proeb no contexto escolar. O Proalfa é mencionado em várias passagens do texto, mas como parte do Simave, visto que o foco da pesquisa foi mesmo o Proeb (ao longo do texto, de 327 páginas, o Proalfa ganha um subtópico, em conjunto com o PAAE, de apenas

---

<sup>79</sup> Intitulada *A influência da elaboração de itens na eficácia das avaliações em larga escala*. Defendida em 2018, no âmbito do Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública da UFJF.

<sup>80</sup> Intitulada *As políticas educacionais do segundo mandato do governo Paulo Hartung e seus impactos no trabalho docente*. Defendida em 2013, no Programa de Pós-Graduação em Educação da Universidade Federal do Espírito Santo.

<sup>81</sup> Possuindo, na verdade, o título *A elaboração e instituição do currículo mínimo de matemática no Rio de Janeiro*.

<sup>82</sup> Intitulada *Uma proposta de integração de ciências e matemática com disciplinas técnicas no curso de geoprocessamento*. Defendida em 2013, no âmbito do Programa de Educação em Ciências e Matemática, do Instituto Federal do Espírito Santo.

<sup>83</sup> Intitulada *Avaliações em larga escala no contexto escolar: percepção de diretores escolares da rede estadual de ensino de Minas Gerais* e defendida em 2016, no Programa de Pós-Graduação em Educação da UFMG. A palavra Proalfa não aparece nem no título, nem como palavra-chave.

uma página). Na mesma linha seguiu a dissertação de mestrado de Patrícia Aparecida da Silva<sup>84</sup>, que busca investigar a influência do Simave sobre as escolas no que diz respeito à busca por melhores resultados na avaliação. A autora também investiga a compreensão do sistema de avaliação por parte dos atores escolares (no sentido do sistema influenciar práticas pedagógicas para fins de treinamento dos alunos), sem fazer uma análise detalhada e específica do Proalfa. A dissertação de Ana Paula de Moura Ferreira Dias<sup>85</sup> aponta os resultados do Proalfa (mais especificamente, a preocupação com o que significavam) como o gatilho para a implementação do Programa de Intervenção Pedagógica/Alfabetização no Tempo Certo (PIP/ATC) na Superintendência Regional de Ensino de Leopoldina. Essa é a mesma temática da dissertação de Carla Cândida da Silva Reis<sup>86</sup>, que tem como objeto de pesquisa o PIP/ATC e analisa os resultados do Proalfa como forma de analisar o referido programa, e da dissertação de Graciele Aparecida Nunes<sup>87</sup>, que analisa a implementação do programa em três escolas estaduais da SRE de Patrocínio (os resultados do Proalfa são apresentados como forma de analisar o efeito do PIP/ATC). Nos três últimos casos, o programa PIP/ATC é o objeto do trabalho e o Proalfa não aparece como centro da pesquisa. Da mesma maneira, a tese de Rayssa Lopes Bastos<sup>88</sup> pesquisa o currículo de Língua Portuguesa em Juiz de Fora/MG e os resultados do Proalfa são apontados como a principal justificativa para a implementação de políticas destinadas à ampliação do tempo escolar (o Proalfa é apresentado em uma nota de rodapé). O PIP/ATC também é o tema de pesquisa da dissertação de Renata Mota Teixeira<sup>89</sup>, na qual o referido programa tem sua implementação analisada em duas escolas estaduais de Juiz de Fora/MG. Os baixos resultados no Proalfa funcionaram como critério de seleção das escolas investigadas e o programa é apresentado, mesmo que timidamente, ao longo do texto (há um subtópico, com três páginas, destinado à apresentação do programa).

---

<sup>84</sup> Intitulada *As influências do sistema mineira de avaliação na construção da qualidade da educação*, defendida em 2016, no Programa de Pós-Graduação em Educação da UFU.

<sup>85</sup> Com título *Os efeitos do programa de intervenção pedagógica no ciclo de alfabetização: o desafio para que todos os alunos leiam e escrevam até os oito anos de idade*. Defendida em 2015, no PPGP/CAEd.

<sup>86</sup> Intitulada *Da agenda à implementação: um olhar sobre as dinâmicas de interação dos agentes implementadores do programa de intervenção pedagógica/Alfabetização na Idade Certa no estado de Minas Gerais*. Defendida em 2013, no PPGP/CAEd.

<sup>87</sup> Defendida em 2015, no PPGP/CAEd, com o seguinte título *Análise da implementação do Programa de Intervenção Pedagógica em uma Superintendência Regional de Ensino (SRE/MG)*.

<sup>88</sup> Tese defendida em 2017, no Programa de Pós-Graduação em Educação da UCP, com o título *A proposta curricular de Língua Portuguesa de Juiz de Fora/MG: reflexões acerca do processo de recontextualização na prática*.

<sup>89</sup> Defesa no PPGP/CAEd, em 2013, com o título *O programa de intervenção pedagógica em duas escolas de Juiz de Fora: sucessos e entraves*.

Seguindo a análise dos trabalhos em que o Proalfa aparece como parte da pesquisa, a dissertação de Denise Alessandra Palhares Diniz Moraes<sup>90</sup> destina-se à investigação dos fatores associados ao desempenho, tema central da pesquisa em eficácia escolar. O trabalho dedicou-se a identificar os principais fatores relacionados a bons desempenhos em uma escola formada por alunos com alta vulnerabilidade social. Os resultados do Proalfa aparecem, juntamente com os do Avalia BH, como critério para aferir a melhoria da qualidade na escola, mas o foco é a identificação dos fatores que levaram a essa melhoria. Os fatores associados ao desempenho são também o tema de pesquisa da dissertação de Rita de Cássia Costa Maia<sup>91</sup>, que analisa os elementos que, no âmbito da gestão, afetam os resultados do Proalfa das escolas estaduais de Patrocínio/MG (há um subtópico, de quatro páginas, destinado à política de avaliação do Proalfa, mas com foco nos resultados do programa para a SRE de Patrocínio). Na mesma linha, seguiram a dissertação de Zely Mattos Rosa<sup>92</sup>, que teve com objetivo analisar os resultados de uma escola de Juiz de Fora/MG no Proalfa à luz das características da gestão, e a dissertação de Nádia Ramos Grisson de Oliveira<sup>93</sup>, que dedicou-se à análise das ações da gestão de uma escola (em Governador Valadares) e sua relação com os resultados do Proalfa (há um tópico destinado a descrever as avaliações no estado de Minas Gerais, com foco nos resultados). A dissertação de Maria Lúcia Viana do Prado<sup>94</sup> tem como objeto de pesquisa a análise dos cursos de capacitação para a apropriação dos resultados do Simave, com foco na SRE de Teófilo Otoni/MG. Mais uma vez, o tema da apropriação de resultados aparece como central, contudo, a dissertação coloca ênfase no processo de capacitação dos atores para a compreensão e uso dos resultados. O Proalfa é mencionado, mas o curso de apropriação diz respeito aos resultados do Simave, incluindo o Proeb, e não apenas aos do Proalfa (o programa recebe um subtópico descritivo de uma página). Os resultados do Proalfa também aparecem como parte da dissertação de Adriana Moreira Diniz<sup>95</sup>, cujo foco é a forma como o caderno de boas práticas recebido pelos diretores de escola de Minas Gerais influencia a relação desses diretores com a

---

<sup>90</sup> Com título *Gestão escolar eficaz: o diferencial de uma escola em contexto de vulnerabilidade social*, foi defendida em 2014, no PPGP/CAEd.

<sup>91</sup> Intitulada *O desempenho de escolas da rede estadual no município de Patrocínio-MG no Proalfa: uma análise dos fatores afetos à gestão escolar de quatro casos comparados*. Defendida em 2013, no PPGP/CAEd.

<sup>92</sup> Com título *A gestão de resultados e a avaliação em larga escala no cotidiano de uma escola da rede municipal de Juiz de Fora*, e defesa em 2013, no PPGP/CAEd.

<sup>93</sup> Defendida em 2017, no PPGP/CAEd, com o título *Análise de ações escolares e resultados do Proalfa: um estudo de caso de uma escola de sucesso em Governador Valadares*.

<sup>94</sup> Intitulada *Análise dos cursos de capacitação para divulgação e apropriação de resultados do Simave/Proalfa e Proeb da Superintendência Regional de Teófilo Otoni/MG*. Defendida em 2014, no PPGP/CAEd.

<sup>95</sup> Defesa no PPGP/CAEd, em 2013. Intitulada *O caderno de boas práticas dos diretores de escola de Minas Gerais, sua relação com a prática docente e as repercussões deste instrumento nos resultados do Proalfa*.

prática docente e os resultados do Proalfa (não há tópico algum destinado a apresentar, descrever ou analisar o programa).

Já a tese de Edilene Mizael de Carvalho Perboni<sup>96</sup> tem como foco o Proalfa. A análise é pautada pelo conceito de governamentalidade e pela análise do discurso, de modo a compreender, em especial, os sujeitos que fazem parte da avaliação. Embora o foco seja o Proalfa, há seleção de uma dimensão específica para a análise do programa, sem um enfrentamento mais amplo dos diferentes elementos dos quais o Proalfa é composto. A dissertação de Ana Cláudia Osório Martins<sup>97</sup> também é dedicada ao Proalfa, mais especificamente, aos impactos dos resultados da avaliação no ambiente escolar, tendo em vista a realidade de cinco escolas da rede estadual de ensino em Juiz de Fora/MG. Mais uma vez, os resultados se apresentam como a faceta de interesse para a pesquisa.

Por fim, dois trabalhos que fazem uma referência instrumental ao Proalfa. A tese de Márcio Moterani Swerts<sup>98</sup> analisa o Dia D da educação em Minas Gerais, a partir da ótica da disputa de poder que o evento põe em cena (a pesquisa destina-se a analisar o sentido da expressão Dia D no âmbito das escolas; o Proalfa é mencionado no corpo no texto apenas para fazer referência à forma como os resultados da avaliação são utilizados para pressionar as ações das escolas). A implementação do PIP/ATC na SRE Zona da Mata é o tema da dissertação de mestrado de Gabriela dos Santos Pimenta<sup>99</sup>, na qual os resultados do Proalfa são apresentados como base para a reorganização do trabalho da equipe na superintendência no ano de 2013 (contudo, não há nenhum tópico destinado a analisar ou a descrever o programa de avaliação).

Outros trabalhos que resultaram da busca não abordam o Proalfa. A dissertação de Sueli Machado Pereira de Oliveira<sup>100</sup> dedica-se à análise da implementação do ensino fundamental de nove anos, tendo como base a visão do aluno. O trabalho (dissertação de mestrado) de Edvânia de Lana Morais Andrade<sup>101</sup> aborda a formação continuada dos diretores escolares no cenário da expansão das avaliações externas.

---

<sup>96</sup> Intitulada Os sujeitos das políticas de avaliação sistêmica da alfabetização em Minas Gerais: uma analítica à luz da governamentalidade. Defendida em 2015, no Programa de Pós-Graduação em Educação da Universidade São Francisco.

<sup>97</sup> Intitulada Proalfa: avaliação e propostas pedagógicas. Defendida em 2013, no Programa de Pós-Graduação da UFJF.

<sup>98</sup> Intitulada O Dia D da educação em Minas Gerais: jogos estratégicos de poder. Defendida em 2016, no Programa de Pós-Graduação em Educação da Universidade São Francisco.

<sup>99</sup> Dissertação defendida no PPGP/CAEd, em 2014, com o título Avaliando a implementação do Programa de Intervenção Pedagógica/Alfabetização no Tempo Certo (PIP/ATC): estudo de caso da SRE Zona da Mata.

<sup>100</sup> Defendida em 2011, no Programa de Pós-Graduação em Educação da PUC-MG, com o título Implantação do ensino fundamental de 9 anos em Minas Gerais: a visão da criança.

<sup>101</sup> Com o título *A formação continuada dos diretores escolares no contexto da política pública das avaliações externas*. Defendida em 2012, no PPGP/CAEd.

Alguns trabalhos que apareceram como resultados da consulta ao banco não puderam ser analisados de forma mais profunda, tendo em vista que foram publicados antes da existência da Plataforma Sucupira. Para esses trabalhos, a análise baseou-se tão somente no título e nas palavras-chaves. Em nenhum deles, o Proalfa figurou no título ou como palavra-chave. A dissertação de mestrado de Maria Inez Barroso Simões<sup>102</sup> dedica-se ao estudo do PIP, o Programa de Intervenção Pedagógica de Minas Gerais. A dissertação de Gisele Costa Maia<sup>103</sup> tem como objeto de pesquisa o sucesso da alfabetização de crianças em casos de reagrupamentos temporários (com foco em duas escolas da Superintendência Regional de Ensino de Montes Claros). A avaliação externa é o tema de pesquisa que dá base à dissertação de Railma Aparecida Cardoso Marinho<sup>104</sup>, com análise específica da política de avaliação na microrregião de Januária/MG. A pesquisa de Margarida Montejano da Silva<sup>105</sup>, uma dissertação de mestrado, dedica-se à formação de professores alfabetizadores, com foco em uma cidade do estado de São Paulo. A dissertação de Lúcia Helena Miranda Bastos<sup>106</sup> versa sobre a gestão educacional e a eficácia escolar, ao passo que o trabalho de Rúbia Mara Schattner<sup>107</sup>, também uma dissertação, diz respeito a um estudo de uma escola eficaz, tendo em vista a avaliação nos anos iniciais do fundamental. O trabalho (uma dissertação) de José Carlos Ferreira Araújo<sup>108</sup> analisa as políticas públicas para a EJA em Belém/PA. A dissertação de Gisele Francisca da Silva Carvalho<sup>109</sup> tem como tema os impactos da avaliação oficial sobre a prática docente. A dissertação de Marlene Dias Pereira Pinto<sup>110</sup> diz respeito a um programa do Rio de Janeiro também chamado Proalfa<sup>111</sup>, mas sem qualquer ligação com o programa de avaliação da alfabetização de Minas Gerais. O

---

<sup>102</sup> Defendida em 2012, com o título O Programa de Intervenção Pedagógica do estado de Minas Gerais (PIP), no âmbito do PPGP/CAEd.

<sup>103</sup> Intitulada Reagrupamentos temporários e o sucesso na alfabetização de crianças: estudo de caso em duas escolas públicas da Superintendência Regional de Ensino de Montes Claros, com defesa em 2012, no PPGP/CAEd.

<sup>104</sup> Com defesa em 2010, no Programa de Pós-Graduação em Educação da PUC-Campinas, o título da dissertação é Políticas públicas de avaliação: a avaliação externa e a realidade educacional da microrregião de Januária/MG.

<sup>105</sup> Defendida em 2000, no Programa de Pós-Graduação em Educação da PUC-Campinas, com o título A questão da formação de professores alfabetizadores: uma experiência com educação de adultos na realidade de Espírito Santo do Pinhal/SP.

<sup>106</sup> Com o título Gestão educacional: uma análise qualitativa dos fatores que geram o êxito da escola, foi defendida em 2012, no PPGP/CAEd.

<sup>107</sup> Intitulada Avaliação em dois momentos nos anos iniciais do ensino fundamental: estudo de caso de uma escola eficaz, foi defendida no PPGP/CAEd, em 2012.

<sup>108</sup> Com defesa em 2006, no Programa de Pós-Graduação em Educação da UFPA, e com o seguinte título As Políticas Públicas para a Educação de Jovens e Adultos no Município de Belém no período de 1997 a 2004.

<sup>109</sup> Defendida na UFSJ, no âmbito do Mestrado em Processos Socioeducativos e Práticas Escolares, em 2010, com o título Avaliação oficial: subsídios para a compreensão do impacto sobre a prática docente.

<sup>110</sup> Com o título O funcionamento cognitivo do idoso e o processo de ensino da alfabetização do Proalfa – reflexões para um ensino melhor. Defendida em 2004, no Programa de Pós-Graduação em Educação da UERJ.

<sup>111</sup> Trata-se do Programa de Alfabetização, Documentação e Informação, programa de extensão da UERJ, que desenvolve, desde 1997, projetos na área de alfabetização de jovens e adultos (para mais, acesso à página do programa por meio do seguinte endereço <http://www.proalfa.uerj.br/conteudo/sobre.html>).



Programa Nacional do Livro Didático, com ênfase nas escolhas feitas por professores para o público formado por jovens e adultos, é o tema da dissertação de Juliane Gomes de Oliveira<sup>112</sup>.

Entre os trabalhos publicados antes da Plataforma Sucupira, há aqueles que fazem menção ao Proalfa no título. É o caso da dissertação de Patrícia Valesca Gomes<sup>113</sup>, que estuda a apropriação dos resultados do programa em duas escolas de Governador Valadares. A mesma linha de pesquisa (apropriação dos resultados) foi trilhada pela dissertação de Cátia Cristina Claudiano Trindade Rosa<sup>114</sup> (com foco no estudo de caso de uma escola com bom desempenho no programa) e pela tese de Karina Fideles Filgueiras<sup>115</sup> (com foco em duas escolas com alunos com baixo desempenho no programa).

Em relação ao Proalfa, portanto, dos 32 trabalhos que retornaram como resultado da busca, temos: i) 17 que abordam o Proalfa como tema (objeto principal do trabalho, como parte do estudo ou com menção ao programa); ii) 2 que possuem *Proalfa* como palavra no título (anteriores à Plataforma Sucupira); iii) 3 que não abordam o Proalfa (sequer como menção); e iv) 10 que não possuem *Proalfa* como palavra no título (anteriores à Plataforma Sucupira). A tabela abaixo apresenta essa distribuição dos trabalhos.

Tabela 4 - Produção de teses e dissertações sobre o Proalfa

Posteriores à Plataforma Sucupira		Anteriores à Plataforma Sucupira		Total
Com menção	Sem menção	Com menção	Sem menção	
17	2	3	10	32
<b>Total com menção</b>				20
<b>Total sem menção</b>				12

Fonte: Banco de teses e dissertações Capes. Tabulação própria.

Chama atenção, ainda, o fato de que, dos 20 trabalhos, 14 fazem referência, mais ou menos central, aos resultados da avaliação, a dimensão que mais recebeu atenção das pesquisas sobre a avaliação do Proalfa no âmbito da pós-graduação. Trata-se, é verdade, de uma dimensão importante para o processo de avaliação, mas não é a única. Tal concentração revela uma preocupação dos pesquisadores com a etapa mais visível do processo (a divulgação dos resultados e o que se faz com eles), deixando de investigar outras dimensões do programa, que ajudariam em sua melhor compreensão.

<sup>112</sup> Defendida em 2011, no Programa de Pós-Graduação em Educação da UFMG, com o título Programa nacional do livro didático para alfabetização de pessoas jovens adultas e idosas: os professores e suas escolhas.

<sup>113</sup> Intitulada Estudo comparado da ação gestora na apropriação dos resultados do PROALFA: análise de dois casos de sucesso em Governador Valadares, com defesa em 2014, no PPGP/CAEd.

<sup>114</sup> Com defesa no PPGP/CAEd, em 2012, e com o título Análises de ações educacionais de uma escola com bom desempenho no PROALFA.

<sup>115</sup> Defendia no Programa de Pós-Graduação em Educação, em 2012, com o título Alunos avaliados com baixo desempenho no PROALFA (MG): análise de duas escolas estaduais de Belo Horizonte.

#### 4.4.3 A produção sobre o Spaece Alfa

Dos três programas de avaliação pesquisados, o Spaece Alfa (com busca pela palavra Spaece, pelos motivos já especificados anteriormente) é o que mais apresentou retorno na busca no banco de teses e dissertações da Capes, com 41 trabalhos. Da mesma maneira que ocorreu com o Proalfa, há trabalhos em que o Spaece-Alfa é objeto da pesquisa, ou faz parte, com maior ou menor grau de centralidade, do estudo. Em outros casos, o Spaece Alfa não é mencionado. Em ambos os casos, os trabalhos podem ser divididos entre aqueles que foram publicados em período anterior à existência da Plataforma Sucupira, levando a uma análise menos profunda, com base no título do trabalho, e aqueles que foram publicados em período posterior à plataforma.

Entre os trabalhos que fazem menção ao Spaece Alfa, temos a dissertação de Olívia Coelho da Silva<sup>116</sup>, que teve como objeto de estudo o processo de formação de leitores e seu objetivo foi a investigação acerca de como ocorre o trabalho pedagógico de professores(as) do ciclo de alfabetização para o desenvolvimento das habilidades relacionadas à leitura nos alunos. O Spaece Alfa aparece de forma muito incidental no texto, sem nenhuma centralidade. Não há tópico próprio destinado a apresentar ou a descrever o programa, que aparece a partir da faceta de seus resultados e como política que influencia, sobremaneira, as ações pedagógicas na escola, a partir da responsabilização que a avaliação prevê: “Essa avaliação externa influencia consideravelmente as ações pedagógicas da escola, uma vez que está atrelada a uma política de premiação” (SILVA, 2016b, p. 105). Incidental é também o papel desempenhado pelo Spaece Alfa na dissertação de Diana Nara da Silva Oliveira<sup>117</sup>, que tem com o objetivo analisar o papel desempenhado por uma escola (como foco de análise no projeto pedagógico e nas práticas docentes) em um assentamento de reforma agrária em Jaguaruana/CE. Mais uma vez, o programa de avaliação aparece apenas a partir de seus resultados, sem que uma descrição ou apresentação do mesmo tenha lugar. É através dos seus resultados que o Spaece Alfa aparece na dissertação de Silvana Mendes Sabino Soares<sup>118</sup>. O estudo tem como objeto de estudo a influência da “cultura dos resultados das avaliações externas” (SOARES, 2017, p.7) sobre as práticas pedagógicas, em especial, no condicionamento do aluno “a uma visão da leitura

<sup>116</sup> Intitulada *Formação do leitor: desafio à prática docente e à avaliação da aprendizagem*. Defendida em 2016, no Programa de Pós-Graduação em Educação da UFC.

<sup>117</sup> Defendida em 2017, no Mestrado em Educação e Ensino, da Universidade Estadual do Ceará, com o título *A educação do/no campo em escola de assentamento de reforma agrária em Jaguaruana/Ceará*.

<sup>118</sup> Com o título *O trabalho pedagógico com a leitura na educação infantil e no ensino fundamental: investigações sobre ensino e avaliação*, a dissertação foi defendida em 2017, no Programa de Pós-Graduação em Educação da UFC.

enfadonha e mecânica” (idem), preocupada com o aprender a ler e com a decodificação, em prejuízo da leitura por fruição e prazer. Trata-se, portanto, de uma visão predominantemente pejorativa da avaliação do Spaece Alfa (não há nenhum tópico específico destinado a explicar ou a descrever o programa).

O trabalho de Helen Cristina Vieira Costa<sup>119</sup> faz uma referência tão superficial ao Spaece Alfa que hesitei em classificá-lo neste grupo. Contudo, como adotei um critério objetivo (a presença ou não de referência ao programa, sem qualificar a eventual referência em termos de profundidade da abordagem), decidi por segui-lo e incluir o trabalho aqui. A referência, a rigor, é ao Spaece, como um todo, e não ao Spaece Alfa em específico, mas o trabalho tem como o foco o 2º ano do ensino fundamental (o que reforça o critério objetivo para sua inclusão aqui). O programa é listado, junto com Pisa e Prova Brasil, como programas que “evidenciam a dificuldade que os alunos têm em produzir textos de qualidade e de compreender o que leem” (COSTA, 2018, p.8). Seu objetivo é investigar o trabalho pedagógico com a produção textual de crianças do 2º ano. Na mesma situação, encontra-se o trabalho de Dogival Alencar da Silva<sup>120</sup>. Sua dissertação de mestrado pesquisou o nível de desenvolvimento do pensamento formal e sua relação com o desempenho em Língua Portuguesa e Matemática no ensino médio. Os resultados de desempenho apresentados dizem respeito ao Spaece, e não ao Spaece Alfa, que aparece apenas como parte de uma breve descrição do sistema Spaece no Ceará. Apesar disso, seguindo os critérios selecionados para a categorização das dissertações e teses, o fato de ter mencionado o Spaece Alfa faz com que o trabalho seja computado como parte deste grupo. Ainda no âmbito dos trabalhos que apenas fazem uma breve referência ao Spaece Alfa está a dissertação de Maria Ideuma Saraiva Amaral<sup>121</sup>, que tem como foco a relação entre a gestão de resultados e as avaliações externas. A autora optou por realizar um estudo de caso em uma escola municipal do Ceará para analisar essa relação. Os resultados do Spaece Alfa são apresentados brevemente, mas a concentração da análise recai sobre os resultados do Spaece.

A dissertação de Ana Paula Pequeno Matos<sup>122</sup> é outro caso de trabalho que apenas cita o Spaece Alfa de modo incidental, tendo em vista que seu objetivo é analisar a forma como a

---

<sup>119</sup> Intitulada *Mão guiando mãos: o trabalho pedagógico com a produção textual de crianças do 2º ano do ensino fundamental no município de Fortaleza*. Defendida em 2018, no Programa de Pós-Graduação em Educação da UFC.

<sup>120</sup> Defendida em 2013, no Programa de Pós-Graduação em Educação da UFC, com o título *Nível do desempenho do pensamento formal e desempenho em Português e Matemática no Ensino Médio*.

<sup>121</sup> *Gestão por resultado e avaliações externas: estudo de caso na EMEIEF Governador César Cals de Oliveira Filho (Maracanaú, 2013-2015)* é o título da dissertação, defendida em 2017, no Mestrado Profissional em Planejamento e Políticas Públicas, da Universidade Estadual do Ceará.

<sup>122</sup> Intitulada *Um estudo de caso sobre a gestão escolar no contexto dos resultados de matemática e português do Spaece*. Defendida em 2017, no PPGP/CAEd.

gestão de uma escola (um estudo de caso) usa os resultados do Spaece para alterar práticas pedagógicas e de gestão<sup>123</sup>. O Spaece Alfa é citado, por exemplo, quando a autora fala sobre as mudanças de desenho do Spaece, em 2008, e sobre o índice de qualidade educacional criado no estado. A dissertação de Francisco Igor Magalhães Mapurunga Bezerra<sup>124</sup> é mais uma na lista dos trabalhos em que o Spaece Alfa aparece superficialmente. O programa é mencionado apenas uma vez em todo o texto, quando o autor explica que o Spaece possui três focos, um deles a alfabetização. O foco do autor é a investigação do impacto do ensino noturno nos resultados do Spaece (estudo de caso com duas escolas cearenses). Soma-se a eles, a dissertação de Roberto Cláudio Bento da Silva<sup>125</sup>, um trabalho dedicado ao tema da apropriação dos resultados da avaliação, no caso, o Spaece, pela gestão escolar (o estudo envolveu duas escolas). O Spaece Alfa é mencionado apenas uma vez, quando o autor fala sobre os três braços do Spaece, incluindo sua avaliação para o ciclo de alfabetização. Uma única menção ao Spaece também é feita na dissertação de Larissa Martins Dantas<sup>126</sup>, quando a autora cita as três vertentes do Spaece atual, incluindo sua avaliação para a alfabetização. O trabalho é sobre os efeitos das avaliações externas sobre a prática docente, com foco no Spaece e no uso de seus resultados em sala de aula. Esse é o mesmo caso (menção única) da dissertação de Francisco Jucivânio Félix de Sousa<sup>127</sup>. A dissertação dedica-se ao estudo das ações de apropriação dos resultados do Spaece pela gestão de uma escola específica, para fins de aprimoramento do acompanhamento pedagógico em matemática. O Spaece Alfa é apenas mencionado quando o trabalho fala das mudanças pelas quais o Spaece passou ao longo de sua história. O foco de apresentação dos resultados fica mesmo com o Spaece. A dissertação de Diva Lima<sup>128</sup> é mais uma a fazer parte desse grupo de trabalhos: apenas cita o Spaece Alfa ao falar dos três focos do Spaece (um deles a alfabetização). O objetivo da dissertação foi compreender a influência do Spaece sobre a gestão pedagógica de uma escola de ensino médio do município de Tauá.

---

<sup>123</sup> Como professor do PPGP/CAEd, com participação em diversas bancas de qualificação e defesa do programa, posso afirmar que estamos diante de um modelo bastante utilizado pelos alunos que se interessam pela avaliação, qual seja, a investigação de como os resultados de uma avaliação externa são apropriados pela escola.

<sup>124</sup> Com título *O impacto do ensino noturno nos resultados do Spaece: o caso de duas escolas do Ceará*. Defendida em 2014, no PPGP/CAEd.

<sup>125</sup> Intitulada *Apropriação dos resultados do Spaece pelos gestores escolares: um estudo de caso envolvendo duas escolas do interior do Ceará*. Defendida em 2014, no PPGP/CAEd.

<sup>126</sup> Defendida em 2015, no Programa de Pós-Graduação em Educação da UFC, com o título *Repercussões da avaliação externa sobre a prática docente: o caso do Sistema Permanente de Avaliação da Educação Básica do Ceará (Spaece)*.

<sup>127</sup> Intitulada *Reflexões sobre a apropriação dos resultados de matemática do Spaece – o caso da Escola do Jabuti*. Defendida em 2016, no PPGP/CAEd.

<sup>128</sup> Com título *O Sistema Permanente de Avaliação da Educação Básica do Ceará (Spaece) e sua influência sobre a gestão pedagógica de uma escola de ensino médio situada no município de Tauá – Ceará*. Defendida em 2015, no Programa de Pós-Graduação em Educação da UFC.

Também sobre o uso de resultados da avaliação educacional, mas pela secretaria de educação, é a dissertação de Linéia Moreira Maciel<sup>129</sup>. Os resultados, mais uma vez, são a faceta apresentada dos programas de avaliação (Spaace e Spaace Alfa), incluindo um subtópico somente para a apresentação dos resultados. O objetivo da dissertação foi analisar a repercussão do Spaace no setor pedagógico da secretaria de educação do município, concluindo pela necessidade de uma apropriação mais qualificada dos resultados por parte dos atores da secretaria.

O Spaace Alfa aparece de modo incidental na dissertação de Rosângela Maia de Freitas<sup>130</sup>, cujo objeto de pesquisa foi o Prêmio Escola Nota Dez e o regime de cooperação que dá base ao prêmio (especificamente, o trabalho dedicou-se a analisar a cooperação, dos pontos de vista pedagógico, financeiro e administrativo entre duas escolas do Ceará). Os resultados do Spaace Alfa são apresentados de modo sumário, mas sem que haja uma descrição ou apresentação do programa. O referido prêmio também foi o objeto de estudo da tese de Maria Océlia Mota<sup>131</sup>. No trabalho, a autora estudou a implementação do prêmio em quatro pares de escolas parceiras no Ceará, com foco na percepção e atuação dos atores implementadores. O Spaace Alfa aparece brevemente, com referência aos resultados das escolas estudadas.

A dissertação de Flávio Alves Pereira<sup>132</sup> também é sobre as políticas de responsabilização oriundas do sistema de avaliação externa do Ceará, com foco nas ações pedagógicas no caso de uma escola em específico. O tratamento dado ao Spaace Alfa é menos apressado do que aquele observado nos dois trabalhos citados no parágrafo anterior. Há um curto subtópico apresentando o Spaace e o Spaace Alfa, mas o foco da abordagem sobre o programa está mesmo em seus resultados (em especial, os do município onde situa-se a escola selecionada para o estudo).

A tese de Andreia Serra Azul da Fonseca<sup>133</sup> tem como objeto de estudo o Paic<sup>134</sup>, particularmente, o efeito do programa de modo a alterar a prática e o planejamento dos

---

<sup>129</sup> Defendida em 2013, no PPGP/CAEd, com o título O uso do Spaace pela secretaria municipal da educação de Quixadá: desafios e possibilidades.

<sup>130</sup> Intitulada O Prêmio Escola Nota Dez e o regime de cooperação estabelecido entre dois municípios cearenses. Defendida em 2018, no Programa de Pós-Graduação em Educação da Universidade Estadual do Rio Grande do Norte.

<sup>131</sup> Com título Entre a meritocracia e a equidade: o Prêmio Escola Nota Dez na percepção e atuação de agente implementadores. Defendida em 2018, no Programa de Pós-Graduação em Educação da PUC-RJ.

<sup>132</sup> Intitulada A incidência das políticas de responsabilização do estado do Ceará nas ações de gestão pedagógica em âmbito municipal – o caso da Escola Maria Nair (IPU-CE). Defendida em 2014, no PPGP/CAEd.

<sup>133</sup> Defendida em 2013, no Programa de Pós-Graduação em Educação da UFC, com o título Avaliação do Programa de Alfabetização na Idade Certa – Paic: reflexos no planejamento e prática escolar.

<sup>134</sup> A busca pela palavra PAIC, no banco de teses e dissertações da Capes, usando os mesmos critérios e filtros utilizados para as demais buscas realizadas para os programas de avaliação da alfabetização, traz 33 trabalhos como resultado (busca realizada no dia 08 de julho de 2019). O programa tem mais visibilidade do que o Spaace Alfa, em especial, por ter servido de modelo para o Paic, em nível nacional.

professores, analisando a influência das formações recebidas e dos materiais disponibilizados, além da presença das avaliações externas. É em relação a esse último ponto que o Spaece Alfa aparece no texto (juntamente com o Spaece), como um dos cinco eixos (o de avaliação externa) do Paic. Embora o programa de avaliação da alfabetização ganhe mais linhas do que em outros trabalhos, o enfoque dado a ele não é diferente: os resultados e sua influência sobre as práticas dos atores escolares.

O trabalho de Lucas Melgaço da Silva<sup>135</sup>, uma dissertação, possui um interesse diferente dos demais até então apresentados. Em parte, seu objetivo aproxima-se com o que move a presente tese. Ele busca realizar uma análise de diferentes sistemas de avaliação em larga escala na área de alfabetização, a partir de um viés comparativo. Para tanto, são analisados, no contexto de um município do Ceará, a ANA, a Provinha Brasil, o Spaece-Alfa, a Provinha Paic e Avaliação Municipal do 1º ano. As comparações entre os programas de avaliação da alfabetização realizadas pelo autor giram em torno das matrizes de referência, da elaboração e da aplicação dos instrumentos, da divulgação de resultados e de seu uso. Da mesma forma, meu objetivo com esta tese é descrever e analisar três diferentes programas de avaliação da alfabetização, com base em um conjunto amplo de critérios selecionados, entre os quais estão todos esses que foram mobilizados por Silva (2016a), contudo, o escopo é diferente, visto que o cenário escolhido por ele para análise, um município, é diferente do caminho que sigo, qual seja, a análise dos programas em si mesmo, com suas características, e não sua execução em um município específico.

A dissertação de Erineuda do Amaral Soares<sup>136</sup> dedica-se a analisar as estratégias que as escolas adotam a partir do momento em que recebem os resultados do Spaece Alfa, concluindo de que as estratégias de uso dos resultados são direcionadas exclusivamente aos alunos de 2º ano, sem que os alunos do 3º ano recebam a mesma atenção.

Dos trabalhos que não abordam, ou mesmo fazem menção ao Spaece Alfa, a tese de Alberto Filho Maciel Maia<sup>137</sup> pesquisa a influência do Projeto Professor Diretor de Turma (PPDT) para a qualidade do ensino em escolas do ensino médio em Fortaleza. A dissertação de

---

<sup>135</sup> Defendida em 2016, no Programa de Pós-Graduação em Educação da UFC, com o título Avaliações em larga escala na alfabetização: contextos no ensino público de um município do estado do Ceará.

<sup>136</sup> Intitulada Sistema Permanente de Avaliação da Educação Básica do Ceará – Alfabetização: estratégias de acompanhamento de estudantes no contexto da prática. Defendida em 2016, no Mestrado Profissional em Gestão Educacional da Universidade do Vale do Rio dos Sinos.

<sup>137</sup> Com o título Avaliação das contribuições do Projeto Professor Diretor de Turma (PPDT) para a qualidade de ensino nas escolas públicas estaduais de ensino médio de Fortaleza. Defendida em 2018, no Programa de Pós-Graduação em Educação da UFC.

Francisco das Chagas Rodrigues de Oliveira<sup>138</sup> destina-se a analisar as ações das equipes de gestão de 4 escolas do Ceará (microrregião Sertões de Crateús) em relação à formação continuada de professores em serviço. O Spaece figura como parte do trabalho, na medida em que o autor esforça-se por analisar os impactos dessas ações sobre os resultados do programa. Contudo, essa análise concentra-se no ensino médio, sem fazer referência ao Spaece Alfa (de todo modo, trata-se de mais um caso em que o programa de avaliação é visto sob a ótica de seus resultados). Os resultados do Spaece para o ensino médio também aparecem como parte da dissertação de Antônio Roberto de Araújo Souza<sup>139</sup>. O foco do trabalho foi analisar as ações da gestão de uma escola, em particular, no que diz respeito à gestão de programas e projetos, e sua influência sobre os resultados de avaliações como o Spaece e o Enem (novamente, os resultados são a faceta aparente do programa de avaliação). A dissertação de Paulo Alexandre Sousa Queiroz<sup>140</sup> teve como objetivo a proposição de uma estratégia metodológica para o ensino do conjunto de números complexos na disciplina matemática (por ser um mestrado profissional, o foco pragmático faz todo sentido). Os resultados do Spaece (e também os do Pisa) são apresentados como ponto de partida para o reconhecimento de problemas crônicos de aprendizagem em matemática, dando base à proposição pretendida pelo autor. Não há referência ao Spaece Alfa.

A dissertação de José Valmir Guimarães de Oliveira<sup>141</sup> teve como objeto de estudo a comparação entre o efeito escola e os resultados do Spaece em oito escolas da rede estadual do Ceará. Como parte dos estudos dedicados aos fatores associados ao desempenho, o trabalho procurou identificar a relação entre variáveis contextuais (como a condição socioeconômica do aluno, as práticas pedagógicas e o clima escolar) e a proficiência média nos testes do Spaece. Seu foco foi o ensino médio, portanto, sem abordar o Spaece Alfa. A dissertação de Rosângela Teixeira de Sousa<sup>142</sup> é mais um trabalho que, ao tratar do Spaece, o fez com foco nos resultados e no ensino médio. A autora propõe-se a comparar os resultados da avaliação institucional das

---

<sup>138</sup> Intitulada *Gestão escolar e formação do professor na educação pública da microrregião Sertões de Crateús/CE*. Defendida em 2013, no PPGP/CAED.

<sup>139</sup> Defendida em 2016, no PPGP/CAED, com o título *A gestão de programa/projetos como fator de sucesso: o caso da escola estadual Francisco Holanda Montenegro do Ceará*.

<sup>140</sup> Com o título *Uma proposta metodológica para o ensino de números complexos: história e prática*. Defendida em 2016, no Mestrado Profissional em Ensino de Ciências e Matemática da UFC.

<sup>141</sup> Intitulada *Análise do Sistema Permanente de Avaliação da Educação Básica do Ceará (Spaece): um estudo comparativo entre o efeito escola e os resultados de proficiência em escolas da rede estadual nos anos de 2012 a 2014*. Defendida em 2016, no Programa de Pós-Graduação em Educação da UFC.

<sup>142</sup> Com o título *Avaliação Institucional nas escolas estaduais do município de Fortaleza: uma análise comparativa dos resultados do Sistema Permanente de Avaliação da Educação Básica do Ceará-Spaece*. Defendida em 2013, no Programa de Pós-Graduação em Educação da UFC.

escolas estaduais em Fortaleza, com base na percepção do professor, com os resultados de desempenho dos alunos da 3ª série do ensino médio no Spaece 2011.

O trabalho de Francisco Ramilson Holanda Luz<sup>143</sup>, uma dissertação, tem como objetivo analisar as ações da gestão de duas escolas de uma Crede cearense no que diz respeito à política de implementação da premiação de computadores para os alunos. A política de responsabilização vinculada ao Spaece é apresentada e analisada, sem qualquer menção ao Spaece Alfa. Já a dissertação de Artemizia Ribeiro Lima Costa<sup>144</sup> trata dos efeitos dos resultados do Spaece (em língua portuguesa) sobre a prática docente no município de Aracati, concluindo, com a pesquisa, que os professores da rede, em sua maioria, usam os resultados da avaliação com o objetivo de melhorar a aprendizagem dos alunos. O foco do trabalho foram os professores de 5º e 9º anos do ensino fundamental.

A dissertação de José Célio Pinheiro<sup>145</sup> estudou o papel desempenhado pelos gestores escolares para o uso dos resultados do Spaece. O trabalho teve como foco a atuação da gestão em três escolas estaduais de Quixadá, analisando as ações destinadas à divulgação e à apropriação dos resultados. O trabalho de Jobenemar Carvalho dos Santos<sup>146</sup> (uma dissertação), por sua vez, destina-se ao estudo dos Seminários Escolares de Apropriação dos Resultados do Spaece, na Crede 11, eventos realizados pela equipe técnica da coordenadoria com o intuito de mobilizar a comunidade escolar para o uso dos resultados da avaliação. Por fim, a dissertação de Dalmário Heitor Miranda de Abreu<sup>147</sup> estudou a influência dos resultados do Spaece, para os anos finais do ensino fundamental, sobre as práticas pedagógicas dos professores de matemática de quatro escolas do município de Fortaleza.

Entre os trabalhos anteriores à Plataforma Sucupira, nenhum traz o Spaece Alfa como parte do título. A dissertação de Jocelaine Regina Duarte Rossi<sup>148</sup> aborda o tema da formação continuada e sua relação com as práticas pedagógicas em alfabetização. A dissertação de

---

<sup>143</sup> Intitulada As ações gestoras que impactam na implementação da política de premiação de microcomputadores para os alunos de duas escolas da rede estadual de ensino do Ceará. Defendida em 2013, no PPGP/CAEd.

<sup>144</sup> Com o título Avaliação do efeito dos resultados do Spaece em língua portuguesa dos alunos do 5º e 9º ano na visão de docentes do município de Aracati/CE. Defendida em 2018, no Programa de Pós-Graduação em Educação da UFC.

<sup>145</sup> Defendida em 2016, no PPGP/CAEd, com o título Gestão dos resultados do Spaece pelas escolas da rede estadual de Quixadá – Ceará.

<sup>146</sup> Intitulada A organização e realização dos Seminários de Apropriação dos Resultados do Spaece na Coordenadoria Regional de Desenvolvimento da Educação – Crede 11. Defendida em 2014, no PPGP/CAEd.

<sup>147</sup> Defendida em 2018, no Programa de Pós-Graduação em Educação da UFC, com o título Avaliação do Spaece nas práticas pedagógicas dos professores de matemática.

<sup>148</sup> Intitulada Entre o estável e o fortuito: a formação continuada em serviço e as rotinas pedagógicas em alfabetização. Defendida em 2010, no Programa de Pós-Graduação em Educação da UFC.



Luciano Nery Ferreira-Filho<sup>149</sup> diz respeito ao uso de TIC's pelos professores da rede estadual do Ceará. O trabalho de Juceli Lima de Souza<sup>150</sup>, também uma dissertação, dedicou-se ao estudo da influência da performance docente e da gestão no desempenho dos alunos em matemática (como o Spaece Alfa não avalia matemática, a análise do desempenho não diz respeito ao programa; além disso, o trabalho foi defendido em período anterior ao da existência do Spaece Alfa). A dissertação de Doris Sandra Silva Leão<sup>151</sup> tem como foco o ensino médio, mais especificamente, o uso de jogos para a recuperação de alunos em matemática, ao passo que a dissertação de Aline Maria Gomes Lima<sup>152</sup> pesquisa os fatores associados à eficácia escolar em escolas municipais de Fortaleza (o Spaece Alfa não prevê a aplicação de questionários contextuais aos alunos e aos pais e as pesquisas com fatores associados ao desempenho, de maneira geral, no âmbito dos sistemas próprios de avaliação no Brasil, tendem a lançar mão justamente dessas informações para esse tipo de investigação).

A tese de doutorado de Dogival Alencar da Silva<sup>153</sup> (sua dissertação foi classificada no grupo dos trabalhos que abordam o Spaece Alfa) não foi disponibilizada ainda no banco de teses e dissertações. Embora tenha sido defendida após o surgimento da Plataforma Sucupira, o trabalho encontra-se indisponível (provavelmente, pela defesa recente, em 2018). Desse modo, o trabalho foi analisado como se tivesse sido publicado antes da plataforma, com base no título. Apesar da ausência de possibilidade de uma consulta mais profunda, o trabalho tem como foco o ensino médio e não a alfabetização, e não faz nenhuma referência ao Spaece Alfa no título ou no resumo. A tese de Alessio Costa Lima<sup>154</sup> encontra-se na mesma situação (o trabalho não está disponível para consulta no banco). De todo modo, pelo título e pelo resumo, nenhuma referência é feita ao Spaece Alfa, visto que o trabalho destina-se a analisar a eficácia escolar no ensino médio, baseando-se em dados do Spaece, sem vínculo com o programa no âmbito da alfabetização.

---

<sup>149</sup> Defendida em 2012, no PPGP/CAEd, com o título O uso de tecnologias da comunicação e da informação pelos professores da rede pública estadual do estado do Ceará.

<sup>150</sup> Defendida em 2000, no Programa de Pós-Graduação em Educação da UFC, com o título Influência da performance docente e da gestão escolar no desempenho dos alunos em matemática (estudo de caso).

<sup>151</sup> Com título Recuperação da aprendizagem na disciplina matemática mediante jogos: experiência com alunos do 1º ano do ensino médio de uma escola pública cearense, foi defendida em 2009, no Programa de Pós-Graduação em Educação da UFC.

<sup>152</sup> Intitulada Fatores associados à eficácia escolar: um estudo de instituições públicas municipais de Fortaleza/CE. Defendida em 2012, no Programa de Pós-Graduação em Educação da UFC.

<sup>153</sup> Defendida em 2018, no Programa de Pós-Graduação em Educação da UFC, com o título Avaliação da aprendizagem dos alunos do ensino médio em diferentes enfoques curriculares: pressupostos, concepções e práticas avaliativas.

<sup>154</sup> Intitulada A eficácia escolar no ensino médio: uma avaliação longitudinal na rede estadual do Ceará. Defendida em 2015, no Programa de Pós-Graduação em Educação da UFC.

O ensino médio também é a etapa de escolaridade sobre a qual debruçou-se a pesquisa que deu base à dissertação de Joílson Pedrosa de Sousa<sup>155</sup>. Seu objetivo foi analisar a influência da metodologia adotada pelo Programa Jovem de Futuro, no Ceará, sobre a aprendizagem em matemática dos alunos da 2ª série do ensino médio em uma escola no estado. O Spaece é citado, a partir de seus resultados, mas sem referências ao ciclo de alfabetização e ao Spaece Alfa. O Spaece, com foco nos resultados do ensino médio, também é mobilizado pela dissertação de Luzia Queiroz Hippolyto<sup>156</sup>, que dedica-se a analisar os resultados dos alunos do 3º ano do ensino médio em matemática e sua relação com a prática pedagógica dos professores. Ainda com destaque para o ensino médio, a dissertação de Francesca Danielle Gurgel dos Santos<sup>157</sup> pesquisou os impactos do Spaece na melhoria do ensino e aprendizagem na referida etapa de escolaridade. Já a dissertação de Ana Geovanda Mourão Rezende<sup>158</sup> estudou o uso da avaliação de desempenho de professores como ferramenta da gestão para a melhoria da prática pedagógica e dos resultados de aprendizagem dos estudantes, sem que qualquer referência seja feita ao Spaece Alfa.

A tabela a seguir apresenta a distribuição das produções sobre o Spaece Alfa, com base nos trabalhos disponíveis no banco de teses e dissertações da Capes.

Tabela 5 - Produção de teses e dissertações sobre o Spaece Alfa

Posteriores à Plataforma Sucupira		Anteriores à Plataforma Sucupira		Total
Com menção	Sem menção	Com menção	Sem menção	
19	11	0	11	41
<b>Total com menção</b>				19
<b>Total sem menção</b>				22

Fonte: Banco de teses e dissertações Capes. Tabulação própria.

Como pode ser observado, dos 41 trabalhos: i) 19 deles mencionam o Spaece Alfa e são posteriores à existência da Plataforma Sucupira; ii) 11 são posteriores à plataforma e não fazem qualquer menção ao programa; e iii) 11 são anteriores à plataforma e não fazem menção ao Spaece Alfa (não há nenhum trabalho anterior à plataforma que faça referência ao programa).

<sup>155</sup> Com título a *Aprendizagem matemática no âmbito do Programa Jovem de Futuro: foco na metodologia entre jovens*. Defendida em 2015, no Mestrado Profissional em Ensino de Ciências e Matemática da UFC.

<sup>156</sup> Intitulada *Avaliação dos resultados do 3º ano do ensino médio em matemática no Ceará e sua repercussão na prática pedagógica dos professores: um estudo descritivo a partir dos testes do Spaece nos anos 2008-2009-2010*. Defendida em 2012, no Mestrado Profissional em Ensino de Ciências e Matemática da UFC.

<sup>157</sup> Intitulada *Impactos gerados pelo Sistema Permanente de Avaliação da Educação Básica do Estado do Ceará (Spaece) na melhoria do ensino e aprendizagem no ensino médio*. Defendida em 2010, no Programa de Pós-Graduação em Educação da UFC.

<sup>158</sup> Defendida em 2012, no PPGP/CAEd, com o título *Avaliação do desempenho docente como ferramenta de gestão para melhorar a prática pedagógica dos professores e os resultados de aprendizagem dos alunos*.

O número de trabalhos que têm como objeto, abordam ou mesmo apenas mencionam o Spaece Alfa é menor do que o número de trabalhos que não o fazem (19 e 22, respectivamente). Vale notar, ainda, que, entre os 19 trabalhos no grupo com menção, 10 deles apenas citam o programa, de modo tão incidental que os trabalhos só foram computados como parte deste grupo pela adoção de um critério objetivo (ter mencionado ou não o programa ao longo do texto). A rigor, os outros nove trabalhos é que adotam uma abordagem menos ilustrativa do Spaece Alfa. Assim, dos resultados obtidos com a busca pelo programa, 46,34% dos trabalhos fazem menção ao Spaece Alfa (destes, menos da metade apenas faz breve referência ao programa, respondendo por 21,95% do total). Complementando a análise, é preciso ressaltar também que, entre os 19 trabalhos com menção ao Spaece Alfa, a maioria deles dedica-se ao tema da apropriação dos resultados, o enfoque preferido pela literatura aqui analisada para a abordagem da avaliação educacional em larga escala.

#### 4.5 ALGUMAS OBSERVAÇÕES

Neste tópico, são apresentadas algumas observações acerca do trabalho de pesquisa realizado com o banco de teses e dissertações da Capes.

##### **4.5.1 Dos trabalhos de pós-graduação e da natureza de sua abordagem**

A análise das publicações no banco de teses e dissertações da Capes, em busca de trabalhos sobre os três programas aqui pesquisados, reforça a conclusão obtida com a análise da produção veiculada pelos três periódicos e a hipótese que deu sustentação a essas buscas: a baixa produção acadêmica sobre a avaliação da alfabetização e sobre os programas estaduais de avaliação da alfabetização. Das 83 teses e dissertações analisadas (todas elas, em princípio, pela própria natureza da ferramenta de busca, deveriam fazer referência aos programas pesquisados), pouco mais da metade (43) estava relacionado, de fato, a algum dos programas de avaliação temas da presente pesquisa (4 para o Paebes Alfa, 20 para o Proalfa e 19 para o Spaece Alfa). Apenas uma delas esforçou-se em fazer um estudo comparativo entre diferentes programas de avaliação (o estadual comparado ao nacional) (Silva, 2016a) e nenhum dos trabalhos comparou programas de diferentes estados. Soma-se a isso o fato de que a maior parte dos trabalhos analisados apenas menciona o programa, sem que ele seja objeto central do estudo ou mesmo sem que receba uma descrição mais detalhada de suas características. Ainda, na maior parte dos casos, os trabalhos,

ao se referirem aos programas, o fazem através de seus resultados. Suponho que a prevalência desse enfoque deva merecer algumas palavras. Mais um detalhe: as dissertações respondem pela maior parte dos trabalhos, já que apenas 15 dos 83 são teses de doutorado.

Evidentemente, os resultados da avaliação são uma parte fundamental do programa. É o objetivo final de um sistema de avaliação, pensando que seu propósito central é dar suporte, através da produção de informações qualificadas, a decisões também qualificadas no cenário educacional, seja no nível de uma turma, de uma escola ou de uma rede de ensino. Para tanto, é preciso que os resultados sejam divulgados e, mais do que isso, provavelmente, o maior desafio atual dos programas de avaliação, que eles sejam utilizados. Isso ajuda a explicar a ênfase dada à apropriação de resultados entre os trabalhos que dedicam-se à avaliação em larga escala. Como vimos, entre as teses e dissertações que trataram dos programas de avaliação da alfabetização, a maior parte fazia referência a esses programas a partir dessa perspectiva, de uso dos resultados. Em uma pesquisa realizada com o banco de dissertações<sup>159</sup> do PPGP/CAEd<sup>160</sup>, da UFJF, mestrado profissional destinado especificamente aos temas relacionados à gestão e à avaliação da educação pública, usando como palavra de busca *apropriação*, há 34 dissertações sobre o tema (apropriação de resultados). Quando a busca é feita pela palavra *interpretação*, são mais 3 dissertações (interpretação de resultados). Um número considerável de trabalhos em um mesmo programa de pós-graduação.

Não se trata, é claro, de não reconhecer que os resultados são importantes ou justamente a parte mais visível do processo. Soma-se à sua própria e reconhecida importância a ênfase dada pela mídia e pela sociedade em geral aos resultados da avaliação. As consequências da avaliação passam pelo destaque dado aos resultados, principal alvo das críticas também. O uso do resultado, como vimos em tópico anterior da tese, é um dos principais elementos ao qual as críticas se destinam. Muitas vezes, todo o sistema de avaliação é julgado pelo uso dado a seus resultados. A rigor, a avaliação é conhecida do público pelos seus resultados e pela forma como os atores educacionais fazem uso deles. Mesmo entre os estudiosos da educação isso não é diferente. A pesquisa em eficácia escolar, tão central para o desenvolvimento dos sistemas de avaliação, ressaltando seu papel heurístico, também contribui para o protagonismo dos resultados. A investigação acerca dos fatores associados ao desempenho coloca em cena, justamente, a identificação dos elementos que afetam, em maior ou menor medida, a

---

<sup>159</sup> Vale ressaltar que a ferramenta de buscas do banco do PPGP/CAEd apresenta as mesmas limitações do que aquela utilizada pelo banco de teses e dissertações da Capes. Por isso, a busca por expressões torna-se ineficaz, devendo ser feita por palavras.

<sup>160</sup> Por meio do seguinte endereço: <http://www.mestrado.caedufjf.net/menu/dissertacoes-defendidas/>, acesso realizado entre os dias 04 e 06 de julho de 2019.

aprendizagem (no âmbito da avaliação em larga escala, encarnada em medidas de proficiência nos testes padronizados).

Sem tirar a relevância dos resultados, resta reconhecer que um programa de avaliação envolve bem mais aspectos do que esse, abrindo espaço para uma investigação mais profunda acerca dos problemas e das virtudes do programa. Os resultados fazem parte dos critérios selecionados e propostos como orientadores da análise comparativa entre três programas de avaliação da alfabetização, que será apresentada no capítulo seguinte da tese. Contudo, é um entre outros critérios. Assim como a avaliação não termina com a produção dos resultados (da medida), visto que extrapola para sua divulgação e, principalmente, para sua apropriação, ela não começa por eles. A escolha pela criação de um programa de avaliação passa por elementos políticos que antecedem a consequência dos resultados, o que justifica seu tratamento como política pública, envolvendo sua especificação por meio de instrumentos normativos e a forma como sua implementação será divulgada e conduzida. Além dos elementos definidores da agenda, a avaliação envolve uma série de outros aspectos técnicos, que vão desde a produção dos testes, passando por sua aplicação, pelo tratamento dos dados e pela seleção dos meios através dos quais eles serão divulgados, até sua apropriação pelas redes e escolas. Em cada etapa, um conjunto de decisões precisa ser tomado, permitindo a observância de inúmeros caminhos potenciais, que levariam a desenhos bastante diversificados de programas de avaliação. Essa miríade de possibilidades oferece ao pesquisador interessado um sem número de objetos de pesquisa, que vão muito além do que dizem os resultados e de como são utilizados. Mais uma vez, sinto necessidade de reforçar: os resultados são muito importantes para a avaliação, apenas não precisam obscurecer outros elementos de interesse. A relação dos pesquisadores com seus temas de pesquisa é permeada por uma série de fatores e ninguém é obrigado a escrever sobre um tema em específico. O que chama a atenção aqui é a desproporção de trabalhos sobre a avaliação da alfabetização quando este tema tem se apresentado como um elemento importante das políticas públicas educacionais dos estados.

Entre os trabalhos que mencionam um dos programas de avaliação aqui analisados, é notório que boa parte deles foi produzida no âmbito do PPGP/CAEd, em especial, no que diz respeito ao Spaece Alfa (há também dissertações sobre o Proalfa). Ao todo, entre os 83 trabalhos de pós-graduação que retornaram da busca no banco da Capes, 32 deles são dissertações defendidas no âmbito do programa. O protagonismo do PPGP na produção de trabalhos sobre avaliação educacional está relacionado, diretamente, ao escopo do mestrado e ao público atendido: trata-se de um programa destinado, especificamente, ao tema da avaliação

e da gestão da educação pública, com o objetivo de atender os atores educacionais das redes de ensino, em especial, gestores e técnicos de secretaria, além de professores e coordenadores pedagógicos. Trata-se de um público que está, majoritariamente, em serviço (trata-se de um mestrado profissional), trabalhando nas escolas ou nas secretarias de educação, e envolvido, diretamente ou indiretamente, com algum sistema estadual ou municipal de avaliação educacional, em geral, conduzido com o apoio do CAEd.

#### 4.5.2 Dos desdobramentos possíveis para a pesquisa no banco da Capes

A busca realizada no banco de teses e dissertações da Capes, conforme já foi afirmado anteriormente, apresentou limitações em virtude das características da ferramenta de busca. Outros caminhos seriam possíveis e a pesquisa poderia beneficiar-se deles. Para efeitos comparativos, seria interessante analisar o número de trabalhos produzidos sobre alfabetização e sobre avaliação, por exemplo, dando um parâmetro para a ordem de grandeza dos trabalhos que retornaram da busca pelos programas de avaliação da alfabetização.

Realizando a busca pela palavra *alfabetização*, sem uso de qualquer filtro (em especial, por área do conhecimento), o retorno é de 6.094 trabalhos. Aplicando o filtro usando os mesmos critérios para a seleção de área do conhecimento utilizados para a busca feita para os programas de avaliação da alfabetização (com a seleção das áreas *educação* – e derivadas, como, por exemplo, *educação de adultos* e *educação especial* – e *ensino* – e derivadas, como, por exemplo, *ensino de ciências e matemática*), o número de trabalhos é de 4.154.

Seria interessante comparar a produção sobre Paebes Alfa, Proalfa e Spaece Alfa com a produção sobre avaliação educacional como um todo e sobre avaliação em larga escala, em específico. Contudo, a ferramenta de buscas da Capes é limitada, em particular, quando se trata de pesquisa com base em expressões e termos compostos por mais de uma palavra. Como ocorreu com a pesquisa por Paebes Alfa e Spaece Alfa, a busca pela expressão *avaliação educacional* não leva a uma pesquisa com base na expressão como um todo, mas sim com base em cada uma das palavras das quais é formada. A busca, nesse caso, traz trabalhos com título ou palavra-chave contendo a palavra *avaliação* ou com a palavra *educacional*, não necessariamente com a expressão *avaliação educacional*, tomada como um todo. O mesmo ocorre para avaliação em larga escala. A solução seria, assim, pesquisar com base em apenas uma palavra, no caso, *avaliação*. Esse caminho mostrou-se igualmente infrutífero, em virtude do enorme contingente de trabalhos relacionados a esse tema. Numa pesquisa sem a aplicação

de qualquer filtro, fazendo a busca pela palavra *avaliação*, a ferramenta deu retorno de 235.096 trabalhos. A avaliação é tema caro para outras áreas do conhecimento, em particular, para a *ciências da saúde* (com destaque para a área médica), *engenharias*, *psicologia* e ciências ligadas à *administração*. Aplicando-se o filtro por área do conhecimento, selecionando *educação* e áreas correlatas (*educação de adultos*, *educação especial*, *ensino*, *ensino de ciências e matemática*, em especial), esse número cai para 12.292, o que ainda inviabiliza uma análise caso a caso.

É importante ressaltar que a análise das teses e dissertações aqui pesquisadas, bem como dos trabalhos publicados nos três periódicos objetos do estudo, não diz respeito à qualidade ou à pertinência do material, das hipóteses levantadas, da metodologia aplicada e das conclusões obtidas por essas pesquisas. A análise recai, exclusivamente, sobre a inclusão de um dos programas de avaliação da alfabetização que interessam a esta tese como objeto central de estudo ou parte dele. O objetivo era verificar o número de trabalhos acadêmicos destinados aos estudos desses programas, partindo da hipótese, até aqui comprovada, de que há um descompasso entre a importância desses programas como políticas públicas no âmbito da alfabetização e o interesse que desperta à pesquisa educacional.

#### **4.5.3 Da invisibilidade da avaliação da alfabetização**

Se observarmos a produção acadêmica sobre a avaliação da alfabetização aqui analisada, podemos afirmar que o tema é praticamente invisível para a literatura educacional. Em contrapartida, os programas de avaliação da alfabetização nos estados têm sido centrais para a tomada de decisões e para o desenho e implementação de políticas públicas, importância em descompasso com a falta de estudos sobre o tema. Afinal, por que a avaliação da alfabetização é invisível? Creio ser necessário levantar algumas hipóteses que ajudem a lançar luz sobre o problema.

- i) um primeiro problema gira em torno da avaliação em geral, não apenas da avaliação da alfabetização, em específico; é bem verdade que, no cenário nacional, a avaliação educacional tem se apresentado como uma política duradoura, que experimentou expansão e desenvolvimento ao longo das últimas duas décadas; isso fez com que muito da resistência inicial em relação à avaliação arrefecesse; contudo, essa resistência ainda existe e não pode ser desprezada; como vimos em tópico anterior, à avaliação são destinadas duras

críticas, que, muitas vezes, por seu radicalismo, buscam obscurecer qualquer virtude que um sistema de avaliação possa ter para as escolas e redes de ensino; se é verdade que, como qualquer outra política, a avaliação não está a salvo de críticas (e isso é essencial para seu aprimoramento, como tem sido ao longo de sua história), é preciso reconhecer que as críticas totais à avaliação, como se ela só trouxesse problemas para a educação como um todo, operando em uma lógica necessariamente contrária àquela que deveria ser a da escola, são vazias de sentido e infrutíferas, trilhando um caminho puramente ideológico, sem considerações de ordem técnica; isso contamina a visão sobre a avaliação em geral, incluindo a avaliação da alfabetização, o que pode fazer com que pesquisadores se interessem menos pelo tema; com uma crítica totalizante já pronta sobre a avaliação, o que resta para concluir sobre o tema? Apenas aqueles dispostos a ir além da crítica já estabelecida aventuram-se em compreender a avaliação mais a fundo, não para defendê-la acriticamente, mas para poder criticá-la em bases mais sólidas e fundamentadas; em relação a essa primeira hipótese, é preciso que eu mesmo levante uma problematização; se é verdade que a pronta crítica ideológica à avaliação ajuda a explicar o desinteresse pelo tema como objeto de estudo, ela o faz apenas em parte; afinal de contas, como apresentado ao longo deste tópico, há um periódico de alta qualidade destinado a publicar trabalhos sobre avaliação educacional (EAE) e muitos artigos sobre o tema têm sido publicados ao longo dos últimos anos (de 2004 a 2017, foram 404 artigos); e, mesmo na EAE, há poucos artigos produzidos sobre avaliação da alfabetização; portanto, a pouca produção sobre o tema exige mais hipóteses para que encontre explicação;

- ii) a ausência de uma avaliação da alfabetização consolidada em âmbito nacional também ajuda a explicar a baixa produção; a Provinha Brasil desempenhou um papel importante para a avaliação da alfabetização no país, em especial, ajudando a superar, entre os professores, a resistência em relação à avaliação em larga escala; no entanto, foi uma avaliação descontínua e não era propriamente uma avaliação externa, pois os professores aplicam os testes, que não usavam a TRI para a produção de medidas e não estabeleciam um processo mais denso de divulgação dos resultados; o mesmo pode ser dito sobre a ANA, que, ainda no início, teve interrupções e ainda não se sabe bem qual o seu futuro; ao contrário,



o Saeb, que não conta com aplicações no ciclo de alfabetização, possui considerável histórico e se consolidou como uma política de avaliação duradoura no âmbito nacional; com base nele, ou pelo menos tendo-o como inspiração, os sistemas próprios de avaliação, municipais e estaduais, puderam se desenvolver (como foco em avaliações, majoritariamente, nos 5º e 9º anos do ensino fundamental e no 3º ano do ensino médio); a ausência de uma avaliação da alfabetização consolidada em nível nacional faz com que o interesse pelos sistemas próprios que contam com a alfabetização fique relativamente restrito ao âmbito local, no cenário onde ocorrem; por isso, os trabalhos de pós-graduação sobre o Paebes Alfa concentram-se na UFES, os do Spaece Alfa na UFC e os do Proalfa nas universidades de Minas Gerais;

- iii) sem um programa nacional, a avaliação da alfabetização ganhou força nos estados e esse é um ponto importante para compreender a produção em torno dela; os estados são os principais entes federados com sistemas próprios de avaliação; como um programa de avaliação é muito custoso (tanto financeira como politicamente), não são muitos os municípios capazes de organizá-los e implementá-los (em geral, programas municipais de avaliação tendem a se concentrar nas capitais); isso faz com que os estados planejem seus programas de avaliação com base em suas prioridades e necessidades, imprimindo-as no desenho do programa; assim, o ensino médio e os anos finais do ensino fundamental (por conta das atribuições constitucionais) tenderam a ser as principais preocupações das redes estaduais e os sistemas de avaliação refletiram isso, como pode ser observado, por exemplo, na história dos três sistemas de avaliação dos quais os programas aqui analisados fazem parte; o ciclo de alfabetização, com o avanço do processo de municipalização, ficou sob a responsabilidade quase exclusiva (em muitos casos, já exclusiva) das redes municipais; com isso, caso existisse, a avaliação da alfabetização deveria ficar a cargo dos sistemas municipais, que, conforme já ressaltado, não foram capazes de se articular e de se desenvolver da mesma maneira que os sistemas estaduais; coube, dessa forma, aos programas estaduais a instituição da avaliação da alfabetização em seus sistemas já existentes, como ocorreu no Ceará e em Minas Gerais, apenas para ficar em exemplos dos programas que estão sendo objeto de estudo, estabelecendo com os municípios um regime de colaboração; o programa

de avaliação é estadual, mas cabe aos municípios a tomada de decisão com base em seus resultados; mesmo com essa instituição, a preocupação dos sistemas estaduais com o ensino médio, em particular, salta aos olhos; a alfabetização entrou no cenário de preocupações quando os resultados da avaliação de outras etapas mostraram que os alunos chegavam aos anos finais do ensino fundamental e mesmo ao ensino médio com sérios problemas de alfabetização (ver novamente Lima et al, 2008); tanto é assim que muitos dos trabalhos que retornaram da busca no banco da Capes diziam respeito ao programa de avaliação do ensino médio ou dos anos finais do fundamental, e não ao programa de avaliação da alfabetização.

Tenho clareza de que qualquer dessas hipóteses, isoladamente, não é capaz de explicar o pouco interesse da literatura pela avaliação da alfabetização. De todo modo, em conjunto, elas podem contribuir para a construção de um cenário que nos ajude a compreender o desajuste entre a produção da literatura educacional sobre o tema e a importância que os programas de avaliação da alfabetização vêm assumindo no âmbito dos estados e municípios.

## **5 PROGRAMAS ESTADUAIS DE AVALIAÇÃO DA ALFABETIZAÇÃO: OS CASOS DE PAEBES ALFA, PROALFA E SPAECE ALFA**

Este capítulo está organizado em quatro tópicos. No primeiro, são apresentadas as justificativas que deram base à seleção dos programas de avaliação da alfabetização a serem pesquisados e à seleção do período de análise. O segundo tópico trata das escolhas metodológicas, em especial, do tipo de pesquisa empreendido, detalhando as entrevistas realizadas. O terceiro tópico, o mais extenso deles, apresenta os critérios selecionados para a análise dos programas, acompanhados da análise efetiva para cada um deles. Por fim, o último tópico apresenta a visão dos atores das secretarias entrevistados acerca da importância dos programas de avaliação da alfabetização para a gestão das redes de ensino.

### **5.1 A ESCOLHA DOS PROGRAMAS E DO PERÍODO DE ANÁLISE**

A ausência de uma avaliação da alfabetização consolidada em âmbito nacional e a incipiência e descontinuidade de iniciativas de mesma natureza nos municípios (como no caso do Avalia BH<sup>161</sup>, por exemplo) fizeram com que os estados ganhassem protagonismo no desenvolvimento e condução de programas de avaliação para essa etapa de escolaridade. Diante disso, uma pesquisa sobre avaliação em larga escala na alfabetização deveria considerar os programas estaduais, pois esse foi um movimento, no Brasil, conduzido pelos estados. Apesar disso, permanece de pé a questão acerca de quais programas deveriam ser analisados. Um primeiro caminho seria analisar todos os programas de avaliação estaduais que avaliam ou avaliaram o ciclo de alfabetização. Contudo, essa alternativa foi rechaçada, principalmente, pelo seguinte motivo: a análise exigida para essa empreitada consumiria muito mais tempo e esforço do que os que foram aqui dedicados, sem que isso trouxesse a garantia de que a pesquisa da avaliação da alfabetização em larga escala ficasse mais completa do que a análise restrita aos principais programas, já que muitos dos sistemas de avaliação que aplicaram testes para alfabetização foram rapidamente descontinuados e se baseavam, substantivamente, nos moldes desses programas principais.

Mas, afinal, quais seriam os programas principais? Um primeiro critério para sua escolha foi a longevidade. Programas que possuíam um histórico, uma linha do tempo

---

<sup>161</sup> Programa da prefeitura de Belo Horizonte, destinado a avaliar a rede municipal da cidade, com testes sendo aplicados do ciclo de alfabetização ao 9º ano do ensino fundamental. O programa aconteceu de 2008 a 2015.

considerável, permitem uma análise mais profunda e progressiva sobre a avaliação da alfabetização, incluindo as mudanças que eventualmente ocorram no desenho da política. Por esse critério, o Saepe, o sistema de avaliação de Pernambuco, deveria ter sido incluído, já que realizou uma primeira avaliação da alfabetização (2ª série/3º ano) ainda no ano 2000, e avalia, desde 2008, de modo contínuo essa mesma série/ano. Apesar disso, as políticas educacionais de Pernambuco e, por consequência, o Saepe, têm como foco o ensino médio, como pode ser observado pelo peso que o estado tem dado para a expansão de sua rede de escolas em tempo integral nessa etapa de escolaridade. Vale ressaltar que isso não é um demérito da política de avaliação do estado, mas uma escolha, uma decisão acerca do que será priorizado. O estado, por exemplo, não apresenta nenhuma política de destaque associada aos resultados do Saepe para o ciclo de alfabetização, além de não ter apresentado mudanças e inovações importantes nos testes de alfabetização o longo do tempo. Por esses motivos, o Saepe não foi incluído na análise.

Se a longevidade é necessária, mas não suficiente, o que mais deveria ser considerado para a seleção dos programas de avaliação? As políticas a eles associadas foi um dos critérios, como veremos a seguir (incluindo as de responsabilização). Isso fez com que tais programas tivessem mais visibilidade (mesmo diante da patente invisibilidade que os caracteriza no âmbito da pesquisa educacional, como vimos no capítulo 3 da tese) no cenário da avaliação educacional. Essa visibilidade, por vezes, é percebida a partir de elemento de difícil mensuração, mas, nem por isso, menos relevantes. A minha experiência de trabalho com a divulgação de resultados dos programas de avaliação executados em parceria com CAEd/UFJF, entre 2011 e 2019, colocou-me, inúmeras vezes, diante de cenários em que esses programas de avaliação apareciam como referência para os atores educacionais de outros estados, tanto no nível das secretarias quanto no das escolas. Explico-me. Anualmente, eu viajava aos estados avaliados em parceria com o CAEd para divulgar os resultados da avaliação do programa daquele estado e era questionado, pelos secretários de educação, técnicos da secretaria, diretores e professores das escolas avaliadas, acerca dos programas de avaliação de outros estados e de seus resultados. As questões giraram em torno do tipo de teste aplicado, da matriz utilizada, das disciplinas e séries avaliadas, de como os resultados eram divulgados etc. Invariavelmente, nesses situações, os mesmos programas despertavam interesse: O Spaece Alfa, do Ceará, o Proalfa, de Minas Gerais, e o Paebe Alfa, do Espírito Santo.

Não tenho registro de quantas vezes fui a estados e municípios divulgar resultados de avaliação e nem de quantas vezes fui questionado sobre os três programas acima mencionados.

O fato é que era um fato tão recorrente que me arrisco a afirmar que nunca participei de uma divulgação de resultados em que uma pergunta sobre um desses programas não tivesse sido feita. Embora essa informação e sua fonte não possam ser validadas pelos critérios científicos atuais, sua existência foi um dos motivos que levaram à escolha dos três programas como objeto de pesquisa. Não poderia, portanto, deixar de mencionar isso ao justificar tal escolha.

O Ceará vinha sendo tomado como referência, principalmente, diante do avanço diagnosticado pelos resultados. O estado vinha apresentando melhorias notáveis nos resultados de alfabetização. O Paic foi tomado como modelo de programa de alfabetização para o Pnaic, do governo federal, e o município de Sobral tornou-se exemplo de combate ao analfabetismo. O Ceará começou a ser entendido como exemplo de um estado que, partindo de baixos desempenhos, era capaz de mudar o cenário do analfabetismo. Sua referência extrapolou o nordeste e o Spaece alfa passou a despertar o interesse de outros estados.

O estado de Minas Gerais, durante quase todo o período analisado (2009-2017), foi tomado como exemplo de bons resultados na alfabetização. Comparado com outros estados do Brasil, o estado mantinha-se com resultados consolidados e considerado em um bom patamar. O Ideb impulsionou ainda mais essa imagem do estado, já que Minas apresentou os melhores resultados no Ideb - Anos Iniciais durante algumas edições do índice. Por tratar-se de um estado muito extenso, populoso e heterogêneo, Minas chamou a atenção do país com bons resultados na alfabetização mesmo diante dessa heterogeneidade econômica, social e cultural. O Proalfa fazia parte dessa arquitetura das políticas educacionais mineiras e também era tomado como referência por outros estados.

O Espírito Santo não despontou ao mesmo tempo em que Ceará e Minas Gerais como referência para a avaliação da alfabetização. Mesmo tendo realizado avaliações no ciclo de alfabetização há quase tanto tempo quanto os dois outros estados, sua preponderância foi construída ao longo do tempo. O Paebs Alfa passou a chamar a atenção não apenas pelos resultados das redes, como ocorreu nos casos de Spaece Alfa e Proalfa, mas pelo programa de avaliação em si. Para os especialistas entrevistados no âmbito desta pesquisa de doutorado, o Paebs Alfa tornou-se a principal referência de programa de avaliação da alfabetização, em virtude de suas características técnicas e da completude sua proposta.

Esse foi mais um dos critérios utilizados para a escolha dos três programas. A experiência de trabalho com o CAEd não me trouxe apenas essas impressões colhidas ao longo das devolutivas, mas também o trabalho conjunto com os especialistas responsáveis pelo desenvolvimento dos programas de avaliação, da construção das matrizes à divulgação dos

resultados, passando pela elaboração dos itens e testes, por sua aplicação e pela elaboração das medidas. De acordo com esses especialistas, os três programas aqui estudados são as principais referências em avaliação da alfabetização que possuímos no Brasil.

Minha atuação no CAEd influenciou também a escolha do período de tempo a ser analisado. Entre 2009 e 2017, temos 9 edições dos programas de avaliação, quase uma década para estudo. Acredito que isso seja tempo suficiente para produzir uma análise mais detida dos programas, de modo a poder analisar eventuais mudanças que tenham ocorrida. Assim, o critério da longevidade estaria atendido. Mas por que o início em 2009? De 2009 a 2017, os três programas selecionados foram desenvolvidos em parceria com o CAEd, o que representava uma grande vantagem em termos de obtenção da informação e permitia uma análise comparativa mais adequada dos programas. Como o intuito era, desde o início, adotar uma perspectiva metodológica qualitativa que pudesse dar voz a atores importantes no processo de avaliação, meu acesso a esses atores era facilitado em virtude do meu tempo de trabalho na instituição e do meu conhecimento em relação às funções que eles ali exerciam. Apesar disso, não foi tarefa simples obter as informações que eu desejava para a produção do trabalho. Na verdade, muitas delas não foram obtidas, me levando a reajustar alguns caminhos previstos pela pesquisa.

Esse é um ponto importante em pesquisas dessa natureza, que se alongam ao longo do tempo. As instituições, e acredito que possa afirmar isso com tal grau de generalidade, não registram suas informações, pelo menos, não no Brasil, esperando que pesquisadores e estudantes de pós-graduação as solicitem para seus trabalhos e teses. As instituições registram aquilo que julgam importante para seus próprios fins e objetivos. Por vezes, as informações disponíveis são aquelas necessárias para o trabalho do pesquisador, mas, na maior parte dos casos, não é isso o que acontece. A informação pode não ter sido registrada ou ter sido registrada de tal forma que sua obtenção torna-se custosa. Ela pode estar pulverizada em diversos documentos, de naturezas distintas, o que levaria a um processo de sua montagem, recolhendo dados aqui e acolá. Para isso, costuma ser indispensável a ajuda de algum membro da instituição que conheça bem esses documentos, saiba localizá-los e localizar as informações neles. Nem sempre isso é possível.

Se 2009 é o ponto de partida, por que 2017 foi tomado como ponto final da análise? A resposta mais direta tem a ver com precaução. Como já mencionei, 9 edições, em minha leitura, é tempo suficiente para realizar uma análise contundente da política. Aliás, esse é um período grande o suficiente para gerar dificuldades para a análise, tendo em vista o que expus nos

parágrafos anteriores, acerca da busca por informação. Quando mais longe você volta no tempo, mais difícil é conseguir as informações desejadas. Soma-se a isso uma característica dos sistemas de avaliação. As edições referem-se a um ano, mas seus efeitos se estendem ao longo do ano seguinte. Isso quer dizer que, por exemplo, a edição de 2017 do Paebes Alfa aplicou testes no próprio ano de 2017, no último trimestre do ano. Discussão de matrizes, elaboração de itens e testes e sua aplicação são ações do programa que ocorrem em 2017. Entretanto, a avaliação não se encerra com sua aplicação. Depois de aplicados, os testes são recolhidos, processados, as bases são tratadas, as medidas de proficiência são produzidas e os resultados divulgados. A divulgação de resultados realizada pelo CAEd em parceria com as secretarias de educação é, em regra, efetivada através de oficinas presenciais e de cursos online. Todas essas etapas posteriores levam tempo e adentram o ano seguinte ao da aplicação. Assim, no exemplo dado, a edição de 2017, na verdade, estende-se pelo ano de 2018. Assim, as edições analisadas foram as de 2009 até 2017, mas, cronologicamente, isso significa que o período total de análise foi de 2009 a 2018. Como a previsão de defesa para a tese era o início do ano de 2020, a ideia foi reservar o ano de 2019 para a análise do material recolhido e escrita do texto propriamente dito, sem que eu precisasse dedicar energia e atenção à análise de mais um ciclo de avaliação, que ainda estaria acontecendo enquanto eu produzia a tese.

## 5.2 METODOLOGIA

A metodologia adotada para a realização da pesquisa foi de natureza qualitativa, lançando mão de pesquisa bibliográfica, pesquisa documental e entrevistas semi-estruturadas (individuais e coletivas) para a coleta de informações. A pesquisa bibliográfica foi o instrumento utilizado, primordialmente, para a construção dos 3 primeiros capítulos, sobre políticas públicas, políticas educacionais, de alfabetização e de avaliação. O 3º capítulo, em específico, envolveu também uma espécie de pesquisa documental, tendo em vista a análise feita em relação aos periódicos e ao banco de teses e dissertações da Capes. Naquele momento, além do conteúdo publicado pelos artigos e trabalhos de pós-graduação, interessava saber acerca da incidência do tema da avaliação da alfabetização nas revistas, nas teses e nas dissertações, de modo a produzir um mapeamento da pesquisa educacional sobre o tema.

O presente capítulo 5, por sua vez, usa a pesquisa bibliográfica de modo incidental, apenas quando necessária para esclarecer algum ponto ou como indicação para um estudo mais aprofundado, tendo sido construído, principalmente, com base na pesquisa documental e nas

entrevistas semi-estruturadas. A pesquisa documental foi importante para a coleta de informações sobre vários dos critérios selecionados para análise, como o desenho da pesquisa (matrizes de referência, padrões de desempenho, linha histórica da avaliação, com séries e disciplinas avaliada etc), legislação de referência, políticas de *accountability* associadas, entre outros.

As entrevistas semi-estruturadas<sup>162</sup> desempenharam, pelo menos, duas funções importantes: i) dar voz a atores envolvidos com os programas de avaliação, de modo que a análise pudesse se basear em algo além dos documentos disponíveis, o que abre espaço para avaliações e opiniões acerca da política de avaliação; e ii) complementar as análises naquilo que a pesquisa documental mostrou-se insuficiente. As entrevistas, de fato, foram muito importantes para essa segunda função, embora alguns pontos ainda tenham ficado em aberto. Sobre eles, voltarei a tratar nas conclusões da tese. Embora com roteiro, as entrevistas nem sempre seguiram exatamente o previsto. À medida que temas conexos aos previstos pelo roteiro foram aparecendo, as entrevistas foram sendo guiadas nessa direção.

Ao todo, foram realizadas 11 entrevistas individuais: 4 com atores educacionais que estiveram na secretaria de educação de Minas Gerais, trabalhando com o Proalfa, ao longo de determinado período de tempo dentro daquele estabelecido para análise; 3 com atores educacionais que estiveram na mesma situação com o Spaece Alfa; 2 com atores que trabalharam com o Paebes Alfa (todos nos nível da secretaria); 1 com a supervisora<sup>163</sup> responsável, no CAEd, pelo processo de divulgação de resultados dos programas de avaliação que têm parceria com a instituição, e 1 com a responsável pela equipe de capacitação dos aplicadores de teste no CAEd. Além delas, foram realizadas 3 entrevistas coletivas: 1 com a equipe de produção de instrumentos de avaliação para a alfabetização do CAEd, da qual participaram três especialistas, incluindo a supervisora de montagem de testes; 1 com dois especialistas do CAEd em avaliação da escrita; e 1 com 5 especialistas da equipe de produção de instrumentos de avaliação para a alfabetização do CAEd, destinada a discutir a avaliação da matemática no ciclo de alfabetização. É importante ressaltar ainda que foram selecionados, inicialmente, 4 atores educacionais de cada programa para a realização das entrevistas. Em Minas Gerais, todos responderam, ao passo que 1 pessoa no Ceará e 2 no Espírito Santo aceitaram participar da pesquisa, mas não responderam às questões propostas em tempo hábil para a escrita da tese. As entrevistas coletivas e a entrevista individual com a responsável pelo

---

<sup>162</sup> Os roteiros das entrevistas estão disponíveis no Anexo IV.

<sup>163</sup> Essa pessoa assumiu o cargo em 2019, função que eu ocupava até então, mas já trabalhava no CAEd, com divulgação de resultados, há mais tempo.



processo de aplicação foram todas realizadas presencialmente, em duas salas de reunião do CAEd, em Juiz de Fora/MG. As demais entrevistas individuais, com os atores das secretarias e com a supervisora de divulgação de resultados do CAEd ocorreram de modo não presencial: as entrevistas foram enviadas e os sujeitos da pesquisa responderam posteriormente às questões propostas, enviadas por meio digital.

Em relação ao Paebes Alfa, as entrevistadas foram: i) Andressa Buss Rocha - foi assessora de Planejamento e Gestão Estratégica da Seduc<sup>164</sup> do Espírito Santo e, entre 2016 e 2018, Subsecretária de Planejamento e Avaliação na secretaria de educação; e ii) Renata da Silva Marinho - foi subgerente de avaliação educacional da Seduc até o início do ano de 2020. O entrevistados para o Proalfa foram: i) Ana Lúcia Gazzola - foi secretária de educação do estado de Minas Gerais entre 2011 e 2014; ii) Maria das Graças Pedrosa Bittencourt – na Seduc de Minas, foi gerente do Programa de Intervenção Pedagógica para o ensino fundamental, entre os anos de 2007 e 2013, sendo consultora para o programa até o ano de 2014; iii) Sueli Pires – foi secretária adjunta de educação do estado de Minas, entre 2012 e 2014; e iv) Raquel Elizabete de Souza Santos – foi subsecretária de desenvolvimento da educação básica do estado de Minas Gerais até 2014. Sobre o Spaece Alfa, os entrevistados foram: i) Maurício Holanda Maia, ex-secretário de educação do estado do Ceará (2014-2016) e ex-secretário adjunto da educação do estado do Ceará (2007-2014); ii) Alessio da Costa Lima, foi orientador da Célula de Avaliação do Desempenho Acadêmico da Secretaria da Educação do Ceará e é ex-presidente da Undime Nacional; iii) Márcia Oliveira Cavalcante Campos, foi uma das idealizadoras do Paic, coordenadora da Cooperação com os municípios na Seduc Ceará, entre 2007 e 2012, e foi secretária adjunta da educação da rede municipal de Fortaleza entre 2013 e 2016. O quadro a seguir sumariza as informações sobre as entrevistas realizadas.

---

<sup>164</sup> Embora a secretaria do Espírito Santo faça referência a si própria como Sedu e a secretaria de Minas como SEE, uso Seduc de modo genérico para todas elas, querendo me referir à secretaria de estado da educação.

Quadro 2 - Entrevistas para a pesquisa, atores e temáticas.

Modalidade de Entrevista	Formato	Entrevistados	Instituição do entrevistado (à época da entrevista ou em virtude de cargo já exercido anteriormente)	Temática central
Coletiva	Presencial	2 especialistas em elaboração de itens para avaliação da alfabetização e 1 especialista responsável pela montagem dos testes	Equipe de elaboração de instrumentos para avaliação da alfabetização - CAEd/UFJF (2 pessoas); Equipe de montagem de testes (supervisora) - CAEd/UFJF	Matrizes de referência, modelos de teste e elaboração de itens para os três programas analisados, com foco em língua portuguesa-leitura
Coletiva	Presencial	4 especialistas em elaboração de itens para avaliação da alfabetização e 1 especialista da equipe de montagem de testes (especialista em matemática)	Equipe de elaboração de instrumentos para avaliação da alfabetização - CAEd/UFJF (4 pessoas); Equipe de montagem de testes (especialista em matemática) - CAEd/UFJF	Matrizes de referência, modelos de teste e elaboração de itens para os três programas analisados, com foco em matemática
Coletiva	Presencial	2 especialistas em avaliação da escrita	Equipe de avaliação da escrita - CAEd/UFJF	A avaliação da escrita nos três programas analisados
Individual	Presencial	Responsável pela capacitação dos aplicadores dos testes	Equipe de aplicação de testes do CAEd/UFJF	Aplicadores e sua capacitação e aplicação dos testes nos três programas analisados
Individual	Não presencial	Ex-Secretária de Educação de Minas Gerais	Secretaria de Estado da Educação de Minas Gerais	Importância do Proalfa para a tomada de decisão na secretaria e para o desenho de outras políticas no estado
Individual	Não presencial	Ex-Subsecretária de desenvolvimento da educação básica de Minas Gerais	Secretaria de Estado da Educação de Minas Gerais	Importância do Proalfa para a tomada de decisão na secretaria e para o desenho de outras políticas no estado
Individual	Não presencial	Ex-Secretária Adjunta de educação de Minas Gerais	Secretaria de Estado da Educação de Minas Gerais	Importância do Proalfa para a tomada de decisão na secretaria e para o desenho de outras políticas no estado
Individual	Não presencial	Ex-Gerente do Programa de Intervenção Pedagógica da Secretaria de Educação de Minas Gerais	Secretaria de Estado da Educação de Minas Gerais	Importância do Proalfa para a tomada de decisão na secretaria e para o desenho de outras políticas no estado
Individual	Não presencial	Ex-Secretário de Educação do Ceará	Secretaria de Estado da Educação do Ceará	Importância do Spaee Alfa para a tomada de decisão na secretaria e para o desenho de outras políticas no estado
Individual	Não presencial	Ex-Coordenador da Célula de Avaliação do Desempenho Acadêmico na Seduc do Ceará	Secretaria de Estado da Educação do Ceará	Importância do Spaee Alfa para a tomada de decisão na secretaria e para o desenho de outras políticas no estado
Individual	Não presencial	Uma das idealizadoras do Paic e ex-Coordenadora da Cooperação com os municípios no estado do Ceará	Secretaria de Estado da Educação do Ceará	Importância do Spaee Alfa para a tomada de decisão na secretaria e para o desenho de outras políticas no estado
Individual	Não presencial	Ex-Subsecretária de planejamento e avaliação da Seduc Espírito Santo	Secretaria de Estado da Educação do Espírito Santo	Importância do Paebes Alfa para a tomada de decisão na secretaria e para o desenho de outras políticas no estado
Individual	Não presencial	Ex-Subgerente de avaliação educacional da secretaria de educação do Espírito Santo	Secretaria de Estado da Educação do Espírito Santo	Importância do Paebes Alfa para a tomada de decisão na secretaria e para o desenho de outras políticas no estado

Fonte: Elaboração do próprio autor.

### 5.3 OS CRITÉRIOS PARA ANÁLISE

Como vimos no capítulo 1 da tese, a criação de tipologias é uma característica da pesquisa sobre políticas públicas. Os tipos ajudam a mapear o campo das políticas, identificando semelhanças entre políticas de diferentes áreas, o que justifica sua reunião em tipos. Os tipos sintetizam, reúnem, permitindo análises mais gerais sobre as políticas, e impedindo que cada política seja tomada como tão idiossincrática que sua comparação com qualquer outra política torna-se inadequada, ou mesmo impossível. Dois exemplos que discutimos são as tipologias de Lowi e de Deubel. A tipologia de Lowi (MAIA; GRANDO, 2016) é demasiadamente generalista e com baixo poder explicativo para os objetivos da tese. Seu uso não levaria à identificação de diferenças mais finas entre as políticas de avaliação da alfabetização dos três estados aqui analisados. A tipologia de Deubel (2009) destina-se mais a orientar a formulação de políticas do que propriamente a analisá-las. Assim como Lowi, ele se interessa pela criação de uma tipologia de toda e qualquer política, de modo que as especificidades de uma política como a avaliação da alfabetização não são captadas. Por conta disso, nenhuma das duas tipologias foram aqui adotadas para a análise dos três programas de avaliação da alfabetização.

O caminho seguido foi outro<sup>165</sup>: selecionar critérios que permitissem um mapeamento mais específico das políticas de avaliação da alfabetização, identificando pontos em comum e diferenças entre os programas em pontos de interesse para a área de avaliação da alfabetização, e não, de modo genérico, para a área de políticas públicas em geral. Esses critérios foram selecionados de modo arbitrário, observando o binômio: o que foi considerado importante para a avaliação da alfabetização/o que foi passível de ser analisado em um esforço dessa natureza (uma tese de doutorado). O reconhecimento do que há de arbitrário na escolha não é um demérito, mas uma contingência. O controle dessa arbitrariedade é dado pela argumentação: a apresentação de cada um dos critérios selecionados é acompanhada pela justificativa de sua escolha, de modo que o leitor/interlocutor possa ser convencido de sua plausibilidade ou, como é parte do fazer científico, direcione argumentos para rechaçá-la. Contribuem para a arbitrariedade da seleção desses critérios a ausência de literatura sobre o tema e a falta de interesse da pesquisa educacional pela avaliação da alfabetização. Talvez essa seja uma das contribuições que a tese possa dar para o campo.

---

<sup>165</sup> Uma discussão mais aprofundada sobre os limites das duas tipologias (Lowi e Deubel) e sobre algumas problematizações acerca da combinação dos critérios propostos pode ser encontrada no Apêndice IV.

Cada critério selecionado serve à construção de um tipo de programa tendo em vista aquele critério. Outro caminho seria reunir os critérios para produzir os tipos de acordo com essa reunião. Os limites desse segundo caminho são apresentados no Apêndice IV. Aqui, para os efeitos da análise, adotamos a ideia de cada critério produz um tipo de programa em relação a ele (critério como elemento gerador do tipo). O quadro a seguir apresenta os critérios selecionados.

Quadro 3 - Critérios para a construção de uma tipologia de políticas de avaliação da alfabetização

Critérios	Base para descrição
1. Alcance da política	Qual é o alcance da política? Sua amplitude. Ele envolve toda a rede de ensino ou apenas uma parte dela?
2. Público-alvo	Qual o público-alvo da política? A quem ela se destina prioritariamente? Quais são os destinatários secundários?
3. Contexto de criação (agenda) e modelo de implementação	Como a agenda em torno da política foi construída? É possível rastreá-la? Há registro desse momento? Quais agentes atuaram decisivamente para que a agenda fosse definida? Como a política foi implementada? O modelo utilizado foi <i>top/down</i> ? <i>Bottom-up</i> ? Houve participação dos <i>bureaucracy street-level</i> ou apenas dos tomadores de decisão de níveis mais altos?
4. Desenho da avaliação	Quais são os instrumentos mobilizados pela política? Quais são as séries e disciplinas avaliadas? Qual é a escala de proficiência? Como é realizada a aplicação? Quais são as matrizes de referência?
5. Divulgação e uso dos resultados (natureza dos canais de divulgação e comunicação)	Os resultados da política são divulgados? Se sim, por que meios e para quem? Como é o acesso a esses resultados? Que tipo de informação é disponibilizada? Quanto tempo depois da avaliação os resultados são disponibilizados? A entrega, a compreensão e a utilização dos resultados são dimensões levadas em consideração pela divulgação?
6. Concepção de alfabetização	Qual é a concepção de alfabetização que dá base à elaboração das matrizes de referência e construção dos testes?
7. Vínculo com outras políticas educacionais e geração da política ( <i>accountability</i> )	A política se relaciona com outras políticas educacionais? Se sim, quais são elas? Há política de responsabilização atrelada aos resultados da avaliação? Se sim, de que tipo? As consequências dessa política são brandas ou severas para seus participantes? À qual geração de avaliação educacional pertence a política?
8. Longevidade (duração e constância)	Qual a constância e duração da política? Que transformações ela sofreu ao longo do tempo?
9. Participação dos municípios	Como é a participação dos municípios no programa? Ela é feita por adesão? Os gastos com a participação no programa são cobertos pelo estado ou o município também tem gastos?
10. Volume e transparência dos recursos aplicados em avaliação	Os recursos são divulgados com transparência? Há prestação de contas para a população em geral? Onde essa divulgação é feita? Qual o volume de recursos destinado à avaliação quando comparado aos gastos gerais dos estados com educação?

Fonte: Elaboração do próprio autor.

Muitos dos critérios propostos aqui para a construção da tipologia poderiam ser aplicados a outras políticas. No entanto, há critérios específicos para os tipos de políticas de avaliação da educação, como é o caso do desenho da avaliação, da forma de divulgação dos resultados e do tipo de *accountability* que mobiliza, e um critério específico para a análise de políticas de avaliação da alfabetização (acerca da concepção de alfabetização que dá base aos testes). Uma das diferenças dessa tipologia em relação às propostas de Lowi e Deubel é a forma de sua construção. Nestes autores, a construção é do abstrato ao concreto, ou seja, os critérios para os tipos são definidos teoricamente e depois aplicados a casos concretos, o que garante sua

amplitude e generalidade. Aqui, os critérios foram selecionados a partir da especificidade da política de avaliação educacional. Se eles podem ser aplicados para outras políticas é uma consequência e não o objetivo de sua construção. Além disso, como são arbitrários, vale lembrar que esses critérios poderiam ser outros, poderiam ser reorganizados de outra forma e recombinaos, dando origem a outros critérios. A pertinência dos critérios aqui selecionados, portanto, ainda está por ser debatida e testada.

É importante ressaltar ainda que os critérios selecionados servem a uma análise comparativa entre os programas, com duas observações: i) o objetivo não é estabelecer o melhor programa entre os três, dando margem a uma espécie de competição entre eles, afinal, cada programa pode ser mais ou menos ajustado ao cenário educacional e às necessidades do estado; e ii) a comparação, encerrada entre os três programas, é sempre relativa. Em relação ao primeiro ponto, não creio que seja possível, ou mesmo que faça sentido, estabelecer um programa como melhor. Eles serão mais ou menos ajustados ao cenário no qual estão inseridos e podem ser mais ou menos complexos de acordo com o critério estabelecido. Se não há, de antemão, como definir qual é o critério mais importante, não faz sentido falar em um programa melhor do que o outro. Isso não impede, contudo, que análises sobre a complexidade e o grau de sofisticação de cada um dos programas em relação aos demais sejam feitas. E isso pode experimentar variações para cada um dos critérios de análise. No que tange ao segundo ponto, a relatividade da comparação diz respeito ao fato de que o padrão para a comparação é sempre o que existe nos outros programas e não um padrão estabelecido como o que deve ser seguido (o que explica também o problema de estabelecer um programa como o melhor). Assim, por exemplo, um programa pode ser considerado como possuidor de um desenho de avaliação mais complexo do que outro, não por existir um padrão de como deve ser o desenho de uma avaliação, mas em virtude da comparação específica entre os programas selecionados.

### **5.3.1 Alcance da política**

Qual é o alcance de uma política pública? Essa pergunta se relaciona com o público-alvo da política, mas não se confunde com ele, pelo menos, não no caso de uma política de avaliação da educação. Qual é o alcance da avaliação? Mais especificamente, em contextos de um sistema de avaliação estadual, quanto da rede faz parte da avaliação? Toda ela ou apenas uma parte? Caso seja apenas uma parte da rede, de qual se trata? É uma amostra da rede? Como ela foi construída? Mais de uma rede participa da avaliação? Essas questões interessam para a

caracterização de uma política de avaliação por diversos motivos. O alcance da política reflete o interesse e os objetivos dos tomadores de decisão e dos planejadores da política. Além disso, seu alcance é influenciado por uma série de fatores: além dos objetivos perseguidos por aqueles que elaboraram a política, os recursos financeiros e as exigências logísticas também atuam como fatores que influenciam a decisão.

No que diz respeito às avaliações da alfabetização, interessa saber quais são as redes que participam da avaliação e se toda a rede participa (a avaliação pode ser censitária ou amostral, por exemplo). O alcance da política, nesses casos, influencia a produção dos resultados do sistema de avaliação, sua representatividade e o que pode ser inferido a partir deles. A participação de diferentes redes de ensino coloca em cena também o tema do federalismo, um dos critérios selecionados para a tipologia. Por outro lado, a escolha de uma avaliação censitária ou amostral afeta o tipo de resultado produzido bem como o público ao qual ele se destina, afetado também a geração da política de avaliação, conforme a classificação proposta por Bonamino e Sousa (2012).

Uma avaliação de amplo alcance, no caso da alfabetização, tendo em vista os sistemas estaduais, objetos da presente pesquisa, envolverá as redes municipais e as estaduais, no caso do processo de municipalização não ter se efetivado completamente no estado. Esse é o caso de Minas Gerais, por exemplo. A rigor, tratando a educação como fenômeno essencialmente público, a despeito de acontecer em instituições privadas, a rede privada também deveria ser avaliada para que considerássemos o alcance da política como amplo. Essa é uma questão delicada no cenário brasileiro, pois envolve um antagonismo, em termos ideológicos e sociológicos, das redes públicas e privadas. Por outro lado, uma avaliação que abarque apenas a rede estadual (é importante lembrar que a análise está sendo feita em relação a sistemas estaduais de avaliação) seria de alcance restrito.

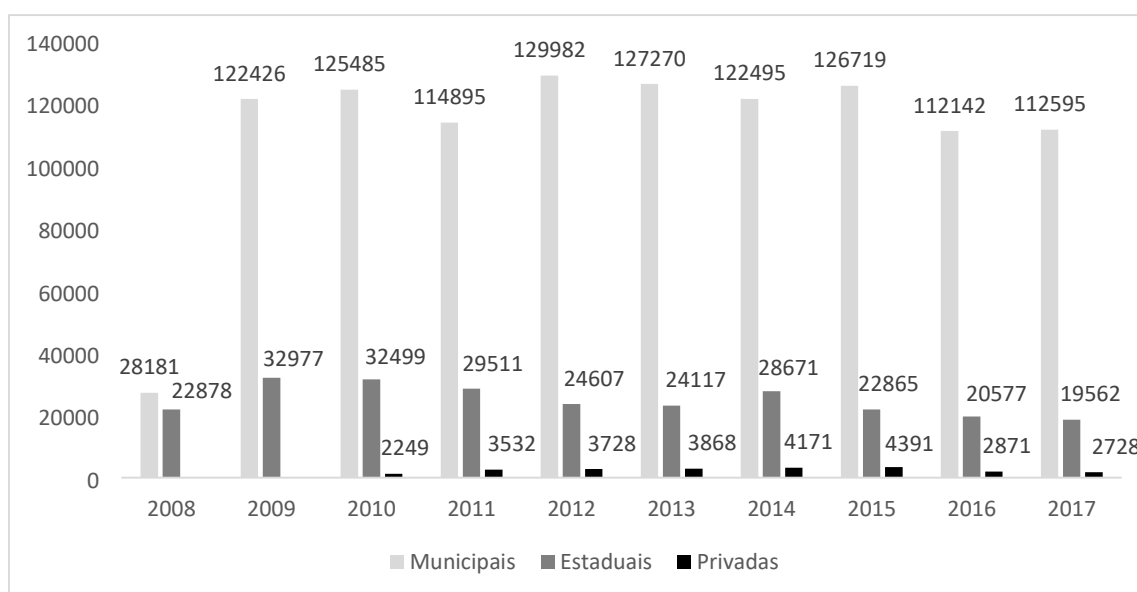
Soma-se à questão da seleção das redes uma eventual seleção dentro da própria rede. É o que acontece quando uma opção é feita entre a avaliação censitária e a avaliação amostral. Podemos considerar como de amplo alcance uma avaliação censitária (levando em consideração os anos escolares avaliados) e de alcance restrito uma avaliação amostral. Isso porque a avaliação censitária permite a produção de resultados por unidade escolar, colocando em cena um tipo de mobilização para uso dos resultados diferente daquela movida pela avaliação amostral. Em caso de amostra, a unidade de análise é a rede de ensino, não as escolas, fazendo com que os resultados sejam produzidos com vistas a informar atores e tomadores de decisão no nível das redes, e não no das unidades escolares.

O Paebes Alfa, desde sua primeira edição, em 2008, avalia escolas da rede estadual e escolas das redes municipais de ensino. Assim, desde sua concepção, o estado propôs um sistema de avaliação que envolvesse, também, as redes municipais. A partir de 2010 (2ª onda), o programa incluiu a avaliação das escolas das redes particulares que desejassem participar.

No âmbito das redes públicas, as avaliações do Paebes Alfa são censitárias. Como o Paebes Alfa avalia as três séries do ciclo de alfabetização, isso significa que, a princípio, todos os alunos do ciclo das redes públicas de ensino são avaliados. É preciso fazer menção, aqui, à peculiaridade das redes municipais. Como os municípios participam do Paebes Alfa por adesão, a avaliação é censitária no âmbito do município (uma vez feita a adesão, todas as escolas do município participam da avaliação), mas as redes municipais não são obrigadas a participar. Quanto às redes particulares participantes, a decisão fica a cargo das escolas.

O gráfico a seguir mostra a evolução da participação das redes, em termos de número de alunos, a partir da edição de 2011 do Paebes Alfa.

Gráfico 1 - Número de alunos participantes do Paebes Alfa por rede – série histórica

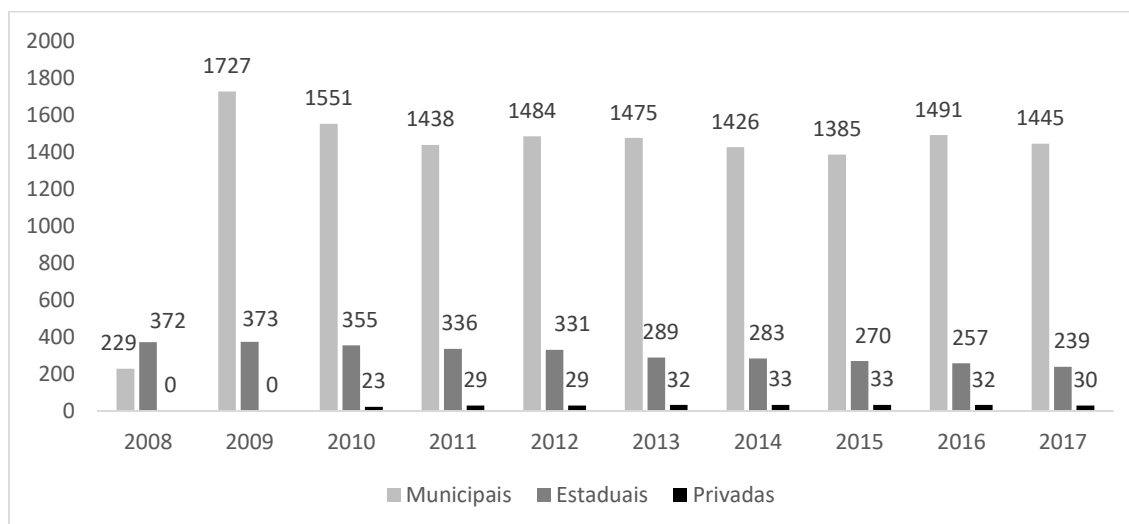


Fonte: Paebes Alfa. Planilhas de resultados. Tabulação própria.

Como podemos ver, a partir da leitura do gráfico anterior, as redes municipais respondem pela maior parte dos alunos que participam das avaliações do Paebes Alfa. A partir de 2009, mais de cem mil alunos das redes municipais foram avaliados a cada ano. Os alunos da rede estadual avaliados variaram anualmente entre dezenove mil e trinta e dois mil. Evidentemente, a rede privada responde por um número menor de participantes: variação entre dois e pouco mais de quatro mil alunos. No gráfico, o número total de alunos envolve todas as séries/anos avaliadas (1º, 2º e 3º anos do ensino fundamental) e todas as aplicações no ano (1ª

e 2ª ondas). Esse arranjo da avaliação ficará mais claro em tópico a seguir, quando o desenho da avaliação será apresentado.

Gráfico 2 - Número de escolas participantes do Paebes Alfa por rede – série histórica



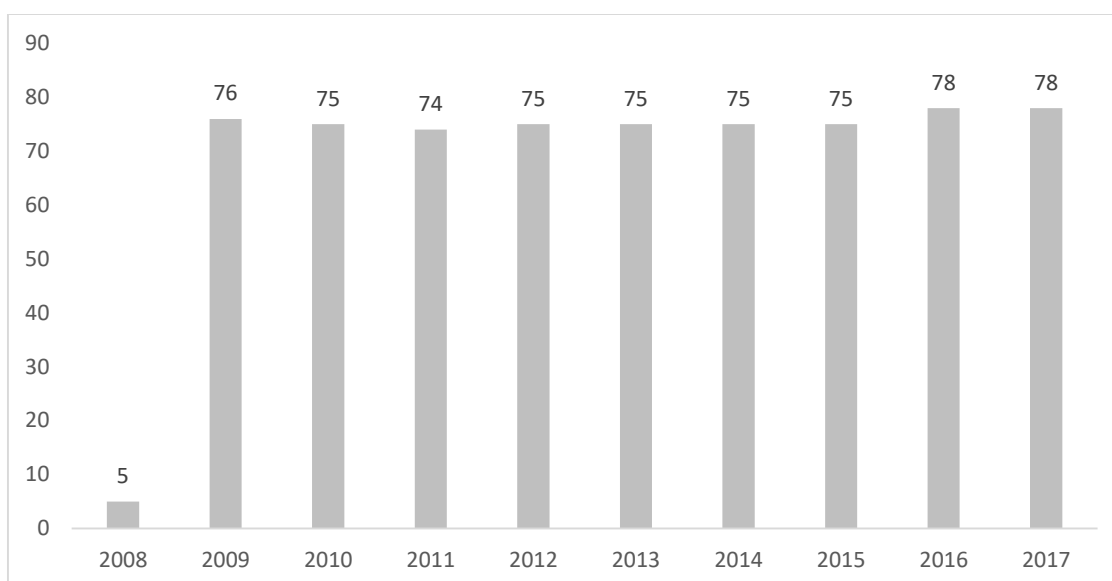
Fonte: Paebes Alfa. Planilhas de resultados. Tabulação própria.

A participação das redes municipais é massiva também em relação às escolas. Como mostra o gráfico anterior, desde 2009, o número de escolas municipais avaliadas pelo Paebes Alfa varia entre 1385 (sendo sua menor participação, em 2015) e 1727 (maior participação, em 2009). O número de 229 escolas, observado em 2008, se explica pela participação de poucos municípios na primeira edição do programa de avaliação. As escolas da rede estadual foram apresentando números cada vez menores a partir de 2009, fruto da intensificação da municipalização no estado. A partir de 2010, as escolas privadas mantêm um número estável de unidades participantes, próximo a 30.

O gráfico a seguir apresenta o número de redes municipais que fizeram adesão ao Paebes Alfa desde sua origem. A primeira edição contou com apenas 5 municípios: Cariacica, Castelo, Serra, Viana e Vila Velha. Em sua segunda edição, a participação já experimentou um grande avanço, com 76 municípios participando do programa. A partir de então, o programa contou com 75 municípios (exceção feita à edição de 2011, que contou com a participação de 74), até o ano de 2015. Nos anos de 2016 e 2017, foram 78 municípios participantes, todos os do estado.



Gráfico 3 - Número de redes municipais participantes do Paebes Alfa – série histórica



Fonte: Paebes Alfa. Planilhas de resultados. Tabulação própria.

O Paebes Alfa é, desse modo, uma política de amplo alcance. Conta com a participação censitária das escolas da rede estadual e das escolas municipais para os municípios que aderem ao programa (é censitária também para os municípios, portanto, visto que todos fazem a adesão desde a edição de 2011). Além disso, diferente da maioria das políticas estaduais de avaliação, o Paebes Alfa faculta a participação a escolas particulares que possuem interesse em participar do programa, estendendo ainda mais seu alcance. Soma-se a isso o fato de avaliar as três séries do ciclo de alfabetização.

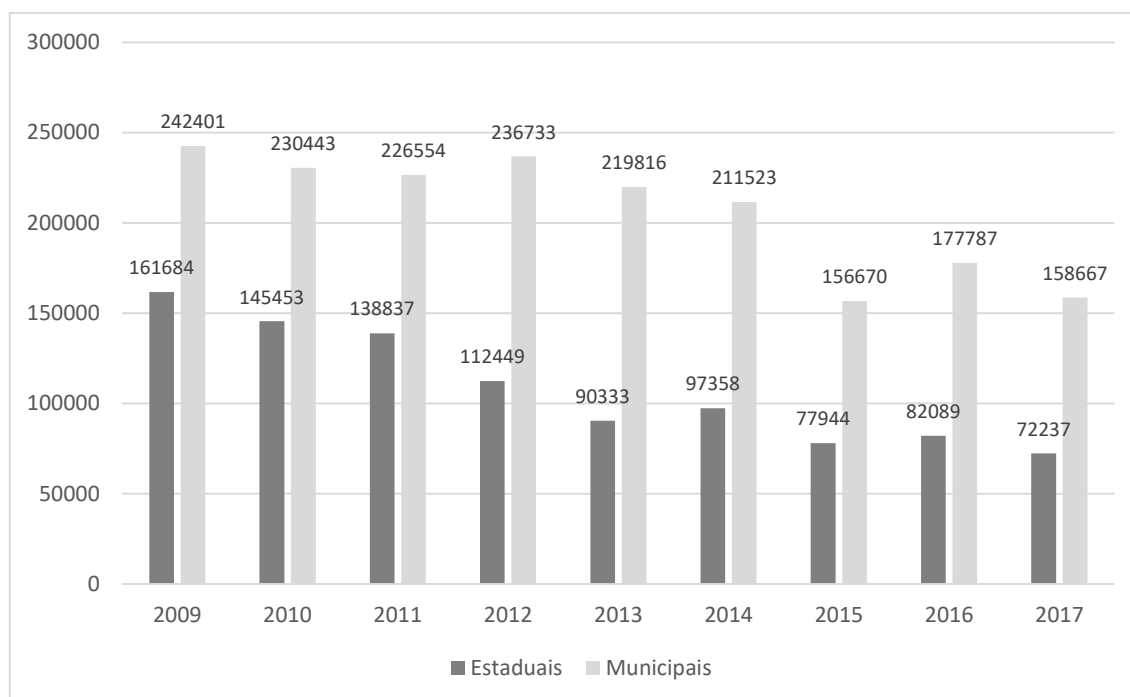
O Proalfa não foi sempre uma avaliação censitária. Em sua primeira edição, em 2005, o programa aplicou uma avaliação amostral para os alunos do 2º ano do ensino fundamental, matriculados apenas na rede estadual. Em 2006, o 2º ano continuou sendo avaliado por amostra, ao passo que foi introduzida uma avaliação censitária do 3º ano. Entre 2007 e 2014, o formato foi o mesmo: avaliação amostral o 2º ano e censitária do 3º, sendo realizada, também uma avaliação amostral do 4º ano do ensino fundamental (incluindo uma avaliação censitária, nesta série, para os alunos de baixo desempenho – mais sobre isso será abordado no tópico sobre o desenho da avaliação). De 2015 a 2017, a avaliação foi censitária, sendo aplicada apenas aos alunos do 3º ano do ensino fundamental. O quadro a seguir sumariza essas informações. No período analisado (entre 2009 e 2017), as redes municipais sempre foram avaliadas, ao passo que as escolas particulares nunca fizeram parte da avaliação.

Quadro 4 - Formato da avaliação por edição do Proalfa – 2009 a 2017

Edição	Formato	Série/ano
2005	Amostral	2º
2006	Amostral	2º
	Censitária	3º
2007 a 2014	Amostral	2º
	Censitária	3º
	Amostral	4º
	Censitária	4º Baixo Desempenho
2015 a 2017	Censitária	3º

Fonte: Proalfa. Planilha de resultados. Tabulação própria.

Gráfico 4 - Número de alunos participantes do Proalfa por rede – série histórica

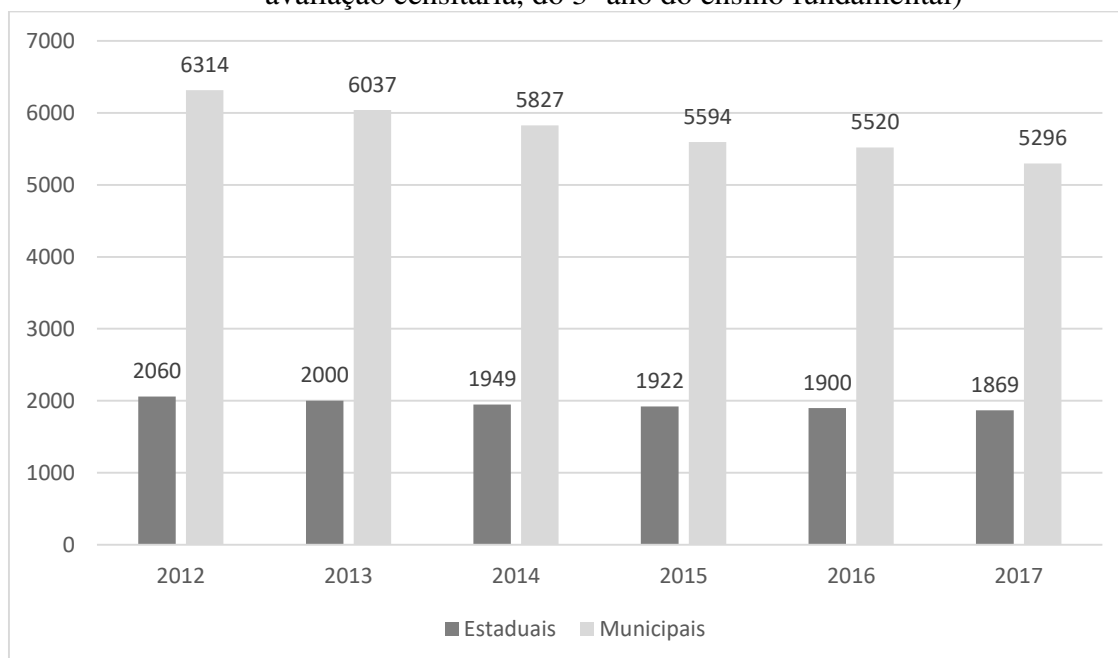


Fonte: Proalfa. Planilha de resultados. Tabulação própria.

O gráfico a anterior apresenta o número de alunos participantes das avaliações do Proalfa por rede, entre 2009 e 2017. Esses números levam em consideração a soma dos alunos em cada uma das séries avaliadas (2º, 3º e 4º anos). Até 2014, como pode ser observado mais de 200.000 alunos foram avaliadas a cada ano nas redes municipais. Nas escolas da rede estadual, foram mais de 100.000 alunos por ano até 2012. A partir da edição de 2015, somente o 3º ano passou a ser avaliado, com a saída das avaliações amostrais de 2º e 4º anos, o que explica a diminuição no número de alunos avaliados.

O Proalfa possui uma linha histórica de resultados do 3º ano do ensino fundamental, entre 2012 e 2017, para 850 municípios (Minas têm, em 2020, 853 municípios). Vale lembrar que, em formato censitário, o programa avaliou apenas essa série/ano. As avaliações amostrais do 2º e 4º anos eram realizadas por SRE – Superintendência Regional de Ensino, de modo que não foram os mesmos municípios e escolas avaliados a cada ano. O número de escolas avaliadas pelo Proalfa, tendo em vista sua avaliação censitária (3º ano), é apresentado pelo gráfico a seguir. Como podemos observar, foram avaliadas quase 2.000 escolas por ano na rede estadual e mais de 5.000 escolas municipais em todas as edições presentes na linha histórica disponível para o Proalfa.

Gráfico 5 - Número de escolas participantes do Proalfa por rede – série histórica (para a avaliação censitária, do 3º ano do ensino fundamental)



Fonte: Proalfa. Planilha de resultados. Tabulação própria.

Diante dessas informações, podemos definir o Proalfa como uma avaliação de alcance intermediário, se a compararmos com o Paebes Alfa. O Proalfa assumiu o formato censitário para o 3º ano do ensino fundamental, ao passo que o 2º foi avaliado de maneira amostral até 2014, quando parou de ser avaliado. O 1º ano nunca foi avaliado pelo programa, embora o 4º tenha sido (por amostra, entre 2007 e 2014). As redes municipais foram avaliadas em todas as edições do período analisado (na verdade, desde 2006), mas as escolas particulares nunca puderam participar. É preciso levar em consideração, nesse ponto, o tamanho da rede de Minas Gerais. O critério sobre o alcance do programa não envolve apenas números brutos. Se fosse assim, o Proalfa teria que ser classificado como de amplo alcance, mas não o Paebes Alfa. A

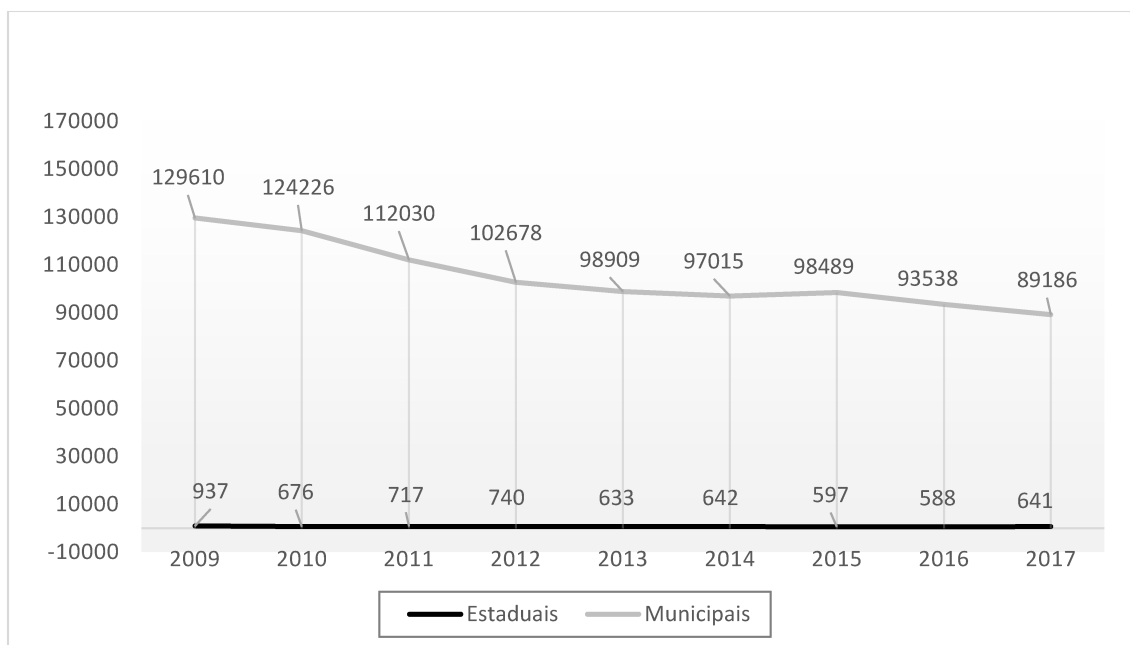
rede pública de Minas Gerais é imensa quando comparado à do Espírito Santo (o Proalfa avaliou no período de 7.000 a 8.000 escolas, ao passo que o Paebes Alfa avaliou, em média, menos de 2.000 escolas – são 850 municípios em Minas e 78 no Espírito Santo). Apesar disso, conforme já informado, a princípio, todos os alunos matriculados no ciclo de alfabetização nas escolas das redes públicas capixabas fazem parte da avaliação e o programa permite a participação, por adesão, de escolas particulares, o que não ocorre em Minas Gerais. Por esse motivo, comparativamente, o Paebes Alfa foi classificado como de amplo alcance, ao passo que o Proalfa foi tratado como de alcance intermediário<sup>166</sup>.

O Spaece Alfa nunca avaliou outra série/ano que não o 2º ano do ensino fundamental. Desde sua criação, em 2007, o programa avalia o 2º ano em formato censitário, tanto na rede estadual quanto nas redes municipais. Desde o início, também, o número de redes municipais participantes é o mesmo: 184 (todos os municípios do estado). O formato estável do programa no ciclo de alfabetização, por um lado, pode ser considerado uma virtude, visto que permite, por exemplo, a criação de uma linha histórica. Contudo, se levarmos em consideração o argumento desenvolvido no capítulo 2 da tese, o de que a avaliação pode ser pensada como um direito, o programa deixa de fora outras etapas de escolaridade do ciclo de alfabetização. O gráfico a seguir apresenta o número de alunos avaliados entre 2009 e 2017 pelo Spaece Alfa para cada rede de ensino. Como podemos notar, o processo de municipalização é avançado no estado, de modo que a rede estadual responde por um número pequeno de alunos no 2º ano do ensino fundamental. As redes municipais avaliaram mais de 100.000 alunos a cada ano até 2012, passando a avaliar entre 90 e 100.000 alunos nos anos seguintes. A rede estadual, por sua vez, avaliou menos de 1.000 alunos a cada edição no período analisado.

---

<sup>166</sup> Mais uma vez, é preciso lembrar da relatividade dessa classificação: intermediário quando comparado ao que foi chamado de amplo alcance para o Paebes Alfa, já que não existe um padrão absoluto do que é amplo alcance para a avaliação.

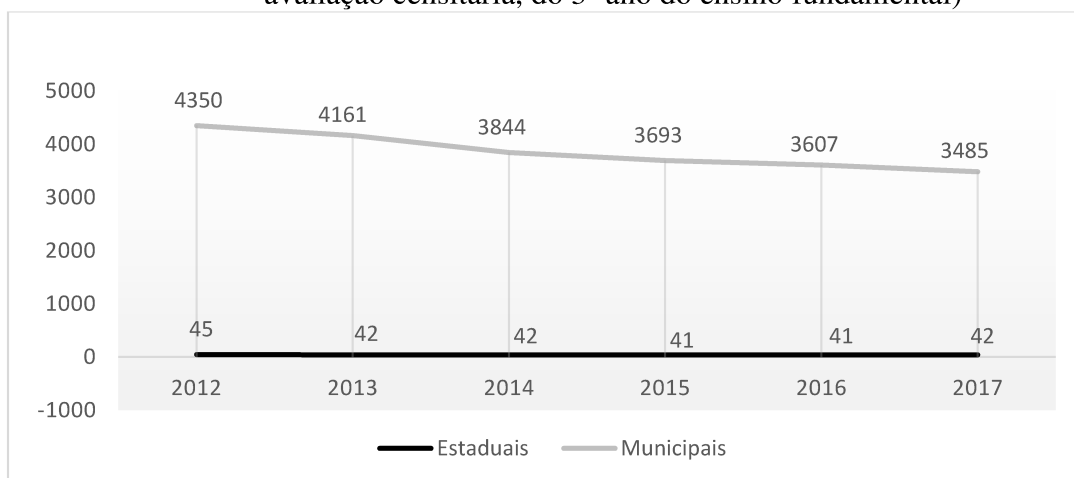
Gráfico 6 - Número de alunos participantes do Spaece Alfa por rede – série histórica



Fonte: Spaece Alfa. Planilha de resultados. Tabulação própria.

O número de escolas avaliadas reflete essa diferença de tamanho entre as redes municipais e a rede estadual no âmbito do ciclo de alfabetização. Como pode ser observado a partir da leitura do gráfico a seguir, na linha histórica entre 2012 e 2017 (assim como ocorre para o Proalfa, a planilha de resultados do Spaece Alfa apresenta linha histórica para o referido período no que diz respeito ao número de escolas avaliadas), entre 3485 e 4350 escolas foram avaliadas nas redes municipais, ao passo que esse número não passou de 42 escolas na rede estadual.

Gráfico 7 - Número de escolas participantes do Spaece Alfa por rede – série histórica (para a avaliação censitária, do 3º ano do ensino fundamental)



Fonte: Spaece Alfa. Planilha de resultados. Tabulação própria.

Diante dessas características, é possível categorizar o Spaece Alfa como de alcance intermediário, na esteira do que foi exposto para o Proalfa. O programa concentrou-se em avaliar o 2º ano do ensino fundamental, de forma censitária, ao longo de todas as edições analisadas, sem ter aplicado testes a nenhuma outra série do ciclo de alfabetização. As redes municipais foram sempre avaliadas, desde o início do programa, mas não houve participação de escolas particulares. Assim como no caso do Proalfa, o número de alunos e de escolas avaliados é bastante expressivo e superior àquele observado no Paebes Alfa. Entretanto, conforme já argumentado, o número absoluto de escolas e alunos avaliados não é o elemento central para definir a amplitude do alcance. Além disso, embora o Proalfa tenha avaliado outras séries, de modo censitário, e essa é uma diferença em relação ao Spaece Alfa, sua descontinuidade faz com que essa experiência não tenha sido aqui considerada suficiente para estabelecer uma distinção significativa em relação a esse critério entre os dois programas.

### **5.3.2 Público-alvo**

Alcance da política e público-alvo são dimensões relacionadas, mas que não se confundem. O público-alvo é aquele para o qual os resultados de uma política são dirigidos. A quem se destina a política? No caso da avaliação, a quem se destinam os resultados da avaliação? Uma política de avaliação da alfabetização tem como resultado as informações que produz a partir dos instrumentos que aplica. Em essência, a entrega de uma política de avaliação são os resultados dessa avaliação. Com quem esses resultados dialogam, afinal?

Se os resultados da avaliação definem seu público-alvo, outra dimensão relacionada é a divulgação de resultados. O processo de divulgação é tão central que, podendo ser incluído como parte do desenho da avaliação, para a análise aqui levada a cabo recebeu um tópico próprio. A divulgação dos resultados variará dependendo do público-alvo da política. Além disso, outro aspecto relacionado ao destinatário da política de avaliação é o tipo de política de responsabilização que pode ser desenvolvida, que se relaciona também, como vimos, ao alcance do programa. Assim, uma política de avaliação pertencerá a uma diferente geração de acordo com aquilo que é possível fazer com seus resultados, em termos de responsabilização, e isso está relacionado com o público-alvo. Uma avaliação censitária abre espaço para uma responsabilização no nível das escolas, o que leva a política a poder ser caracterizada, dependendo da existência de fato da responsabilização, como pertencente à segunda ou à terceira geração de avaliação, conforme a tipologia proposta por Bonamino e Sousa (2012).

No âmbito de uma política de avaliação da alfabetização, podemos pensar em público-alvo imediato, ou direto, e mediato, ou indireto. Isso ocorre porque os resultados da avaliação podem ser direcionados a diferentes públicos e com diversos objetivos. Essa definição é influenciada pelo alcance da política. Uma avaliação amostral de escolas, por exemplo, não tem como público-alvo direto os diretores e professores da escola.

Imediata e diretamente, podemos afirmar que o público-alvo direto das políticas de avaliação da alfabetização são os tomadores de decisão: gestores de rede (dirigentes, superintendentes – e equivalentes – e técnicos das secretarias), gestores de escola (incluindo aqui a equipe de gestão – coordenadores pedagógicos e equivalentes) e professores. Em relação a estes últimos, é preciso fazer uma ressalva. Os sistemas de avaliação da alfabetização se baseiam em uma metodologia de produção de medidas de aprendizagem que exige determinados requisitos. Um deles diz respeito ao número mínimo de alunos para a produção de resultados. A princípio, os sistemas de avaliação em larga escala tinham como principal destinatário os tomadores de decisão no âmbito das redes de ensino (as secretarias de educação). Essa é uma perspectiva que ainda permanece em vigor. Muitas vezes, as secretarias de educação, no nível macro, são pensadas como as únicas destinatárias dos resultados da avaliação educacional, como se as avaliações, por elas implementadas, fossem pensadas única e exclusivamente para elas. Contudo, e a história de desenvolvimento e consolidação da avaliação educacional é tributária desse ponto, os gestores escolares passaram a ser destinatários centrais das políticas de avaliação. A secretaria recebe informações para a tomada de decisão no âmbito da rede, assim como os gestores escolares recebem informações para a tomada de decisão no âmbito da escola. A produção de resultados no nível das turmas, por sua vez, não permite excluir o professor como destinatário dos resultados da avaliação, para a tomada de decisão no âmbito da turma. O grande objetivo de um sistema de avaliação educacional é dar informações que sirvam como suporte para a tomada de decisões educacionais, conferindo, assim, maior legitimidade a elas.

Se o objetivo da avaliação é produzir informação qualificada como subsídio para decisões, o público-alvo direto de uma política de avaliação da alfabetização é aquele que usará tal informação para tomar decisões no âmbito do ciclo de alfabetização. Isso inclui os dirigentes municipais e suas equipes, os gestores de escolas (e suas equipes de gestão) que oferecem o ciclo de alfabetização e os professores alfabetizadores.

O público-alvo indireto é formado por aqueles que se beneficiam da política de forma mediata. É o caso dos alunos matriculados nos anos iniciais e que são avaliados através dos

testes. As informações produzidas pelas avaliações têm como horizonte último a melhoria da aprendizagem. Sendo assim, os alunos são beneficiários mediatos da política de avaliação da alfabetização, pois, através dela, a aprendizagem pode ser potencializada. No caso dos alunos, isso fica claro: os resultados da avaliação não produzem mudanças imediatas na educação; se a aprendizagem pode ser potencializada e o ensino aprimorado, isso exigirá outras políticas e ações para tanto, levando tempo para que os efeitos sejam percebidos, o que justifica sua mediaticidade. O fato de ser mediato não quer dizer, entretanto, que o destinatário seja secundário. O caso do aluno é paradigmático: a política de avaliação serve, em última instância a ele, à potencialização de sua aprendizagem, mas os resultados são direcionados, imediatamente, àqueles que devem agir para promover essa potencialização.

Além dos alunos, os pais e responsáveis por eles também se beneficiam dos resultados da política de avaliação, compondo o público-alvo mediato. Eles podem tomar decisões sobre as escolas nas quais desejam matricular seus filhos com base em resultados da avaliação (o que é uma forma de *accountability*) ou exigir melhorias e cobrar por mudanças nas escolas onde seus filhos já estão matriculados. Somam-se aos pais e responsáveis, a comunidade escolar e a sociedade civil como um todo, como destinatários mediatos da avaliação da alfabetização. Os resultados, divulgados e trazidos à tona para o público em geral, podem levar a uma reflexão da sociedade sobre o ensino oferecido e sobre a aprendizagem dos alunos, levando a exigências por melhorias na educação pública, por exemplo. Nesse caso, o que está em jogo é um funcionamento esperado para sociedades democráticas e participativas, nas quais a população atua no sentido de exigir mudanças por parte dos governos. Na prática, o funcionamento dessa engrenagem enfrenta um conjunto de obstáculos para ser posta em cena. De todo modo, para a sociedade, incluindo a comunidade escolar, as políticas de avaliação são instrumentos de prestação de contas do Estado àqueles que financiam tais serviços. Por fim, podemos acrescentar como público-alvo indireto os pesquisadores e estudiosos da educação, que ganham, com a avaliação, uma fonte ainda pouco explorada, tendo em vista suas inúmeras possibilidades e o quantitativo de dados, de informações educacionais, que ajudam a testar e confirmar hipóteses e a compreender melhor o fenômeno educacional.

Paebes Alfa, Proalfa e Spaece Alfa possuem um público-alvo semelhante. Todos eles se destinam, prioritariamente, à entrega dos resultados para os tomadores de decisão no âmbito das redes e das escolas. O fato dos três programas terem sido avaliados em parceria com a mesma instituição de pesquisa ao longo do período analisado contribui para essa semelhança. A decisão sobre o público-alvo cabe à secretaria de educação, mas os recursos de



disponibilização dos resultados, desenvolvidos, para os três programas, pelo CAEd/UFJF, permitem acessos específicos a diferentes atores.

O público-alvo imediato dos três programas é formado por: i) secretário de educação e equipe de avaliação da secretaria (recebem acesso a todos os resultados da avaliação, desde o nível do aluno até o nível da rede); ii) superintendentes/coordenadores regionais (recebem os resultados de todo o projeto até o nível das escolas, chegando ao nível dos alunos das escolas de sua superintendência); iii) secretários municipais de educação (recebem os resultados de todo projeto até o nível das escolas, chegando ao nível dos alunos das escolas de seu município); iv) gestores escolares (recebem os resultados de todo o projeto, chegando ao nível dos alunos de sua escola); v) professores (recebem os resultados de todo o projeto, chegando ao nível dos alunos de sua turma). Os destinatários mediatos são os pais, a comunidade escolar, a sociedade civil, os pesquisadores. Eles têm acesso livre aos resultados de todo o projeto, até o nível da escola. Assim, os resultados das escolas e dos níveis de agregação superiores a elas (municípios regionais e todo o projeto) são abertos ao público. Os resultados das turmas e dos alunos são acessíveis apenas aos atores envolvidos com as escolas (professores e diretores) e aqueles superiores na hierarquia educacional (coordenadores de regional e secretários de educação). O quadro a seguir resume essas informações.

Quadro 5 - Atores com acesso aos resultados por nível de agregação – Paebes Alfa, Proalfa e Spaece Alfa

<b>Nível de agregação do resultado</b>	<b>Atores que possuem acesso</b>
Alunos	Professor da turma, diretor da escola, coordenador da regional (escola estadual), secretário municipal (escola municipal), secretário de educação do estado
Turmas	Professor da turma, diretor da escola, coordenador da regional (escola estadual), secretário municipal (escola municipal), secretário de educação do estado
Escolas	Aberto ao público
Municípios (resultados das escolas municipais)	Aberto ao público
Regionais (resultados das escolas estaduais)	Aberto ao público
Projeto	Aberto ao público

Fonte: Portais de Paebes Alfa, Proalfa e Spaece Alfa. Elaboração do próprio autor.

Para o Spaece Alfa e o Paebes Alfa os resultados estão disponíveis, em seus portais de divulgação<sup>167</sup>, desde a edição de 2008 e 2009, respectivamente, ao passo que o Proalfa disponibiliza em seu portal os resultados a partir de 2013<sup>168</sup>. O que poderia diferenciar o público-alvo desses programas são as restrições de acesso aos resultados. Por exemplo, o programa poderia divulgar o resultado de cada escola somente para si própria, de modo que uma escola não seria capaz de ver o resultado da outra. As únicas restrições, entretanto, ocorrem nos níveis da turma e do aluno, de modo a preservar a identidade dos mesmos. O acesso a esses resultados ocorrem por meio de login e senha, enviados às escolas. O público-alvo, portanto, não é um critério que diferencia os três programas de avaliação da alfabetização analisados.

Os canais de divulgação dos resultados, que serão abordados em tópico próprio, nos ajudam a perceber que os destinatários principais da política de avaliação são os tomadores de decisão, a despeito dos dados serem públicos a partir do nível da escola. Os materiais de divulgação, por exemplo, são pensados para os tomadores de decisão – a estrutura dos materiais impressos (com revistas específicas para a gestão do sistema, a gestão da escola e para os professores), sua linguagem, os exemplos dados, a preocupação em todos eles com o uso dos resultados para dar base ao planejamento, a lógica de navegação para os filtros no portal de divulgação (que pressupõe um conhecimento mínimo de como uma rede de ensino é hierarquicamente organizada). Para aqueles que não conhecem a estrutura de uma rede de ensino, pode ser mais complicado ter acesso aos dados, embora eles estejam lá disponíveis. Da mesma forma, pesquisadores interessados em estudar os dados da avaliação podem sentir alguma dificuldade de encontrar os dados reunidos no formato em que desejam (em particular, as bases de dados), visto que as informações são destinadas aos tomadores de decisão, e não diretamente a eles. Quanto aos alunos, uma observação. Embora tenha havido uma iniciativa, durante algum tempo, nos três programas, de divulgar os resultados para os pais, através de um material próprio, ela foi descontinuada, por não ter sido bem-sucedida (esse ponto voltará a ser abordado no tópico sobre a divulgação dos resultados), o fato é que os alunos têm acesso aos resultados através de seus professores e da gestão da escola. Evidentemente, isso é pressuposto, visto que não há pesquisas que mostram como os resultados chegam aos alunos. A lógica de navegação para acesso aos resultados das escolas no portal de divulgação não é orientada para alunos matriculados no ciclo de alfabetização e os alunos e seus pais não recebem uma senha para acesso aos resultados do alunos especificamente.

---

<sup>167</sup> Os endereços dos portais de Paebes Alfa, Proalfa e Spaece Alfa são, respectivamente: <http://www.paebes.caedufjf.net/>; <http://simave.educacao.mg.gov.br/>; <http://www.spaece.caedufjf.net/>

<sup>168</sup> Última consulta em janeiro de 2020.

### 5.3.3 Contexto de criação (agenda) e modelo de implementação

Com base no ciclo de políticas de Ball, conforme vimos no capítulo 1 da tese, a definição da agenda é um elemento importante para os estudos de políticas públicas. Do ponto de vista político, a agenda é a etapa do ciclo que mais tem recebido atenção dos pesquisadores de políticas públicas. A pesquisa sobre a agenda, entretanto, apresenta uma série de dificuldades, entre elas, o problema de rastrear o início de uma política que já pode estar em execução há muito tempo. Além disso, pode haver uma distância temporal considerável entre a definição da agenda (ou, pelo menos, o início do processo de definição) e a implementação da política, de modo que muitas informações importantes para o entendimento desse processo sejam perdidas.

Um primeiro passo para saber como a política teve sua agenda construída é saber se há registro desse momento. Caso não haja, os atores que participaram da agenda são fontes plausíveis para consultar sobre a construção desse momento, embora isso acarrete nova série de dificuldades. Outra possibilidade é realizar um estudo sobre o contexto político e social no âmbito do qual a política foi proposta e desenvolvida. Mesmo que não existam documentos que registrem especificamente esse momento de discussão e debate para a criação de uma política, cenário bem provável, a análise do contexto político pode ajudar a lançar luz sobre a construção da agenda. Apesar dessas dificuldades, a agenda permanece sendo um elemento importante para a compreensão de uma política pública.

A existência de um documento que reúna os principais pontos de construção da agenda, inclusive, é um elemento caracterizador da política. Ele indica a preocupação dos elaboradores da política em registrar os principais argumentos e forças presentes no momento de sua criação. Trata-se de um registro histórico significativo, que abre a possibilidade de analisar e reavaliar a política a partir das forças que estiveram presentes em sua origem e da mudança dessas forças ao longo do tempo. Como já afirmado, um documento como esse pode não existir.

Em termos do contexto de criação e definição da agenda, a política pode ter sido formado por forças convergentes, alinhadas, que se uniram e se articularam para a criação da política. Nesse cenário, há pouco conflito na criação da política e seus objetivos e diretrizes são produtos de uma visão tendente à homogeneização das ideias e concepções. Um subtipo de política com agenda formada sem conflitos ideológicos é aquela em que a criação é dada pelo próprio Estado, de maneira unilateral, sem a participação efetiva de outros grupos e atores sociais, ainda que convergentes. Vale notar que, mesmo nesse caso, de criação por parte do Estado, os conflitos internos não estão descartados, visto que o Estado é composto por forças

dissonantes. Mesmo assim, para a definição das forças, o que interessa mais diretamente é a participação de outros atores, além do Estado.

Por outro lado, é possível que a agenda tenha sido construída com base em conflitos e com a participação de grupos sociais antagônicos e divergentes. Nesse caso, de forças divergentes, a agenda é formada com base na heterogeneidade ideológica dos participantes, recebendo influência de forças distintas. Evidentemente, uma agenda construída com base nesse cenário difere significativamente da agenda definida por forças convergentes. A política, nesse caso, terá características distintas, pois assumirá valores também distintos, que, de uma maneira ou de outra, conseguiram se acomodar e chegar a um consenso mínimo para a criação da política. Essa divergência manifestada na construção da agenda, certamente, exercerá efeitos sobre as demais fases da política, que, de uma forma ou de outra, permanecerá sempre tensionada.

No âmbito da avaliação da alfabetização, uma agenda convergente, por exemplo, terá sido formada pelo consenso em relação aos objetivos da avaliação dessa etapa de escolaridade, bem como no que diz respeito à ideia do que seja um aluno alfabetizado. A concepção de alfabetização terá alinhamento entre os grupos participantes na criação da política. No caso da criação unilateral por parte do Estado, a concepção de alfabetização será aquela adotada oficialmente (sem excluir conflitos internos que por ventura existam).

Caso a agenda tenha sido construída com base em conflitos ideológicos e em posições divergentes, a concepção de alfabetização da política de avaliação terá sido forjada com base em acomodações. Isso provavelmente significará que um dos grupos envolvidos na definição terá sua concepção de alfabetização refletida pela política. Além disso, os objetivos da avaliação e a forma como os resultados deverão ser utilizados também terão sido alvo de discordâncias e acomodações para a definição da agenda.

Em um caso ou em outro (forças convergentes ou divergentes), é uma tarefa muito delicada e complexa analisar quais são as forças envolvidas na definição da agenda e quais delas exerceram mais influência. Em caso de documentos disponíveis para a consulta, uma análise retórica dos argumentos produzidos por cada grupo de força poderia ser realizado, ou até mesmo um enfoque na análise do discurso, mas isso pressupõe a improvável existência de registros dessa natureza e as dificuldades metodológicas que acompanham esse tipo de pesquisa.

Mesmo reconhecendo a importância da agenda como etapa da política, a análise aqui oferecida sobre este tema é limitada. Ela se restringe a rastrear o contexto (ou, pelo menos, seus principais elementos) que deu origem à política de avaliação da alfabetização nos estados,

buscando apontar o que motivou as secretarias a desenvolver um programa de avaliação para a etapa.

Outra etapa importante para o ciclo de políticas proposto por Ball é a implementação. Assim como a agenda, a implementação tem sido o objeto de estudo de muitas pesquisadores de política públicas. A implementação está associada à agenda, visto que, conforme já mencionado, a forma como a agenda foi construída influencia significativamente a maneira como a política será implementada. Por isso, aqui, foi feita a opção por reunir esses dois aspectos no mesmo critério. A contexto de criação da política, como se verá, está diretamente associadas à forma como ela foi implementada nos três estados analisados.

No que diz respeito à implementação, podemos realizar uma análise semelhante àquela proposta para a agenda. Uma política pode ser implementada com ou sem a participação dos atores aos quais se destina ou dos quais precisa para sua execução. De maneira geral, os modelos que envolve a participação de outros atores, que não apenas os tomadores de decisão, são chamados de *bottom-up*, ao passo que um modelo de implementação sem tal envolvimento é chamado de *top/down*<sup>169</sup>. A relação com a construção da agenda não é linear, mas pode ser pensada da seguinte forma: uma agenda construída unilateralmente tende a gerar uma implementação também unilateral, sem que os atores que deveriam estar envolvidos com a política façam parte dessa etapa. Por outro lado, uma agenda definida com base em forças divergentes pode levar a dois diferentes cenários: i) a divergência manifestada na construção da agenda pode ter gerado um conflito ainda maior entre os grupos, fazendo com que o grupo que perdeu mais com o debate fique afastado da implementação; ii) por outro lado, a divergência pode ter gerado no grupo vencedor (prioritário na definição da agenda) a necessidade de se aproximar dos demais grupos, buscando construir com eles maiores consensos, o que faria com que esses grupos participassem da implementação da política.

A participação de diferentes grupos na implementação da política, em especial daqueles que abarcam os atores diretamente envolvidos com sua execução, é tratado como um fator fundamental para que a política seja bem-sucedida, alcançando os objetivos previstos. A rigor, o envolvimento desses atores em todas as fases da política garante o conhecimento sobre as ações previstas, aumentando as chances delas serem bem executadas. O não envolvimento, por outro lado, tende a ser visto como um fator de novos conflitos e, muitas vezes, de rechaço à política. Isso foi o que aconteceu com a avaliação educacional em larga escala no Brasil, ao

---

<sup>169</sup> Para mais sobre essas categorias, rever novamente o capítulo 1 da tese, onde essa discussão é apresentada com mais minúcia.

longo de seu processo de implementação e desenvolvimento, desde os anos 1990. Ao não envolver atores essenciais ao processo, como diretores de escola e professores, a avaliação foi alvo de um conjunto de críticas, nem sempre justificadas, como vimos no capítulo 3, e de uma resistência que ofereceu obstáculos à sua plena compreensão e execução (como veremos, esse aspecto influencia um dos critérios selecionados para a tipologia, qual seja, os efeitos colaterais). A despeito dessa perspectiva, não há evidências de que um modelo de implementação tenha mais chances de sucesso do que o outro. Do mesmo modo, não há informações precisas sobre até que ponto deve acontecer a participação (e de quais atores) para que seja tratada como suficiente. O envolvimento dos atores na implementação de uma política permanece sendo uma defesa de princípio, não necessariamente ancorada em resultados.

No âmbito da política de avaliação da alfabetização, interessa investigar se a implementação do sistema de avaliação contou ou não com a participação da chamada *street-level bureaucracy*, ou seja, de gestores de escolas que oferecem o ciclo de alfabetização e dos professores alfabetizadores. Além deles, quem mais foi consultado para a implementação da política? Especialistas em políticas de avaliação da alfabetização, pesquisadores de universidades e centros de pesquisa? Ou o sistema de avaliação da alfabetização foi implementado apenas por técnicos da secretaria de educação?

Conforme apresentado por Segatto (2011), o Espírito Santo foi um dos estados brasileiros que, na década de 2000, mais encarnaram a ideia de uma reforma administrativa abrangente. Marcado por escândalos nas administrações dos anos 1990, o estado começou, em 2003, uma reforma de grande monta. A secretaria de educação foi uma das pioneiras a realizar a reforma, baseada em uma cultura voltada para resultados (SEGATTO, 2011, p.72). Conforme aponta a pesquisa de Segatto, muito disso deve-se ao perfil mais técnico do secretário de educação naquele momento, Haroldo Côrrea Rocha. Essa cultura de resultados favoreceu o desenvolvimento de um sistema de avaliação educacional no estado, o que abriu espaço para a proposição de um sistema de bônus por desempenho. Aliás, o argumento fundamental de Segatto é esse: o contexto das reformas administrativas no âmbito educacional abriu espaço para a disseminação do gerencialismo e, como consequência, das avaliações em larga escala. Sem o sistema de avaliação, dificilmente o bônus poderia ser proposto, conforme reconheceu, em entrevista para Segatto, a ex-Assessora de Planejamento e Gestão da Seduc, Andressa Rocha, que também foi entrevistada no âmbito desta tese. Aliás, a própria criação da Assessoria de Planejamento e Gestão na Seduc, em 2007, foi um passo importante para a criação de um sistema próprio de avaliação. Para uma secretaria que decide se organizar em torno de

evidências e resultados, é necessário um sistema de avaliação que os forneça. A primeira avaliação do Paebes Alfa, em parceria com o Ceale, aconteceu já no ano seguinte, em 2008.

A reforma na administração da educação do Espírito Santo baseou-se nos exemplos de outros estados, como São Paulo, Pernambuco, Ceará e Minas Gerais (SEGATTO, 2011). Sendo assim, é difícil imaginar que os programas de avaliação da alfabetização desses dois últimos estados, por exemplo, não tenham servido como inspiração para a criação do Paebes Alfa. No âmbito da avaliação educacional, José Francisco Soares, professor da UFMG, que também já havia dado consultoria para São Paulo e é referência na área de avaliação educacional, foi contratado como consultor para o estado.

Para Côco (2014), um importante fator para a criação do Paebes Alfa foi o Termo de Compromisso Todos Pela Educação, formalizando o compromisso dos estados em atingir as metas pactuadas e estabelecidas pelo Decreto nº 6.094/2007. Além de estabelecer o foco na aprendizagem, estabelecendo resultados concretos a serem alcançados, uma das metas propõe a alfabetização das crianças até, no máximo, os 8 anos de idade, prevendo aferição de resultados por exame periódico específico. Paulo Hartung assumiu o governo em 2003, com propostas de mudanças na administração do estado, conforme já ressaltado. Em 2006, é publicado o Plano de Desenvolvimento do Espírito Santo 2025, que estabeleceu, entre outros aspectos, a transparência dos resultados anuais de cada escola. É nesse contexto de reforma administrativa, da influência do Compromisso Todos Pela Educação e da influência de outros estados que o Paebes Alfa é criado. Diferentemente dos casos de Proalfa e Spaece Alfa, o programa de avaliação da alfabetização do Espírito Santo não foi criado pela influência de resultados ruins experimentados pelas redes públicas do estado.

A principal justificativa para a criação do Proalfa<sup>170</sup>, em 2005, de acordo com o Boletim Pedagógico do programa para a edição de 2009 (MINAS GERAIS, 2009), foi a necessidade de acompanhamento dos efeitos da política de ampliação do ensino fundamental sobre a qualidade da educação. Em 2004, foi implementado o ensino fundamental de 9 anos em Minas Gerais. A finalidade principal dessa ampliação era a potencialização da aprendizagem dos alunos e sua principal modificação dizia respeito ao ciclo de alfabetização, já que os alunos passariam a ingressar mais novos no 1º ano, a partir dos 6 anos. Soma-se a isso o estabelecimento, pela Seduc de Minas Gerais, de sua meta prioritária: a alfabetização das crianças aos oito anos de idade. Diante disso, o Proalfa surgiu como parte das avaliações do estado destinadas a acompanhar os efeitos dessa política de ampliação do ensino, sendo incorporado ao Simave em

---

<sup>170</sup> Iniciativa pioneira da avaliação da alfabetização nos estados no Brasil.

2006. Seu objetivo principal era fornecer informações qualificadas sobre o sistema de ensino do estado, de modo a orientar a construção de estratégias para o alcance da meta de alfabetização estabelecida pelo estado.

O surgimento do Proalfa, portanto, encarna a função do sistema de avaliação de funcionar como política de avaliação dos efeitos de outras políticas (no caso de Minas Gerais, o ensino fundamental de 9 anos), passando a fornecer informações para a tomada de decisões. Martins (2013a) recolhe, nas entrevistas que realizou sobre o início do Proalfa, um terceiro indutor para sua criação (além da meta de alfabetizar as crianças aos 8 anos e do ensino fundamental de 9 anos): os resultados do Saeb para o estado, que diagnosticaram à época resultados muito abaixo dos esperados para leitura no 4º ano do ensino fundamental (naquele momento, avaliado no âmbito do Proeb também). De acordo com a entrevista que Martins realizou com uma representante do Ceale, responsável pelas primeiras avaliações do Proalfa, esses resultados levaram ao questionamento sobre as relações entre o baixo desempenho em leitura e problemas na alfabetização das crianças (Martins, 2013a, p. 65-66), o que abriu espaço para a proposta de uma avaliação para o ciclo de alfabetização que ajudasse a compreender melhor esses resultados. A ideia de que os resultados do Saeb para o estado aparecem como fator que impulsionou a criação do Proalfa aparece também em Filgueiras (2013). Os dados da avaliação, portanto, ao mostrarem um cenário preocupante para as secretarias de educação, serviram como base para a criação de um sistema de avaliação da alfabetização em Minas.

De acordo com Alessio Costa Lima, ex-Coordenador da Célula de Avaliação do Desempenho Acadêmico na Seduc do Ceará, um conjunto de fatores explicam a criação do Spaece Alfa: i) em 2007, a eleição de Ciro Gomes para o governo do estado levou a professora Izolda Cela à secretaria estadual de educação – ela havia sido secretária de educação em Sobral, que tinha se tornado o município de referência, em todo o Brasil, para o combate ao analfabetismo; ii) da experiência de Sobral, a nova equipe da secretaria de educação trouxe a concepção de que a alfabetização infantil deveria ser concluída no 2º ano do ensino fundamental e não no 3º ano; iii) uma pesquisa realizada pelo Comitê Cearense para a Eliminação do Analfabetismo Escolar, conduzida pela Assembleia Legislativa do estado com o apoio da Unicef, Inep e Undime Ceará (entre outros parceiros), com base em dados do IBGE (2000) e do Saeb (2003), apontou um quadro preocupante de analfabetismo no estado (em mais de 50 municípios, as taxas de analfabetismo variavam entre 20 e 50%, CEARÁ, 2006); e iv) com base nos dados desse relatório, a nova secretária de educação do estado iniciou o desenho do Paic – Programa de Alfabetização na Idade Certa. Segundo o entrevistado, tudo isso criou um



ambiente favorável à inclusão de uma avaliação externa da alfabetização no já existente sistema de avaliação do estado, o Spaece, que avaliava, àquele momento, 4ª e 8ª séries do ensino fundamental e o 3º ano do ensino médio. Essa avaliação passou a ser chamada de Spaece Alfa e sua primeira aplicação ocorreu ainda no ano de 2007. A escolha do 2º ano para aplicação dos testes e não do 3º (como na experiência da avaliação censitária do Proalfa em Minas Gerais) justifica-se pela experiência de Sobral. Além disso, o Spaece Alfa assumiu um formato censitário, destinado às redes municipais e estadual.

O Spaece Alfa surgiu, portanto, no âmbito de um conjunto de medidas propostas pela política destinada à alfabetização no estado a partir de 2007. Ele foi um dos elementos da resposta que o estado deu às taxas de analfabetismo que apresentava. Não é possível compreender o programa fora do Paic. As ações de formação continuada de professores e o pacto de acordo entre os entes federados para o combate ao analfabetismo tinha como objetivo direto produzir efeitos sobre a alfabetização dos alunos e isso deveria ser mensurado, para o monitoramento dos resultados da política. O Spaece Alfa foi apresentado como o instrumento dessa mensuração. Seus resultados serviriam para avaliar a capacidade do Paic de surtir efeitos. Além disso, não é possível compreender o Spaece Alfa também sem considerar o cenário que a avaliação em larga escala já experimentava naquele momento. No âmbito nacional, os resultados do Saeb 2003, referentes à 4ª série do ensino fundamental, apontavam que 71,6% das crianças do Ceará estavam nos padrões Muito Crítico (30,4%) e Crítico (41,2%) em língua portuguesa (CEARÁ, 2006). Os dados da avaliação foram um dos motivos para a criação do Paic, que levou, por sua vez, à criação do Spaece Alfa.

Uma análise mais detida sobre as forças divergentes nesse processo seria importante e instigante, mas é difícil de ser feita diante dos elementos que reuni. Contudo, é possível afirmar que, a despeito da existência e da capacidade de mobilização dessas forças, o Spaece Alfa é fruto de uma política que reuniu muitas forças convergentes em torno do compromisso com a erradicação do analfabetismo no estado. O Comitê Cearense para a Eliminação do Analfabetismo Escolar, por exemplo, contava com a participação de diferentes forças: Inep, Undime, Unicef, Assembleia Legislativa, Seduc do Ceará, secretários de educação de algumas cidades cearenses (entre elas, Sobral), representantes de universidades (UFC, UECE, UNIFOR), membros do Conselho Estadual de Educação. A existência consolidada do Spaece também é um fator que pode ter ajudado a reduzir a resistência para a criação do Spaece Alfa.

No que diz respeito ao modelo de implementação, podemos perceber, a partir dos contextos de criação dos programas, semelhanças nos três casos. O modelo de implementação

de Paebes Alfa, Proalfa e Spaece Alfa aproxima-se mais do *top/down* do que do modelo *bottom-up*. Porém, penso ser necessário relativizar aqui o significado que *top/down* pode assumir. A tradução literal da expressão, de cima para baixo, pode sugerir que o processo de implementação não apenas foi gestado e conduzido pelo governo e seus órgãos representativos, mas foi, necessariamente, autoritário, impositivo e excludente, deixando de fora, conscientemente, atores que poderiam participar do processo. Nessa leitura, o modelo assume uma perspectiva negativa e pejorativa, devendo ser preterido, como princípio, pelo modelo antagônico, que prevê a participação e o envolvimento desses atores, de modo que a oposição entre autoritarismo e democracia seja correlata daquela entre *top/down* e *bottom/up*. Nessa linha de raciocínio, outras oposições da mesma natureza podem ser pensadas, como aquela entre Estado (necessariamente autoritário e impositivo) e sociedade (necessariamente livre e plural). Assim, a partir de uma lógica reducionista, tudo o que é feito e conduzido pelo Estado é entendido como pernicioso e autoritário, visto que a sociedade não é envolvida no processo de condução de seu próprio futuro. Uma linha de interpretação da história do Brasil, que remonta a Ruy Barbosa e Tavares Bastos (BRANDÃO, 2005), verá o Estado como um de nossos maiores obstáculos para o desenvolvimento de uma sociedade organizada, democrática e participativa, ao passo que uma linha contrária, que remonta a Oliveira Vianna, verá no Estado nosso caminho para a modernidade.

O ponto é que, no contexto das políticas públicas, por definição, o Estado é sempre um ator envolvido. Evidentemente, ele não é obrigado a conduzir o processo todo sozinho e, sob minha perspectiva, há uma diferença importante a ser considerada: a condução pelo Estado na implementação de uma política não é sinônimo de autoritarismo ou de necessária exclusão dos atores que serão afetados diretamente pela política. Além disso, dependendo do tipo da política que será implementada, poderá ser necessário envolver também atores que não serão diretamente afetados pela política, mas não técnicos e especialistas cuja ação poderá ser decisiva para seu sucesso.

No caso da avaliação educacional, conforme vimos no capítulo 2, a história de sua implementação no Brasil, com o Saeb, baseou-se no não envolvimento dos atores educacionais escolares, o que contribuiu para a resistência que a avaliação experimentou nos anos seguintes (e ainda experimenta). Apesar disso, essa condução unilateral do Estado, se pode ser criticada como princípio, não pode apagar a contribuição que a avaliação educacional teve no Brasil nas últimas décadas.

Os sistemas estaduais seguiram um caminho semelhante, inclusive pelo fato do Saeb ter sido sua principal referência. Nos três casos aqui analisados, com as informações disponíveis sobre esse processo, podemos afirmar que os programas de avaliação da alfabetização foram pensados, desenvolvidos e implementados por meio da condução dos estados, através das secretarias de educação. Não há registros de uma participação efetiva dos atores escolares no processo (professores, diretores, coordenadores pedagógicos, funcionários, alunos etc), através de consultas públicas, reuniões etc. Contudo, isso não significa que essa condução pelo poder público tenha sido unilateral e puro ato de vontade. Em nenhum dos três casos, a secretaria de educação conduziu o processo sozinha. Em todos eles, consultores e instituições de pesquisa foram mobilizados. Como vimos, Minas Gerais iniciou o Proalfa estabelecendo uma parceria com o Ceale, da UFMG, dando sequência ao programa em parceria com o CAEd/UFJF. A criação do Peabes Alfa envolveu consulta à experiência de outros estados (incluindo Ceará e Minas) e a participação de Chico Soares, pesquisador da UFMG, além da parceria com o Ceale, seguida da parceria com o CAEd. A criação do Proalfa baseou-se, particularmente, num movimento de mobilização de diversas entidades estaduais e nacionais para o combate ao analfabetismo. Embora o envolvimento do que a literatura chama de *street-level bureaucracy* não tenha ocorrido em nenhum dos estados analisados, não parece ser ajustado dizer que a ação da secretaria foi autoritária e unilateral, pura e simplesmente. As parcerias e consultorias firmadas, mesmo que de maneira menos incisiva do que poderia ser defendido, funcionam como limites técnicos a um agir político baseado puramente na vontade e no poder. Entretanto, reconhecer isso, vale ressaltar, não implica negar a importância que o envolvimento dos atores educacionais poderia ter para o processo de implementação da política de avaliação da alfabetização.

Uma última observação. É complexo definir, com precisão, o período de implementação de uma política. Os programas de avaliação que estamos aqui analisando, por exemplo, ocorrem anualmente. Todos os anos, é como se a política de avaliação fosse implementada novamente. Evidentemente, com o passar do tempo, o programa pode se consolidar e não é como se os atores partissem do zero para sua implementação. De todo modo, é possível compreendermos a implementação como um fenômeno que pode se alongar no tempo. Se é assim, à medida que o tempo passa, a implementação da política pode experimentar a participação de novos atores. Esse argumento foi apresentado no capítulo 2 da tese. No caso da avaliação, à medida que professores e gestores fazem uso dos resultados, a política ganha força para se consolidar. A cada ano, não estariam eles contribuindo para a implementação do programa? Alguns dos

entrevistados para a tese, como veremos adiante, afirmam que esse uso da avaliação como parte das ações das escolas é um dos motivos que explicam a longevidade dos programas aqui analisados. O uso dos resultados é parte fundamental da política de avaliação. Sem sua consolidação na rede, poderíamos falar que a política foi, de fato, implementada?

### **5.3.4 Desenho da avaliação**

Este critério gerador de tipos é específico para políticas de avaliação. A avaliação educacional pode assumir diferentes desenhos, de acordo com os elementos que a compõem. Na verdade, esse é um critério extremamente abrangente, englobando uma série de aspectos que poderiam, cada um deles, levar ao surgimento de novos critérios. Aqui, a decisão foi por reuni-los sob um critério único, o desenho da avaliação.

É importante ressaltar que, como política pública, a avaliação educacional é, em última instância, uma decisão eminentemente política. Apesar disso, entre os critérios selecionados para análise, o desenho da avaliação é o mais técnico deles. A rigor, todos os critérios aqui escolhidos possuem uma dimensão política e uma dimensão técnica. A diferença está na incidência de cada uma dessas dimensões. Quando o alcance da política é definido, assim como seu público-alvo, a existência de uma política de responsabilização, a forma como os municípios participam do programa de avaliação etc, fica evidente a dimensão política da decisão. No entanto, uma vez selecionado um caminho, o aspecto técnico passa a ser critério para as demais escolhas. Por exemplo, se o objetivo da secretaria de educação do estado é atrelar os resultados da avaliação a uma política de responsabilização para as escolas, a decisão sobre o alcance da política deve ser por um programa censitário. Avaliar de modo censitário exige a observância de determinados procedimentos técnicos, que não podem ser simplesmente alterados por força da decisão política sem que isso tenha consequências. O desenho da avaliação envolve uma decisão política, como os demais critérios, mas a incidência da dimensão técnica fica mais patente.

Mas afinal, quais são os elementos presentes no desenho da avaliação. O que está sendo aqui analisado não representa um rol taxativo, mas, antes, exemplificativo. De todo modo, os aspectos seguintes foram considerados pertinentes para a análise do desenho da avaliação dos programas: o tipo de avaliação; as séries/anos e as disciplinas avaliadas; as matrizes de referência; as escalas de proficiência e os padrões de desempenho; a composição dos cadernos; a aplicação dos testes. Todos eles elementos envolvem o trabalho de especialistas e decisões de

caráter técnico. Outros pontos poderiam ser considerados. A divulgação dos resultados da avaliação, por exemplo, poderia fazer parte da análise sobre o desempenho da avaliação. Contudo, a decisão foi por conceder a ela um tópico próprio de análise, tendo em vista a centralidade dos resultados para a avaliação educacional.

As séries e as disciplinas avaliadas são, em regra, uma decisão dos tomadores de decisão. Como o que está em análise são sistemas estaduais de avaliação, a decisão é da secretaria de educação do estado. Apesar disso, há, no contexto da avaliação educacional no Brasil, séries e disciplinas com maior histórico de avaliação. Quanto às séries, a terminalidade da etapa pesa na escolha. Dessa maneira, as séries finais dos anos iniciais (5º ano) e dos anos finais do ensino fundamental (9º ano), bem como o última série do ensino médio (3ª série), são avaliadas por boa parte dos sistemas de avaliação educacional (seja em nível nacional, seja no nível dos estados). O argumento apresentado é o de que os resultados apresentados por essas séries/anos permitem uma análise de maior alcance, referente a todas as séries/ano daquela etapa. Na avaliação da alfabetização, que série/ano seria avaliada? Todo o ciclo? Apenas o 3º ano, seguindo a lógica da terminalidade do ciclo? O 2º ano, de modo a seguir as orientações mais recentes do MEC, ou políticas estaduais no mesmo sentido, sobre a idade esperada para que um aluno esteja alfabetizado?

Seguindo o mesmo raciocínio aplicado às séries/anos, a escolha das disciplinas avaliadas também é uma prerrogativa da secretaria de educação do estado, mas, da mesma forma, há disciplinas com maior histórico de avaliação. É o caso de língua Portuguesa e matemática, disciplinas tratadas como representativas da lógica que preside o aprendizado de outras disciplinas, respectivamente, na área das ciências humanas e na das ciências exatas e da natureza. Ainda, a língua portuguesa pode ser avaliada em diferentes dimensões. Historicamente, a leitura foi a dimensão que recebeu maior atenção dos sistemas de avaliação. Em seguida, a escrita também passou a ser avaliada. Hoje, além da leitura e da escrita, há a avaliação da oralidade, mais especificamente, da fluência em leitura, experiência que já tem sido realizada por algumas redes no Brasil. No que diz respeito às disciplinas, portanto, resta saber quais são avaliadas pelo sistema de avaliação da alfabetização (apenas língua portuguesa ou também matemática) e, no caso de língua portuguesa, quais são as dimensões avaliadas.

Em relação ao tipo de avaliação, o sistema pode colocar em cena avaliações diagnósticas, formativas ou somativas (ou mesmo um conjunto cumulativo desses diferentes tipos), de acordo com os objetivos esperados para a avaliação e para o uso de seus resultados. A decisão por um ou outro tipo de avaliação influenciará os resultados que serão produzidos, o

tipo de divulgação que será feita, a proposta de uso dos resultados, a construção das matrizes de referência e dos testes, a forma e a data da aplicação. De maneira geral, as avaliações diagnósticas são realizadas no início do período letivo, ao passo que as somativas acontecem ao fim do período letivo. As avaliações formativas ocorrem ao longo do período letivo, mas a data de sua aplicação depende do número de avaliações formativas previstas.

As matrizes de referência podem dar mais ênfase a um conjunto de habilidades do currículo do que a outra. Em língua portuguesa, por exemplo, as habilidades de decodificação podem ser mais enfatizadas, respondendo pela maior parte da matriz. Essa é uma escolha, como já afirmado, que depende dos tomadores de decisão e dos objetivos que perseguem com a política de avaliação. Ela revela, ainda, a concepção de alfabetização adotada pelo estado e encarnada pelo programa, algo que será discutido em tópico próprio. Nesse caso, a análise deve ser feita caso a caso, observando a matriz para cada edição da avaliação do programa.

A escala de proficiência é outro elemento da avaliação que pode fazer com que os sistemas apresentem diferenças uns em relação aos outros. A escala de proficiência da alfabetização pode ser própria do sistema, ou seja, desenvolvida especificamente para ele, sem estabelecer parâmetros com base em outras escalas. Por outro lado, as escalas podem ter base em outros sistemas de avaliação, permitindo uma comparabilidade que o primeiro caso, das escalas próprias, não autoriza.

As diferenças podem estar também nos modelos de caderno adotados para cada um dos programas e na forma como ocorre a aplicação. O aplicador pode ser o próprio professor da escola, que pode aplicar o teste para a própria turma para a qual leciona ou para turmas de outros professores; outra possibilidade é a figura do aplicador externo, contratado e capacitado pelo sistema de avaliação especificamente para essa função. Ainda, os testes podem ser lidos pelo aplicador ou respondidos pelo aluno sem essa leitura de auxílio. Uma pesquisa detalhada sobre a influência (e magnitude) desses dois aspectos sobre o resultados dos alunos nos testes ainda está por ser realizada. Como hipóteses: i) a de que o professor da turma auxilia, consciente ou inconscientemente, os alunos a responderem às questões do teste, exercendo um efeito de aumento nos resultados médios; e ii) a de que a leitura do teste pelo aplicador beneficia os alunos, também exercendo um efeito de aumento nos resultados médios.

### 5.3.4.1 Tipo de avaliação, séries e disciplinas avaliadas

O Spaece Alfa avaliou, ao longo do período analisado, de forma censitária, os alunos do 2º ano do ensino fundamental das redes municipais e estadual. A escolha do 2º ano, como vimos em tópico anterior, está relacionada à perspectiva adotada pelo Paic, um dos indutores para o surgimento do programa, de que os alunos deveriam estar alfabetizados aos 8 anos. O Spaece Alfa é uma avaliação somativa, aplicada ao fim de cada ano, com o objetivo de fornecer os subsídios necessários para o planejamento do ano seguinte. Em toda sua história, o programa dedicou-se à avaliação da leitura, nunca tendo avaliado escrita e matemática.

O Proalfa apresentou um desenho menos estável<sup>171</sup> do que aquele que caracteriza o Spaece Alfa. O 3º ano do ensino fundamental foi avaliado censitariamente ao longo de todo o período analisado. No entanto, entre 2007 e 2014, o programa avaliou, por amostra, o 2º e o 4º ano do ensino fundamental. A extensão da avaliação ao 4º amplia o entendimento da alfabetização como ciclo, um progresso que se estende ao longo das séries/anos. Além disso, no mesmo período, o Proalfa avaliou, de modo censitário, os alunos do 4º ano com baixo desempenho (BD). A avaliação do baixo desempenho era aplicada aos alunos que, na avaliação censitária (3º ano) do ano anterior não haviam consolidado as habilidades previstas para o período de escolaridade. O objetivo dessa avaliação era permitir uma intervenção pedagógica pontual e específica, de modo a elevar o desempenho desses alunos com dificuldades ao padrão recomendado de leitura. Essa avaliação do baixo desempenho em Minas é um exemplo interessante de como os resultados da avaliação podem ser usados para pensar em novas ações pedagógicas que auxiliem os alunos com as dificuldades diagnosticadas pelo próprio teste. Assim como no Spaece Alfa, adotou avaliações somativas, aplicadas ao fim do ano letivo.

A partir da edição de 2015, apenas o 3º ano foi avaliado, mesmo com a meta estabelecida pelo estado de alfabetizar as crianças até os 8 anos de idade. Desde sua primeira edição, em 2005, até 2016, o Proalfa avaliou apenas a leitura. A edição de 2017 trouxe uma mudança significativa no que estava sendo avaliado: foram aplicados testes de escrita e matemática, além dos de leitura.

---

<sup>171</sup> Estabilidade, aqui, não tem implicações valorativas. Ser estável não significa melhor, nesse caso, apenas indica a manutenção das características principais e do formato, sem experimentar alterações. Apesar disso, é possível identificar vantagens e desvantagens nessa estabilidade. A ausência de alterações significativas permite a construção de uma linha histórica, ao passo que não fornece aos tomadores de decisão informações sobre outras séries e disciplinas. Em realidade, não são fenômenos excludentes: é possível manter as características da avaliação já existente, construindo a linha histórica, e, somado a isso, acrescentar novas séries e disciplinas.

O Paebes Alfa, desde 2009, avaliou as três séries do ciclo de alfabetização. Ao longo de todo o período analisado, leitura e escrita foram avaliadas. Matemática passou a ser avaliada a partir de 2010. Assim como o Spaece Alfa, nesse quesito, trata-se de uma avaliação com formato estável. Apesar disso, essa é uma proposta de avaliação que difere significativamente tanto de Spaece Alfa quanto de Proalfa. Avaliar as três séries do ciclo significa que a avaliação ocorre a despeito da série/ano estabelecida como aquela em que o processo de alfabetização dos alunos deveria ser concluído. Isso permitiu ao estado e aos municípios acompanhar de forma muito mais minuciosa e detalhada o aprendizado de seus alunos no ciclo. As redes municipais e estadual foram sempre avaliadas e o programa ainda permite a adesão de escolas particulares, outra diferença em relação aos programas de Minas Gerais e Ceará.

Entre 2009 a 2015, o Paebes Alfa aplicou avaliações diagnósticas e somativas, no começo e no fim do ano letivo, respectivamente. Essas duas avaliações no mesmo ano eram conhecidas por ondas (1ª e 2ª ondas). No período analisado, a 1ª onda era aplicada apenas para o 1º ano, com o intuito de analisar como os alunos estavam chegando, em termos de aprendizagem, nas escolas das redes públicas para o início do ensino fundamental. A 2ª onda era aplicada para as três séries do ciclo. As edições de 2016 e 2017 contaram apenas com a aplicação da avaliação somativa (a 2ª onda) para as três séries. O quadro a seguir sumariza todas essas informações para os três programas.



Quadro 6 - Tipos de avaliação, disciplinas e séries avaliadas – por edição e programa

Programas	Características	2009	2010	2011	2012	2013	2014	2015	2016	2017	
Spaace Alfa	Tipo de avaliação	Somativa	Somativa	Somativa	Somativa	Somativa	Somativa	Somativa	Somativa	Somativa	
	Séries/anos avaliados	2º censitário	2º censitário	2º censitário	2º censitário	2º censitário	2º censitário	2º censitário	2º censitário	2º censitário	
	Disciplinas avaliadas	Leitura	Leitura	Leitura	Leitura	Leitura	Leitura	Leitura	Leitura	Leitura	
Proalfa	Tipo de avaliação	Somativa	Somativa	Somativa	Somativa	Somativa	Somativa	Somativa	Somativa	Somativa	
	Séries avaliadas	2º e 4º amostrais, 4º BD, 3º censitário	2º e 4º amostrais, 4º BD, 3º censitário	2º e 4º amostrais, 4º BD, 3º censitário	2º e 4º amostrais, 4º BD, 3º censitário	2º e 4º amostrais, 4º BD, 3º censitário	2º e 4º amostrais, 4º BD, 3º censitário	2º e 4º amostrais, 4º BD, 3º censitário	3º censitário	3º censitário	3º censitário
	Disciplinas avaliadas	Leitura	Leitura	Leitura	Leitura	Leitura	Leitura	Leitura	Leitura	Leitura, escrita e matemática	
Paebes Alfa 1ª onda	Tipo de avaliação	Diagnóstica	Diagnóstica	Diagnóstica	Diagnóstica	Diagnóstica	Diagnóstica	Diagnóstica	-	-	
	Séries/anos avaliados	1º censitário	1º censitário	1º censitário	1º censitário	1º censitário	1º censitário	1º censitário	-	-	
	Disciplinas avaliadas	Leitura e escrita	Leitura, escrita e matemática	Leitura, escrita e matemática	Leitura, escrita e matemática	Leitura, escrita e matemática	Leitura, escrita e matemática	Leitura, escrita e matemática	-	-	
Paebes Alfa 2ª onda	Tipo de avaliação	Somativa	Somativa	Somativa	Somativa	Somativa	Somativa	Somativa	Somativa	Somativa	
	Séries avaliadas	1º, 2º e 3º censitários	1º, 2º e 3º censitários	1º, 2º e 3º censitários	1º, 2º e 3º censitários	1º, 2º e 3º censitários	1º, 2º e 3º censitários	1º, 2º e 3º censitários	1º, 2º e 3º censitários	1º, 2º e 3º censitários	
	Disciplinas avaliadas	Leitura e escrita	Leitura, escrita e matemática	Leitura, escrita e matemática	Leitura, escrita e matemática	Leitura, escrita e matemática	Leitura, escrita e matemática	Leitura, escrita e matemática	Leitura, escrita e matemática	Leitura, escrita e matemática	

Fonte: Paebes Alfa 2017, Revista do Sistema – Redes Municipais. Tabulação própria.

### 5.3.4.2 Matrizes de referência

As matrizes de referência são documentos fundamentais para a avaliação em larga escala. Elas são a base para a construção dos itens e a montagem dos testes. As matrizes são construídas com base no currículo local, o que significa que, por mais que possam ser compostas por habilidades semelhantes, a rigor, são específicas de cada estado. Elas variam também por disciplina e etapa de escolaridade. De modo simplificado, podemos dizer que as matrizes definem o que será avaliado pelos testes. Cada habilidade da matriz é descrita a partir de um verbo central (resolver, ler, compreender, identificar, conhecer, distinguir etc), cuja fonte remonta à taxonomia de Bloom. Os descritores associam o conteúdo curricular a operações cognitivas, indicando as habilidade que serão avaliadas através dos itens do teste. A reunião de descritores associados em torno de habilidades semelhantes forma uma competência (por exemplo, *Recuperação de informações no contexto de práticas sociais de leitura* é uma

competência presente na matriz de referência do Proalfa nas edições de 2008 e 2009, envolvendo três diferentes descritores), ao passo que a reunião de competências associadas forma um tópico (por exemplo, na matriz de 2010 e 2011 do Paebes Alfa, para leitura, o tópico 1, *Apropriação do sistema de escrita*, é composto por duas competências, *Identificação de letras do alfabeto* e *Uso adequado da página*)

É importante ressaltar que a matriz de referência não se confunde com o currículo. Muitas críticas direcionadas à avaliação em larga escala baseiam-se na ideia de que a matriz de referência é transformada, na prática educacional das escolas, no próprio currículo. Esse uso da matriz como currículo é reducionista e não é uma característica imposta pelo instrumento de avaliação, embora seja um uso que tenha sido feito, equivocadamente, dele. A matriz tem como fonte o currículo, mas é mais restrita do que ele. Ela só contempla as habilidades que podem ser avaliadas através de testes padronizados, compostos por questões de múltipla escolha. Diante disso, nem todas as habilidades previstas pelo currículo podem integrar a matriz, por limitações metodológicas relativas ao formato do teste.

A decisão sobre a matriz de referência, ou seja, sobre o que será avaliado pelo teste, é da secretaria de educação, por meio da equipe de sua equipe de avaliação. Apesar disso, a parceria das secretarias com instituições especializadas, fez com que, historicamente, para os três programas aqui analisados, a construção das matrizes ocorresse de forma conjunta, entre a equipe da secretaria e o CAEd/UFJF. Segundo a entrevista coletiva<sup>172</sup> realizada com as especialistas em elaboração de itens de alfabetização e montagem de testes do CAEd, o processo de elaboração das matrizes sempre culmina com a validação das equipes das secretarias, mas os especialistas do CAEd foram participando cada vez mais desse processo de elaboração. Em relação aos itens e aos testes, o mesmo processo pôde ser observado. No âmbito do Spaece Alfa e do Proalfa, os itens e os testes são validados com frequência pelas equipes técnicas da secretaria. Em relação ao Paebes Alfa, essa validação pela secretaria tornou-se menos frequente ao longo do tempo.

Sobre os itens, todos eles, na alfabetização, são compostos por 4 alternativas de respostas. Até a edição de 2013 do Spaece Alfa e de 2015 do Paebes Alfa, por força de contrato, um quantitativo de itens, especificado em contrato, elaborados pelos professores das redes de ensino deveriam integrar os testes. Os itens eram elaborados pelos professores após a realização de oficinas de elaboração realizadas no estado e ministradas pelos especialistas do CAEd.

---

<sup>172</sup> Entrevista com duas especialistas em elaboração de itens para a alfabetização e com a supervisora da equipe de montagem de testes, que possui ampla experiência com a elaboração de itens, realizada no dia 14 de maio de 2019.

Depois de elaborados pelos professores, os itens eram validados pela equipe de especialistas do CAEd antes de integrarem os testes. Esse processo, embora interessante sob o ponto de vista da capacitação dos professores para a elaboração de itens, deixou de ser realizado pelos estados em virtude da necessidade, cada vez maior, de testes mais ajustados à população, o que o trabalho dos especialistas era capaz de garantir em menor tempo.

Observando as matrizes de leitura<sup>173</sup> para os três programas<sup>174</sup>, é possível perceber que, em geral, elas apresentam um grau considerável de estabilidade, tendo em vista que experimentaram apenas mudanças pontuais em algumas edições. No que tange à leitura, o Paebes Alfa apresentou uma matriz única como base para o ciclo de alfabetização (abarcando as três séries avaliadas) até o ano de 2011. A partir da edição de 2012, houve uma especificação em relação aos níveis referentes a cada um dos descritores e a definição de quais deles eram cobrados para cada série avaliada. Esse formato permaneceu para 2013. Entre 2009 e 2010, houve uma mudança na matriz, que contava com três tópicos sem especificação de competências e passou a contar com 5 tópicos e 12 competências (além disso, o número de descritores passou de 25 para 30). Entre 2014 e 2017, as matrizes foram divulgadas já com a separação por série avaliada e permaneceram sem alterações, contando com cinco tópicos: i) reconhecimento de convenções do sistema alfabético; ii) apropriação do sistema alfabético; iii) leitura: compreensão, análise e avaliação; iv) usos sociais da leitura e da escrita; e v) produção escrita.

As matrizes do Paebes Alfa de 1º, 2º e 3º anos, para língua portuguesa, são compostos pelo mesmo conjunto de descritores: as matrizes de cada série representam subconjuntos desse conjunto maior. De acordo com as especialistas entrevistadas, as diferenças estão presentes nas classes dos itens e em suas dificuldades. A expertise que foi sendo desenvolvida pelo CAEd com a avaliação educacional influenciou a maneira como os itens foram sendo elaborados e a forma como a matriz passou a ser representada nos testes. Segundo as entrevistadas, o aprimoramento do processo de elaboração de itens culminou, em 2015, com o desenvolvimento de padrões técnicos para a elaboração.

Ainda de acordo com elas, a matriz de avaliação da alfabetização mais completa é a do Paebes Alfa, visto que cobre melhor o currículo que lhe dá base do que as matrizes dos outros dois programas.

---

<sup>173</sup> Escrita e matemática possuem tópicos próprios de análise, apresentados a seguir.

<sup>174</sup> As matrizes de referência de Paebes Alfa, Proalfa e Spaece Alfa encontram-se nos Anexos I, II e III, respectivamente.

O Proalfa, para leitura, também apresentou uma matriz única para as séries avaliadas (o 3º ano censitário, como referência, e 2º e 4º anos amostrais). Como base para a produção de testes, contudo, alguns descritores da matrizes não eram utilizados no 4º ano. De maneira geral, a matriz do Proalfa sofreu poucas alterações. Algumas competências foram reordenadas, passando de 5 a 3, mas, em essência, foram: *apropriação do sistema de escrita; leitura: compreensão, análise e avaliação; e usos sociais da leitura e da escrita* (em 2017, acoplou os descritores para escrita). A competência *Reconhecimento de convenções do sistema alfabético* apresentava descritores mais ajustados ao primeiro e ao segundo ano, o que ficou evidente à medida que os resultados do 3º ano foram melhorando, de modo que foi retirada da matriz para o 3º ano.

Entre 2012 e 2017, a matriz de referência do Spaece Alfa para leitura, no 2º ano do ensino fundamental é a mesma, organizada em torno de dois eixos, com 19 descritores no total: i) apropriação do sistema de escrita (habilidades relacionadas à identificação e ao reconhecimento de aspectos relacionados à tecnologia da escrita); e ii) leitura (habilidades relacionadas à leitura de palavras, de frases e de textos). A matriz para as edições de 2010 e 2011 é a mesma, apresentando apenas uma modificação em relação àquela de 2012: ela possui um descritor a menos (*Reconhecer o gênero discursivo*). A matriz de 2009, em relação às demais, apresenta diferenças na nomenclatura utilizada pelos descritores, embora as habilidades avaliadas sejam, essencialmente, as mesmas: por exemplo, identificar finalidade de textos de diferentes gêneros passou a identificar o propósito comunicativa em diferentes gêneros; além disso, a palavra decodificação passou a não ser mais utilizada na matriz (decodificar palavras deu lugar a ler palavras). De modo geral, portanto, é possível afirmar que a matriz de referência para o teste de leitura no 2º ano no Spaece Alfa permaneceu avaliando, em essência, as mesmas habilidades ao longo do período analisado.

#### 5.3.4.3 Escalas de proficiência, padrões de desempenho, cadernos de teste e aplicação

A escala de proficiência é um intervalo numérico que tem como objetivo traduzir medidas em diagnósticos qualitativos e pedagógicos sobre o desempenho escolar. Ela é organizada por níveis (de 25 pontos, em regra), que indicam o grau de consolidação das habilidades previstas pelas matrizes de referência. Para as demais etapas de escolaridade, as escalas de proficiência, em geral, seguem o mesmo padrão, definido pelo Saeb. Contudo, para a alfabetização, cada programa tem definido sua própria escala. É o que acontece com Paebes

Alfa, Proalfa e Spaece Alfa. Nenhuma das três escalas são comparáveis. As escalas de Paebes Alfa e Proalfa variam de 0 a 1.000 pontos, ao passo que a escala do Spaece Alfa varia de 0 a 500 pontos.

Os padrões de desempenho são categorias definidas a partir da escala de proficiência e de seus cortes numéricos, agrupando um conjunto de níveis de acordo com as metas educacionais estabelecidas por cada estado. A seleção de cortes numéricos na escala é o que define os padrões de desempenho. Se uma secretaria de educação, diante de uma escala de proficiência, opta por selecionar como recomendado um conjunto de habilidades representado por determinados níveis da escala, esse será o padrão recomendado na avaliação. Os cortes dos padrões de desempenho variam por etapa e disciplina avaliada. Considerando a avaliação de leitura, os padrões de desempenho não são os mesmos para os três programas aqui avaliados.

O Spaece Alfa, para a avaliação da leitura do 2º ano, trabalhou, ao longo do período analisado, com 5 padrões de desempenho: i) Não alfabetizado (até 75 pontos); ii) Alfabetização incompleta (de 75 a 100 pontos); iii) Intermediário (de 100 a 125 pontos); iv) Suficiente (de 125 a 150 pontos); e v) Desejável (a partir de 150 pontos).

Para a avaliação censitária do Proalfa, no 3º ano do ensino fundamental, até 2014, existiam três padrões de desempenho: i) Baixo (até 450 pontos); ii) Intermediário (de 450 a 500 pontos); e iii) Recomendado (a partir de 500 pontos). A partir da edição de 2015, um quarto padrão foi criado, o Avançado, e o novo arranjo ficou desta forma: i) Baixo (até 450 pontos); ii) Intermediário (de 450 a 550 pontos); e iii) Recomendado (de 550 a 650 pontos); e iv) Avançado (a partir de 650 pontos). A criação desse quarto padrão ocorreu por meio de uma decisão da Seduc Minas, juntamente com o CAEd, com o objetivo de cobrir melhor as habilidades avaliadas, conferindo maior precisão à distribuição dos alunos pelos grupos de desempenho.

No Espírito Santo, para leitura, o Paebes Alfa adotou padrões específicos para cada série avaliada (os nomes são os mesmos): i) Abaixo do Básico (1º ano – até 400 pontos); (2º ano – até 500 pontos); (3º ano – até 600 pontos); ii) Básico (1º ano - entre 400 e 500 pontos); (2º ano – entre 500 e 600 pontos); (3º ano – entre 600 e 650 pontos); iii) Proficiente (1º ano – até 600 pontos); (2º ano – entre 600 e 700 pontos); (3º ano – entre 650 e 750 pontos); iv) Avançado (1º ano – a partir de 600 pontos); (2º ano – a partir de 700 pontos); (3º ano – a partir de 750 pontos).

Os modelos de caderno de teste podem variar por programa, disciplina e série avaliada. O quadro a seguir mostra a variação dos modelos por programa ao longo do período analisado. O número de modelos de caderno relaciona-se com o número de itens que o programa pretende

aplicar na edição, o que, por sua vez, varia em função do número de itens novos para a avaliação no estado e do quantitativo de itens já utilizados que ainda podem ser usados. O número de itens por caderno pode ser influenciado por custos, se os testes forem impressos, como é o caso dos três programas.

Quadro 7 - Modelos e itens por caderno de teste – por edição e por programa

Programas	Séries – disciplinas	Cadernos e itens	2009	2010	2011	2012	2013	2014	2015	2016	2017
Spaace Alfa	2º EF – Leitura	Modelos de caderno	2	12	12	12	12	12	12	10	10
		Itens por caderno	24	24	24	24	24	24	24	24	25
Proalfa	2º EF – Leitura	Modelos de caderno	36	36	36	36	36	36	-	-	-
		Itens por caderno	21	21	21	21	21	21	-	-	-
	3º EF – Leitura	Modelos de caderno	4	4	4	4	4	4	16	16	16
		Itens por caderno	28	28	28	28	28	28	20	20	20
	3º EF – Matemática	Modelos de caderno	-	-	-	-	-	-	-	-	16
		Itens por caderno	-	-	-	-	-	-	-	-	-
	4º EF – Leitura	Modelos de caderno	28	28	28	28	28	28	-	-	-
		Itens por caderno	30	30	30	30	30	30	-	-	-
Paebes Alfa	1º, 2º e 3º EF – Matemática	Modelos de caderno	-	8	8	8	8	8	8	8	8
		Itens por caderno	-	24	24	24	24	24	24	24	24
	1º, 2º e 3º EF – Língua Portuguesa	Modelos de caderno	8	8	12	12	12	12	12	12	12
		Itens por caderno	24	25	25	25	25	25	25	25	25

Fonte: Caed/UFJF. Equipe de testes. Tabulação do autor.

Em relação à aplicação dos testes, em entrevista<sup>175</sup> com a especialista do CAEd responsável pela capacitação dos aplicadores, ela afirma que há diferenças entre os três programas. No Spaace Alfa, os aplicadores eram externos até 2017. A partir de então, os aplicadores passaram a ser voluntários, incentivados a participarem da avaliação pela secretaria. Essa decisão ocorreu, segundo a entrevistada, em virtude de corte de custos. De acordo com ela, quando os aplicadores são externos, contratados e capacitados pelo CAEd, há menos perda de informações ao longo do processo. No Paebes Alfa e no Proalfa, os aplicadores são os próprios professores da rede. Nesses casos, o CAEd orienta as secretarias para que os professores não apliquem nas turmas pelas quais são responsáveis, de modo a evitar qualquer tipo de influência nos resultados.

De acordo com a entrevistada, anualmente, o CAEd prepara um material específico com orientações para a aplicação dos testes. No caso do Proalfa, esse material, por vezes, não chegou de modo adequado aos professores que aplicam os testes. No caso do Paebes Alfa, há a figura do monitor contratado para auxiliar na organização da aplicação, o que ajuda a diminuir ruídos.

<sup>175</sup> Realizada no dia 03 de junho de 2019.

No dois programas, Paebes Alfa e Proalfa, os professores não são remunerados para aplicarem os testes, entendidos como parte da responsabilidade dos atores escolares.

Os procedimentos de aplicação são estabelecidos pelo CAEd. Eventualmente, as secretarias de educação fazem sugestões e ajustes. Um dos procedimentos indicados pelo CAEd é o de que a aplicação deve começar quando 80% dos alunos previstos para a turma já estiverem na sala. Segundo a entrevistada, no caso do Spaece, por conta da força da política de premiação no estado, em regra, os aplicadores esperam todos os alunos chegarem para dar início à aplicação. Em relação à organização geral da aplicação, ela é feita de maneira conjunta entre o CAEd e as secretarias.

#### 5.3.4.4 A avaliação da escrita<sup>176</sup>

O embate entre a perspectiva do letramento e a crítica à ênfase nos elementos fonéticos da língua como base para o ensino não faz olvidar a importância da escrita no processo de alfabetização. Em um caso ou em outro, não importando a perspectiva da alfabetização que se adote, a escrita é tida como uma faceta central do processo. Sendo assim, adotando a ideia de que a avaliação é um direito e está associada à qualidade da educação (como vimos no capítulo 2 da tese), defender uma avaliação sistêmica no âmbito da alfabetização deveria incluir a defesa também da avaliação da escrita. Contudo, a despeito da lógica que dá suporte a esse raciocínio, do ponto de vista metodológico e logístico, não é tão simples transformar a crença de que a defesa de um sistema de avaliação da alfabetização inclui a escrita em realidade. Uma coisa é conceber essa ideia em termos de monitoramento da qualidade da educação em todas as suas facetas. Outra é desenvolver instrumentos capazes de fazer isso, incluindo a logística para sua aplicação e correção e os custos que isso acarreta.

A avaliação em larga escala, conforme vimos, experimentou grande desenvolvimento no Brasil a partir da introdução da TRI como base para a produção dos itens, testes e medidas de desempenho (em meados dos anos 1990). Os testes padronizados de múltipla escolha tornaram-se o modelo seguido pelos sistemas de avaliação no Brasil. A avaliação da escrita colocou em cena alguns problemas: i) o da correção de questões abertas para um número muito grande de alunos; ii) como consequência, o problema do tempo de devolução dos resultados

---

<sup>176</sup> Para recolher informações mais detalhadas sobre a escrita, foi realizada uma entrevista com os dois membros da equipe de escrita que, no CAEd/UFJF, eram responsáveis pela elaboração dos itens, montagem dos testes e capacitação dos corretores da escrita à época (a entrevista foi realizada no dia 28 de maio de 2019). Ambos já trabalhavam no CAEd há, pelo menos, oito anos, e acompanharam de perto o desenvolvimento da avaliação da escrita nos Paebes Alfa e no Proalfa.

para as redes (lembrando que o objetivo central da avaliação educacional é dar suporte para a tomada de decisões); e iii) o de como produzir uma medida de proficiência para questões abertas. Se as questões de escrita envolvem correção, ao contrário das questões objetivas, que estão baseadas na correspondência direta entre as respostas e os gabaritos das questões, fica evidente o tema da influência da subjetividade dos corretores. Como é impossível que um único corretor fique responsável pela correção de todos os itens de todos os testes (e mesmo assim, a questão da diferença de correção permaneceria de pé, diante do fator temporal), a necessidade de muitos corretores abre espaço para que muitos pontos de vista diferentes sejam utilizados para realizar a correção, mesmo diante de chaves de correção bem estabelecidas. Como lidar com o problema da leniência ou rigidez dos corretores? A dupla (ou mais) correção é uma saída. Ela atenua o problema, mas não o resolve. Soma-se a isso outro fator de dificuldade: um grande número de corretores aumenta significativamente os custos de um sistema de avaliação e o elemento econômico é um dos principais critérios a serem analisados quando a decisão sobre a implementação de um programa de avaliação está em jogo.

A TRI, se já havia dado respostas satisfatórias para a produção de medidas educacionais oriundas de testes objetivos padronizados, ainda precisava ser testada e ajustada para os casos de itens abertos, como no caso da escrita. Isso exigiu considerável esforço de pesquisa, em especial, para a construção de uma escala de escrita e para a produção das medidas de proficiência.

Diante disso, a decisão sobre a avaliação da escrita não passa somente pela crença de que um sistema de avaliação da alfabetização deve avaliar todas as dimensões do processo. Outros aspectos precisam ser analisados, como os recursos e a tecnologia disponíveis para tanto.

O Spaece Alfa nunca avaliou a escrita, ao passo que o Paebes Alfa avalia de modo contínuo desde a edição de 2009. No período analisado pela tese, o Proalfa realizou avaliação da escrita somente na edição de 2017. A Provinha Brasil, realizada em 2008, foi apontada, pelos dois entrevistados da equipe de escrita do CAEd/UFJF, como o principal fator de influência para a realização de avaliações de escrita no âmbito dos sistemas próprios. O Saers, no Rio Grande do Sul, por exemplo, aplicou um ditado para os alunos do 2º ano do ensino fundamental nas edições de 2007 e 2008, sem fazer uso, no entanto, da TRI.

Apesar de ter realizado uma avaliação da escrita apenas em 2017, a análise das matrizes de referência do Proalfa entre 2009 e 2017 pode confundir o pesquisador. Até o ano de 2012, a construção das matrizes de referência do programa ficou a cargo do Ceale/UFMG, em uma parceria com a Seduc/MG. Essa construção contou com a previsão, desde o início, de algumas



habilidades de escrita, caso a Seduc decidisse pela realização da avaliação em momento posterior (o que só ocorreu em 2017). Como é possível observar no anexo II, a matriz do Proalfa contou com dois descritores de escrita desde as edições de 2008 e 2009: i) escrever palavras; e ii) escrever frases/textos. Na matriz de referência de 2007, aparece um terceiro descritor, além desses dois: demonstrar conhecimentos sobre a escrita do nome.

O Paebes Alfa é, dentre os três, o programa com uma avaliação da escrita consolidada. Desde sua edição de 2009, contou com instrumentos dedicados à sua avaliação. No anexo II, são apresentadas, na íntegra, as matrizes do programa, com os descritores de escrita para cada ano. Como pode ser observado, em 2009, assim como na matriz do Proalfa, havia dois descritores para escrita: escrever palavras e escrever frases/textos. Na edição de 2010 do Paebes Alfa, a matriz sofreu alterações, passando a prever três descritores e escrita: i) escrita de palavras; ii) escrita de frases; e iii) produzir texto, formato que foi mantido até a edição de 2017. Trata-se, portanto, de um matriz estável para escrita. Segundo os entrevistados, o mesmo pode ser dito para a chave de correção dos testes, que não sofreu modificações significativas ao longo do período analisado.

O que mais mudou em relação aos testes de escrita do Paebes Alfa, de acordo com os dois entrevistados, foi a elaboração dos itens. Essa mudança está relacionada à alteração na parceria da secretaria com o Ceale/UFG. Até 2011, o Ceale era responsável pela elaboração das matrizes, dos itens e dos testes. O CAEd/UFJF era responsável por aplicar e corrigir os testes de escrita. A partir da 2ª onda de 2011, o CAEd passou a também elaborar os itens e testes, e não apenas a aplicá-los e corrigi-los. Com essa mudança, a Seduc deu ao CAEd autonomia e liberdade para pesquisa na área de avaliação da escrita.

Em relação aos itens, uma mudança significativa foi na escolha das palavras e dos suportes imagéticos. A ideia passou a ser não trabalhar mais com figuras, mas com imagens reais (fotografias de objetos reais). Além disso, passou-se a priorizar imagens mais limpas, com menos informações secundárias, o que poderia desviar a atenção do aluno e confundi-lo. Os comandos dos itens também mudaram, ficando mais diretivos e claros. A autonomia dada pela Seduc ao CAEd a partir da edição de 2012 do Paebes Alfa, fez desenvolver a pesquisa com a avaliação da escrita no próprio CAEd. A Seduc não aprova os testes de escrita, desde então, ficando os mesmos integralmente sob a responsabilidade dos especialistas do CAEd. Por conta disso, o centro criou, em 2013, uma equipe própria e especializada em avaliação da escrita (até então, os testes de escrita eram elaboradas pela equipe de língua portuguesa, responsável também pelos itens e testes de leitura). Esse foi um grande passo para a avaliação da escrita nos

programas que possuíam parceria com o CAEd. A especialização da equipe de elaboração de itens e testes de escrita fez com que houvesse um salto de qualidade nos instrumentos.

Outro efeito da criação de uma equipe especializada pode ser observado no processo de correção. A partir de 2013, os critérios para selecionar os corretores dos testes de escrita passaram a ser mais rigorosos. Desde então, os corretores precisam ser especialistas, com titulação nas áreas de Letras, Pedagogia ou Normal Superior. Eles são selecionados através de processo seletivo (com análise de currículo, prova escrita de conhecimentos linguísticos, prova de correção e análise de experiência docente), e, em regra, atuam em sala de aula (experiência docente de, pelo menos, três anos, é desejável no âmbito do processo seletivo). Os corretores são capacitados, presencialmente, para o processo de correção e acompanhados por supervisores (selecionados, também através de processo seletivo, entre os corretores). Todos ficam sob a supervisão geral da equipe de escrita do CAEd. Soma-se a isso o desenvolvimento de uma plataforma própria para correção, feita exclusivamente online. Isso deu agilidade e segurança ao processo de correção. Depois da ação dos corretores, os supervisores, usando a plataforma, fazem uma segunda correção, com uma amostra por corretor (2% do total corrigido por ele), de modo a garantir a qualidade da correção no processo. Essas mudanças promoveram um aumento na qualidade de todo o processo que envolve a avaliação da escrita. O Paebes Alfa foi o grande laboratório onde isso aconteceu.

A avaliação duradoura e consolidada da escrita (em nenhuma de suas edições, o programa deixou de avaliá-la) é um dos elementos que, segundo os entrevistados, especialistas do CAEd/UFJF nessa área, fazem do Paebes Alfa um programa de avaliação da alfabetização mais completo, abrangente e complexo que os demais (Proalfa e Spaece Alfa). De acordo com os entrevistados, o Paebes Alfa costuma antecipar tendências na área de avaliação de escrita (o teste de escrita aplicado na avaliação do Saethe, na rede municipal de Teresina, em 2015, por exemplo, foi influenciado pelo modelo do Paebes Alfa).

Em realidade, a edição do Paebes Alfa 2009 pode ser considerada um marco para a avaliação da escrita em larga escala no Brasil<sup>177</sup>. Ela usa a TRI desde o início, com base em uma escala e em uma linha histórica considerável. Em 2008, houve um piloto de avaliação da escrita do estado, mas sem o uso de metodologia ajustada para construir uma escala e produzir uma linha histórica. Da mesma forma, em Minas Gerais, sob o nome de Simavinho, o Proalfa realizou uma experiência com a avaliação da escrita em 2005, 2006 e 2007, contudo, não houve

---

<sup>177</sup> Para uma análise mais detalhada acerca da avaliação da escrita no Paebes Alfa, em especial, no que diz respeito aos resultados, à ancoragem dos itens e aos testes, ver Silva, 2013b.

produção de medida de proficiência, em virtude da não utilização da TRI na ocasião. Nestas três edições, o Proalfa produziu resultados de escrita apenas com base na TCT – Teoria Clássica dos Testes, apresentando o percentual de acerto dos itens. No teste de 2007, por exemplo, foram aplicados três itens de escrita, dois versando sobre a escrita de palavras (bruxa e caminhão) e um sobre a escrita de uma pseudopalavra (tapoja).

Em 2017, o teste de escrita do Paebes Alfa, influenciado pela ANA, sofreu algumas modificações, passando a ser menos extenso do que aquele aplicado em edições anteriores. Os itens avaliaram a escrita de palavras e textos. As frases não eram avaliadas isoladamente, sendo o tópico frasal analisado no âmbito da própria construção textual. A avaliação de escrita para o Proalfa neste mesmo ano também seguiu esse formato.

Até 2015, os testes de escrita do Paebes Alfa eram muito dependentes do suporte imagético. Compostos por 5 itens, eles eram compostos por itens que avaliavam i) cópia de frase, ii) escrita de palavras (por ditado ou com base em imagem; podendo ter dois itens); iii) escrita de frases (também por ditado ou com base em imagem, podendo ter dois itens); e iv) escrita de texto (lista ou narrativa a partir de sequência de imagens; bilhete ou convite, somente no 3º ano do ensino fundamental). Todo o texto era povoado por imagens, sem espaço para a escrita autônoma. De acordo com os dois entrevistados, a ANA foi importante para mudar esse modelo de teste.

A partir de 2015, os testes de escrita do Paebes Alfa começaram a dar mais atenção aos itens de produção autônoma, baseados em um tema para livre desenvolvimento do aluno. Na primeira edição da ANA, em 2013, um dos itens do teste, baseado em imagem, levou, de acordo com os entrevistados, a problemas para o desenvolvimento do texto por parte dos alunos. Na edição de 2014 da ANA, no lugar da imagem, foi colocada uma frase inicial, com ponto de partida para a escrita do texto, mas, na avaliação dos especialistas, isso também atrapalhou a construção do texto. Outras mudanças que puderam ser observadas no Paebes Alfa a partir de 2015 foram: i) a retirada da lista de palavras do teste; e ii) a supressão de gêneros textuais como o convite e o bilhete, que já não estavam mais sendo usados pelos alunos – muitos sequer conheciam esses gêneros).

A ANA trouxe um design diferente para o teste de escrita: uma única página de teste, com 2 itens sobre a escrita de palavras e 1 item para a produção de texto. Tanto o Paebes Alfa quanto o Proalfa (2017) usam duas páginas de teste (frente e verso). No Prolafa, são 4 itens (sem avaliação da escrita de frase), ao passo que no Paebes Alfa são 5 itens (2 palavras, 2 frases e 1 texto).

Quanto aos resultados, os especialistas entrevistados afirmaram que (tendo como base o Paebes Alfa, que possui linha histórica): i) eles estão ajustados à população avaliada, os resultados estão melhorando e, com isso, os testes estão ficando mais difíceis; ii) eles ainda precisam ser melhor utilizados pelas redes; e iii) não é possível usá-los para dar um diagnóstico absoluto sobre se o aluno está ou não alfabetizado. Em relação ao primeiro ponto, cumpre notar que os testes de escrita não possuem pré-teste e é natural que os testes experimentem esse aumento de dificuldade. À medida que o teste de escrita torna-se parte da avaliação, a rede vai se preparando mais para sua realização, os alunos passam a acertar mais itens, os resultados melhoram e, para que continue produzindo informações relevantes, é necessário que a avaliação aumente o grau de dificuldade do teste.

No que diz respeito ao uso dos resultados, a escrita ainda possui menos visibilidade do que os resultados de leitura. A apropriação de resultados, baseada não só na disseminação dos mesmos, mas na capacitação dos profissionais da rede para seu uso, dá mais ênfase aos resultados de leitura do que aos de escrita. Esse foco da devolutiva nos resultados de leitura contribui para dicotomizar leitura e escrita, como processos estanques, impedindo uma análise, muito mais produtiva e adequada, conjunta dessas duas dimensões.

Isso nos leva ao terceiro ponto. A avaliação da escrita, no molde em que é aplicada e isoladamente, não permite um diagnóstico absoluto acerca da condição de alfabetizado do aluno. Para tanto, é necessário cruzar os dados de escrita com os de leitura, de modo a produzir uma visão mais completa acerca do processo de alfabetização. A referência teórica aqui é a Emília Ferreiro e Ana Teberosky, a *Psicogênese da língua escrita* (2007). Embora escrita e leitura sejam compostas por habilidades diferentes, a alfabetização é um processo complexo que envolve os dois conjuntos de habilidades, não podendo ser observado apenas por um deles.

Em relação aos padrões de desempenho, no Paebes Alfa, são 4 os padrões para escrita, distribuídos em uma escala de proficiência variando de 0 a 1000 pontos: Abaixo do Básico (abaixo de 600 pontos) – onde estão alocados os alunos que apenas começaram a desenvolver habilidades relativas à apropriação do sistema de escrita, tanto aquelas que dizem respeito a aspectos gráficos quanto aquelas que se referem a elementos sonoros (consciência fonológica); Básico (entre 600 e 650 pontos) – representando alunos que começam a escrever, por exemplo, frases ditadas curtas (no padrão sujeito-verbo-complemento) ou não ditadas, com base em uma imagem; Proficiente (entre 650 e 750 pontos) – alunos que escrevem frases com observância de algumas regras ortográficas e sem problemas de hipossegmentação (uma palavra na outra) e textos de gêneros mais familiares a eles; e Avançado (acima de 750 pontos) – no qual estão

alocados os alunos que são capazes, por exemplo, de produzir uma história a partir de uma cena ou de uma sequência de imagens, usando os principais elementos de uma narrativa. Esses padrões nunca foram alterados ao longo das avaliações de escrita do Paebes Alfa. Já para o Proalfa, em sua edição de 2017, os padrões são parecidos aos do Paebes Alfa, com uma pequena diferença: o padrão Proficiente varia de 650 a 800 pontos, fazendo com o que o padrão Avançado caracterize os alunos com média de proficiência acima de 800 pontos (e não 750, como no Paebes Alfa).

#### *5.3.4.5 A matemática na avaliação da alfabetização<sup>178</sup>*

No que tange à avaliação da matemática no ciclo de alfabetização, o cenário observado para os três programas é semelhante àquele que foi descrito para a avaliação da escrita. O Spaece Alfa nunca avaliou matemática, em nenhuma de suas edições. O Proalfa realizou sua primeira avaliação da matemática em 2017, aplicada aos alunos do 3º ano do ensino fundamental. Já o Paebes Alfa não apenas avalia matemática desde sua edição de 2010, como nunca deixou de avaliar nas edições seguintes e aplicou testes nessa área do conhecimento para alunos dos 1º, 2º e 3º anos do ensino fundamental.

De acordo com as especialistas entrevistadas, a negligência das secretarias de educação com a alfabetização matemática, que se manifesta, entre outras coisas, na ausência de avaliação desses conhecimentos em dois programas importantes de avaliação da alfabetização (Proalfa e Spaece Alfa), é correlata da negligência que existe em relação a esse tema na própria academia. A palavra alfabetização já remete ao vínculo com a língua materna, sem que a matemática seja imediatamente identificada com o termo, mesmo que se trate, também, de um tipo de linguagem. Somam-se a isso outras duas questões: i) o aumento dos custos da avaliação (aumento do número de questões no teste, aumento o número de impressões, por exemplo, além do aumento do número de itens para produção); e ii) o fato dos resultados em língua portuguesa serem tão preocupantes a ponto de obliterarem a preocupação com a alfabetização matemática. A capacidade de ler está relacionada com a capacidade de aprender em qualquer outra disciplina, entendimento este que leva a estabelecer a resolução dos problemas de alfabetização na língua como prioritários. Apesar dessa perspectiva, a pesquisa educacional ainda carece de

---

<sup>178</sup> Para coletar informações sobre a avaliação da matemática no âmbito dos três programas de avaliação da alfabetização, foi realizada uma entrevista coletiva (no dia 18 de junho de 2019) com cinco especialistas em alfabetização que integravam, à época, a equipe de alfabetização para a produção de itens e testes do CAEd/UFJF.

evidências sólidas acerca da relação entre os aprendizados nessas duas áreas do conhecimento. Na verdade, mesmo se aceitamos a relação proposta acima (capacidade de ler estando relacionada com a capacidade de aprender em outras disciplinas), é importante ressaltar que, conforme afirma uma das entrevistadas, especialista em matemática, o não leitor apresenta capacidade de aprender conceitos matemáticos mesmo ainda sem saber ler. Para fins de produção dos itens e testes no âmbito dos programas avaliados, os resultados de matemática são tratados como independentes dos resultados de língua portuguesa, quando os testes são guiados pelo aplicador.

Ainda de acordo com a opinião das especialistas do CAEd/UFJF, o Pnaic contribuiu para dar mais visibilidade à avaliação da alfabetização matemática, tendo em vista que seu material de formação, a princípio, dava o mesmo peso à alfabetização em língua materna e àquela em matemática. Contudo, se essa equivalência é percebida no material, o mesmo não pode ser dito a respeito do uso que foi feito dele pelas universidades, que deram ênfases distintas a seus conteúdos, em geral, dando mais atenção para a língua materna<sup>179</sup>.

Como o Paebes Alfa é o único dos três programas a possuir uma história com a avaliação da alfabetização matemática, é sobre ele que falaremos mais neste tópico. Até a edição de 2012, as matrizes de referência para matemática do Paebes Alfa eram produzidas pelo Ceale/UFMG através de uma parceria com a Seduc/ES. A partir da edição de 2013, as matrizes, itens e testes ficaram sob responsabilidade do CAEd/UFJF, o que se manteria nas edições seguintes do programa. A parceria com o CAEd levou a uma primeira reformulação da matriz de referência de matemática do Paebes Alfa, a pedido da própria secretaria de educação. Entre as mudanças, destaca-se a redução dos descritores destinados à resolução de problemas com o uso das operações básicas. Até então, cada operação (soma, subtração, multiplicação e divisão) possuía um descritor na matriz. Esses 4 descritores foram reunidos em dois (adição e subtração; multiplicação e divisão), abrindo espaço na matriz para outras habilidades. Com um foco menos exacerbado no sistema de numeração, outras áreas ganharam mais destaque, como grandezas e medidas (ler as horas, por exemplo, foi uma habilidade que ganhou espaço).

As cinco especialistas entrevistadas acreditam que uma das limitações da matriz de matemática do Paebes Alfa até 2017 estava no fato de que havia uma mesma matriz para 1º, 2º e 3º anos. As mesmas habilidades eram previstas, como pode ser observado através da análise

---

<sup>179</sup> Em minha já referida experiência com o Pnaic, os resultados da avaliação, quando abordados nas formações, davam mais ênfase à língua portuguesa, inclusive como demanda dos professores alfabetizadores, que, em sua maioria oriundos dos cursos de Pedagogia ou correlatos (como o Normal Superior), não tinha afinidade com a matemática.

das matrizes em anexo (I), e o que mudava era a dificuldade dos itens para cada série/ano. Com o tempo, o que pôde ser observado quanto aos resultados é o teste estava ficando mais fácil para o 3º ano. Para contornar isso, em 2017, foram construídas matrizes específicas para cada ano avaliado, mas apenas a partir da edição de 2018 é que essa mudança foi efetivada no programa. A edição de 2017 do Proalfa, a primeira a avaliar matemática, apresenta uma matriz para o 3º ano do ensino fundamental já mais ajustada ao público dessa série/ano. Ela possuía algumas habilidades mais complexas do que a matriz do Paebes Alfa para a mesma edição, principalmente, em números e operações (exigindo uma apropriação maior, por parte do aluno, do sistema de operações e de cálculo).

O Paebes Alfa possui um diferencial em relação ao Proalfa, no que tange à avaliação da alfabetização matemática. O programa contou, entre 2010 e 2015, com a oferta, para os professores das escolas avaliadas, de oficinas de elaboração de itens (realizadas a cada dois anos). A ideia foi montar um banco de itens feitos pelos professores, que passaram a integrar, depois de validados pela equipe de especialistas em elaboração de itens do CAEd, os testes aplicados pelo programa.

Ao longo de todo o período analisado, as equipes das secretarias (tanto as do Espírito Santo quanto as de Minas Gerais) não validavam os itens produzidos pelo CAEd, que possuía autonomia, dada pelas próprias secretarias, para a elaboração dos itens e testes (as matrizes sempre são validadas pelas secretarias). Isso acontecia, entre outros fatores, muito em virtude da ausência de especialistas em alfabetização matemática na área de avaliação das duas secretarias. As estruturas dos itens e dos testes de alfabetização matemática não foram alterada no período. Os itens possuíam 4 alternativas de resposta. No Paebes Alfa, os testes contavam com 96 itens, divididos em 16 blocos, contando com 6 itens cada. Cada teste era composto por 4 blocos, num total de 24 itens (respondidos por cada aluno). No Proalfa, em sua edição de 2017, os testes contaram com 80 itens, divididos em 8 blocos, cada um com 10 itens. Os alunos respondiam a 20 itens cada, divididos em 2 blocos. Nos dois casos, os itens possuíam 5 níveis de dificuldade, de acordo com a série/ano a que pertenciam. A aplicação do teste de matemática ocorria sempre um dia após a aplicação do teste de língua portuguesa.

Assim como ocorreu no caso da escrita, a história da avaliação da matemática no ciclo de alfabetização, em particular, no caso do Paebes Alfa, confunde-se, em parte, com a história do próprio CAEd ao longo do período. Até 2013, os itens de alfabetização matemática eram elaborados como parte das atividades da equipe de matemática do CAEd, responsável pelos itens e testes de todas as séries/anos (da alfabetização ao ensino médio). Segundo as

entrevistadas, elaborar itens para o ciclo de alfabetização no âmbito da equipe de matemática foi um grande desafio ao longo das primeiras edições do Paebes Alfa, tendo em vista a formação de seus integrantes, não especializados em alfabetização matemática. A partir de 2013, um passo foi dado, com a integração de uma pedagoga na equipe de elaboração de itens de matemática, que passou a ser responsável pelos itens para o ciclo de alfabetização. Já para a edição de 2017, nova mudança: as demandas de testes para matemática na alfabetização foram integradas à equipe de alfabetização, que passou a ser responsável pela elaboração de itens nas duas disciplinas, língua portuguesa-leitura e matemática.

Ao longo da entrevista, as especialistas afirmaram que um dos desafios da avaliação da alfabetização matemática está no uso dos resultados. Conforme já afirmado anteriormente, eles recebem menos atenção do que os resultados em língua portuguesa, particularmente, os que dizem respeito à leitura. No âmbito da divulgação dos resultados, contudo, o Paebes Alfa cuidou, ao longo de todas as suas edições, de contar com especialistas em alfabetização matemática em suas capacitações. Quanto aos resultados, o Paebes Alfa apresentou melhoras ao longo do tempo, mas é preciso analisá-las, em especial, no 3º ano do ensino fundamental, à luz do que já afirmamos anteriormente: contanto com uma mesma matriz para os três anos da alfabetização, os testes foram ficando mais fáceis para os alunos do 3º ano ao longo do tempo.

A avaliação da matemática da alfabetização ressenete-se também de outra lacuna: a ausência de uma definição oficial do que significa ser alfabetizado em matemática. Essa ausência afeta as secretarias de educação e também instituições de pesquisa como o CAEd/UFJF, por exemplo, que não encontram orientações oficiais sobre o conceito de alfabetização matemática. A academia tampouco tem se apresentado como capaz de auxiliar nessa questão, visto que é permeada por diferentes ideias e perspectivas acerca desse conceito. Diante disso, os testes não são os instrumentos adequados e capazes de dizer, de modo oficial, o que é ser alfabetizado em matemática. Eles informam sobre as tarefas que os alunos são capazes de realizar na área. Contudo, a descrição do padrão adequado na avaliação tem sido tomado como uma referência para dizer o que significa a alfabetização matemática, diante da ausência de uma definição oficial. Estar no padrão adequado em matemática no 3º ano do ensino fundamental implica, de maneira geral, o conhecimento das quatro operações básicas (soma, subtração, multiplicação e divisão) e o domínio de duas delas (soma e subtração), além do conhecimento acerca do sistema de numeração e apresentar noções espaciais.

Em relação aos padrões de desempenho, o Paebes Alfa sempre teve os mesmos padrões em todas as edições entre 2010 e 2017, com 4 padrões, distribuídos em uma escala variando de



0 a 1000 pontos. Para o 1º, o 2º e o 3º ano, os padrões têm o mesmo nome, mas com cortes diferentes: Abaixo do Básico (1º ano – até 350 pontos); (2º ano – até 400 pontos); (3º ano - até 450 pontos); Básico (1º ano – de 350 a 450 pontos); (2º ano\_ de 400 a 500 pontos); (3º ano - de 450 a 550 pontos); Proficiente (1º ano - de 450 a 550 pontos); (2º ano – de 500 a 600 pontos); (3º ano - de 550 a 650 pontos); e Avançado (1º ano – a partir de 550 pontos); (2º ano - partir de 600 pontos); (3º ano - a partir de 650 pontos). Conforme já informado, em 2017, o Proalfa realizou sua primeira avaliação da matemática, apenas para o 3º ano. Os padrões são os seguintes: Baixo (até 500 pontos); Intermediário (de 500 a 550 pontos); Recomendado (de 550 a 675 pontos); e Avançado (a partir de 675 pontos). É importante anotar que, embora as duas escalas variem de 0 a 1000 pontos e possuam quatro padrões, elas não são iguais, de modo que esses resultados não são diretamente comparáveis.

### **5.3.5 Divulgação e uso dos resultados (natureza dos canais de divulgação e comunicação)**

A divulgação dos resultados responde pela etapa mais visível da avaliação. Os resultados são a entrega de um programa de avaliação. Todas as etapas pelas quais a avaliação é composta desaguam em seus resultados. Eles são a razão de ser da avaliação, que, conforme já afirmamos, tem como função primordial fornecer informações qualificadas para o processo de tomada de decisões. Sua importância para o processo avaliativo faz com que eles sejam analisados como critério autônomo.

Os resultados dizem respeito ao público-alvo da avaliação. Seu público, conforme vimos, é composto pelos destinatários de seus resultados. Se o público é formado por atores diferentes, que assumem posições distintas no processo educacional, a forma como os resultados são divulgados e o material produzido para dar base a seu entendimento, interpretação e uso devem atender a essas especificidades. Para o critério de divulgação dos resultados, serão analisados os canais utilizados para a divulgação e os materiais destinados aos diferentes atores.

Uma observação é importante. O público que receberá os resultados da avaliação é uma decisão das secretarias de educação. Uma vez tomada essa decisão, contudo, aumenta a incidência de critérios técnicos para definir quais são os melhores canais de divulgação, qual linguagem deve ser utilizada e que materiais surtem mais efeitos para cada um dos atores que compõem o público-alvo. Para o período analisado, nos três programas, o CAEd/UFJF foi uma

instituição parceira, que ficou responsável pela elaboração dos materiais de divulgação de resultados, de modo que os três programas apresentam muitas semelhanças quanto a seu processo de divulgação de resultados.

Uma primeira característica comum aos três programas diz respeito ao processo de disponibilização do material. Ele era montado pela equipe de divulgação de resultados do CAEd, mas, de acordo com a coordenadora da equipe<sup>180</sup>, todos os materiais de divulgação são validados pelas equipes de avaliação das secretarias. Nada é publicado sem o aval delas.

Os três programas possuem um portal de divulgação de resultados, contendo informações e materiais sobre cada um deles. Os portais não são específicos dos programas de alfabetização, antes, referem-se aos sistemas de avaliação dos estados: no Espírito Santo, o portal<sup>181</sup> fornece informações sobre Paebes e Paebes Alfa; no Ceará, o portal<sup>182</sup> apresenta informações sobre Spaece e Spaece Alfa; em Minas Gerais, o portal<sup>183</sup> é do Simave, com informações para Proeb e Proalfa. As informações e materiais disponíveis em cada um deles variam significativamente.

No portal do Paebes, relativamente ao Paebes Alfa, estão disponíveis para download, a coleção de revistas de divulgação de resultados de 2013 a 2017, embora o Paebes Alfa tenha tido sua primeira edição em 2008. Há materiais de apoio disponíveis desde o ano de 2009 (exceto para o ano de 2010), como materiais de suporte das oficinas, vídeos de apropriação de resultados e até mesmo itens dos testes. Os resultados por escola estão disponíveis para qualquer usuário, a partir da edição de 2009 (exceção feita aos resultados de 2010). Os resultados das escolas particulares participantes só podem ser acessados por meio de login e senha, enviados especificamente para cada uma das escolas, de modo que tenham acesso apenas a seus próprios resultados. O site fornece acesso aos resultados de TCT, mas apenas para as escolas, com login e senha específicos para cada uma delas.

Para todas as edições disponíveis no portal (de 2013 a 2017), as coleções do Paebes Alfa apresentam uma mesma estrutura: i) Revista do Sistema, destinada aos secretários de educação e equipes da secretaria, com resultados agregados por superintendência/regional de ensino, com enfoque na apropriação dos resultados para a tomada de decisão no nível da rede; o programa oferece uma revista para cada tipo de rede (estadual e municipais); ii) Revista do Gestor, destinada ao diretor, com o resultado de sua própria escola e com foco na apropriação dos resultados nesse nível; iii) Revista do Professor, destinada aos professores alfabetizadores, com

---

<sup>180</sup> Entrevista realizada em 6 de novembro de 2019.

<sup>181</sup> <http://www.paebes.caedufjf.net/>

<sup>182</sup> <http://www.spaece.caedufjf.net/>

<sup>183</sup> <http://simave.educacao.mg.gov.br/>

vistas ao uso pedagógico dos resultados (o Paebes Alfa oferecia uma revista específica, com os resultados de língua portuguesa e matemática, para cada série avaliada – 1º, 2º e 3º anos). Para a edição de 2017, está disponível também um sumário executivo, reunindo todos os dados da avaliação em uma mesma publicação. Durante todo o período analisado, o programa realizou, também, anualmente, oficinas presenciais de divulgação e apropriação de resultados, com um grupo de gestores das redes de ensino selecionados pela secretaria de educação.

No portal do Simave, estão disponíveis os resultados do Proalfa a partir da edição de 2013 (embora a primeira edição do programa tenha ocorrido em 2005). Os resultados por escola estão disponíveis para qualquer usuário. Os resultados de TCT (percentual de acerto por descritor) também são disponibilizados, mas, assim como ocorre para o Paebes Alfa, só podem ser acessados por meio de login e senha, enviados a cada uma das escolas.

As coleções de revistas de divulgação dos resultados do programa estão disponíveis para as edições de 2008 a 2014. Em 2009 e 2010, o programa contava com duas revistas de divulgação: i) uma Boletim de Resultados, com os resultados gerais do programa, tendo como foco a secretaria de educação e as superintendências de ensino; e ii) um Boletim do Professor, com foco no professor e apresentando os dados da avaliação censitária do 3º ano do ensino fundamental. Para a edição de 2011, o formato mudou: i) uma Revista do Gestor, com foco nos resultados gerais do programa e das superintendências de ensino, publicada em dois diferentes volumes, um para a rede estadual e outro para as redes municipais; e ii) uma Revista Pedagógica, como foco nos resultados do 3º ano do ensino fundamental para os professores. É interessante notar como, até esse momento, os gestores escolares ainda não haviam recebido um material próprio para a apropriação de resultados em suas escolas. Os resultados por escola foram divulgados apenas no site e a Revista do Gestor servia indistintamente aos gestores da rede e gestores de escola. Em 2012 e 2013, o programa contou com 4 revistas: i) Revista do Sistema, voltada para a secretaria de educação e superintendências, com dados para as duas redes (estadual e municipais) na mesma publicação; ii) Revista da Gestão, voltada para os gestores escolares, com dados de sua escola; iii) Revista Pedagógica, voltada para os professores; e iv) Boletim Contextual, com os indicadores produzidos a partir das questões que compuseram os questionários contextuais aplicados juntamente com os testes. Para 2014, esse mesmo formato foi mantido, mas sem a disponibilização da Revista Contextual. Até 2014<sup>184</sup>,

---

<sup>184</sup> A partir de 2014, houve uma mudança considerável na avaliação educacional em Minas, especialmente, em relação à importância para a gestão da rede que ele havia assumido até aquele ano. Tal mudança, para ser melhor compreendida, exige um esforço próprio de pesquisa, de modo a observar os motivos que levaram a essa mudança e no que, especificamente, ela consiste.

foram realizadas, anualmente, oficinas de apropriação de resultados no estado, no mesmo formato daquelas realizadas no âmbito do Paebs Alfa.

O portal do Spaece permite acesso livre aos resultados por escola desde a edição de 2018. Não há link disponível no portal para acesso aos resultados de TCT, embora eles cheguem às mãos da secretaria de educação por meio de planilhas específicas. Há, para 2016 e 2017, vídeos de divulgação de resultados disponíveis para download. As coleções de revistas de divulgação de resultados estão disponíveis desde a primeira edição do programa, em 2008.

No que tange às revistas, o Ceará apresenta uma diferença em relação a Minas e Espírito Santo. O Spaece possui uma única coleção para todo o sistema, sem que exista uma coleção própria para o Spaece Alfa. Dentre as revistas publicadas pelo Spaece, uma delas sempre foi referente ao Spaece Alfa. Trata-se do Boletim do Professor – Alfabetização (nomenclatura adotada em 2016). Em 2009 e 2010, ela era chamada de Boletim de Resultados da Escola – 2º ano do ensino fundamental. Em 2011, passou-se a chamar Boletim Pedagógico – 2º ano do ensino fundamental, nome adotado até a edição de 2015. Apesar da nomenclatura, ele era voltado, em todas as edições, para o professor.

Em 2012, a coleção do Spaece publicou, além da revista destinada ao professor, mais duas revistas referentes ao Spaece Alfa, uma Revista do Sistema, voltada para as secretarias e gerências regionais de ensino, e uma Revista da Gestão, voltada para os gestores escolares, ambas com resultados específicos para a alfabetização. Esse formato foi também adotado para as edições de 2013 e 2014 do programa. De 2015 a 2017, a coleção de revistas do Spaece publicou, especificamente para o Spaece Alfa, apenas o Boletim Pedagógico para o 2º ano do ensino fundamental. Em relação às oficinas de apropriação de resultados, o Spaece as realiza anualmente, com público designado pela secretaria de educação.

Para os três programas, as oficinas de apropriação de resultados, como sugere o próprio nome, eram dedicadas a apresentar os resultados da edição atual do programa e formas de como eles poderiam ser usados como base para o planejamento. As oficinas eram ministradas por especialistas do CAEd e o público era designado pelas secretarias (de modo geral, gestores e professores das escolas da rede). No caso do Paebs Alfa, por conta da avaliação das disciplinas avaliadas, as oficinas contavam com especialistas não apenas em língua portuguesa, mas também em matemática.

O material de divulgação dos três programas apresentou, ao longo das edições, os mesmos destinatários principais: os gestores das redes (secretários, equipe da secretaria e superintendentes); os gestores escolares (em Minas, receberam material específico a partir da

edição de 2012); e os professores. As revistas específicas para cada uma desses públicos contavam com linguagem e informações ajustadas aos seus interesses. Nas revistas para os professores, por exemplo, sempre havia um conjunto de itens do teste analisados pedagogicamente, ao passo que a revista do sistema (para os gestores da rede) apresentava os dados agregados por superintendência/regional de ensino. Embora os resultados por escola sejam abertos a qualquer usuário do portal, não há um material específico que explique a avaliação educacional e a forma de ler os resultados com os mesmos detalhes e com o mesmo cuidado daquele oferecido aos destinatários principais. Da mesma forma, não há um material destinado aos alunos e a seus responsáveis. Por experiência de trabalho nessa área do CAEd, sei que houve, em algumas edições dos três programas, um boletim destinado aos pais dos alunos, mas, por conta da proteção à identidade dos que fazem os testes, esse material era impresso e entregue a cada uma das escolas, sem divulgação por meio do portal. De todo modo, essa ação foi descontinuada e a avaliação ainda apresenta o desafio de se comunicar melhor com os alunos e seus responsáveis.

### **5.3.6 Concepção de alfabetização**

Conforme apontado anteriormente, um dos elementos mais importantes de uma política de avaliação são as matrizes de referência. Baseadas no currículo de cada estado, elas representam uma seleção do que será avaliado pelos testes. Trata-se de um documento que envolve decisões políticas e técnicas. A escolha sobre o que será avaliado é sempre uma decisão da secretaria de educação, de acordo com as suas necessidades e prioridades. Contudo, a avaliação em larga escala, pela metodologia adotada, oferece limitações a essa escolha, de modo que nem tudo o que está previsto no currículo é passível de ser avaliada por um teste padronizado de múltipla escolha.

No âmbito da avaliação da alfabetização, uma das questões mais importantes para a análise dos programas diz respeito à concepção de alfabetização encarnada pelos testes. Conforme já discutimos anteriormente, no capítulo 2, o campo da alfabetização no Brasil é caracterizado por uma tensão entre, de um lado, os defensores de um processo de alfabetização centrado no texto e baseado no construtivismo, e, de outro, aqueles que defendem que a atenção aos elementos fonêmicos é fundamental para o desenvolvimento das habilidades necessárias para o aluno ser considerado alfabetizado. A superação dessa tensão pode ser encontrada na proposta de Magda Soares (2004), de que alfabetização (o ensino e a aprendizagem do sistema

alfabético) e letramento (o uso da leitura e da escrita nas práticas sociais), são processos distintos, porém complementares. A crítica dos construtivistas, como Gontijo (2014), é de que a ideia de letramento coloca em cena um falso consenso, afinal, ela própria é contra a ideia de que o aspecto fonético é central para o processo de alfabetização.

A perspectiva de Soares ganhou adeptos e ficou conhecida como alfabetizar letrando: as habilidades destinadas à apropriação do sistema alfabético são fundamentais para o processo de alfabetização e não se constituem como um obstáculo à formação de leitores, mas, antes, contribuem, decisivamente para tanto.

Nos programas de avaliação da alfabetização, interessa saber qual é a concepção de alfabetização que dá base à construção das matrizes de referência e se foram alteradas ao longo do tempo. A análise das matrizes de Paebes Alfa, Proalfa e Spaece Alfa mostram que, em todas as edições dos programas para o período analisado, a mesma concepção de alfabetização esteve presente: a do alfabetizar letrando. Conforme vimos em tópico anterior, o Ceale/UFMG esteve presente nas primeiras edições de Proalfa e Paebes Alfa, inclusive, elaborando a matriz de referência em parceria com as secretarias de educação. Magda Soares é uma das fundadoras do Ceale, o que certamente influenciou a concepção de alfabetização que deu base às matrizes. Quando, posteriormente, o CAEd assumiu a elaboração da matrizes para os dois programas (em Minas e no Espírito Santo), a perspectiva do alfabetizar letrando foi mantida.

Em todas as matrizes dos três programas analisados, entre 2009 e 2017, havia descritores relacionados à apropriação do sistema alfabético e ao reconhecimento de suas convenções (envolvendo competências como a identificação de letras do alfabeto, uso adequado da página, reconhecimento da palavra como unidade gráfica, aquisição da consciência fonológica). Identificar as direções da escrita, o número de sílabas de uma palavra, identificar sílabas canônicas e não-canônicas, diferenciar as letras de outros símbolos gráficos, reconhecer as letras do alfabeto são habilidades que sempre estiveram presentes nas matrizes de referência de Paebes Alfa, Proalfa e Spaece Alfa.

O material de divulgação do Paebes Alfa faz referências explícitas e diretas ao processo de alfabetização e letramento. Nas revistas de divulgação de resultados de 2014, por exemplo, o texto fala que “é necessário que esse processo de alfabetização seja realizado junto com o processo de letramento.” (ESPÍRITO SANTO, 2014, p.64). Um dos tópicos, da mesma revista, é intitulado Alfabetizando e letrando através dos gêneros textuais. No material de divulgação do ano seguinte, 2015, Magda Soares aparece como referência explícita em um dos tópicos da revista, intitulado *Gêneros discursivos: considerações sobre os usos sociais da leitura e da*

*escrita* (ESPÍRITO SANTO, 2015, p.65). A Revista do Professor para o Paebes Alfa 2017 apresenta como um novo indicador daquela edição perfis de alfabetização e letramento (ESPÍRITO SANTO, 2017).

Referências explícitas ao processo de alfabetização e letramento também são encontradas nos materiais de divulgação de resultados do Ceará e de Minas Gerais. Na edição de 2014, o Boletim Pedagógico do Spaece Alfa defende a necessidade do processo de alfabetização ocorrer junto com o professor de letramento (CEARÁ, 2014). Em 2015, o Boletim Pedagógico do Spaece Alfa faz referência explícita à perspectiva da alfabetização e letramento proposta por Magda Soares (CERÁ, 2015, p.47). Na edição de 2017, fala-se na importância de explorar a relação entre imagem e texto escrito para construir o significado como um recurso importante para o processo de desenvolvimento da alfabetização e do letramento. (CEARÁ, 2017, p.30).

Em Minas, na Revista Pedagógica da edição de 2012, Magda Soares aparece como referência explícita para os processos avaliados pelo Proalfa, a partir da defesa de que o ingresso da criança no mundo da escrita “deve ocorrer simultaneamente pela aquisição do sistema convencional da escrita – a alfabetização – e pelo desenvolvimento de habilidades de uso desse sistema em atividades de leitura e escrita, nas práticas sociais que envolvem a língua escrita – o letramento” (MINAS GERAIS, 2012, p.11). Na edição de 2013, a Revista Pedagógica afirma que o programa “possibilita um diagnóstico completo e consistente dos processos de alfabetização e letramento no estado” (MINAS GERAIS, 2013, p.8). Na mesma revista, os conceitos de avaliação e letramento são apontados como os “marcos teóricos que fundamentam o Programa de Avaliação da Alfabetização” (idem, p.9).

A adoção de uma mesma concepção de alfabetização nos ajuda a compreender as semelhanças entre as matrizes de referência, observadas para os três programas analisados.

### **5.3.7 Vínculo com programas relevantes e geração da avaliação (*accountability*)**

Conforme já ressaltamos anteriormente, um programa de avaliação tende a encarnar a função de uma política que avalia outras políticas. Por essa perspectiva, o vínculo entre um programa de avaliação e outros programas educacionais (mas não somente e necessariamente) é quase pressuposto. A despeito disso, o vínculo com outros programas pode ser investigado a partir de diferentes enfoques. Um deles se relaciona com o tipo de política com a qual a política de avaliação estabelece vínculo. Trata-se de outra política de avaliação? Se sim, como essa

articulação ocorre? Como a política de avaliação da alfabetização é influenciada ou influencia outras políticas de avaliação? Os resultados dessas diferentes políticas de avaliação podem estar associadas e sua divulgação pode ser conjunta ou complementar, o que seria um exemplo de articulação.

A política de avaliação pode estar associada a políticas educacionais de outra natureza. Em relação à alfabetização, como a política de avaliação se articula com a formação inicial de professores alfabetizadores no estado? E com a formação continuada desses profissionais? A formação de gestores também sofre influência da política de avaliação? Essas e outras questões podem mover o interesse da pesquisa quanto a este critério.

Podem ser investigadas, ainda, as relações entre os programas de avaliação da alfabetização e políticas de responsabilização, um dos temas mais polêmicos e espinhosos no âmbito da avaliação educacional. De acordo com o que já foi discutido no capítulo 2 da tese, as políticas de responsabilização atreladas aos resultados da avaliação atraem boa parte da crítica direcionada a ela. Aqui, interessa saber se os resultados dos três programas de avaliação da alfabetização analisados estão vinculados a alguma política de responsabilização do estado. Em caso afirmativo, quais são as características dessas políticas? Suas consequências são brandas ou severas? Assumem um caráter positivo, de bonificação, ou negativo, de punição? Conforme vimos também no capítulo 2, Bonamino e Souza (2012) reconhecem três gerações de avaliação educacional no Brasil, de acordo com o tipo de consequências gerado por seus resultados. Com base na tipologia das autoras, esperamos classificar Paebs Alfa, Proalfa e Spaece Alfa de acordo com a geração de avaliação à qual pertencem.

É importante notar que um programa de avaliação, se é importante para a política educacional, estabelecerá vínculos, diretamente ou indiretamente, com outras programas educacionais. Seus efeitos podem impactar toda a política educacional do estado, caso os resultados da avaliação sejam utilizados, por exemplo, para tomar decisões em relação a investimentos prioritários. O programa que sofre cortes orçamentários ou mesmo é cancelado por conta de uma decisão dessa natureza é influenciado pelo programa de avaliação, mesmo sem estabelecer vínculos diretos com ele. Não é esse tipo de vínculo que nos interessa aqui. O objetivo é analisar a relação do programa de avaliação da alfabetização com programas representativos no estado, relevantes para a política de educação como um todo.

Cumprе ressaltar, de saída, que os três programas de avaliação da alfabetização fazem parte de um sistema de avaliação constituído por outros programas. Quando o Paebs Alfa surgiu, em 2008, o Paebs também avaliou alunos do 1º ano do ensino médio. Em 2009, o 5º e



o 9º anos do ensino fundamental foram acrescentados à avaliação do 1º ano do médio e, a partir da edição de 2010, o Paebes passou a avaliar também o 3º ano do ensino médio, variando, ano a ano, a avaliação de outras séries do médio, como o 2º ano. O Paebes avalia, desde sua primeira edição, conteúdos de língua portuguesa e matemática, mas passou a avaliar, a partir de 2011, Ciências da Natureza e Ciências Humanas, para as séries avaliadas no ensino médio, em anos alternados. Em 2016, um novo programa de avaliação foi integrado ao sistema do Paebes, a Avaliação Diagnóstica da Aprendizagem, o Paebes Trimestral, ou simplesmente Paebes TRI. O objetivo dessa avaliação, aplicada somente para o ensino médio, é dar suporte pedagógico ao professor em sala de aula, assumindo um caráter de avaliação formativa. Para tanto, o Paebes TRI não faz uso da Teoria da Resposta ao Item (como seu nome pode equivocadamente sugerir), divulgando os resultados pela TCT, disponibilizando o percentual de acerto por descritor para os professores, para que possam identificar as principais dificuldades de seus alunos.

O Proalfa foi criado em 2005, mas o Simave – Sistema Mineiro de Avaliação da Educação Pública já existia desde 2000. O Proeb – Programa de Avaliação da Rede Pública de Educação Básica integra o Simave desde sua primeira edição, avaliando estudantes de 5º e 9º anos do ensino fundamental e do 3º ano do ensino médio, das escolas estaduais e municipais de Minas (o programa teve edições de 2000 a 2003, foi descontinuado e voltou em 2006; até 2014, avaliou as séries/anos mencionadas anteriormente; a partir de 2015, passou a avaliar de modo alternado, o 5º e o 9º anos, em anos pares, e o 7º ano do ensino fundamental e o 1º ano do ensino médio nos anos ímpares, avaliando sempre o 3º ano do ensino médio). O Simave possuía ainda outro programa de avaliação, o PAAE – Programa de Avaliação da Aprendizagem Escolar, que disponibilizava um banco de itens para os professores aplicarem a seus alunos de modo a subsidiar o planejamento e propiciar intervenções pedagógicas. O PAAE tinha como foco o ensino médio e os anos finais do fundamental, e destinava-se somente aos professores da rede estadual, responsáveis por aplicar os testes. O programa foi implantado gradativamente entre 2005 e 2009, sendo institucionalizado em 2010.

No Ceará, quando o Spaace Alfa foi criado em 2007, o Spaace já possuía uma história desde 1992, sendo um dos sistemas próprios de avaliação mais antigos do país. Até 1998, aplicou testes para 4ª e 8ª séries do ensino fundamental. Retornou em 2001, com testes para a 8ª série e a 3ª série do ensino médio, até a edição de 2013, depois da qual o 4º ano passou a ser novamente avaliado.

Essa história pregressa da avaliação em Minas e Ceará pode ter contribuído para reduzir a resistência das redes e seus atores em relação à criação de um programa de avaliação da

alfabetização. Além disso, possivelmente, preparou os tomadores de decisão e implementadores do programa para que lidassem com os obstáculos já enfrentados quando da criação do Spaece e do Simave.

O importante a destacar é o fato de que os programas de avaliação da alfabetização dos três estados fazem parte de um sistema de avaliação mais abrangente, que cobre outras etapas e disciplinas. Isso significa que a expertise alcançada em um programa pode ser compartilhada com os demais. As equipes específicas das secretarias de educação não são responsáveis apenas pelo programa de avaliação da alfabetização, mas sim pela avaliação como um todo, o que contribui para esse processo. Na entrevista para a tese com Alessio da Costa Lima, por exemplo, ele afirma que mudanças na divulgação dos resultados do Spaece Alfa levaram a mudanças também no Spaece, mostrando a interação entre os programas.

Além de pertencerem a sistemas mais amplos, os programas de avaliação analisados possuem vínculos com políticas de responsabilização. No Ceará, o Prêmio Escola Nota Dez relaciona-se diretamente com o Spaece Alfa (apresentando vínculo também com o Spaece, em função dos resultados do 5º e do 9º)<sup>185</sup>. Ele foi criado pela lei 14.371, de 2009 (disciplinado pela lei 14.580/2009, com regulamentação pelo decreto 29.896/2009), revogada em 2011 pela lei 15.052 (regulamentada pelo decreto 30.797, do mesmo ano). Em 2015, nova legislação sobre o prêmio foi promulgada, a lei 15.923 (regulamentada pelo decreto 32.079/2016), permanecendo ainda hoje em vigor.

Em um contexto de gestão por resultados, o objetivo principal do prêmio é valorizar a gestão educacional de caráter pedagógico, com foco na aprendizagem do aluno. De acordo com a descrição no portal da secretaria de educação do estado, o prêmio assume um caráter indutor para as escolas melhorarem seus resultados. Para o ciclo de alfabetização, ele é concedido às 150 escolas com melhores resultados no IDE-Alfa<sup>186</sup> (Índice de Desempenho Escolar para o 2º ano do ensino fundamental), no valor de R\$ 2.000,00 por aluno avaliado. Os critérios estabelecidos pela lei para o recebimento do prêmio são: i) ter, no momento da aplicação do Spaece Alfa, pelo menos, 20 alunos matriculados no 2º ano do ensino fundamental regular; ii) ter participação de 90% ou mais no Spaece Alfa; e iii) obter média de, pelo menos, 8,5 no IDE-Alfa (numa escala que váia até 10).

---

<sup>185</sup> Outro prêmio que o estado possui é o *Aprender pra Valer*, destinado ao ensino médio.

<sup>186</sup> O IDE é formado por três elementos: i) a proficiência da escola, convertida para uma escala de 0 a 10; ii) a taxa de participação na avaliação; e iii) o fator de ajuste para universalização do aprendizado, que recebe um valor percentual de 0 a 100%, a partir da distribuição dos alunos pelos padrões de desempenho estabelecidos pelo Spaece Alfa.

Para 5º e 9º anos, as escolas 150 escolas com menores resultados no IDE-5 e IDE-9 recebem uma contribuição financeira como estímulo para a melhoria, no valor de R\$ 1.000,00 por aluno avaliado. Além disso, cada uma delas é adotada por uma escola premiada, que deve fornecer auxílio técnico e pedagógico, ajudando a implementar um plano de melhoria dos resultados. A legislação prevê que as escolas premiadas de 5º e 9º anos receberão 75% do valor do prêmio, sendo os outros 25% condicionados à observância de dois critérios: pelo menos, manter seus resultados no Spaece e apoiar a escola adotada (parceira). Já as escolas apoiadas recebem 50% do valor, ficando o restante condicionado à elevação de seu IDE para no mínimo 5. É interessante notar que essa política de parceria entre as escolas não é adotada para o ciclo de alfabetização.

O Prêmio Escola Nota 10 é uma política de *accountability* que transfere recursos para as escolas premiadas (no caso, do 5º e 9º anos, também para as escolas com menores resultados). Vinculado aos resultados da avaliação, ele faz com que o Spaece Alfa tenha consequências mais fortes (*high-stakes*) para as escolas, simbólicas e materiais. No critério adotado por Bonamino e Sousa (2012) para definir as três gerações de políticas de avaliação no Brasil, podemos afirmar que o Spaece Alfa, desde a instituição do prêmio, em 2009, é uma política de avaliação de terceira geração. Araújo et al (2019) reconhecem a associação entre a instituição do prêmio e a melhoria dos resultados do Ceará no ciclo de alfabetização. Apesar disso, veem consequências nefastas dessa política de responsabilização, em especial, por associar a qualidade da educação somente aos resultados de desempenho nos testes padronizados, levando a uma busca a qualquer custo pela melhoria desses resultados. Já para Calderón et al (2015), o prêmio assume um caráter indutor importante no que diz respeito à cooperação entre as escolas.

O Spaece Alfa está associado, também, ao programa mais significativo da política educacional do Ceará, o Programa de Alfabetização na Idade Certa (Paic). Criado em 2007, o Paic atravessou três ciclos de gestão do governo estadual, apresentando uma longevidade admirável para uma política educacional no contexto brasileiro. Na leitura de Maurício Holanda, ex-secretário de educação do Ceará, um dos entrevistados para a tese, a longevidade do Paic impressiona tanto quanto os resultados que produziu para o processo de alfabetização dos alunos. O Spaece Alfa foi criado no mesmo ano e assumiu, de imediato, a função de “termômetro” do Paic (conforme expressão usado por Alessio Lima, outro dos entrevistados para a tese). As relações entre o Spaece Alfa e o Paic são, portanto, umbilicais: eles surgem juntos, no mesmo cenário e para lidar com o mesmo problema. O contexto é aquele já narrado anteriormente: o compromisso assumido por diversas entidades, entes, órgãos e instituições

para eliminar o analfabetismo no Ceará (Paic e Spaece Alfa foram gerados sob a influência do Relatório do Comitê Cearense para a Eliminação do Analfabetismo Escolar). Os resultados que mantêm o Paic como uma política contínua no Ceará são atestados pelo Spaece Alfa, o que explica sua importância e longevidade.

O objetivo aqui não é fazer uma análise do Paic<sup>187</sup>, apenas apresentar seu vínculo com o programa de avaliação. Como esse vínculo é muito estreito e forte, a importância concedida ao Paic acaba por reverberar na importância atribuída ao Spaece Alfa. O Paic pode ser definido como um programa de intervenção pedagógica, cujo objetivo é garantir que toda criança esteja alfabetizada ao fim do 2º ano do ensino fundamental. Ele surge em torno de 5 eixos interdependentes: i) alfabetização; ii) educação infantil; iii) literatura infantil; iv) gestão municipal; e v) avaliação externa. A avaliação ocupa um papel tão importante na concepção do programa que recebeu um eixo próprio, assim como a cooperação com os municípios. A importância do Paic no estado do Ceará está ligada à melhoria progressiva nos resultados dos alunos na alfabetização, atestada pelo Spaece Alfa. Isso fez com que o programa extrapolasse os limites estaduais, servindo como base para o Pnaic, que, a exemplo do Paic, também contava com um eixo próprio para a avaliação. A ideia era que a ANA pudesse desempenhar para o Pnaic um papel semelhante àquele desempenhado pelo Spaece Alfa para o Paic. No entanto, ANA e Pnaic são políticas descontinuadas.

Em Minas Gerais, os resultados do Proalfa foram o ponto de partida para a criação do Programa de Intervenção Pedagógica – Alfabetização na Idade Certa, o PIP, um dos programas educacionais de maior duração e mais conhecidos no estado ao longo do período analisado. Os resultados da edição de 2006 do Proalfa evidenciavam um quadro em que mais da metade dos alunos do 3º ano do ensino fundamental apresentavam problemas de alfabetização, não tendo desenvolvido plenamente as habilidades de leitura prevista para a série. Como reação a esse quadro, foi criado o PIP, com algumas ações acontecendo ainda ao final do ano de 2006 (SIMÕES, 2012). O PIP assumiu, assim como o Compromisso Todos Pela Educação, lançado também em 2006, como seu horizonte principal, a alfabetização dos alunos até os oito anos de idade. O programa previa a produção de material didático e a formação continuada dos agentes pedagógicos, além da estipulação de metas para os resultados de aprendizagem. Ele definiu também escolas estratégicas e buscou um acompanhamento mais fino dos alunos com baixo desempenho. Como vimos, o Proalfa contou, entre 2007 e 2014, com uma avaliação censitária, para o 4º ano, dos alunos com baixo desempenho no 3º ano na edição anterior. Foi proposta a

---

<sup>187</sup> Para mais sobre o Paic, ver, por exemplo, Marques et al (2009), Gusmão e Ribeiro (2011) e Fonseca (2013).

elaboração de um plano de intervenção pela escola, levando em consideração os resultados dos Proalfa (o Dia D – o Dia da Avaliação, um evento anual em toda a rede para a apresentação e discussão dos resultados da avaliação, mostra a centralidade que o programa de avaliação tinha para o PIP).

O trabalho das escolas com os dados do Proalfa foi intensificado com o PIP, que ajudou a desenvolver e a consolidar uma cultura de avaliação no estado, valorizando decisões baseadas em evidências (esse é a opinião das quatro entrevistadas sobre o Proalfa, conforme veremos em tópico adiante). Os resultados no ciclo de alfabetização em Minas começaram a experimentar melhorias significativas, atribuídas às ações do PIP nas escolas (é a leitura de Simões, 2012, e de Ana Lúcia Gazzola, ex-Secretária de Educação de Minas e entrevistada para a tese). Da mesma forma que seus resultados foram a base para a criação do PIP, o Proalfa estava diretamente articulado a esse programa por servir como o instrumento de monitoramento de seus efeitos. Inicialmente destinado ao ciclo de alfabetização, na leitura de Ana Lúcia Gazzola, os resultados positivos do programa levaram à sua expansão para outras etapas de escolaridade. O programa foi descontinuado em 2014.

Com base na tipologia proposta por Bonamino e Sousa (2012), Franco e Calderón (2017) identificam no Simave, ao longo de sua história, as três gerações de avaliação. Entre 1992 e 2002, o sistema de avaliação em Minas (o Simave, com esse nome, foi criado em 2000) pertenceu à primeira geração da avaliação, com instrumentos para fins de diagnóstico e sem consequências para as escolas. Entre 2003 e 2006, a segunda geração: a reforma administrativa do estado, levada a cabo pelo primeiro mandato de Aécio Neves como governador, foi marcada pelo Acordo de Resultados, um grande pacto entre diversas entidades e órgãos do estado que previa o estabelecimento de metas, com base em indicadores predeterminados, e o pagamento de prêmio por produtividade aos servidores em virtude do atingimento das metas de desempenho. O Simave tornou-se um instrumento de gestão ainda mais importante nesse contexto, tendo em vista que os indicadores educacionais desempenhavam um papel importante para o acordo. Os resultados de Proeb e Proalfa foram utilizados como referência “para ordenar as ações do governo no campo educacional” (FRANCO; CALDERÓN, 2017, p.148).

A terceira geração da avaliação em Minas, conforme proposta de Franco e Calderón (2017), ocorreu no período entre 2007 e 2013, envolvendo parte do período analisado pela tese. O Acordo de Resultados deu lugar ao Estado para Resultados e a qualidade da educação, mensurada pelo Simave, passou a ser vinculada a um prêmio por produtividade para os

servidores. Embora a concepção do prêmio tenha sido gestada ainda na fase do Acordo de Resultados, seu pagamento começou a ser efetuado a partir de 2007.

Estabelecidas as metas e uma vez atingidas, os servidores passaram a receber uma bonificação, em dinheiro, em virtude do desempenho das escolas no Simave. As metas diziam respeito ao percentual de alunos no padrão Recomendado pelo Proeb e pelo Proalfa, à proficiência média nos testes, às taxas de distorção idade-série e ao fluxo escolar (FRANCO; CALDERÓN, 2017). Essa política de bonificação foi cancelada em 2014.

No Espírito Santo, conforme relata Andressa Rocha, em entrevista para a tese, a avaliação está associada a outros programas da secretaria, como a formação de professores. Contudo, não há no estado, para o período analisado, um programa tão representativo para a política educacional como o Paic, no Ceará, e como foi o PIP, em Minas. Simões (2012) compara o PIP, o Paic e um programa chamado Direito de Aprendizado, no Espírito Santo. Este último, à semelhança dos outros dois, foi criado a partir do que diziam os resultados de avaliação: um número significativo de alunos com problemas de aprendizagem no ciclo de alfabetização. Tendo como ponto de partida os resultados do Paebes Alfa, foi elaborado um Guia de Elaboração para a intervenção pedagógica, dividido em duas partes: i) orientações estratégicas para a escola, que envolvia, entre outras coisas, o estímulo à participação da família, e ii) orientações metodológicas, compostas por sequências didáticas elaboradas por professores da rede estadual. Os resultados do Paebes Alfa serviam ao monitoramento dos efeitos do programa. O programa não existe mais na secretaria<sup>188</sup> e não chegou a se tornar representativo. Outro programa da Seduc do Espírito Santo é o PAES – Pacto pela Aprendizagem no Espírito Santo, que tem foco na aprendizagem e é organizado em três eixos de trabalho. O portal da secretaria estadual de educação não especifica suas ações. Como o PAES foi criado em 2018, ele está fora do período de análise da tese, motivo pelo qual não foi considerado.

Como no Ceará, o Espírito Santo possui uma política de bonificação consolidada, existente desde 2010, o Bônus Desempenho (criado pelo Lei Complementar nº 504/2009), o que faz com que possamos considerar o Paebes Alfa uma avaliação de terceira geração. O Bônus Desempenho é um prêmio em dinheiro, sem vínculo à remuneração mensal, concedido aos profissionais da educação do estado, anualmente, e calculado a partir de indicadores individuais e coletivos. Seu objetivo é funcionar como estímulo para a melhoria da qualidade da educação

---

<sup>188</sup> Em consulta a uma das entrevistadas, Renata Marinho, ela desconhecia o programa, que já havia sido cancelado antes de sua chegada à equipe de avaliação da secretaria. Encontrei um dos Guias do programa, para o ano de 2010, mas as buscas não me levaram a outro material.

no estado, usando os resultados da avaliação como base para concessão. Os servidores vinculados à Seduc podem receber até 1 salário a mais por ano, de acordo com o desempenho da unidade em que estão alocados. O bônus é calculado a partir de dois critérios: i) um coletivo – o IMU (Indicador de Merecimento da Unidade), que leva em conta os resultados no Paebes (incluindo o Alfa), consolidados no IDE (Índice de Desenvolvimento das Escolas Estaduais do Espírito Santo); o IDE considera os resultados dos testes (aprendizagem), o nível socioeconômico dos estudantes e o nível do ensino ofertado; e ii) um individual, obtido por meio do Indicador de Contribuição ao Desempenho (ter exercido suas atividades em uma mesma unidade por, pelo menos, 122 dias no ano) e do Fator de Valorização da Assiduidade (descontos no bônus em virtude de ausências registradas).

### **5.3.8 Longevidade (duração e constância)**

Uma das principais características atribuídas às políticas públicas em educação no Brasil é sua inconstância, sua fluidez e fragilidade diante de mudanças no cenário político no qual estão inseridas. De forma geral, a mudança de governo leva a mudanças nas políticas educacionais. As políticas de Estado respondem pelo minoria dos casos, em um contexto no qual as políticas de governo dominam o cenário da educação. A busca por uma identidade governamental na área de educação, central para as pretensões políticas dos candidatos ao governo, faz com que as políticas educacionais sejam caracterizadas por sua pouca durabilidade, sem que o tempo de maturação para o alcance de seus efeitos seja observado. Em educação, os efeitos de uma política não costumam ser imediatos, exigindo períodos de tempo mais longos para que possam ser observados. Se as políticas não têm duração mínima para que os efeitos apareçam, elas são julgadas, de modo precipitado, como inadequadas.

Diante disso, a duração de uma política ao longo do tempo é um fator central para sua efetividade. A longevidade de políticas educacionais costuma não ser a regra, mas as políticas de avaliação, no Brasil, têm apresentado uma duração não observada para outras políticas. De acordo com o que foi apresentado no capítulo 2, desde o início dos anos 1990, com a criação do SAEB, a avaliação educacional não apenas se manteve como uma política importante no cenário educacional do país como se expandiu e se desenvolveu, a despeito de mudanças governamentais. Os sistemas estaduais e municipais de avaliação são produto desse processo de expansão.

Para a análise dos sistemas estaduais aqui levada a cabo, interessa saber qual a longevidade da política de avaliação da alfabetização, a existência de eventuais interrupções e o motivo pelo qual elas aconteceram. A longevidade e constância da política de avaliação influenciam todos os outros critérios selecionados para a construção da tipologia. Os objetivos da avaliação podem ter mudado ao longo do tempo, afetando também o desenho da avaliação e outras de suas características. Vínculos com outras políticas podem ter sido rompidos e novos vínculos criados. Mais ou menos recursos podem ter sido destinados para a avaliação. O mesmo pode ser dito para o vínculo com a legislação, para o tipo de federalismo colocado em cena pela política, para a articulação com ações de *accountability* e para tipo de divulgação de resultados proposto pelo sistema de avaliação.

Assim, para fins da definição de tipos, no que tange à longevidade, uma política de avaliação da alfabetização pode ser longa, com uma série histórica de mais de cinco anos, ou breve, com menos de cinco anos de duração. Cinco edições é tempo suficiente para o estabelecimento de uma linha histórica considerável. Esse é um corte arbitrário, mas como os governos têm duração de quatro anos, a ideia foi estabelecer um corte temporal para definir o que seria ou não longo em termos de política de avaliação no âmbito dos estados. Sendo longa, a política pode apresentar dois subtipos, com ou sem interrupções. No caso das interrupções da política, interessa saber os motivos pelos quais elas ocorreram (financeiros, técnicos, políticos etc) e as consequências desses intervalos para a política.

O Paebes Alfa pode ser considerado uma política de avaliação longa, de acordo com o critério estipulado para tanto. O programa teve início em 2008, conforme afirmado anteriormente. Desde então, não houve uma interrupção sequer no Paebes Alfa, o que faz com que, além de longa, seja contínua. O programa passou por alguns ajustes, em especial, no que diz respeito ao desenho da avaliação, mas sem alterações significativas em seu formato (podemos dizer, também, que é uma política estável). Ao todo, até 2017, foram dez edições consecutivas do Paebes Alfa, uma política de avaliação que atravessou governos e equipes na secretaria de educação. Na visão de Andressa Rocha, ex-Assessora de Planejamento e Gestão da Seduc do Espírito Santo, em entrevista para a tese, a longevidade do Paebes Alfa pode ser atribuída à continuidade, nas diferentes gestões à frente da secretaria de educação do estado, da ideia de que a mensuração dos resultados da ação pedagógica ao longo do tempo é fundamental para o planejamento de melhorias através desses resultados, “considerando que o avanço de todos os estudantes é importante e merece atenção”. De acordo com Renata Marinho, ex-Subgerente de Avaliação Educacional da secretaria, também em entrevista para a tese, os



principais motivos para a continuidade do programa foram suas características técnicas, próprias de um sistema de avaliação bem estruturado e planejado (a entrevistada define o programa como robusto), o que fizeram com que o Paebes Alfa se tornasse uma referência para os outros estados do país, e o fato das redes de ensino terem entendido a proposta da avaliação da alfabetização, fazendo uso dos resultados ao longo das edições.

Da mesma forma, o Spaece Alfa pode ser considerado uma política longa, constante e estável. Sem sofrer alterações significativas em seu formato, o programa realizou sua primeira edição em 2007. Até 2017, foram onze edições do programa, sem qualquer interrupção. Segundo Alessio da Costa Lima, ex-Orientador da Célula de Avaliação da Seduc do Ceará e um dos entrevistados no âmbito desta teste, a continuidade da política educacional do estado, a despeito das mudanças no governo, é um dos elementos principais para a solidez assumida pelo Spaece Alfa, acompanhada de seu vínculo com a política de responsabilização implementada pelo estado, com o desenvolvimento de uma cultura de avaliação no Ceará e com a preponderância que os resultados do Spaece Alfa ganharam na mídia e nos eventos políticos no estado. Para Márcia Campos (em entrevista para a tese), uma das idealizadoras do Paic, a consciência dos diferentes governos acerca da importância da avaliação como instrumento de gestão, possibilitando intervenções para melhorar a qualidade da educação, foi um dos principais fatores para a constância experimentada pelo Spaece Alfa. A continuidade da política educacional no Ceará é outro elemento destacado por ela, em especial, a cooperação com os municípios e o compromisso de combate ao analfabetismo, o que influencia a posição de centralidade assumida pela avaliação. Soma-se a isso a melhoria dos resultados experimentada pelo estado no ciclo de alfabetização, fazendo com que o monitoramento seja percebido como um fator indispensável para o acompanhamento dessa melhoria.

O Proalfa foi o pioneiro em avaliação da alfabetização no Brasil. Outras iniciativas já haviam ocorrido antes, inclusive, no próprio Espírito Santo, em 2000, mas elas foram isoladas, sem uma metodologia rigorosa e sem continuidade. O Proalfa teve sua primeira edição em 2005, de maneira amostral, para escolas estaduais, e, a partir de 2006, foi integrado ao Simave, sendo realizado de maneira constante desde então. Entre 2005 e 2017 foram três edições consecutivas, sem interrupções. O desenho da avaliação foi alterado ao longo do tempo, como vimos. A avaliação amostral de 2º e 4º ano aconteceu até 2014, depois, foi descontinuada. Apesar disso, a avaliação censitária do Proalfa, para o 3º ano do ensino fundamental, é longa e contínua (desde 2006). Segundo Maria das Graças Pedrosa Bittencourt, ex-Gerente do PIP e uma das entrevistadas para a tese, a continuidade do Proalfa pode ser atribuída, em primeiro lugar, à

própria importância da avaliação educacional para a tomada de decisões voltadas à transformação da realidade escolar, como mostra seu vínculo com o PIP, que foi fundamental, em sua leitura, para a melhoria dos resultados de aprendizagem no estado. Para Raquel Elizabete, ex-Subsecretária de Desenvolvimento da Educação Básica de Minas, em entrevista para a tese, o Proalfa tornou-se longo, primeiramente, pelo cuidado, no âmbito do programa, em reduzir “as naturais resistências e tensões (*relativas à avaliação educacional, em específico, e a qualquer mudança na educação em geral*) e criar um ambiente de confiança favorável às mudanças” (acréscimo nosso, em itálico). A qualidade técnica do programa foi outro ponto a seu favor, mas, segundo ela, o grande passo para consolidação do Proalfa foi a utilização dos resultados da avaliação pelas escolas e redes de ensino. A utilização dos resultados pelos atores escolares para o planejamento das escolas fez com que o programa passasse a fazer parte das ações rotineiras das unidades escolares.

Na opinião de Sueli Pires (em entrevista para a tese), ex-Secretária Adjunta de Educação de Minas, o fundamental para a longevidade do Proalfa foi o envolvimento dos atores no processo. Segunda, ela “a rede percebeu e incorporou o Proalfa à sua política pedagógica como um programa ou instrumento de melhoria da qualidade do ensino”. A melhoria dos resultados no ciclo de alfabetização ao longo do tempo, fez, em sua opinião, com que o programa ganhasse mais credibilidade e visibilidade, consolidando a cultura de avaliação em todo o estado, em especial, pela participação das redes municipais. Para Ana Lúcia Gazzola, ex-Secretária de Educação de Minas Gerais, é impossível “avançar num sistema educacional sem a existência de avaliações consequentes e transparentes”. Há uma relação direta entre o monitoramento e o tempo. Acompanhar a evolução (ou a não evolução) de uma rede de ensino pressupõe instrumentos de mensuração que se prolonguem no tempo. O movimento de melhoria da educação necessita do estabelecimento de metas e, sem avaliação, não há metas possíveis de serem estabelecidas. A educação precisa se pautar em evidências e é isso o que a avaliação propicia. Para ela, portanto, o principal motivo para a longevidade do Proalfa está na própria importância da avaliação e na função que exerce como instrumento de gestão.

A longevidade dos três programas aqui analisados influencia outras de suas características. Programas longevos estão, por exemplo, mais propensos a possuir políticas de responsabilização. Em geral, espera-se que o programa de avaliação esteja consolidado em um cenário educacional, em certa medida, para que uma política de responsabilização seja a ele atrelada (essa é uma expectativa, não uma regra). Da mesma forma, programas mais longevos tendem a experimentar um grau mais elevado de institucionalização nas secretarias de

educação, manifestada, entre outros fatores, na existência de equipes próprias de avaliação. É o caso de Paebes Alfa, Proalfa e Spaece Alfa. As secretarias de Espírito Santo, Minas Gerais e Ceará contam com equipes específicas para o trabalho com a avaliação educacional, o que é um indicativo da institucionalização da avaliação como prática recorrente de gestão da rede. A pesquisa em avaliação se beneficiaria de um estudo sobre essas equipes: as equipes possuem autonomia na secretaria? Que tipo de decisões estão sob sua responsabilidade? A que órgão estão vinculadas? Como foram constituídas? Como recrutam os atores que as compõem? Qual é a especialização e experiência desses atores com a avaliação educacional? Como é o trabalho em parceria com instituições de pesquisa que apoiam a execução de sistemas de avaliação, na visão desses atores? Um estudo dessa natureza ainda está por ser feito.

### **5.3.9 Participação dos municípios**

As avaliações educacionais dos sistemas próprios, em particular as dos sistemas estaduais, como é o caso dos três sistemas aqui analisados, se destinam a produzir informações sobre as redes públicas daquele ente federado. Desde a Constituição Federal de 1988 e o avanço dos processos de municipalização da educação, o que vemos é uma demanda necessária de integração entre as diferentes redes de ensino e entre os entes federados, em especial os estados e os municípios. A responsabilidade precípua dos municípios pela educação infantil, ciclo de alfabetização e anos iniciais do ensino fundamental coloca em movimento uma dinâmica através da qual o aluno que hoje é da rede municipal passará a ser da rede estadual em um futuro próximo. Assim, à rede do estado interessa a qualidade do ensino oferecido pelas redes municipais, pois isso influenciará decisivamente o desempenho e a trajetória escolar de seus alunos.

A cooperação entre os entes federados, dessa forma, para que seja efetiva, deve ir além dos princípios políticos do federalismo. A cooperação deve se orientar por questões técnicas e a preocupação que a alimenta é pedagógica, não apenas política. O que está em jogo é a qualidade do ensino ofertado, de um lado, como garantia de um direito e como exigência de um compromisso público, de outro, como imperativo técnico e pedagógico com vistas à potencialização da aprendizagem dos alunos.

O desenho de um federalismo cooperativo em educação encerra ainda mais sentido quando falamos do ciclo de alfabetização. Trata-se de uma etapa nevrálgica para toda a escolaridade do aluno, afetando, decisivamente, suas chances de sucesso escolar (e também

fora da escola). O analfabetismo em idade escolar, daqueles que permanecem na escola, mas não estão plenamente alfabetizados, é, hoje, no Brasil, um dos grandes problemas pedagógicos a serem enfrentados. Como regra, o ciclo de alfabetização é de responsabilidade dos municípios. Uma política de avaliação da alfabetização, em nível estadual, deve, necessariamente, envolver os municípios, sob pena de ser inadequada, a não ser que o estado ainda responda por boa parte da oferta das matrículas nessa etapa de escolaridade. De todo modo, mesmo nesses casos, em que o processo de municipalização da alfabetização ainda não foi inteiramente completado, as redes municipais são as grandes responsáveis, em montante, pela oferta do ciclo. É o que ocorre no Espírito Santo, Minas Gerais e Ceará. Um sistema estadual de avaliação da alfabetização não deve olvidar esse fato. As redes municipais precisam integrar a política de avaliação da alfabetização do estado.

A forma como esse processo de integração será realizado pode variar substantivamente, colocando em cena diferentes caminhos para essa cooperação federalista. Os municípios podem participar diretamente e com protagonismo das decisões acerca do programa de avaliação (como a definição de matrizes, das etapas e das disciplinas a serem avaliadas, por exemplo), ou simplesmente aderirem ao programa desenhado e implementado pelo estado. Mesmo nesse último caso, a participação dos municípios nas decisões que envolvem o programa pode ser prevista e estimulada pelo estado (com feedbacks anuais estabelecidos sobre o programa, em canais próprios e institucionalizados para tanto). Em relação aos custos da avaliação, os municípios podem ficar responsáveis por eles, parcial ou integralmente, ou eles podem ser cobertos totalmente pelos estados, de modo a incentivar a participação das redes municipais na avaliação.

Como vimos, para os três programas, os municípios respondem pela maior parte dos alunos avaliados. Nos três casos, para o período analisado, logo nas primeiras edições, os municípios foram envolvidos no processo de avaliação do ciclo de alfabetização proposto e conduzido pelo estado. Não foi possível rastrear, a partir das decisões de pesquisa da tese (pesquisa documental e entrevistas), qualquer canal institucionalizado de participação efetiva dos municípios nas decisões que envolvem o programa. Nos três programas, as decisões permanecem nas mãos da secretaria de educação do estado. Isso não exclui, contudo, a existência de uma colaboração mais efetiva de um ou outro município através de canais não institucionalizados. A título de exemplo, podemos imaginar que, pela história recente, eventuais considerações sobre o Spaece Alfa vindas da secretaria municipal de educação de Sobral tenham peso sobre as decisões relativas ao programa tomadas pela secretaria de educação do

estado. Hipóteses dessa natureza abrem um novo caminho de pesquisa sobre a cooperação entre estados e municípios no âmbito da avaliação da alfabetização.

A participação dos municípios no Paebes Alfa é feita por adesão. O programa é conduzido pelo estado e os municípios podem aderir ao formato proposto. Eles são estimulados a participar, por interesse do estado, mas não há mecanismos formais para compeli-los a integrar o programa de avaliação. Uma vez que a adesão ocorre, os municípios se submetem às regras estabelecidas pelo estado. Eles não podem tomar decisões sobre como conduzir a aplicação dos testes ou escolher as turmas e escolas que serão avaliadas. O Paebes Alfa, como vimos, é uma avaliação censitária. Como tal, se a rede municipal faz a adesão, todas as escolas deverão ser avaliadas. Os testes são elaborados para o programa como um todo, de forma que as matrizes são as mesmas, independentemente da rede avaliada (em uma avaliação conduzida pelo município, por exemplo, a secretaria de educação poderia decidir sobre quais séries e disciplinas seriam avaliadas, bem como sobre o conteúdo da matriz de referência). Assim, a decisão que cabe ao município é participar ou não da avaliação.

O mesmo pode ser dito no que diz respeito às escolas particulares que decidem participar do processo, fenômeno que singulariza o Paebes Alfa em relação aos outros dois programas aqui analisados. Nesse caso, é evidente, não estamos diante de uma cooperação federativa, visto que apenas o estado figura como ente federativo na parceria. De todo modo, cabe afirmar que as escolas particulares também não possuem nenhum poder de decisão quanto às diretrizes e características do Paebes Alfa. Assim como com os municípios, sua participação é facultada, e mesmo estimulada, mas obedecendo ao que foi estipulado pelo estado.

Se a participação dos municípios é por adesão, isso não significa que não haja trocas entre os entes. Mais especificamente, transferência de conhecimento e tecnologia educacional, além de suporte para todas as fases da avaliação. No termo de cooperação (ver Anexo V), documento enviado pelo estado para que os municípios possam aderir ao Paebes Alfa e ao Paebes, são estabelecidas as condições da cooperação. Como o próprio nome do documento já diz, cooperação entre os entes é técnica. A cláusula terceira do documento afirma explicitamente que não haverá qualquer tipo de repasse de recursos financeiros para o desenvolvimento do programa de avaliação. Apesar disso, todos os custos com a avaliação são de responsabilidade do estado, sem que o município arque com qualquer despesa.

O documento deixa claro, também, quais são as atribuições de cada um dos entes. Além de fornecer os instrumentos para a avaliação, o estado deve oferecer aos municípios informações técnicas sobre o Paebes Alfa, além de assegurar que todas as escolas dos

municípios aderentes participem de fato do programa. Somam-se a isso a capacitação da figura do coordenador municipal, responsável pela aplicação da avaliação nos municípios, e a entrega e capacitação para a análise dos resultados da avaliação, ações de responsabilidade do estado (essas ações compõem a divulgação de resultados do programa, conforme será visto em tópico próprio).

Aos municípios, por sua vez, cabe a organização da avaliação em seus domínios. Isso implica na nomeação de um dos técnicos da secretaria municipal para que atue como interlocutor diante da secretaria estadual de educação. Além disso, os municípios devem compor uma coordenação municipal, formada por um técnico ou mais, que se responsabiliza por todas as ações do Paebes Alfa no município (a figura do coordenador municipal deve, ainda, participar das capacitações oferecidas pela secretaria estadual). Outras ações de responsabilidade do município são a organização de reuniões de sensibilização para o bom andamento da avaliação (o documento prevê a participação de pais e responsáveis pelos alunos, corpo docente das escolas, equipe técnica da secretaria e outros profissionais da rede municipal de ensino) e a designação de profissionais para que sejam aplicadores dos testes nas escolas e a capacitação desses aplicadores.

Além dessas ações de responsabilidade do município, outras duas merecem destaque. O termo de cooperação atribui ao município a responsabilidade pelo envolvimento da rede escolar e da equipe da secretaria municipal no que se refere à análise e à disseminação dos resultados, reconhecendo o papel fundamental que essa etapa da avaliação desempenha. Ainda, prevê o município como responsável pelo uso dos resultados da avaliação para a definição e implementação de políticas públicas, cujo horizonte é a melhoria da qualidade do ensino.

A cláusula quarta do termo de cooperação técnica trata do acompanhamento das ações. A referida cláusula afirma que a secretaria de educação do estado exerce a função gerencial de supervisão durante o período de duração da cooperação, assegurando a seus agentes a possibilidade de intervenção para a reorientação das ações, caso isso seja necessário. Em seu parágrafo único, a cláusula prevê ainda que os municípios devem dar livre acesso, a qualquer lugar e tempo, aos servidores da secretaria de educação do estado que estiverem exercendo as atividades de supervisão da política de avaliação, em relação a todos os atos relacionados ao Paebes Alfa.

Para o Proalfa e o Spaece Alfa, a condução da cooperação com os municípios em relação à avaliação apresenta algumas semelhanças com o que está previsto para o Paebes Alfa. Nos dois casos, os custos com a avaliação também são exclusivos da secretaria de educação do

estado. Não há dúvidas de que esse é um dos fatores de estímulo para a participação dos municípios.

Em Minas Gerais, o município faz adesão ao programa e, ano a ano, a secretaria de educação solicita aos municípios que já fizeram a adesão a base de dados com as escolas e os alunos (organizados por turma) da rede matriculados nas séries avaliadas pelo Proalfa. A adesão é realizada uma única vez, sem que haja necessidade de renovação formal a cada edição do programa. Todos os anos, as Superintendências Regionais de Ensino (SRE) realizam uma capacitação para a condução das avaliações nos municípios, contando com a presença de um representante de cada secretaria municipal sob a coordenação daquela SRE. Na ocasião, os pacotes com os testes impressos são distribuídos para os representantes das secretarias municipais, já com a especificação dos pacotes para cada escola da rede. A aplicação é feita por cada município, responsável por recolher os testes e devolvê-los à superintendência, de modo que sejam encaminhados posteriormente para a etapa de leitura e processamento informações, conduzida pelo CAEd/UFJF. Não há nenhum mecanismo institucionalizado para compelir a participação das redes municipais do Proalfa (como vimos, o programa teve início em 2005, mas a rede municipal de Juiz de Fora, por exemplo, só aderiu ao programa em 2009), embora a participação seja desejada e estimulada pela secretaria estadual de educação.

No Ceará, os municípios também assinaram um termo de adesão para a participação no Spaece Alfa. A rigor, eles não são obrigados a participar. Contudo, o estado apresenta uma peculiaridade, que diferencia sobremaneira a participação dos municípios no programa daquela observada no Espírito Santo e em Minas Gerais. Em primeiro lugar, a secretaria de educação do Ceará criou uma coordenação de Cooperação com os Municípios, o que já demonstra o nível de envolvimento que pretendia desenvolver entre os entes federados. A proposta do Paic, como vimos, pressupunha esse maior envolvimento. Os municípios, mais do que estimulados a participar, faziam parte de um cenário mais amplo, o do compromisso com a eliminação do analfabetismo no estado, uma parceria que envolveu um conjunto abrangente de órgãos, entes e instituições. Em um contexto como esse, é possível imaginar que a participação, embora facultada, sinalize mais do que uma simples decisão nas mãos do município, mas, antes, a reafirmação de seu compromisso com a erradicação do analfabetismo. Não seria exagero afirmar, portanto, que o contexto da política educacional do Ceará, desenvolvido a partir de 2007, gerou para os municípios mais do que a possibilidade de adesão ao Spaece Alfa. Trata-se de um compromisso com a política educacional do estado, uma maneira de compelir sua participação no programa de avaliação.

No âmbito dessa política do estado, um fator que compele ainda mais os municípios a participarem do Spaece Alfa refere-se à distribuição da cota-parte do ICMS. A lei estadual 14.023, de 17 de dezembro de 2007, alterou a lei 12.612/1996, e estabeleceu, em seu artigo 1º, inciso II, que a distribuição da cota-parte do ICMS seria calculada com peso de 18% em função do IQE (Índice Municipal de Qualidade Educacional), composto pela taxa de aprovação dos alunos do 1º ao 5º ano do ensino fundamental e pelos resultados de desempenho dos alunos do 2º e 5º anos no sistema de avaliação do estado, o Spaece.

O sistema de transferências do ICMS estabelecido pela Constituição Federal de 1988, em seu artigo 158, estipula que 25% do valor total arrecadado pelo estado deve ser transferido aos municípios. Destes 25%, 75% figuram como parcela impositiva, distribuída pelo critério do VAF (Valor Fiscal Adicionado). Os 25% restantes compõem uma parcela autônoma, a ser distribuída de acordo com os critérios estabelecidos por cada estado através de lei estadual. No Ceará, os 25% da parcela autônoma passaram a ser distribuída, a partir da referida lei de 2007, com base em indicadores de qualidade em três áreas consideradas prioritárias: educação, saúde e meio ambiente. Contudo, fica claro o peso atribuído ao indicador de qualidade da educação, que responde por 18% dos 25% da parcela autônoma (5% em função do índice de qualidade da saúde e 2% em função do índice de qualidade do meio ambiente). De acordo com Irff e Petterine (2011), a mudança da lei de ICMS do Ceará gerou um impacto sobre alguns indicadores da educação que não foi observada, por exemplo, nos indicadores de saúde.

Segundo Brandão (2014), o rateio do ICMS demonstra a valorização e o peso que a política pública em educação tem no Ceará. Em Minas Gerais, por exemplo, a educação também figura como critério para a divisão do ICMS, mas em proporção bem menor (peso de 2%). E o Spaece Alfa exerce um papel fundamental nessa engrenagem, visto que seus resultados são utilizados para compor o IQE. Diante disso, podemos afirmar que a participação dos municípios no programa de avaliação da alfabetização não é por adesão da mesma maneira que ocorre em Minas Gerais e no Espírito Santo. Ela é compelida pelo contexto da política educacional do estado e pela lei que estipula a distribuição da cota-parte do ICMS.

### **5.3.10 Volume e transparência dos recursos aplicados em avaliação**

Um dos temas mais espinhosos e complexos no campo da educação é o financiamento. A temática, pela importância que possui num contexto como o brasileiro, em que os recursos são escassos e alvos de constantes embates, disputas e conflitos, ganha em complexidade



quando envolve o tema da avaliação em larga escala. Sistemas de avaliação, segundo narrativa já apresentada no capítulo 2, sofreram muitas críticas ao longo do tempo (e ainda sofrem). Ao elemento ideológico, inflado pelo desconhecimento acerca dos reais limites e potencialidades do instrumento de avaliação, somou-se o aspecto financeiro. Um sistema de avaliação exige um considerável montante de investimentos. É necessário o desenvolvimento de itens para a composição dos testes, o que envolve um significativo conjunto de profissionais capacitados para tanto. É preciso aplicar os testes, imprimi-los (quando é o caso) e distribuí-los para as escolas. Os aplicadores precisam ser capacitados e remunerados. A seguir, as medidas são produzidas, por profissionais também especializados, e os resultados são divulgados através de um conjunto de canais, que exigem, por sua vez, profissionais gabaritados. A avaliação movimenta um número substantivo de profissionais especializados e exige uma operação nada trivial de logística.

Tantos recursos destinados à avaliação contribuíram para as críticas que recebeu. Os críticos, nesse ponto, argumentam que tal investimento poderia ser destinado para outros setores da educação, em especial, para a melhoria da remuneração de professores e da infraestrutura das escolas. É evidente que as preocupações com a remuneração dos profissionais da educação e com as condições das escolas são justificadas e têm razão de ser. Contudo, o problema da remuneração dos professores, por exemplo, não pode ser atribuído diretamente a um programa de avaliação. A remuneração sempre foi uma questão de tensão e sempre se apresentou como uma demanda constante da categoria, mesmo antes da existência de sistemas de avaliação em larga escala no Brasil. Além disso, os recursos destinados à avaliação podem ser oriundos de fontes externas ao governo, como organismos internacionais e instituições nacionais do terceiro setor, que os destinam para fins determinados e específicos, como a avaliação. Um levantamento mais profundo e detalhado sobre o financiamento de sistemas próprios de avaliação educacional no Brasil, em especial, os estaduais, ainda está por ser feito, especificando as fontes e os volumes dos recursos, bem como as condições da parceria e as contrapartidas exigidas pelos órgãos financiadores. A rigor, uma pesquisa dessa natureza, sobre o financiamento da avaliação educacional no país, se beneficiaria não apenas da análise dos cenários em que os recursos para a avaliação são oriundos de organismos externos ao governo, mas também dos casos de financiamento pelo próprio governo, investigando a origem desses recursos (a partir da tributação, colaboração entre os entes federados, transferência de recursos etc).

Em relação aos recursos destinados à avaliação, a análise aqui levada a cabo para os três programas se debruçou sobre o volume dos recursos destinados à avaliação em comparação

com o montante destinado para a educação como um todo, tendo em vista o que é divulgado pelos estados através de seus portais da transparência. A análise restringiu-se aos gastos em educação oriundos das ações das secretarias de educação de cada estado, sem levar em consideração despesas com educação oriundas de outros órgãos do governo.

O portal da transparência<sup>189</sup> do governo do estado do Espírito permite consultas das despesas do estado por ano e por órgão do executivo, de modo que é possível obter informações, para todo o período analisado, sobre os gastos gerais da secretaria de educação. Existem vários outros filtros que permitem uma consulta mais detalhada dos gastos, especificando ações, programas, categorias econômicas etc. O gastos gerais da secretaria de educação, a cada ano, foram obtidos sem maiores dificuldades. A dificuldade surge quando busca-se os gastos específicos com o programa de avaliação. No filtro *Ação*, a busca por *Paebes Alfa* e por *Paebes* não dá resultados. A busca por *avaliação* leva aos seguintes resultados: *Avaliação da educação no Espírito Santo; Avaliação da educação básica; Monitoramento e avaliação de políticas públicas; Realização do Censo Escolar, avaliação de alunos, estudos e pesquisas da educação*. Qualquer uma dessas rubricas, ou a soma de todas elas, poderia estar relacionada aos gastos com o sistema de avaliação do Espírito Santo, de modo que o portal da transparência não fornece as informações necessária para realizar a comparação aqui pretendida. Assim, os gastos específicos com o programa de avaliação do estado foram obtidos através dos contratos estabelecidos entre a secretaria de educação e o CAEd/UFJF, publicados no Diário Oficial do estado.

Aqui, três observações são necessárias: i) para os gastos totais da secretaria, foram usados os valores empenhados, ou seja, o valor do orçamento reservado para fazer frente aos compromissos assumidos com credores e fornecedores, e não os valores liquidados ou pagos (o intuito era verificar quanto do orçamento estava comprometido com a avaliação); ii) para os gastos destinados à avaliação, foram usados os valores previstos no contrato com o CAEd, sem levar em consideração posteriores supressões ou aditivos; e iii) as informações sobre gastos específicos com o Paebes Alfa não estavam disponíveis para todo o período, de modo que a tabela a seguir refere-se aos gastos totais com a política de avaliação (envolvendo o Paebes e o Paebes Tri), o que permitiu construir a linha histórica. Em relação a esse último ponto, o argumento não fica prejudicado, visto que a ideia era verificar se o montante destinado à política de avaliação de fato faz jus à crítica de que os custos da avaliação são excessivamente onerosos.

---

<sup>189</sup> <https://transparencia.es.gov.br>

As informações sobre o Paebes Alfa, para os anos disponíveis, são apresentadas no corpo do texto (a mesma lógica foi aplicada para o Proalfa e o Spaece Alfa). Em relação ao ponto ii, é preciso ressaltar que, para determinados períodos, o contrato firmado entre a secretaria de educação e o CAEd tinha como objeto todos os programas de avaliação do estado (e não apenas o Paebes Alfa), de modo que não foi possível separar os gastos específicos de cada programa. A tabela a seguir apresenta os gastos totais da secretaria de educação, os gastos com a política de avaliação e o percentual que estes últimos representam em relação àqueles.

Tabela 6 - Gastos totais e com a política de avaliação da secretaria estadual de educação do Espírito Santo – 2009 a 2017

Tipo de gasto/ano	Gastos em Reais									
	2009	2010	2011	2012	2013	2014	2015	2016	2017	Total
Política de avaliação	3.576.758,81	3.851.830,62	4.143.773,24	2.859.370,53	3.892.652,65	4.253.928,52	4.394.048,30	8.322.913,80	8.614.875,21	43.910.151,68
Gasto total da secretaria	547.042.551,97	765.376.947,73	843.482.843,61	876.778.166,40	955.882.376,00	1.142.840.610,19	945.528.495,18	934.497.083,61	1.006.419.106,82	8.017.848.181,51
Percentual de gastos com a avaliação	0,65	0,50	0,49	0,33	0,41	0,37	0,46	0,89	0,86	0,55

Fonte: Portal da transparência do estado do Espírito Santo e Diário Oficial do estado. Tabulação própria.

Como podemos observar por meio da leitura da tabela, entre 2009 e 2017, o Espírito Santo investiu R\$ 43.910.151,68 em sua política de avaliação, ao passo que os gastos totais da secretaria de educação, para o mesmo período, foram de R\$ 8.017.848.181,51. Considerando o período todo, os gastos com a política de avaliação respondem por 0,55% do total de despesas da secretaria. No ano em que esse percentual foi mais elevado, em 2016, ele respondeu por 0,89% do total. É possível notar também que, a partir de 2012, os investimentos em avaliação aumentaram ano a ano no estado, chegando a R\$ 8.614.875,21, em 2017, o maior valor absoluto com avaliação no período analisado. Acontece que os gastos gerais da secretaria também aumentaram em 2017, quando comparados a quase todos os anos anteriores (com exceção do ano de 2014).

É preciso pontuar que: i) o período de 2009 a 2011 é abarcado por um contrato de três anos com o CAEd, com especificação dos gastos por programa (é possível identificar os gastos com o Paebes Alfa<sup>190</sup> e com o Paebes); ii) o período de 2012 a 2015 possui um contrato único, para as 4 edições da avaliação, sem especificar gastos específicos para o Paebes Alfa e para o Paebes, de modo que a linha histórica que poderia apresentar os gastos específicos para o Paebes Alfa fica impossibilitada; e iii) para o período de 2016 a 2019, há um contrato único que

<sup>190</sup> As despesas com o Paebes Alfa, para os anos de 2009, 2010 e 2011 foram, respectivamente: R\$ 1.551.172,48; R\$ 1.332.457,18; e 1.526.946,32.

especifica as despesas com o Paebes e o Paebes Alfa (em conjunto) e com o Paebes Tri. Ainda uma observação. A variação dos gastos com a avaliação tem como principal determinante o número de alunos avaliados (outras variáveis são a impressão ou não dos testes e o pagamento de aplicadores), de modo que se um programa possui maiores gastos do que o outro, muito provavelmente é em virtude do número de alunos que avalia.

O portal da transparência do estado de Minas Gerais<sup>191</sup> permite uma consulta ampla acerca dos gastos do governo ao longo dos anos. É possível selecionar as despesas por órgão, programas, ação, fonte dos recursos, função etc. Para os fins da pesquisa, interessava ter acesso aos gastos do estado com a avaliação educacional, mais especificamente, com o Proalfa, comparando os valores com os gastos gerais com a educação no estado. O sistema de filtros do portal é bastante simples e eficaz, o que permitiu a consulta, de 2009 a 2017, das despesas com a avaliação educacional. Contudo, não há especificação dos gastos para o Proalfa e para o Proeb, mas para a avaliação educacional como um todo (o que representa uma vantagem para o usuário em relação ao portal do Espírito Santo). Selecionando a secretaria de educação como órgão para consulta, o usuário do portal deve selecionar o filtro para *Ação*, de modo a encontrar a informação desejada para os gastos com a avaliação. Ao longo do período analisado, o nome da ação referente à avaliação foi alterado, o que pode confundir o usuário. Entre 2009 e 2011, a ação está registrada como *Avaliação da educação básica (Proeb e Proalfa)*. Entre 2012 e 2016, o registro é Simave, com uma divisão: *Simave – Ensino Fundamental* e *Simave – Ensino Médio*. Esse é mais um complicador para quem está fazendo a consulta. O Simave – Ensino Fundamental envolve o Proalfa e o Proeb para o 5º e 9º anos, ao passo que o Simave - Ensino Médio envolve o Proeb aplicado nesta etapa de escolaridade. Essa divisão separa o gastos da avaliação aplicada ao ensino fundamental daqueles referentes à avaliação para o ensino médio, diferente da prestação de contas apresentada para o triênio 2009-2011, mas, a despeito disso, permanece sem especificar os gastos com o Proalfa. Para o ano de 2017, a ação é denominada *Avaliação Educacional para Redução de Desigualdades*, sem divisão entre o ensino fundamental e o médio.

Nos gastos gerais da secretaria de educação, de 2009 a 2012, o valor total refere-se a gastos com educação, exclusivamente. Em 2013 e 2014, a secretaria de educação dividiu seus gastos em duas rubricas: educação, e desportos e lazer. Os valores que integram a tabela adiante, para esses dois anos, referem-se apenas à rubrica educação. E ter 2015 e 2017, apenas a rubrica educação consta nos gastos gerais da secretaria.

---

<sup>191</sup> <http://www.transparencia.mg.gov.br/>

A tabela<sup>192</sup> a seguir apresenta os gastos com educação e com a avaliação educacional no estado de Minas entre 2009 e 2017. Vale notar que o valor das despesas refere-se ao valor empenhado pelo estado.

Tabela 7 - Gastos totais e com a política de avaliação da secretaria estadual de educação de Minas Gerais – 2009 a 2017

Tipo de gasto/anos	Gastos em Reais									Total
	2009	2010	2011	2012	2013	2014	2015	2016	2017	
Política de avaliação	9.543.466,44	11.204.431,28	13.905.824,78	10.963.480,51	16.376.571,87	15.671.509,07	12.270.705,77	14.271.432,27	12.060.124,04	116.267.546,03
Gastos totais da secretaria	4.538.262.218,69	5.089.760.809,77	6.393.022.401,37	6.965.497.200,33	7.758.937.750,02	8.551.091.311,99	8.721.572.438,82	9.973.931.957,80	10.633.652.584,13	68.625.728.672,92
Percentual de gastos com avaliação	0,21	0,22	0,22	0,16	0,21	0,18	0,14	0,14	0,11	0,17

Fonte: Portal da Transparência de Minas Gerais e Diário Oficial do estado. Tabulação do autor.

Como podemos observar a partir da leitura da tabela anterior, entre 2009 e 2017, foram investidos R\$ 116.267.546,03 de reais no sistema de avaliação em Minas Gerais, para um total de R\$ 68.625.738.672,92. 2013 foi o ano com maior investimento, R\$ 16.376.571,87. Sem desprezar o volume de recursos destinados à avaliação, e reconhecendo que é uma política que possui despesas consideráveis, em função dos motivos já expostos, quando comparamos o valor destinado à política de avaliação em Minas Gerais com o montante total gasto pela secretaria de educação, não parece que a despesa com a avaliação é tão relevante a ser apontada como um gasto muito elevado para o estado, que poderia estar sendo investido em outras áreas mais importantes. Os gastos com a política de avaliação nunca chegaram próximos a 1% do total de despesas da secretaria de educação. Longe disso, nos anos em que pôde ser observado maior percentual de gastos com a avaliação, em 2010 e 2011, ele representou apenas 0,22% do total de despesas da secretaria. Observando os gastos de todo o período, a avaliação responde por apenas 0,17% do total.

Minas, ao contrário do Espírito Santo, diminuiu o percentual de gastos com a avaliação. Em 2017, o último ano do período analisado, o percentual de gastos com avaliação representou 0,11% do total de despesas da secretaria de educação do estado.

No estado do Ceará, o portal Ceará Transparente<sup>193</sup> disponibiliza informações sobre as despesas do governo em relação a todos os seus órgãos e secretarias. O sistema de filtros permite ter acesso às despesas específicas da secretaria de educação, ano a ano. Há filtros para órgãos, programas do governo, fontes de recurso etc. É possível ter acesso aos gastos gerais da

<sup>192</sup> A opção pelo uso da tabela aqui, e não pelos gráficos, deve-se à grandeza dos números que envolvem as despesas em Minas, o que prejudicaria a visualização, e aos baixos percentuais representados pelos gastos em avaliação, o que inviabilizou a construção de um gráfico.

<sup>193</sup> <https://ceartransparente.ce.gov.br/?locale=pt-BR>

secretaria de educação, mas não aos gastos específicos com a avaliação do Spaece Alfa (ou mesmo do Spaece). Em um dos filtros, é possível ter acesso a uma rubrica chamada *Desenvolvimento das ações de avaliações, estudos e pesquisas educacionais*. Embora seja genérica e não especifique de que tipo de ações se trata, ela é a única informação com referência explícita à avaliação disponível no site. Contudo, ela está localizada nas ações relativas ao ensino médio, de modo que não é uma informação que podemos usar com confiança para realizar a análise pretendida. Não há rubricas que levem o nome de Spaece ou Spaece Alfa.

Os valores totais dos gastos da secretaria de educação foram obtidos por meio do portal da transparência do estado, mas não há informações disponíveis ali sobre os anos de 2010 e 2011, de modo que a linha histórica fica comprometida. As informações sobre os gastos com avaliação, assim como foi feito para o caso do Espírito Santo, foram obtidas por meio dos contratos celebrados, no período analisado, entre a secretaria de educação e o CAEd, publicados no Diário Oficial do estado (mais uma vez, os valores referem-se ao que está previsto no contrato, sem levar consideração eventuais supressões e aditivos). Os valores de 2010 e 2011 referentes aos programas de avaliação do estado foram disponibilizados na tabela a seguir, mas não foram computados (por falta de informação em relação aos gastos gerais da secretaria de educação) para estabelecer os gastos totais com a política de avaliação.

Tabela 8 - Montante e percentual dos gastos da secretaria de educação do estado do Ceará com a política de avaliação educacional – 2009 a 2017

Tipo de gasto/ano	Gastos em reais									
	2009	2010	2011	2012	2013	2014	2015	2016	2017	Total
Política de Avaliação	7.440.095,58	7.440.095,58	9.876.086,06	9.876.086,06	9.876.086,06	9.876.086,06	15.737.572,54	10.062.311,89	10.802.156,21	73.670.394,39*
Gastos totais da secretaria	3.139.142.617,47	-	-	2.075.030.610,55	2.225.539.568,31	2.395.034.276,44	2.335.653.935,46	2.621.242.805,43	2.722.303.906,46	17.513.947.720,12
Percentual gasto com avaliação	0,24	-	-	0,48	0,44	0,41	0,67	0,38	0,40	0,42

Fonte: Portal da Transparência do Ceará e Diário Oficial do estado. Tabulação do autor.

\*Diz respeito aos valores de 2009 somados aos valores do período entre 2012 a 2017, sem os valores de 2010 e 2011.

Conforme informa a tabela anterior, os gastos do Ceará com os programas de avaliação para o período (sem contabilizar 2010 e 2011) foram de R\$ 73.670.394,39, enquanto os gastos totais da secretaria foram de R\$ 17.513.947.720,12. Em todo o período as despesas com a avaliação responderam por 0,42% do total de despesas da secretaria de educação. O ano de 2015 representou o maior gasto bruto e o maior percentual de gastos com a avaliação no estado (R\$ 15.737.572,54 e 0,67%).

Para o período entre 2008 e 2010, houve um contrato único com o CAEd, com a especificação dos gastos com o Spaece Alfa (R\$ 7.180.662,99 para os três anos). Um novo

contrato foi estabelecido para abarcar o período de 2011 a 2014, divisando as despesas da avaliação para o 2º e o 5º anos juntos e para o 9º ano e o ensino médio juntos. A seguir, foi firmado um contrato exclusivo para o ano de 2015, sem dividir os gastos por programa. Para o período de 2016 a 2019, um único contrato foi firmado, mais uma vez, sem dividir os gastos.

A análise dos gastos com a avaliação educacional para os três programas nos permite perceber que os portais da transparência, embora informem os gastos totais das secretarias de educação e sejam capazes de especificar os gastos através de filtros próprios, não estabelecem distinções entre os diferentes programas de avaliação levados a cabo pela política de avaliação do estado. Em Minas Gerais, não fomos capazes de, através do portal, dividir os gastos com o programa de avaliação da alfabetização daqueles que dizem respeito à política de avaliação educacional como um todo, ao passo que, no Ceará e no Espírito Santo, mesmo a informação sobre os gastos com avaliação em geral não estava disponível (pelo menos, não com uma rubrica que levasse esse nome).

Ponto importante diz respeito ao volume dos gastos com avaliação. No Espírito Santo, as despesas com a avaliação apresentaram, para o período analisado, mas seu percentual máximo em relação às despesas da secretaria de educação foi de 0,89%. Em Minas Gerais e no Ceará o percentual máximo gasto com a avaliação chegou a 0,22% e 0,67%, respectivamente. Em nenhum dos casos, esse percentual chegou a 1% do total de gastos da secretaria. Podemos afirmar que, proporcionalmente, os gastos com avaliação não comprometem o orçamento da educação nos estados. A crítica generalizante de que a avaliação mobiliza um volume de recursos que poderia estar sendo investido em outros setores da educação precisa ser relativizada, por três motivos principais: i) os percentuais variam de estado a estado, de modo que uma crítica genérica à avaliação como excessivamente onerosa ao orçamento não se aplica indistintamente a qualquer cenário; ii) os recursos destinados à avaliação podem ser oriundos de fontes externas ao governo, interessadas, justamente, em financiar a avaliação (e não aplicar os recursos em outro setor), de modo que não são retirados recursos de uma área da educação para que sejam aplicados em avaliação; a pesquisa em avaliação se beneficiaria de uma análise mais aprofundada sobre a fonte de recursos para a avaliação (tema que não abordamos aqui); e iii) mesmo nos casos de financiamento pelo próprio estado, as informações produzidas pela avaliação educacional, apesar de custosas, podem ser entendidas como um instrumento de gestão fundamental, sem o qual é muito difícil realizar uma gestão séria e profissional da rede, de modo que vale a pena investir recursos em avaliação, a fim de evitar gastos desnecessários por conta de um planejamento que não é baseado em evidências (esse foi, por exemplo, um dos

pontos aventados por Ana Lúcia Gazzola, ex-Secretária de Educação de Minas e uma das entrevistadas para a tese, para justificar a longevidade do Proalfa).

\*\*\*

As características de cada um dos programas, de acordo com os critérios selecionados para a análise, são apresentadas de modo sumarizado pelos três quadros a seguir, que dizem respeito, respectivamente, a Paebes Alfa, Proalfa e Spaece Alfa.

Quadro 8 - O Paebes Alfa segundo os critérios selecionados para análise

<b>Crítérios</b>	<b>Descrição para o programa</b>
1. Alcance da política	Ampla alcance, avaliando as redes municipais e estadual, além de facultar a participação a escolas particulares; a avaliação foi censitária, para todas as séries avaliadas, em todas as edições do programa; todos os alunos do ciclo de alfabetização nas redes públicas são, a princípio, avaliados.
2. Público-alvo	Prioritariamente, tomadores de decisão, como secretários de educação, técnicos das secretarias, superintendentes, diretores e professores; contudo, os resultados, a partir do nível da escola, são públicos, abertos a qualquer usuário.
3. Contexto de criação (agenda) e modelo de implementação	Reforma administrativa do estado, como foco na gestão por resultados; influência do Compromisso Todos Pela Educação e dos modelos de reforma de outros estados, entre eles, Minas Gerais e Ceará. Implementação top/down, conduzida pela secretaria de educação.
4. Desenho da avaliação	Avaliação censitária de 1º, 2º e 3º anos, em leitura, escrita e matemática
5. Divulgação e uso pedagógico dos resultados (natureza dos canais de divulgação e comunicação)	Coleção de revistas de divulgação própria para o Paebes Alfa; uma Revista Pedagógica específica para cada ano de escolaridade avaliado. Foco em gestores de rede, gestores escolares e professores. Oficinas anuais de apropriação de resultados com público designado pela secretaria de educação.
6. Concepção de alfabetização	Alfabetizar letrando.
7. Vínculo com outras políticas educacionais e geração da política ( <i>accountability</i> )	Com outro programa de avaliação (Paebes), com uma política de bonificação (Bônus Desempenho); terceira geração de avaliação.
8. Longevidade (duração e constância)	Longeva (10 edições entre 2008 e 2017) e contínua (sem interrupções).
9. Participação dos municípios	Por adesão; todos os custos são cobertos pelo estado, incluindo a avaliação das escolas particulares participantes.
10. Volume e transparência dos recursos aplicados em avaliação	No período analisado, os gastos com a avaliação representaram, em média, 0,55% do total de gastos da secretaria de educação.

Fonte: Elaboração pelo próprio autor.



Quadro 9 - O Proalfa segundo os critérios selecionados para análise

<b>Crítérios</b>	<b>Descrição para o programa</b>
1. Alcance da política	Alcance intermediário, avaliando as redes municipais (850 municípios com linha histórica) e estadual; formato censitário para o 3º ano do ensino fundamental durante todo o período e amostral para 2º e 4º anos até 2014; não há participação de escolas particulares.
2. Público-alvo	Prioritariamente, tomadores de decisão, como secretários de educação, técnicos das secretarias, superintendentes, diretores e professores; contudo, os resultados, a partir do nível da escola, são públicos, abertos a qualquer usuário.
3. Contexto de criação (agenda) e modelo de implementação	Implementação do ensino fundamental de 9 anos no estado (avaliar os efeitos da política); foi estabelecida como meta prioritária do estado a alfabetização das crianças até os 8 anos; resultados do Saeb apontavam baixo de desempenho em leitura para os alunos do 4º ano do ensino fundamental. Implementação top/down, conduzida pela secretaria de educação.
4. Desenho da avaliação	Avaliação censitária do 3º ano do ensino fundamental, somente em leitura até 2016 e, em 2017, em leitura, escrita e matemática; avaliação amostral do 2º e 4º anos entre 2007 e 2014, com avaliação censitária no 4º ano para os alunos de baixo desempenho.
5. Divulgação e uso dos resultados (natureza dos canais de divulgação e comunicação)	Até 2014: oficinas anuais de apropriação de resultados com público selecionado pela secretaria de educação; coleção própria de revistas para o Proalfa, com foco nos gestores das redes, gestores de escola (a partir de 2012) e professores.
6. Concepção de alfabetização	Alfabetizar letrando.
7. Vínculo com outras políticas educacionais e geração da política ( <i>accountability</i> )	Com o sistema de avaliação do estado (Simave); já esteve relacionado com a política de intervenção pedagógica de mais destaque no estado entre 2006 e 2014 (PIP) e com uma política de bonificação (até 2014); no período analisado, já foi de segunda e de terceira geração.
8. Longevidade (duração e constância)	Longeva (13 edições entre 2005 e 2017) e contínua (sem interrupções).
9. Participação dos municípios	Por adesão; todos os custos são cobertos pelo estado.
10. Volume e transparência dos recursos aplicados em avaliação	No período analisado, os gastos com a avaliação representaram, em média, 0,17% do total de gastos da secretaria de educação.

Fonte: Elaboração pelo próprio autor.

Quadro 10 - O Spaace Alfa segundo os critérios selecionados para análise

<b>Crítérios</b>	<b>Descrição para o programa</b>
1. Alcance da política	Alcance intermediário, avaliando as redes municipais (todos os 184 municípios do estado) e estadual; formato censitário para o 2º ano do ensino fundamental durante todo o período; não há participação de escolas particulares.
2. Público-alvo	Prioritariamente, tomadores de decisão, como secretários de educação, técnicos das secretarias, superintendentes, diretores e professores; contudo, os resultados, a partir do nível da escola, são públicos, abertos a qualquer usuário.
3. Contexto de criação (agenda) e modelo de implementação	Como ação articulada ao Paic, em resposta às altas taxas de analfabetismo no estado, divulgadas pelo relatório do Comitê Cearense para a Eliminação do Analfabetismo. Implementação top/down, conduzida pela secretaria de educação.
4. Desenho da avaliação	Avaliação censitária do 2º ano, em leitura.
5. Divulgação e uso dos resultados (natureza dos canais de divulgação e comunicação)	Coleção de revistas específicas voltadas para gestores da rede, gestores de escola e professores. Como regra: Revista Pedagógica referente ao Spaace Alfa como parte da coleção geral do Spaace. Oficinas de apropriação de resultados anuais com público designado pela secretaria de educação do estado.
6. Concepção de alfabetização	Alfabetizar letrando.
7. Vínculo com outras políticas educacionais e geração da política ( <i>accountability</i> )	Com outro programa de avaliação (Spaace), com a principal política educacional do estado (Paic) e com uma política de bonificação (Prêmio Escola Nota 10); terceira geração de avaliação.
8. Longevidade (duração e constância)	Longeva (11 edições entre 2007 e 2017) e contínua (sem interrupções).
9. Participação dos municípios	Por adesão, mas, os municípios são compelidos a participar pelo compromisso assumido no âmbito da política educacional e pela lei que estabelece a distribuição do ICMS; todos os custos da avaliação são cobertos pelo estado.

Critérios	Descrição para o programa
10. Volume e transparência dos recursos aplicados em avaliação	No período analisado, os gastos com a avaliação representaram, em média, 0,42% do total de gastos da secretaria de educação

Fonte: Elaboração pelo próprio autor.

#### 5.4 A VISÃO DOS ATORES DAS SECRETARIAS DE EDUCAÇÃO SOBRE A IMPORTÂNCIA DOS PROGRAMAS PARA A GESTÃO DAS REDES.

A avaliação educacional, conforme vimos no capítulo 2, tem como principal função fornecer informações para dar suporte à tomada de decisões e ao desenvolvimento de políticas educacionais. Encarnando essa função, a avaliação assume protagonismo para a gestão, seja das redes de ensino, seja das escolas. Diante disso, julgamos ser importante ouvir o que tomadores de decisão têm a dizer sobre o papel que Paebes Alfa, Proalfa e Spaece Alfa desempenham e desempenharam para a gestão educacional de seus estados. Todos os entrevistados fizeram parte das equipes das secretarias de educação em algum momento do período analisado, estando em contato com o programa de avaliação da alfabetização. Eles foram questionados sobre três aspectos: i) o papel desempenhado pelo programa no âmbito das ações da secretaria; ii) a contribuição do programa para a melhoria da qualidade da alfabetização no estado; e iii) os fatores que ajudam a explicar a longevidade do programa (o roteiro das entrevistas está disponibilizado no Anexo IV). Sobre este último aspecto, algumas das opiniões dos entrevistados já foram apresentadas em tópico anterior.

Evidentemente, nem todos os atores que estiveram à frente desses programas de avaliação foram aqui entrevistados. Um mapeamento mais aprofundado das opiniões dos tomadores de decisão sobre os programas exigiria uma pesquisa mais abrangente, com a oitiva de mais personagens que fizeram parte dessa história (por exemplo, selecionando atores de diferentes gestões e com diferentes funções nas secretarias ao longo do período analisado). Apesar disso, acreditamos que os atores entrevistados para a tese fornecem uma visão importante sobre os programas, no que diz respeito ao seu uso como instrumento de gestão.

Quando perguntado sobre a contribuição do programa para a melhoria da qualidade da alfabetização no estado, Alessio da Costa Lima<sup>194</sup> afirmou que o Spaece Alfa foi uma das “ações mais expressivas da nova política educacional”, que se instaurava no estado naquele momento (início em 2007), deixando claro o compromisso do Ceará com a alfabetização ao fim do 2º ano

<sup>194</sup> Ex-Orientador da Célula de Avaliação do Desempenho Acadêmico da Secretaria de Educação do Ceará.

do ensino fundamental. Mesmo não estando sob a responsabilidade da COPEM – Coordenadoria de Cooperação com os Municípios, que gerenciava o Paic em todas as suas vertentes, Alessio afirma que “o Spaece Alfa foi usado como termômetro” para a principal política educacional do estado (o programa de avaliação como instrumento de avaliação dos efeitos de outras políticas). O que contribuiu para esse protagonismo foi, de acordo com o entrevistado, o próprio desenho da avaliação trazida pelo Spaece Alfa: os resultados, divulgados em diferentes níveis de agregação, permitiram chegar às turmas e aos alunos, dando novo potencial ao sistema de avaliação pela possibilidade de diagnosticar dificuldades específicas de cada criança, através da combinação da divulgação dos resultados com base na TRI (proficiência média e distribuição dos alunos por padrões de desempenho) e na TCT (percentual de acerto por descritor). Dessa forma, os resultados se aproximaram dos professores, que identificavam as dificuldades e a aprendizagem de seus alunos de acordo com os erros e acertos por descritor. Segundo Alessio, o mapeamento oferecido pelos resultados de TCT propiciaram um diagnóstico mais pedagógico para os professores, tanto no que diz respeito às turmas quanto no que tange aos alunos. Isso pôde ser observado nos relatórios pedagógicos divulgados pelo programa, que trouxeram maior riqueza de informações para gestores e professores. De acordo com ele, a linguagem e a forma de comunicar os resultados trazidas pelo Spaece Alfa fizeram com que os professores utilizassem de fato os resultados da avaliação. Todas essas características são condizentes com uma política educacional que objetivava eliminar o analfabetismo no estado. Para isso, o acompanhamento de cada aluno era importante. Para Alessio, o Spaece Alfa inaugurou um novo ciclo de avaliação no estado, com consequências, inclusive, para o Spaece (as mudanças trazidas pelo Spaece Alfa foram levadas para o Spaece).

O entrevistado atribui a, pelo menos, cinco conjuntos de fatores a longevidade do programa: i) a importância dada pelos gestores, do governo do estado aos atores da secretaria de educação, à política educacional proposta e ao “uso da avaliação educacional como estratégia efetiva de acompanhamento, monitoramento e intervenção” – a política de avaliação percebida como integrante e fundamental da política educacional do estado como um todo; ii) a nova dimensão trazida pelo Spaece Alfa, ampliando o uso da avaliação em larga escala no estado às ações do professor, ajudou a consolidar a cultura de avaliação no estado do Ceará; os resultados divulgados de forma mais palatável e didática (o uso de cores específicas para cada padrão de desempenho, por exemplo, tornou-se parte da linguagem e do discurso sobre avaliação e aprendizagem no estado) favoreceu o entendimento de todos os atores educacionais acerca de seu significado (incluindo os prefeitos, importantes atores nesse processo, tendo em vista que

as redes municipais respondiam pela maioria dos alunos no ciclo de alfabetização); iii) o uso político dos dados da avaliação, que começaram a ocupar espaços na mídia e a aparecer em grandes eventos políticos – os municípios do Ceará começaram a ser “pintados” com as cores do padrão de desempenho que representava a média de proficiência da rede no Spaece Alfa, e um mapa do estado com essas cores começou a ser usado como instrumento de monitoramento da alfabetização no estado ano a ano -, levou a um maior controle social sobre a educação, fazendo com que a sociedade monitorasse a alfabetização mais de perto através desses resultados; de acordo com Alessio, a cada ano, as expectativas com os resultados dos municípios aumentavam, levando a uma pressão sobre os prefeitos para que seus municípios ficassem “verdes” (o verde escuro representa o padrão de desempenho Desejável); iv) a política de premiação adotada pelo estado também ajudou a dar visibilidade e importância aos resultados do Spaece Alfa (premiação financeira às 150 escolas com melhores resultados – o Prêmio Escola Nota 10); além disso, o “apadrinhamento” de uma escola apoiada por uma premiada contribuiu para a difusão de boas práticas, pedagógicas e de gestão, entre as escolas de todo o estado (isso para o 5º e 9º anos); além disso, houve alteração da lei estadual, que passou a condicionar um percentual dos recursos da redistribuição do ICMS aos resultados da avaliação nos municípios, o que levou os prefeitos a priorizarem de fato a alfabetização, incluindo, entre as ações, uma escolha mais cuidadosa de seus secretários de educação, com um perfil mais técnico e menos político; e v) a continuidade da política educacional no estado, um dos fatores mais relevantes a se observar; mesmo com as mudanças de gestão, a política educacional cearense esteve alicerçada sobre os mesmos pontos, sendo o combate ao analfabetismo um dos principais deles; com isso, a política de avaliação da alfabetização não apenas permaneceu, mas se fortaleceu no estado.

Conforme relata Márcia Oliveira Cavalcante Campos<sup>195</sup>, a avaliação teve um papel importante no Ceará: o de resgatar o compromisso prioritário da escola com o ensino e a aprendizagem. O cuidado das escolas com a aprendizagem de seus alunos foi avivado pela forma como a avaliação passou a ser usada, contribuindo para gerar e sendo gerado por uma cultura de avaliação no estado. A seriedade com que os resultados passaram a ser considerados pelas redes e escolas demonstra o grau de enraizamento dessa cultura de avaliação. Fazendo jus a sua função primordial, a avaliação passou a dar fundamento e base para as decisões sobre a política educacional do estado.

---

<sup>195</sup> Uma das idealizadoras do Paic, foi coordenadora da Cooperação com os municípios na Seduc Ceará.

Para ela, o Spaece como um todo é importante para a gestão da educação no estado, mas o Spaece Alfa, em específico, assumiu um protagonismo em virtude da cooperação com os municípios. Ele foi uma das peças na engrenagem maior da política educacional do estado, estando associado com a formação continuada dos professores, com a produção de material didático e com a política de premiação no estado, por exemplo. Essa integração da avaliação com a política de alfabetização foi fundamental: os resultados não vinha soltos, mas conjugados com uma série de ações que caminhavam na mesma direção e com os mesmos objetivos. Segundo Márcia Campos, um dos trunfos do Spaece Alfa foi a forma didática como suas informações chegaram às escolas e aos municípios. Isso pôde ser percebido na própria modificação da linguagem: os termos descritor, matriz, competências passaram a ser usados por professores e diretores (esse didatismo também foi ressaltado por Alessio da Costa Lima). Os resultados foram sendo incorporados às práticas de trabalho das escolas e das redes. De acordo com ela, o melhor trabalho com os resultados da avaliação no Ceará foi no ciclo de alfabetização. Posteriormente, o exemplo desse trabalho foi levado também para o uso dos resultados do ensino médio.

Na visão de Márcia, a educação é um dos principais focos do governo do Ceará e isso pode ser percebido nos eventos anuais para a entrega do Prêmio Escola Nota 10. Há uma mobilização de todo o estado em torno dessa entrega. Os prefeitos e secretários de educação se mobilizam para comparecer ao momento da premiação. A avaliação é indispensável nesse processo. Assim como Alessio, ela chama atenção para a importância assumida pela compreensão do que significa os padrões de desempenho do Spaece Alfa no estado: secretários de educação e escolas desejam estar no “verde escuro”, o padrão que representa o nível desejável na alfabetização. Isso passou a ter um significado muito grande no estado.

A sucessão de governos sem alteração substantiva na política de educação, na leitura da entrevistada, ajuda a explicar a longevidade do Spaece Alfa. Ele foi mantido como uma ferramenta central para a gestão centrada nos resultados levada a cabo pelo estado. Segundo ela, ficou evidente, para todo o estado, que a avaliação é um instrumento importante para ajudar a compreender os problemas enfrentados pela educação, abrindo espaço para intervenções destinadas a solucioná-los.

Na visão de Maurício Holanda Maia<sup>196</sup>, o Paic é o grande destaque da política educacional do estado do Ceará, tanto por sua longevidade (foi criado em 2007), quanto pelos

---

<sup>196</sup> Gestor na Secretaria de Educação do Estado do Ceará (Seduc-CE) entre os anos 2007 e 2015 (ex-secretário de educação do estado do Ceará entre 2014 e 2016).

resultados que alcançou. E é justamente neste ponto que fica evidente a importância do Spaece Alfa: ele é o instrumento de monitoramento da principal política educacional do estado. O Spaece Alfa faz jus a uma das funções esperadas da política de avaliação: que ela seja capaz de avaliar os efeitos das demais políticas educacionais. Se o Paic é, conforme afirma Maurício, o principal fator que explica a expressiva melhoria dos resultados não só da alfabetização, mas também do 5º ano, o Spaece Alfa é o instrumento capaz de atestar e monitorar essa melhoria. A longevidade e importância do Paic ajudam a explicar a longevidade e a importância do Spaece Alfa.

Para Renata da Silva Marinho<sup>197</sup>, a importância do Paebes Alfa se revela, primordialmente, nos diagnósticos que é capaz de fornecer. Em um contexto de gestão por resultados, a avaliação torna-se central. A cooperação com os municípios é outro fator importante: a maioria dos alunos avaliados pelo Paebes Alfa pertence às redes municipais e o estado tem o compromisso de cooperar com elas no que diz respeito à melhoria da aprendizagem. De acordo com Renata, a riqueza e o detalhamento dos resultados produzidos pelo Paebes Alfa, por possibilitarem análises minuciosas e um monitoramento contínuo da aprendizagem, ajudam a explicar a centralidade que o programa assumiu para a gestão.

Em sua visão, o programa contribuiu significativamente para a melhoria da qualidade da alfabetização no estado, já que deu suporte a um conjunto de intervenções pedagógicas planejadas a partir de seus resultados. A avaliação passou a integrar o trabalho das escolas, que começaram, de fato, a usar seus resultados como base para o planejamento de suas ações.

De acordo com Andressa Buss Rocha<sup>198</sup>, o Paebes Alfa tornou-se um programa estratégico no Espírito Santo. O alcance das metas dos planos nacional e estadual de educação, referentes à alfabetização na idade recomendada, passa pelo monitoramento oferecido pela avaliação. Além disso, seus resultados passaram a dialogar diretamente com a atuação dos professores em sala de aula, fornecendo evidências para o planejamento. No âmbito da secretaria de educação, os resultados do Paebes Alfa passaram a orientar ações para a formação de professores, por exemplo.

Para Andressa, os resultados do programa de avaliação fizeram com que o trabalho da secretaria de educação ficasse focado na aprendizagem dos alunos, levando à elaboração de estratégias para a correção de rotas, o que contribuiu decisivamente para a melhoria da qualidade da alfabetização no estado. As sucessivas gestões estaduais, desde a criação do

---

<sup>197</sup> Ex-Subgerente de avaliação educacional da Seduc do Espírito Santo.

<sup>198</sup> Ex-Assessora de Planejamento e Gestão Estratégica da Seduc do Espírito Santo.

programa, concederam à avaliação o mesmo papel estratégico, de fornecer evidências para a tomada de decisões, o que explica a longevidade experimentada pelo Paebs Alfa.

Segundo Raquel Elizabete Santos<sup>199</sup>, o maior desafio da educação em Minas Gerais era transformar a rede pública em um sistema de alto desempenho educacional. Para tanto, o primeiro passo era resolver os problemas de aprendizagem os alunos no ciclo de alfabetização. Minas partiu de um diagnóstico preocupante dos alunos em relação à leitura. O Proalfa surgiu como uma ação fundamental para ajudar a reverter esse quadro. Em sua leitura, o programa cresceu em importância com a implementação do PIP. Os resultados do Proalfa foram a base de implementação do PIP e, posteriormente, o programa de avaliação foi potencializado pelas ações do programa de intervenção. Além do acompanhamento individual dos alunos, o PIP buscava desenvolver uma cultura de avaliação no estado, de tomada de decisão com base em evidências. A exigência do PIP para uso dos resultados da avaliação como norte para o planejamento fez com que o Proalfa assumisse maior protagonismo.

De acordo com Raquel, o Proalfa desempenhou um papel primordial para a melhoria da qualidade da alfabetização em Minas. Sua associação com o PIP é responsável por isso. Até 2014, o PIP foi a principal política educacional voltada para a melhoria da aprendizagem no estado. O programa de avaliação ajudou a definir as regiões e escolas prioritárias no estado, que apresentavam maiores dificuldades para o alcance de suas metas (em especial, a de que toda criança deveria estar alfabetizada até os oito anos de idade). No entanto, o maior desafio do PIP estava em concretizar sua proposta de chegar à sala de aula. Para tanto, o programa se baseava na ideia de que os resultados educacionais só podem ser efetivamente alterados tendo como base uma intervenção direta no trabalho realizado em sala de aula. Na leitura de Raquel, foi isso que o Proalfa possibilitou: “atendeu à necessidade crescente por informações mais precisas e seguras, que orientassem a gestão do sistema educacional e o trabalho pedagógico das escolas”. Ainda segundo ela, a longevidade do programa está associada à compreensão de sua importância por parte dos atores educacionais. E isso foi obtido através da forma como o programa de avaliação foi implementado, de modo a reduzir a resistência dos atores em relação à avaliação.

No entendimento de Maria das Graças Pedrosa Bittencourt<sup>200</sup>, não é possível desvincular uma análise do Proalfa de sua relação com o PIP. Os resultados do Proalfa foram o ponto de partida para a criação do programa de intervenção. De acordo com ela, o Proalfa “foi

---

<sup>199</sup> Ex-Subsecretária de desenvolvimento da educação básica do estado de Minas Gerais.

<sup>200</sup> Ex-Gerente do Programa de Intervenção Pedagógica - PIP para o ensino fundamental.

um divisor de águas na história das políticas públicas educacionais em Minas Gerais no período de 2006 a 2014”. Ele permitiu garantir aos alunos o direito de serem alfabetizados no tempo certo. Os resultados de 2006 do Proalfa apontaram que cerca de 51% do total de alunos do 3º ano do ensino fundamental não haviam consolidado habilidades básicas de leitura, não estando plenamente alfabetizados. Esses resultados levaram a Seduc de Minas a estabelecer a alfabetização como prioridade, visto que é muito difícil corrigir problemas de aprendizagem relacionados à alfabetização depois que o aluno já terminou o ciclo. As informações produzidas pelo Proalfa, portanto, foram responsáveis pela criação de um dos programas “mais exitosos” de intervenção pedagógica em Minas Gerais. Dando base para as ações do PIP, o Proalfa contribuiu decisivamente, na visão da entrevistada, para a melhoria da qualidade da alfabetização observada no estado. Ainda, a importância que o programa de avaliação teve para a melhoria dos resultados é o grande motivo pelo qual ele se tornou longo.

Os resultados do Proalfa de 2006 também são apontados por Ana Lúcia Gazzola<sup>201</sup> como um marco na gestão da alfabetização em Minas. Eles deram base à decisão da secretaria de educação do estado de concentrar seus esforços no ciclo de alfabetização. Essa prioridade foi encarnada no PIP, que, na visão de Ana Lúcia, não poderia ter existido sem o Proalfa. Avaliação e intervenção pedagógica se associaram de maneira muito estreita na política educacional de Minas a partir daquele momento, já que as ações do PIP se assentaram, fundamentalmente, nos diagnósticos produzidos pelo Proalfa. Ele se tornou importante por ter sido capaz de encarnar a proposta de toda avaliação, qual seja, dar base, por meio de seus diagnósticos, a uma intervenção ajustada às necessidades identificadas pelos instrumentos avaliativos. Por não existir de forma isolada, mas vinculado a políticas de intervenção, o Proalfa se fortaleceu.

Ana Lúcia afirma não ter dúvidas sobre o papel desempenhado pelo Proalfa para a melhoria da qualidade da alfabetização no estado. Ele integrou, de forma essencial, uma política educacional mais ampla, destinada, justamente, a promover essa melhoria, revelada nos resultados do IDEB – Anos Iniciais, que apontou Minas como o estado com os melhores resultados para a etapa. Os diagnósticos detalhados trazidos pelo programa, com resultados divulgados até o nível do aluno, foram essenciais para apoiar uma política de intervenção como o PIP, que tinha como objetivo, justamente, de chegar à sala de aula. A longevidade do programa, segundo ela, repousa, em primeiro lugar, na necessidade da gestão se basear em evidências para a tomada de decisões.

---

<sup>201</sup> Secretária de educação do estado de Minas Gerais entre 2011 e 2014.



Para Sueli Pires<sup>202</sup>, uma das características marcantes do Proalfa, que ajudam a explicar sua importância, foi a cooperação entre estado e municípios para a efetivação de uma avaliação da alfabetização em Minas. Soma-se a isso o fato das escolas terem se comprometido com o uso dos resultados da avaliação, compreendendo sua importância para a política educacional do estado como um todo. Esse é um dos fatores que ajudam a entender a consolidação da cultura de avaliação no ciclo de alfabetização em Minas, um dos fatores que levaram à longevidade do programa. Além disso, o vínculo com outros programas, como os demais programas de avaliação no âmbito do Simave e o PIP, contribuem para explicar a importância assumida pelo Proalfa, que contribuiu, em sua visão, “de maneira decisiva e direta para os avanços da qualidade de alfabetização em Minas Gerais”.

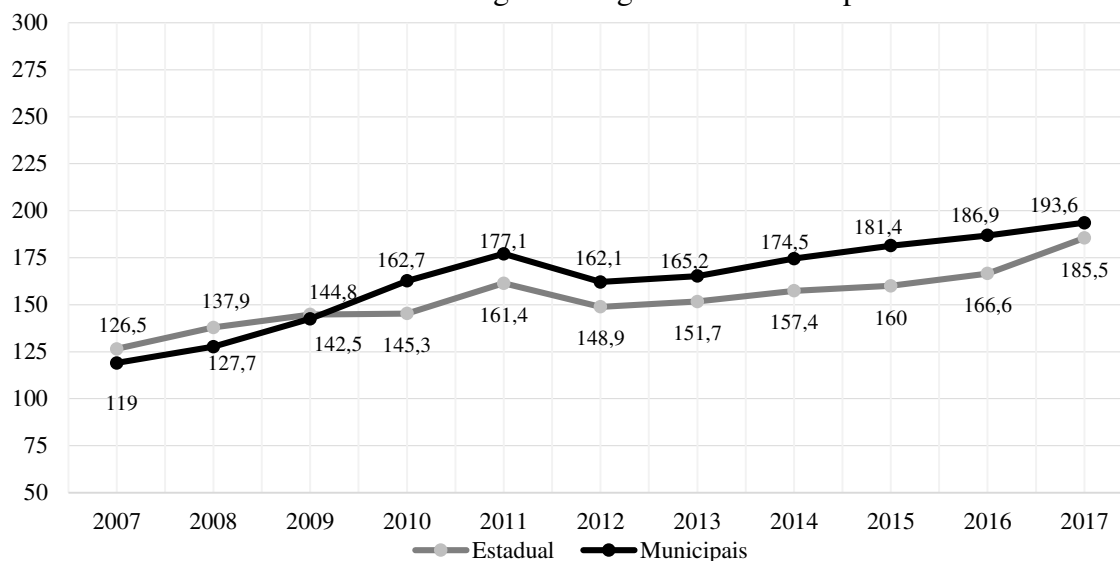
Como pode ser observado, a opinião dos atores que já estiveram nas secretarias de educação dos três estados convergem em uma série de pontos. Todos reconheceram a importância que os programas de avaliação da alfabetização tiveram para a gestão das redes, para a tomada de decisão com base em evidências e para a melhoria dos resultados de alfabetização. Em Minas, o Proalfa foi o ponto de partida para a criação de um dos programas educacionais mais relevantes do período analisado. No Ceará, o Spaece Alfa surge, como parte de outras ações conjuntas, a partir dos resultados de uma avaliação (Saeb) e para monitorar os efeitos do principal programa da política educacional do estado (Paic). No Espírito Santo, o Paebs Alfa é entendido como um programa estratégico. Todos os entrevistados reconhecem que a longevidade dos programas está associada à sua importância para a gestão. Em todos eles, o uso dos resultados pelas escolas e redes de ensino foi apontado como fator determinante para que o programa de avaliação fosse central e duradouro, exercendo papel determinante para a melhoria dos resultados na alfabetização.

A análise da proficiência média das redes ao longo do tempo, para os três programas, ajuda a entender o que esses atores querem dizer quando se referem à melhoria na alfabetização. No Ceará, por exemplo, observando a série histórica, as proficiências médias experimentaram aumentos gradativos, tanto na rede estadual quanto nas redes municipais. Como podemos notar a partir da leitura do gráfico a seguir, em 2007, primeira edição do programa, a proficiência média da rede estadual era 126,5, chegando a 185,5 em 2017 (de suficiente a desejável). Nas redes municipais, a melhoria foi ainda mais expressiva: de 119, em 2007, a 193,6, em 2017 (de intermediário a desejável).

---

<sup>202</sup> Secretária adjunta de educação do estado de Minas, entre 2012 e 2014.

Gráfico 8 - Proficiência média – Língua Portuguesa – 2º EF – Spaece Alfa –Linha histórica

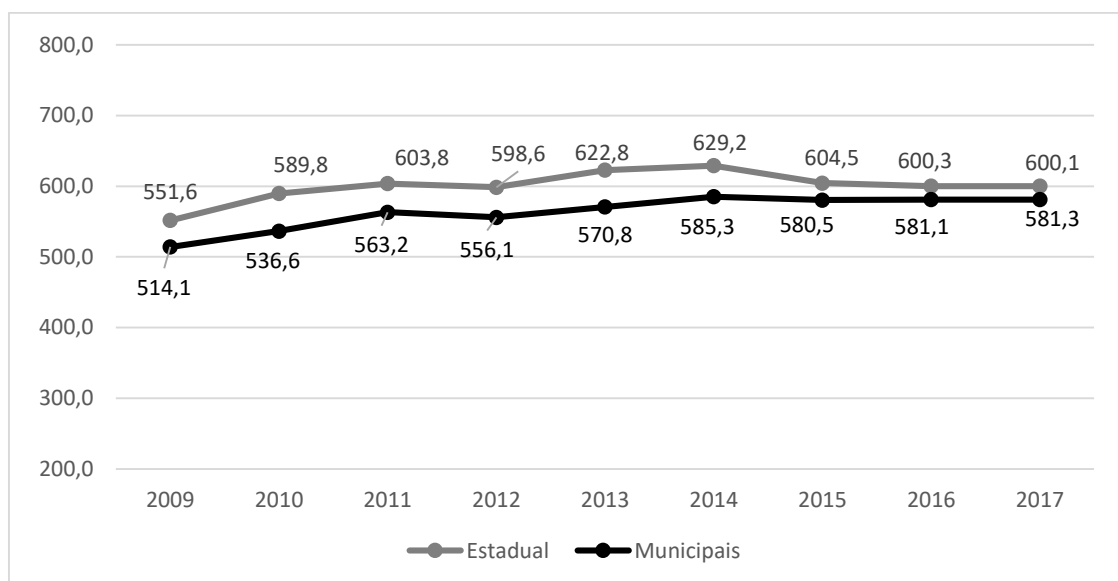


Fonte: Spaece Alfa. Planilha de resultados. Tabulação própria.

Em Minas Gerais, como pode ser observado a partir da leitura do gráfico a seguir, os resultados das redes estadual e municipais, para língua portuguesa, no 3º ano do ensino fundamental<sup>203</sup>, também apresentaram melhorias ao longo do período analisado, com destaque para as redes municipais, que passaram de 514,1, em 2009, e 581,3, em 2017 (seguindo os novos padrões de 2015, de intermediário a recomendado. Na rede estadual, o avanço foi em menor proporção, de 551,6, em 2009, para 600,1, em 2017, variação dentro do padrão recomendado. É interessante observar que as melhorias, para as duas redes, foram crescentes até o ano de 2014, quando houve uma queda nos resultados. Isso é condizente com tudo o que foi dito na análise dos critérios para o Proalfa. A partir de 2015, a avaliação, segundo os critérios aqui adotados para análise, perdeu a centralidade que havia tido até então e programas voltados à intervenção pedagógica para a melhoria dos resultados foram cancelados.

<sup>203</sup> Matemática foi avaliada apenas em 2017, de modo que não possui linha histórica. As avaliações de 2º e 4º anos foram encerradas em 2014 e eram amostrais. Por conta disso, o foco da divulgação dos resultados do Proalfa sempre foi o 3º ano do ensino fundamental, avaliado de modo censitário.

Gráfico 9 - Proficiência média – Língua Portuguesa – 3º EF – Proalfa –Linha histórica



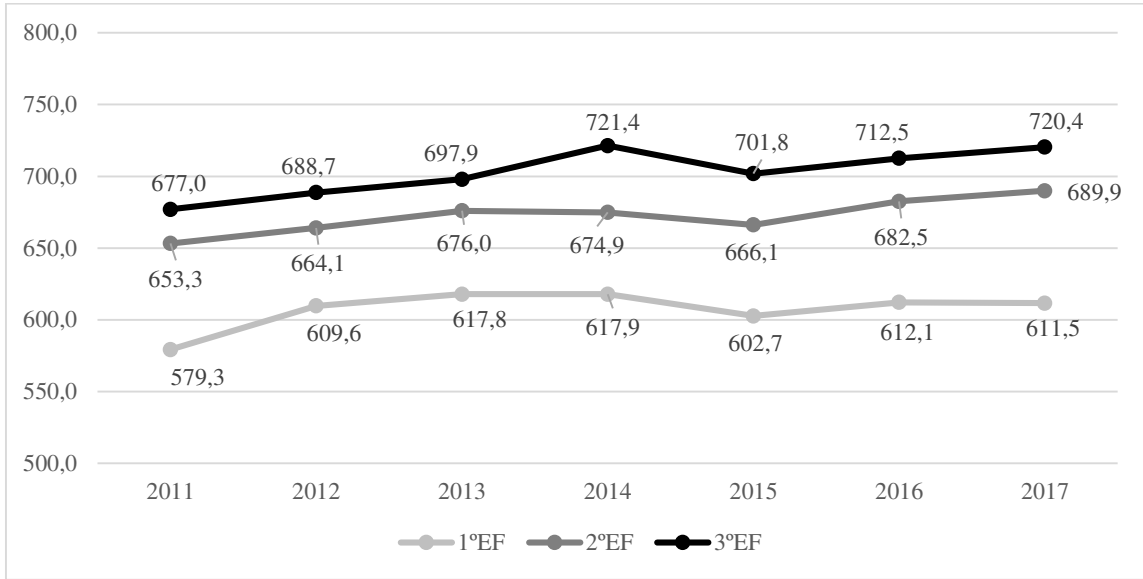
Fonte: Proalfa. Planilha de resultados. Tabulação própria.

Os gráficos a seguir apresentam os resultados<sup>204</sup> da linha histórica para o Paebes Alfa. Na planilha oficial de resultados divulgada em 2017, a linha tem início em 2011. Para leitura, os resultados apresentaram melhorias para todos os anos de escolaridade avaliados, com destaque para o 3º ano do ensino fundamental (na rede estadual, de 677, em 2011, a 720,4, em 2017 – uma variação dentro do padrão proficiente). Nas redes municipais, o movimento é o mesmo: a proficiência média do 3º ano passa de 659,3, em 2011, para 699 em 2017 (variação dentro do padrão proficiente).

Os resultados de escrita seguem a mesma direção que os de leitura. Todas as séries avaliadas apresentam melhorias nos resultados para o período analisado e elas são maiores no 3º ano do ensino fundamental: em relação à rede estadual, a proficiência média passa de 663,9, em 2011, para 714,7, em 2017 (variação no padrão proficiente); a proficiência das redes municipais sai de 657,9, em 2011, para 700,3, em 2017 (variação no padrão suficiente).

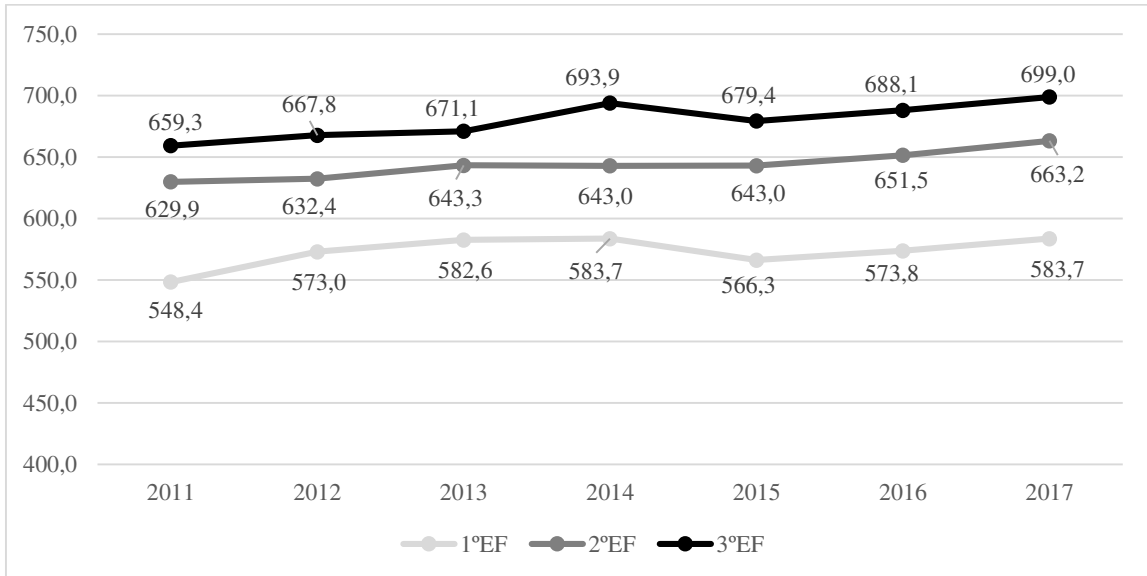
<sup>204</sup> Foram considerados os resultados das avaliações somativas, realizadas ao fim de cada período letivo, chamadas, no âmbito do Paebes Alfa, de 2ª onda. Até 2015, o programa realizou avaliações diagnósticas no começo do período letivo (1ª onda).

Gráfico 10 - Proficiência média – Leitura – 1º, 2º e 3º EF – Paebes Alfa – Rede estadual - Linha histórica



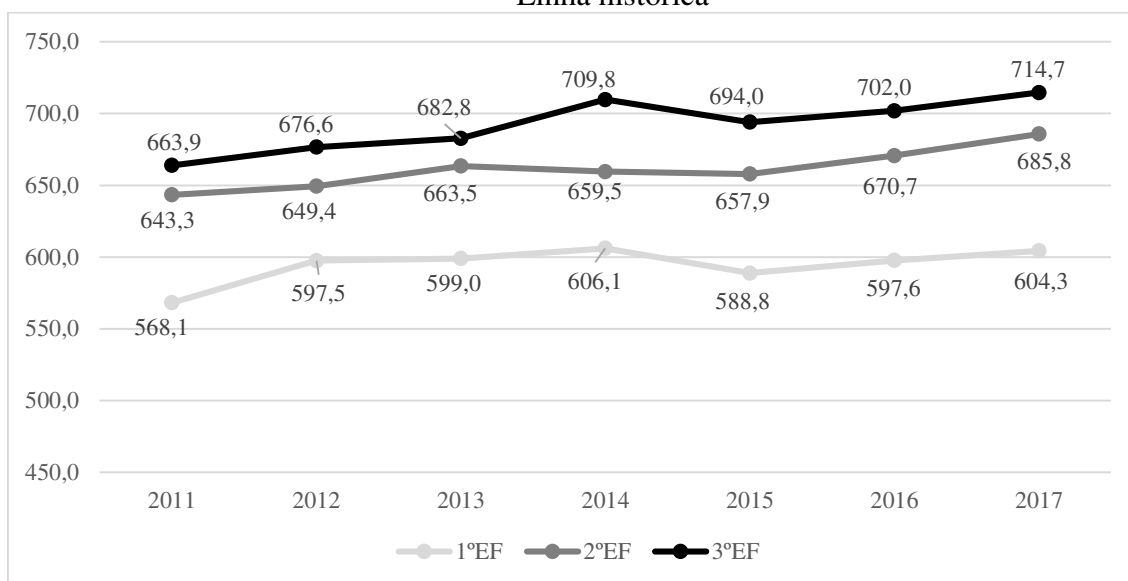
Fonte: Paebes Alfa. Planilha de resultados. Tabulação própria.

Gráfico 11 - Proficiência média – Leitura – 1º, 2º e 3º EF – Paebes Alfa – Redes municipais - Linha histórica



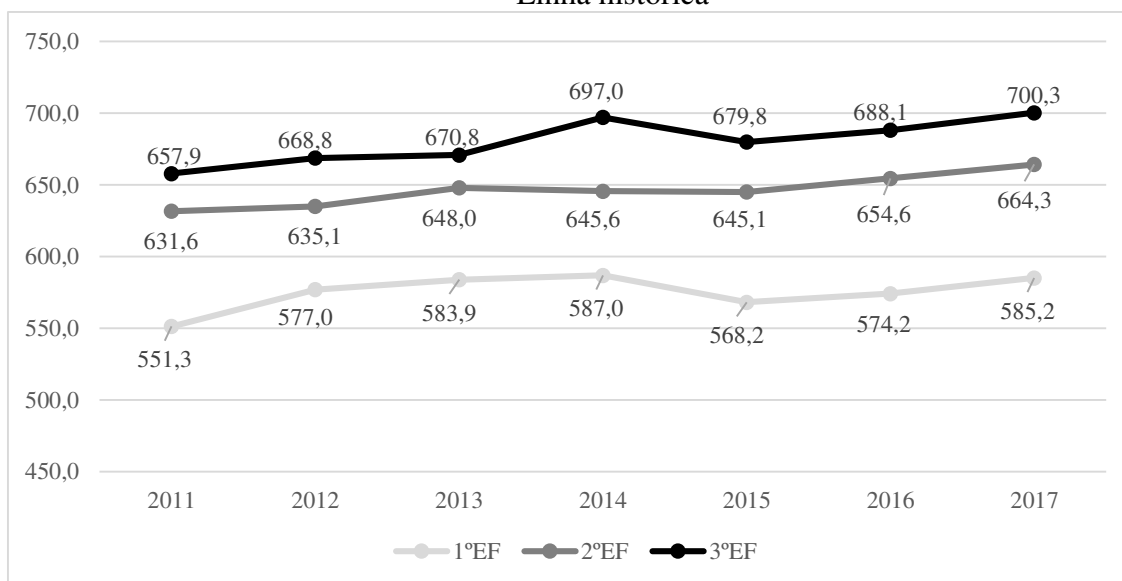
Fonte: Paebes Alfa. Planilha de resultados. Tabulação própria.

Gráfico 12 - Proficiência média – Escrita – 1º, 2º e 3º EF – Paebes Alfa – Rede estadual -  
Linha histórica



Fonte: Paebes Alfa. Planilha de resultados. Tabulação própria.

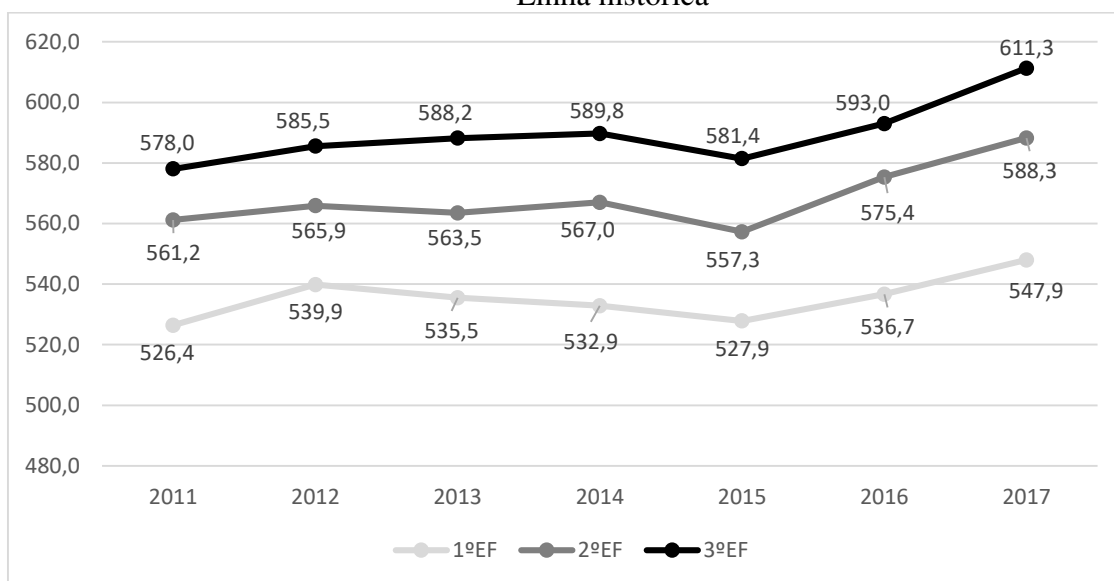
Gráfico 13 - Proficiência média – Escrita – 1º, 2º e 3º EF – Paebes Alfa – Redes municipais -  
Linha histórica



Fonte: Paebes Alfa. Planilha de resultados. Tabulação própria.

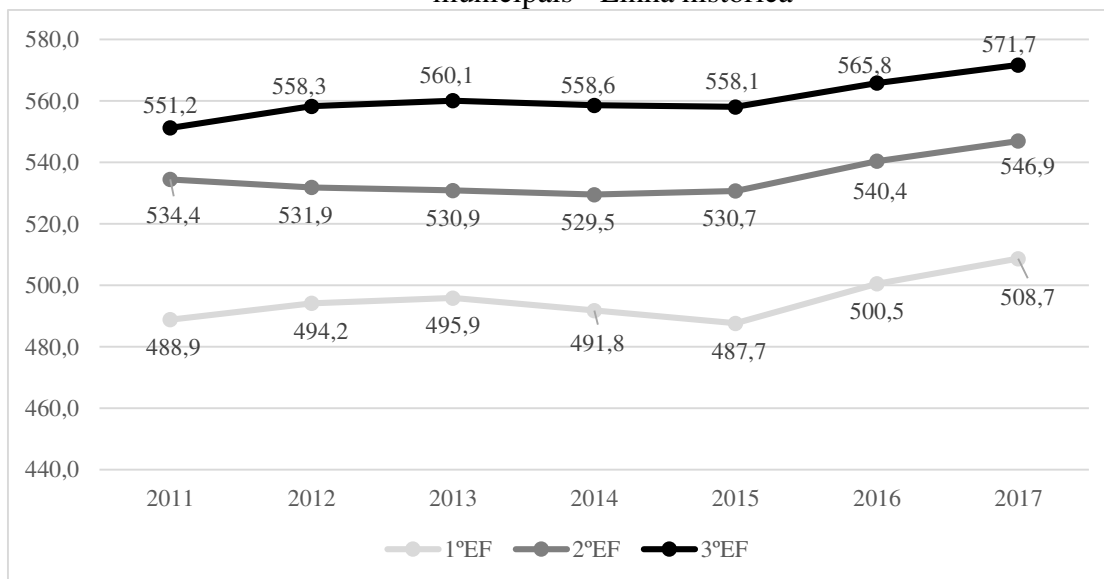
Para matemática, termos absolutos, a melhoria dos resultados é mais tímida do que aquela observada para língua portuguesa. Mesmo assim, o movimento observado tem a mesma direção: ao longo do período analisado, todas as séries melhoram suas médias de proficiência, com destaque para o 3º ano do ensino fundamental – na rede estadual, a proficiência média passa de 578, em 2011, para 611,3, em 2017 (variação no padrão proficiente); nas redes municipais, a proficiência sai de 551,2 e chega a 571,7, entre 2011 e 2017 (variação no padrão proficiente).

Gráfico 14 - Proficiência média – Matemática – 1º, 2º e 3º EF – Paebes Alfa – Rede estadual - Linha histórica



Fonte: Paebes Alfa. Planilha de resultados. Tabulação própria.

Gráfico 15 - Proficiência média – Matemática – 1º, 2º e 3º EF – Paebes Alfa – Redes municipais - Linha histórica



Fonte: Paebes Alfa. Planilha de resultados. Tabulação própria.

Esse aumento na proficiência média no ciclo de alfabetização é o que alimenta, no âmbito dos três programas, a visão otimista adotada pelos atores que fizeram parte das secretarias de educação aqui entrevistados, em relação aos efeitos da política educacional de seus estados e à importância da avaliação da alfabetização para monitorar os efeitos de processo.

## 6 CONSIDERAÇÕES FINAIS

O ponto de partida desta tese, que justifica a existência do capítulo 1 tal como foi escrito, é o de que a avaliação educacional em larga escala é uma política pública. Essa consideração, a princípio não problemática, esconde uma questão. Afinal, a avaliação educacional ocupa um lugar *sui generis* entre as políticas públicas educacionais, em virtude de sua dupla dimensão. De um lado, pode ser pensada como política pública em si mesma, dotada de relativa autonomia, pelo menos do ponto de vista analítico (como objeto de estudo para o pesquisador de políticas públicas), composta por diferentes fases (agenda, implementação, resultados etc), como qualquer outra política. De outro, um instrumento usado como avaliação de outras políticas públicas educacionais, tendo como pressuposto a ideia de que toda política educacional deveria produzir efeitos na aprendizagem, mensurada, por sua vez, por um instrumento de avaliação. Assim, a avaliação educacional seria uma política pública destinada a avaliar o efeito de outras políticas sobre a aprendizagem. Nesse caso, o instrumento seria usado como a fase de avaliação de outras políticas. Trata-se de uma discussão não resolvida pela tese, mas a apresentação dessa questão pode abrir espaço para um estudo mais detido acerca da natureza da avaliação educacional como política pública, o que poderia levar a uma discussão mais ampla sobre o próprio conceito de política pública.

De todo modo, para os fins da tese, entender a avaliação como política pública significa compreendê-la como um fenômeno que, prolongado no tempo, permite sua análise a partir de diferentes dimensões ou etapas, desde a definição da agenda até seus resultados. Uma característica do campo de pesquisa das políticas públicas é a busca pela construção de tipologias que permitam a categorização e classificação das políticas segundo critérios selecionados, o que interessava diretamente à proposta de analisar programas estaduais de avaliação da alfabetização. No entanto, as propostas tipológicas na literatura sobre políticas públicas eram demasiadamente genéricas e não ofereciam grandes vantagens para analisar a especificidade de programas de avaliação da alfabetização. O caminho escolhido, então, foi a propositura de critérios de análise mais ajustados às características desses programas.

Esses critérios, conforme apresentados no capítulo 5, não são taxativos. O que espero deles é que sejam úteis para análises futuras. Contudo, a depender do enfoque da análise, eles podem ser reduzidos ou ampliados. Sua combinação pode gerar a criação de tipos de programas de avaliação da alfabetização, assim como cada um dos critérios, sozinho, é capaz de gerar diferentes tipos (por exemplo, quais são os tipos de programa no que diz respeito ao critério

alcance da política?). Um esforço dessa natureza não foi empreendido pela tese, que se ateve à propor os critérios para analisar os programas selecionados.

Mas antes de chegar a essa análise, foi necessário apresentar as características gerais das políticas educacionais, das políticas de alfabetização e das políticas de avaliação em larga escala. Essa contextualização teve como objetivo identificar a posição ocupada pelas políticas de avaliação da alfabetização no cenário geral das políticas educacionais, mas também nos cenários específicos das políticas de alfabetização e de avaliação. A hipótese em que me baseava, muito em virtude de minha experiência profissional no CAEd/UFJF por quase uma década, desde a escolha do tema da pesquisa, era de que havia um descompasso entre a importância assumida pela alfabetização no cenário das políticas educacionais e a falta de interesse da pesquisa educacional pela avaliação da alfabetização. Essa hipótese precisava ser comprovada e, ao mesmo tempo, uma possível explicação foi levantada: tal falta de interesse, uma vez identificada, poderia estar relacionada, entre outras coisas, à resistência experimentada pela avaliação em larga escala. Por isso, a escrita do capítulo 3.

A busca pela confirmação da existência do referido descompasso ocupou o capítulo 3. Como um tema tão importante para as políticas educacionais como a alfabetização é acompanhado de tão pouco interesse por sua avaliação? A pesquisa educacional praticamente ignora o tema. À resistência à avaliação somam-se a falta de uma avaliação nacional consolidada para o ciclo (já que a experiência da ANA é recente e oscilante) e a ausência do apontamento sobre qual caminho seguir para realizar uma pesquisa dessa natureza. Parte da literatura sobre políticas públicas é crítica em relação a pesquisas que descrevem programas, mas, no que diz respeito à avaliação da alfabetização, faltam estudos de toda natureza, incluindo aqueles com objetivos descritivos.

O capítulo 5 buscou lidar com essa lacuna, sem a pretensão de preenchê-la. A ideia é oferecer uma proposta de pesquisa para analisar programas de avaliação da alfabetização, em especial, pela definição de critérios para a análise. A descrição dos programas faz parte desse processo, que também envolve a análise das características dos programas a partir dos critérios definidos. A análise assume função comparativa, na medida em que semelhanças e diferenças entre os programas são identificadas. No corpo do capítulo 5, as características, tendo em vista cada um dos critérios propostos, foram apresentadas para cada um dos programas. O quadro 11, a seguir, reúne, de modo sumarizado, tais características, abrindo espaço para a comparação entre os programas, realizada na sequência.



Quadro 11 - Comparação entre os três programas segundo os critérios selecionados<sup>205</sup>

<b>Crítérios</b>	<b>Paebes Alfa</b>	<b>Proalfa</b>	<b>Spaace Alfa</b>
1. Alcance da política	Amplio	Intermediário	Intermediário
2. Público-alvo	Tomadores de decisão	Tomadores de decisão	Tomadores de decisão
3. Contexto de criação (agenda) e modelo de implementação	Reforma administrativa do estado - top/down	Ensino fundamental de 9 anos; baixo desempenho no Saeb - top/down	Paic - top/down
4. Desenho da avaliação	Todo o ciclo, leitura, escrita e matemática	3º ano, leitura	2º ano, leitura
5. Divulgação e uso pedagógico dos resultados (natureza dos canais de divulgação e comunicação)	Portal aberto; revistas para as escolas; oficinas com público selecionado	Portal aberto; revistas para as escolas; oficinas com público selecionado	Portal aberto; revistas para as escolas; oficinas com público selecionado
6. Concepção de alfabetização	Alfabetizar letrando	Alfabetizar letrando	Alfabetizar letrando
7. Vínculo com outras políticas educacionais e geração da política ( <i>accountability</i> )	Paebes; Bônus Desempenho - terceira geração	Simave, PIP (2006-2014), bonificação (até 2014) - desde 2014, segunda geração	Paic, Spaace, Prêmio Escola Nota 10 - terceira geração
8. Longevidade (duração e constância)	Longeva e contínua	Longeva e contínua	Longeva e contínua
9. Participação dos municípios	Por adesão	Por adesão	Por adesão, mas municípios compelidos a participar
10. Volume e transparência dos recursos aplicados em avaliação	Média de 0,55% do total de gastos da secretaria	Média de 0,17% do total de gastos da secretaria	Média de 0,42% do total de gastos da secretaria

Fonte: Elaboração pelo próprio autor.

Pela leitura do quadro, é possível perceber que os três programas apresentam muitas semelhanças. Eles se baseiam nas mesma concepção de alfabetização, o alfabetizar letrando, que se manifesta, em suas matrizes de referência, na previsão de habilidades relacionadas a elementos técnicos da língua, como o reconhecimento da palavra como unidade gráfica e a aquisição da consciência fonológica. Possuem um sistema de divulgação de resultados muito semelhante, com portais próprios por meio dos quais os resultados das escolas são abertos a qualquer usuário. Além disso, destinam revistas específicas para diferentes atores educacionais (professores, diretores, agentes das secretarias de educação), com resultados em diferentes níveis de agregação, e oficinas anuais para apropriação de resultados, com público definido pelas

<sup>205</sup> Devo a confecção deste quadro-resumo à sugestão da professora Beatriz de Basto Teixeira, a quem agradeço.

secretarias (em geral, além dos técnicos das secretárias, gestores escolares). Isso faz com que o público-alvo desses programas seja, em regra, o mesmo: prioritariamente, tomadores de decisões. Todas essas características estão relacionadas ao fato dos três programas estabelecerem parcerias com a mesma instituição, o CAEd/UFJF, o que ajuda a explicar essas semelhanças. As decisões sobre público-alvo e forma de divulgação dos resultados, por exemplo, cabem às secretarias de educação, mas o trabalho em parceria com o CAEd/UFJF contribui para que as semelhanças sejam observadas.

Todos os programas são longevos e contínuos, observando o período selecionado para análise. Da mesma forma, todos eles foram implementados por meio da condução das secretarias de educação de cada estado, sem uma participação efetiva dos atores educacionais das redes. Essa é uma característica observada também no sistema nacional de avaliação. Esse modelo top/down de implementação da política não é necessariamente condenável se observarmos seus efeitos: considerando que a avaliação é um instrumento útil e importante para as redes, a ausência de comando e tomada de iniciativa por parte das secretarias poderia fazer com que os programas jamais fossem criados. Por outro lado, isso não elimina a pertinência das críticas que podem ser direcionadas à ausência de envolvimento dos demais atores, em especial, os escolares.

Os três programas são semelhantes, também, no percentual de recursos que destinam à avaliação. Nos três casos, esse percentual responde por menos de 1% do total de gastos da secretaria de educação, o que nos leva a refletir sobre o custo-benefício da política, já que a avaliação tem sido apontada pelos tomadores de decisão como um elemento fundamental para a condução da política educacional como um todo e seus gastos não. Proporcionalmente, o sistema de avaliação do Espírito Santo é o que mais destina recursos à avaliação (0,55%), entre os três programas analisados.

Ainda no âmbito das semelhanças, é possível destacar a participação dos municípios na avaliação. Conduzidos pelos estados, os três programas colocam em cena uma manifestação das parcerias no âmbito do federalismo brasileiro. A maior parte dos alunos avaliados no ciclo de alfabetização, nos três estados, pertence às redes municipais, que não possuem custos ao aderirem ao programa de avaliação. A única diferença está no Ceará: embora seja por adesão, a participação dos municípios é compelida pela lei que define a distribuição do ICMS, mecanismo que não está presente em Minas Gerais e Espírito Santo.

Em relação ao vínculo com outras políticas e à geração de avaliação, embora existam semelhanças, é possível identificar algumas diferenças significativas entre os programas. Os

três estados apresentaram, ao longo do período analisado, vínculo entre seus programas de avaliação da alfabetização e políticas de responsabilização. Embora com características específicas e regras próprias, essas políticas de *accountability* eram encaradas em bonificações com base em resultados. O Espírito Santo e o Ceará apresentaram políticas contínuas dessa natureza. Em Minas, por sua vez, a política de bonificação foi descontinuada a partir de 2014. No que diz respeito ao vínculo com outras políticas educacionais, por meio da análise aqui empreendida, é possível sustentar que o Spaece Alfa, no Ceará, apresenta um laço mais forte. O programa é quase indissociável das ações do Paic, de modo que falar sobre o programa de avaliação da alfabetização, quase invariavelmente, significa abordar o Paic. A mesma força de vínculo não é encontrada no Paebes Alfa e no Proalfa, embora o PIP tenha sido muito vinculado ao programa de avaliação em Minas até o ano de 2014.

As semelhanças são inúmeras, mas é pelas diferenças que podemos proceder a uma melhor comparação entre os três programas. O objetivo aqui não é estabelecer uma hierarquia entre eles, afirmando que um programa é melhor do que os demais. Apesar disso, em virtude de algumas características, podemos identificar programas mais complexos, e mesmo mais completos, a depender do enfoque adotado. Partindo da defesa da avaliação como um instrumento importante para as redes de ensino, e mesmo com um direito, conforme apresentada no capítulo 3 da tese, o Paebes Alfa seria, entre os três programas apresentados, aquele que apresenta o maior nível de complexidade. Isso se deve à análise de dois dos critérios propostos: o alcance e o desenho da avaliação.

Enquanto Spaece Alfa e Proalfa destinaram-se a avaliar a leitura (a ampliação do Proalfa quanto a esse aspecto só ocorreu em 2017), o Paebes Alfa, desde o início, dedicou-se a avaliar leitura, escrita e matemática, compreendendo um gama mais ampla das habilidades previstas para o ciclo de alfabetização. O programa foi inovador ao avaliar a escrita em formato de uma avaliação sistêmica, além de contemplar a matemática, normalmente secundarizada na discussão sobre alfabetização.

Além de avaliar mais disciplinas, o Paebes Alfa não se destinou à avaliação de uma única série, mas de todo o ciclo. O Spaece Alfa avaliou sempre o 2º ano, ao passo que, censitariamente, o Proalfa avaliou o 3º ano (com avaliações amostrais de 2º e 4º anos, durante um período). Soma-se a isso a participação facultada às escolas particulares, o que não ocorre em nenhum dos outros programas analisados. Isso faz com que o alcance do programa seja mais amplo do que os demais, avaliando, proporcionalmente, mais alunos no ciclo de alfabetização do que os outros dois estados. Os técnicos especialistas do CAED/UFJF entrevistados para a

tese corroboram essa leitura: em sua percepção, do ponto de vista técnico (as matrizes, o ajuste dos testes à população, a avaliação da escrita e da matemática), o Paebes Alfa é o programa mais completo dos três analisados.

É importante ressaltar, novamente, que isso não significa que o Paebes Alfa é melhor do que os demais, ou que é mais ajustado às necessidades de sua rede. O que está em análise aqui são os elementos que compõem o desenho da avaliação. No caso do Paebes Alfa, eles são mais numerosos do que o que se observa em Proalfa e Spaece Alfa.

\*\*\*

Ao longo do percurso de escrita da tese, um conjunto de temas foi aparecendo e seus desdobramentos pareceram entusiasmantes a mim. A seguir, exponho algumas dessas ideias para reflexão.

## 1

Apesar de parte da literatura sobre políticas públicas criticar o excesso de trabalhos que apenas descrevem as políticas e programas, sem estabelecer relações com o contexto político mais amplo no qual têm lugar, principalmente o internacional, no caso dos programas de avaliação da alfabetização não é isso o que observamos. Chama a atenção a escassez de trabalhos que procuram descrever esses programas, apresentando suas principais características e como elas se diferenciam e se aproximam de programas de mesma natureza em diferentes contextos. Um dos objetivos da tese é lidar com essa lacuna, analisando as características de três relevantes programas de avaliação da alfabetização nos estados brasileiros.

Com isso, o intuito não é desconsiderar a importância do contexto macropolítico para a compreensão da política pública analisada. No entanto, é preciso reconhecer também que, quando a crítica aponta a falta de uma abordagem mais ampla do contexto político, ela se refere, em regra, à necessidade de estabelecer uma crítica à influência do contexto internacional sobre políticas locais, ao avanço da última geração do capitalismo sobre os países em desenvolvimento, como o Brasil, apontados, geralmente, como vítimas desse processo de expansão do capital. Os aspectos culturais, enraizados em cada contexto, não são reconhecidos, num registro como esse, como elementos que obrigam essa expansão capitalista a ajustes, recuos e limites. Se a expansão da avaliação educacional é um fenômeno em escala mundial, isso não significa que ocorre da mesma forma e com a mesma característica em todos os lugares.

A cooperação entre estados e municípios no âmbito da avaliação da alfabetização no Brasil não encontra equivalente nos Estados Unidos, por exemplo, tomado em regra como nossa principal fonte de influência no assunto.

Se a expansão da avaliação em larga escala não é sempre igual, os programas de avaliação também não são todos iguais, como pode sugerir uma crítica mais apressada à política. Um das maneiras que a crítica encontra de diminuir as virtudes que um sistema de avaliação pode ter é generalizá-lo, apontando-o como cópia de programas com objetivos análogos em outros contextos. Todos são vistos como encarnações da meritocracia e como exportações, para o mundo da escola, de uma lógica mercadológica externa a ele. Assim, todos os programas de avaliação não são apenas iguais, mas igualmente nefastos e perniciosos.

Entretanto, o que a tese procurou mostrar é que a análise mais detida das características dos programas de avaliação da alfabetização leva a reconhecer semelhanças e diferenças entre eles. Criados em cenários políticos semelhantes (a presença do elemento macropolítico), nos quais a gestão por resultados nos estados estava em alta, eles, contudo, apresentam trajetórias diferentes e revelam decisões, técnicas e políticas, também diversas.

Se a concepção da alfabetização que dá suporte aos testes é a mesma entre eles, o mesmo não pode ser dito sobre o desenho da avaliação. Minas Gerais estabelece o 3º ano como o foco de suas ações, ao passo que, no Ceará, o foco é o 2º ano. No Espírito Santo, os três anos do ciclo são avaliados. Se no Ceará e em Minas Gerais, as atenções da avaliação recaem sobre a leitura, o Espírito Santo avalia também escrita e matemática. Observando o desenho da avaliação, os especialistas do CAEd entrevistados para a tese apontam o Paebes Alfa como um programa mais abrangente e complexo do que Proalfa e Spaece Alfa, justamente por suas escolhas quanto às séries e disciplinas avaliadas.

A avaliação das redes municipais é uma preocupação para os três programas. Se ela funciona por adesão em Minas Gerais e no Espírito Santo, no Ceará, é mais do que isso. Os municípios estão inseridos em um compromisso de erradicação do analfabetismo com grau de vinculação diferente daquele observado nos outros dois estados. Da mesma forma, a articulação do Spaece Alfa com outras políticas e ações, como o Paic e o Prêmio Escola Nota 10, coloca o programa numa posição estratégica para toda a política educacional do estado. Proalfa e Paebes Alfa também são considerados estratégicos para a política educacional de seus estados, mas seria um equívoco acreditar que isso ocorre da mesma maneira do que aquela observada no Ceará. Em Minas, a articulação do Proalfa com uma política educacional representativa aconteceu com o PIP, mas ela foi descontinuada. Já no Espírito Santo, a articulação do Paebes

Alfa não ocorre com nenhuma política estadual que tenha tido destaque ou relevância para outros estados, como ocorreu com o Paic, por exemplo.

Reconhecer essas diferenças e semelhanças não tem objetivo de hierarquizar os programas, apontando o melhor e o pior deles. Sendo diferentes, eles podem ser mais ou menos ajustados à realidade em que estão inseridos. Isso não impede, contudo, de juízos como aqueles feitos pelos especialistas do CAEd sobre o grau de complexidade dos programas. Isso não significa que o Paebes Alfa é um programa melhor do que o Spaece Alfa, apenas que eles são diferentes em determinados aspectos. Da mesma forma, o Spaece Alfa não é um programa mais importante do que o Proalfa por ter uma vínculo com uma política estadual que virou referência nacional e alimentar uma política duradoura de *accountability*.

O fato de serem diferentes, contudo, não esconde suas semelhanças. Mais do que isso, não impede que ações levadas a cabo no contexto de um deles sirvam como influência para os demais. A criação do Paebes Alfa foi influenciada pela existência de Proalfa e Spaece Alfa.

## 2

No capítulo 4 da tese, foi apresentado um mapeamento, em 4 importantes canais de divulgação da pesquisa em educação, acerca de seu interesse pela avaliação da alfabetização. Como vimos, a atenção que a pesquisa educacional presta ao tema é quase inexistente. Ao mesmo tempo, os programas de avaliação tornaram-se peças centrais para a política educacional dos estados. Como vimos no capítulo 5, eles estão associados a importantes ações da política local. Soma-se a isso o fato de serem considerados, pelos atores educacionais que ocuparam, ao longo do período de análise, posições vinculadas à tomada de decisões, estratégicas e fundamentais para a gestão das redes. Se eles são considerados tão importantes para esses atores, como explicar o completo desinteresse que caracteriza a pesquisa em educação?

Mais do que a resistência relacionada à avaliação, que, certamente, contribui para esse fato, existe um descompasso entre a pesquisa em políticas públicas educacionais e as políticas públicas executadas na prática. Esse descompasso desdobra-se em diferentes binômios de tensão: pesquisa acadêmica/prática profissional; teoria/prática; concepção/ação. O que a teoria diz sobre os tomadores de decisão no âmbito das políticas públicas, muitas vezes, caminha em sentidos diferentes daqueles atribuídos pelos próprios tomadores de decisão. Se os tomadores de decisão ocupam um papel importante para a pesquisa da política pública, e eles dizem que os programas de avaliação da alfabetização são estratégicos para a condução da política educacional como um todo, por que a pesquisa não se interessa por esses programas? Ou ela

não acredita em seus pressupostos teóricos (que dizem que os tomadores de decisão são importantes peças para a compreensão das políticas públicas), caso em que eles devem ser revistos, ou ela acredita, mas, seletivamente, escolhe o tipo de opinião que deseja ouvir e dar voz, de modo a manter de pé suas crenças sobre como uma política pública de manual deve ser.

### 3

A concepção de alfabetização colocada em cena pelos programas estaduais de avaliação analisados pode ser entendida como um dos fatores que alimentam a resistência em torno da avaliação e se associa, também, à pouca atenção prestada pela pesquisa em educação em relação a esses programas. O alfabetizar letrando, encarnado, nas matrizes de referência, em descritores que avaliam habilidades relacionadas à aquisição do sistema de escrita e ao desenvolvimento da consciência fonológica, é alvo de críticas dos construtivistas e defensores de um aprendizado global do texto, para quem a avaliação em larga escala é, em regra, também um problema. A adoção de Magda Soares como referência para a construção das matrizes dos testes e para a própria concepção do que deve ser avaliado ajuda a compreender, por exemplo, a ojeriza de autoras como Gontijo (2014), para quem tanto a avaliação em larga escala quanto a perspectiva do letramento são obstáculos a um processo virtuoso de alfabetização.

Se a avaliação em larga escala não tivesse, no Brasil, uma história anterior à avaliação sistêmica do ciclo de alfabetização, interessaria à pesquisa reconstruir a influência de uma crítica sobre a outra: a resistência à avaliação teria se desenvolvido por conta da oposição à ideia do alfabetizar letrando, por ela encarnada, ou a oposição ao alfabetizar letrando teria se desenvolvido por conta da avaliação em larga escala. Como a história da avaliação em larga escala para a alfabetização é recente no país e a resistência à avaliação é antiga, parece ser plausível entender que a resistência já existente em relação à avaliação foi potencializada pela concepção de alfabetização que os programas adotam.

### 4

A defesa da avaliação como direito foi uma das ideias apresentadas ao longo da tese. Isso aparece no capítulo 3, de maneira mais aprofundada, e no capítulo 5, no tópico sobre o alcance do programa de avaliação. A avaliação como direito coloca em cena a questão do número de avaliações aplicadas aos alunos. Se levarmos ao limite a ideia de que a avaliação é um direito, todas as séries e disciplinas deveriam ser avaliadas. Soma-se a isso a compreensão de que a avaliação tem como objetivo último servir de base para a tomada de decisões,

fornecendo informações qualificadas (evidências) que deem suporte a ações legitimadas. Entretanto, a defesa de que a avaliação é um direito não leva, necessariamente, ao excesso de avaliações. A questão passa a ser outra. Como fazer avaliações mais inteligentes e eficazes? Como desenvolver instrumentos de avaliação cada vez mais acurados, capazes de cobrir mais conteúdos curriculares sem sobrecarregar os alunos? Esse é um dos desafios permanentes da pesquisa em avaliação.

A TRI é um instrumento importante para conseguirmos respostas a essas perguntas. A princípio, o modelo teórico da TRI pressupõe que seria possível estimar a proficiência do respondente do teste com base em uma única questão. Mesmo sendo uma hipótese, essa é uma ideia poderosa, que norteia a pesquisa sobre o tema. Se chegarmos a desenvolver um banco de itens extenso e qualificado o suficiente para que, com testes compostos por poucos itens, sejamos capazes de produzir diagnósticos precisos e confiáveis para todas as disciplinas e etapas de escolaridade, a questão sobre como produzir avaliações mais inteligentes e eficazes começa a encontrar um caminho de resposta. Não se trata de uma defesa da avaliação a qualquer custo, mas de procurar as melhores soluções para produzir diagnósticos confiáveis da aprendizagem sem sobrecarregar os alunos.

## 5

Em relação à fase de implementação da política, é preciso chamar atenção para um fato acerca dos programas de avaliação da alfabetização aqui analisados. Uma das críticas que a avaliação em larga recebe diz respeito ao fato de ser uma política desenhada e implementada de cima para baixo, pelos governantes, sem a participação dos atores educacionais e escolares que, na prática, irão executar as ações da política. Nessa perspectiva, a avaliação é vista como criada pela e para a secretaria de educação, para atender a seus interesses imediatos. Um desses interesses, segundo outra crítica que se desdobra a partir dessa, é passar uma imagem de que a educação ofertada por determinada rede apresenta uma qualidade muito superior àquela que de fato possui. Assim, a avaliação seria um instrumento de propaganda (enganosa) dos governos. Vimos, no capítulo 3, que os resultados, na verdade, estão longe de apresentar esse poder. Para qualquer etapa de escolaridade, eles, antes, são veículos de preocupação do que de propaganda favorável. A despeito disso, a imagem de que a avaliação é algo distante da escola e de seus atores permanece. De fato, como também vimos, a história do Saeb caracteriza-se por uma implementação conduzida pelo governo, sem envolvimento direto, quanto à definição da agenda, desenho e implementação, dos atores escolares (que, por isso, são aqueles que mais



apresentaram resistência à avaliação em larga escala ao longo de sua história no Brasil). Do mesmo modo, conforme apresentado no capítulo 5, nos três programas de avaliação da alfabetização aqui analisados, a forma de implementação foi parecida, conduzida pelas secretarias de educação.

Esse cenário pode levar a um entendimento, em minha leitura, equivocado, de que a avaliação serve apenas aos propósitos (questionáveis sempre, conforme aponta a crítica) de quem está à frente das secretarias de educação. Como instrumento, a avaliação pode realmente servir a diferentes propósitos, cuja virtude será variável. Apesar disso, as secretarias de educação não controlam esse processo como a crítica sugere. Em primeiro lugar, a avaliação educacional é uma área de pesquisa, que tem se expandido cada vez mais no Brasil (de acordo com o que foi apresentado no capítulo 3). Como tal, ela está sujeita às características do campo científico, o que inclui limitações. Embora o conjunto de pesquisadores nessa área ainda seja inferior ao de outras áreas relevantes na pesquisa educacional (como currículo, formação de professores, gestão, práticas pedagógicas etc), a avaliação em larga escala já conta com um corpus de pesquisadores, incluindo uma associação nacional que os representam<sup>206</sup>, que atuam como limitações aos usos desajustados das avaliações. Essa é uma relação complexa: como instrumento, a avaliação permite os mais diferentes usos, o que envolve, na maioria dos casos, decisões políticas acerca de como os resultados serão utilizados, mas, ao mesmo tempo, existe um campo científico que, atento a tais usos, produz discussões e críticas sobre eles. Por mais que essa relação entre os atores políticos e os atores científicos associados à avaliação educacional (como política pública e como área de pesquisa) ainda não seja estreita a ponto de gerar um debate produtivo e constante, as discussões levadas a cabo pelos pesquisadores da área servem como limites a decisões políticas sobre a avaliação. Nem tudo será plenamente aceito.

Outra limitação à ideia de que a avaliação é simplesmente um instrumento de propaganda na mão dos governos vem da composição das secretarias de educação. Embora isso tenha mudado ao longo do tempo, já que as secretarias, cada vez mais, contam com equipes destinadas especificamente ao trabalho com a avaliação educacional, elas ainda carecem de um grupo sólido de profissionais especializados no tema. Conforme já apresentado anteriormente (também no capítulo 3), a avaliação educacional envolve muitas áreas diferentes e é composta por muitas etapas distintas. As secretarias têm avançado na constituição de equipes destinadas a interpretar e a divulgar os resultados das avaliações, funcionando como uma dobradiça entre a interpretação dos resultados e a decisão sobre o seu uso por parte dos tomadores de decisão

---

<sup>206</sup> ABAVE – Associação Brasileira de Avaliação Educacional.

posicionados nos níveis mais altos da hierarquia educacional, contudo, não possuem ainda especialistas capazes de conduzir uma avaliação em larga escala do começo ao fim (definição de matrizes, elaboração de itens, construção de modelos de teste, aplicação, processamento de dados, construção de escalas, produção de medidas, capacitação para leitura e uso dos resultados etc). Isso leva ao estabelecimento de parcerias com instituições especializadas, que atuam, muitas vezes, como pontes entre a pesquisa acadêmica sobre avaliação e a oferta de serviços para o atendimento das necessidades dos estados. Nos três programas de avaliação da alfabetização que investigamos, para o período analisado, a condução e a execução da política são realizadas em parceria com uma instituição dessa natureza, um centro de pesquisa de uma universidade pública (CAEd/UFJF).

De acordo com o que foi apresentado no capítulo 5, a atuação do CAEd no âmbito dos três programas foi fundamental para uma série de mudanças que ajudam a compreender o desenvolvimento desses programas de avaliação ao longo do tempo. O mesmo pode ser dito para a atuação do Ceale/UFMG, que também participou da execução dos programas. A história do CAEd, como instituição de pesquisa, em muitos pontos, confunde-se com a história dos próprios programas e, por sua vez, com a história da avaliação em larga escala da alfabetização no Brasil, em especial, no que diz respeito aos elementos técnicos analisados (desenho da avaliação, construção das matrizes e das escalas, elaboração dos itens, avaliação da escrita e da matemática etc). Esse tipo de parceria coloca em cena as limitações ao uso das avaliações em virtude do que está sendo discutido pela pesquisa especializada. Desse modo, perde força o argumento da crítica de que a avaliação é um instrumento, pura e simplesmente, a serviço dos interesses propagandísticos das secretarias de educação.

Ainda um ponto importante a ser ressaltado. A análise levada a cabo pela tese chama atenção para o papel dessas instituições de pesquisa na área de avaliação em larga escala, algo que ainda não foi devidamente enfrentado pela pesquisa em educação. Esse tipo de parceria mobiliza, pelo menos, duas discussões importantes e, ao mesmo tempo, negligenciadas: i) a relação entre pesquisa acadêmica e decisão política, uma encarnação das tensões entre teoria e prática, investigando como a teoria sobre o tema é capaz de influenciar decisões políticas na área; e ii) o federalismo que esse tipo de parceria mobiliza, quando as instituições pertencem a diferentes entes federativos (no caso específico do CAEd, uma universidade federal atuando em parceria de estados e municípios). A pesquisa em avaliação educacional e a pesquisa educacional em geral se beneficiaram se o interesse dos pesquisadores começasse a ser direcionado também para essas questões.

## 6

As considerações do ponto anterior me levam a estas. Se instituições como o CAEd e Ceale desempenham um papel importante na efetivação de programas de avaliação no Brasil, seria importante e coerente dar voz a quem, normalmente, não é ouvido numa pesquisa sobre avaliação educacional: aqueles que constroem as matrizes, elaboram os itens, montam os cadernos, aplicam os testes, produzem as medidas, elaboram os materiais de divulgação e capacitação para uso dos resultados. Todos esses atores são, em regra, invisíveis, embora sejam especialistas e exerçam um papel tão crucial para a avaliação. Eles também têm uma história a contar. Uma história do desenvolvimento das tecnologias de avaliação. Isso, inclusive, poderia ter sido muito mais bem explorado pela tese. Contudo, seu objeto primário não era esse. Esse foi um ganho, penso, no caminho metodológico adotado. Um bônus que, com a reflexão posterior, que desagua neste conjunto de considerações finais, foi muito bem-vindo. A ideia inicial era ouvi-los por saber do manancial de informações e da expertise que possuem, o que seria fundamental (e, realmente, foi) para a história que eu desejava narrar e registrar. Mais do que isso, porém, essa escolha levou à percepção de que a história da avaliação educacional poderia ser contada pelo olhar também desses atores, em geral, olvidados. Isso não significa alçá-los a uma posição mais importante do que os demais para a compreensão da avaliação em larga escala no Brasil. Os tomadores de decisão, envolvidos com o processo de desenho e implementação das políticas (e, aqui, também ouvidos), os professores e diretores (a *street-level bureaucracy*), os alunos e seus responsáveis, assim como os pesquisadores da avaliação e os críticos a ela, todos eles formam a constelação de visões que compõem o cenário da avaliação educacional. Ouvi-los todos é o que produzirá uma visão verdadeiramente global do fenômeno educacional colocado cena pela avaliação.

Embora creia nisso, não pude efetivar um projeto dessa monta na tese. Desse modo, reconheço seus limites. Trata-se de uma análise documental acompanhada das leituras de determinados atores, com posições bem definidas: especialistas de uma instituição pública de pesquisa que conduz sistemas de avaliação em conjunto com os entes federados e tomadores de decisão que estiveram, no âmbito das secretarias, envolvidos com os programas analisados. Professores e diretores não foram envolvidos, assim como os alunos. Aliás, essa talvez seja uma das grandes lacunas acerca da avaliação educacional em larga escala, mesmo entre os críticos. Os alunos são tão invisíveis quanto os especialistas das instituições de pesquisa como o CAEd, embora muitos falem em nome deles. É comum, por exemplo, a manifestação da preocupação, em nome dos alunos, do excesso de testes e avaliações, mas não há pesquisas consolidadas que

estudem e deem voz aos alunos sobre isso. O processo de implementação da avaliação em larga escala no Brasil é criticado por não ter ouvido professores e diretores de escola, mas pouco se fala dos alunos. Onde estão as pesquisas que mostram as opiniões dos alunos sobre a avaliação? Uma vez que os sistemas de avaliação foram se consolidando e a resistência de professores e diretores foi diminuindo, quais foram as ações, das secretarias e das escolas, para explicar aos alunos a importância das avaliações? Onde elas estão documentadas? Quem se interessou por elas? Outra lacuna da pesquisa em avaliação diz respeito ao nível de comprometimento dos alunos com os testes padronizados, o que, certamente, impacta no resultado (em que nível, não sabemos, afinal, não há pesquisas sobre isso), e está relacionado com o grau de informação e esclarecimento do aluno acerca dos objetivos da avaliação e do papel que ele exerce nesse cenário.

A despeito das reconhecidas lacunas da tese em relação aos atores que dela fizeram parte como sujeitos da pesquisa, acredito que uma das contribuições deste trabalho, do ponto de vista metodológico, tenha sido chamar a atenção para os especialistas das instituições que apoiam e executam os programas de avaliação no Brasil. Eles tomam decisões e realizam ações cruciais para a política de avaliação, mas, normalmente, não são ouvidos quando a avaliação se torna um objeto de pesquisa. É como se os resultados da avaliação falassem por si só e dissessem tudo o que é preciso saber sobre a avaliação. Esse fetichismo do resultado, que esconde como ele é produzido e todas as decisões anteriores que levaram a ele, impede que a avaliação seja vista em toda sua extensão e complexidade. Se os resultados são importantes, por representarem a entrega da política, aquilo que dá base para a realização de seu objetivo (dar suporte a decisões qualificadas), as vozes de quem faz com que a política aconteça não podem ser negligenciadas, já que contam a história por trás desses resultados.

## 7

Falta uma pesquisa mais detalhada sobre a influência da forma como os resultados são divulgados sobre seu entendimento. A apropriação dos resultados permanece sendo um dos principais nós para a avaliação. Um dos obstáculos para a sua efetivação diz respeito ao entendimento acerca do que os dados dizem e das conclusões que autorizam. Como fazer uso daquilo que não é compreendido? Esse é um problema essencialmente de linguagem e tradução. A avaliação foi (e ainda é) acusada, diversas vezes, de se valer de uma linguagem autorreferenciada, fechada, de difícil acesso para os não especialistas. Como caixa-preta, sua linguagem seria entendida por um conjunto restrito de especialistas, em regra, matemáticos e

estatísticos, que reduzem, ao sintético dos números, a complexidade inarredável da realidade escolar e educacional. É como se a avaliação tivesse sido criada para esconder, não para mostrar. A compreensão dos resultados é parte fundamental da possibilidade de seu uso adequado. Cobra-se que os resultados sejam utilizados como parte das ações rotineiras de planejamento das redes e escolas, mas é preciso que eles sejam plenamente compreendidos pelos atores, sob pena do aparecimento de decisões que se baseiam em conclusões indevidas, por conta de informações mal interpretadas.

Ainda está por ser desenvolvido, à maneira como a pesquisa em eficácia escolar faz com os fatores associados ao desempenho, um inventário dos elementos que afetam a compreensão dos resultados da avaliação pelos atores educacionais. Nesse rol, que, imagino, será amplo, uma parte será dedicada aos elementos gráficos e linguísticos. Os resultados são divulgados por canais visuais e textuais. Há uma série de elementos que influenciam em seu entendimento, desde a linguagem utilizada até os recursos visuais de apoio, incluindo os tipos de gráficos etc. A entrevista com Alessio da Costa Lima, apresentada no capítulo 5, me fez pensar sobre isso. Ele afirmou que um dos elementos que fizeram com que o Spaece Alfa tenha se consolidado no estado estava relacionado à compreensão de seus resultados pelos atores, incluindo os prefeitos (ponto importante para uma política educacional que se baseia na cooperação com os municípios), e que um dos aspectos importantes para isso foi a forma “mais didática e palatável” com que as informações passaram a ser divulgadas. Um desses aspectos foram as cores utilizadas pelos padrões de desempenho, de modo que seu entendimento passou a ser ampliado e compartilhado em todo o estado. Lembro-me de, ainda no âmbito de minhas atividades no CAEd, ter divulgado resultados do Spaece (incluindo o Alfa) no estado e ver um auditório lotado com gestores escolares que vestiam uma camisa com a cor do padrão de sua escola. A despeito de outras considerações que possam ser feitas em relação a isso, o fato é que Alessio tinha razão: as cores dos padrões passaram a ser uma referência direta e imediata acerca da aprendizagem.

Essa é uma área de design instrucional que poderia ser desenvolvida em torno da pesquisa em avaliação educacional. Detalhes podem fazer diferença para uma melhor compreensão dos resultados, o que pode levar, por sua vez, a seu uso mais adequado. Comunicar-se pressupõe o desejo de ser compreendido e os elementos que contribuem para essa compreensão não podem ser tratados como meramente acessórios. A forma como os dados são comunicados é tão importante quanto seu conteúdo.

## 8

Embora a forma como os argumentos foram aqui apresentados possa levar a essa interpretação, o objetivo da tese não é realizar uma defesa cega da avaliação. Espero que isso tenha ficado claro ao longo do texto. É mais fácil, entretanto, encontrar críticas e defesas radicais da avaliação do que uma análise mais detalhada de suas virtudes e limites. Entendo a avaliação como um instrumento. Como tal, ela possui limites. Também como tal, ela está sujeita aos mais diferentes tipos de uso, que devem ser julgados a partir das intenções de quem os mobiliza. Estou convencido de que, usada para produzir diagnósticos sobre a aprendizagem dos alunos, ela é um instrumento importante para a gestão das turmas, escolas e redes. A avaliação não é um fim em si mesma, embora, em alguns cenários, ela possa ser assim utilizada. Sua instrumentalidade se revela no fato de que, servindo à tomada de decisão, seu objetivo maior é contribuir para a potencialização da aprendizagem dos alunos.

## 9

Devo ao professor Manuel Palácios essa última consideração. A análise dos três programas estaduais de avaliação da alfabetização levam a uma reflexão sobre o papel dos entes federados no processo de expansão e consolidação da avaliação no Brasil. A tese é um reconhecimento desse papel. Ela não versa sobre a ANA, por exemplo. Ao buscar contar uma parte que considero importante da história da avaliação da alfabetização no país, foi nos programas estaduais que cheguei.

Os estados são entes protagonistas desse processo. Os programas estaduais de avaliação da alfabetização são anteriores à criação de uma avaliação nacional para a etapa. A avaliação da escrita e da matemática, por exemplo, são realizações dos programas estaduais anteriores ao surgimento da avaliação nacional. Os programas estaduais impulsionaram o desenvolvimento da avaliação da alfabetização no Brasil. Soma-se a isso o exercício da cooperação federalista colocado em cena pelos estados. A cooperação com os municípios para a realização da avaliação é uma manifestação desse federalismo. O caso do Ceará talvez seja o mais exemplar em relação a isso. A cooperação entre estado e municípios para a erradicação do analfabetismo é o elemento integrador de toda a política educacional do estado há mais de dez anos, o que inclui, também, seu programa de avaliação da alfabetização. Em Minas e no Espírito Santo, essa cooperação também é colocada em cena, mesmo em níveis diferentes. Um sistema nacional de avaliação, centrado nas ações da União, provavelmente, não seria capaz de colocar em cena esse nível de articulação e cooperação.

A existência de programas estaduais de avaliação reforça a soberania e a autonomia dos estados para tomar decisões sobre a educação ajustadas às necessidades locais. O estado decide o que é prioridade, o que deseja avaliar de seu currículo estadual, quais séries precisam de um monitoramento maior e por quais motivos. Se a avaliação serve como suporte para a tomada de decisão, o estado deve ser capaz de dizer de quais informações precisa para que essas decisões sejam tomadas e com qual periodicidade. Por isso, os sistemas estaduais de avaliação se expandiram no Brasil (como vimos no capítulo 3). Além disso, cabe aos estados a decisão sobre a articulação das avaliações com suas políticas educacionais e sobre seu grau de importância como instrumento de gestão. Os programas de avaliação estaduais, nesse sentido, são formas de encarnar o princípio federalista no âmbito da educação. Isso ganha contornos especiais para a alfabetização, já que os municípios são os agentes centrais nesse processo.

## 10

Fui provocado a pensar sobre a tese por detrás desta tese<sup>207</sup>. Talvez o motor para a escrita de um trabalho como este, destinado à análise de três programas de avaliação da alfabetização, seja mesmo a própria defesa da avaliação em larga escala. O capítulo 2 da tese dá pistas dessa intenção. Se ela não está declarada abertamente como um dos objetivos da tese, também não está tão profundamente enraizada no texto a ponto de não poder ser reconhecida. Essa intenção não foi deliberadamente omitida. Contudo, somente depois de provocado é que ela pareceu claramente diante de meus olhos. A defesa da avaliação sempre foi uma constante ao longo de minha atuação no CAEd/UFJF, à frente da divulgação dos resultados de sistemas próprios de avaliação no Brasil. Muitas vezes, essa defesa era o ponto de partida para que um diálogo sobre os resultados propriamente ditos pudesse ser estabelecido. Não se trata, contudo, de uma defesa cega da política e do instrumento. A avaliação tem problemas, mas também tem virtudes. Não estou disposto a abrir mão do reconhecimento de suas vicissitudes para estabelecer uma defesa dogmática do instrumento. Do mesmo modo, não estou disposto a abrir mão do reconhecimento de suas virtudes para fazer coro ao discurso, ainda dominante nas universidades, de seu desserviço à educação.

---

<sup>207</sup> Pelo professor Eduardo Magrone, a quem agradeço.

**11**<sup>208</sup>

A tese baseia-se em uma visão da avaliação que busca, analiticamente, diferenciar o instrumento de seu uso. As críticas mais ferozes à avaliação costumam tomar seu uso como um elemento capaz de contaminar todo o instrumento. Assim, a avaliação não prestaria nenhum serviço à educação. Em minha leitura, isso se baseia numa visão equivocada da avaliação, que, diante de um uso inadequado da política e do instrumento, está disposta a sacrificar toda a ideia por trás da avaliação. Por isso, conforme reconhecido no ponto anterior, minha defesa da avaliação é uma espécie de ponto de partida para possibilitar o debate sobre o tema: olhando para o uso, parto do princípio que a avaliação pode ser bem utilizada e contribuir para o debate educacional e para a melhoria da educação.

Para construir um argumento dessa natureza, o bom uso da avaliação repousou na ideia de racionalidade técnica: instrumento dotado de rigor metodológico, produção de informações qualificadas para a tomada de decisões, ações baseadas em evidências, prestação de contas do Estado diante da sociedade, produção de indicadores vinculados à ideia de qualidade da educação, garantia de direitos etc. Todas essas ideias, que foram tratadas, em muitos casos, como pertencentes, como algo inerente, ao polêmico e criticado gerencialismo em educação, em particular o uso de evidências para a tomada de decisão e a necessária prestação de contas de sua eficiência por parte do Estado, integram essa dimensão da avaliação dotada de racionalidade técnica.

Se esse argumento foi, diversas vezes, importante para produzir uma defesa da avaliação como instrumento técnico, estabelecendo aqui uma distinção para lá de problemática com a política, a dimensão política da avaliação, que repousou, também diversas vezes, em seu mau uso pelos agentes educacionais, precisa ser melhor considerada. A avaliação, por mais rigor técnico que possua, não está blindada em relação ao mau uso político. O conhecimento de suas características técnicas permanece sendo, em grande medida, privilégio de poucos agentes educacionais, na maior parte das vezes, acadêmicos e lotados nas esferas mais altas das hierarquias educacionais. Uma ação política necessária é levar esse conhecimento técnico, que informaria os atores educacionais acerca das possibilidades e limites da avaliação, em particular, no que tange a seus resultados, para fora dos gabinetes das secretarias de educação. O conhecimento técnico estaria a serviço, assim, do bom uso político.

Isso evitaria, por exemplo, os excessos relacionados aos resultados: variações nos resultados de escolas e redes, insignificantes estatística e pedagogicamente, são tratadas, pela

---

<sup>208</sup> Uma discussão nessa direção foi proposta pela professora Alicia Bonamino, a quem devo agradecimentos.



mídia, pelas secretarias e pelos próprios atores educacionais, como razões para celebrações desmedidas e para incorporações do fracasso igualmente desajustadas. Decisões muito sérias, como trocas de superintendências, por exemplo, são tomadas com base em resultados que foram mal interpretados. O conhecimento técnico preciso acerca dessas informações evitaria seu mau uso político. Essa seria uma forma, inclusive, de contribuir para arrefecer a resistência, ainda numerosa, acerca da avaliação educacional.

## REFERÊNCIAS

- AFONSO, Almerindo Janela. **Avaliação educacional: regulação e emancipação**. São Paulo: Cortez, 2000.
- AFONSO, Almerindo Janela. **Para uma conceptualização alternativa de accountability em educação**. Educação e Sociedade, Campinas, v. 33, n. 119, p. 471-484, jun. 2012.
- ALBUQUERQUE, E. B. C.; FERREIRA, A. T. B. **Programa Nacional do Livro Didático (PNLD): mudanças nos livros de alfabetização e os usos que os professores fazem desse recurso em sala de aula**. Ensaio: Avaliação e Políticas Públicas em Educação. Rio de Janeiro, vol. 27, n. 103, abr/jun, 2019.
- ALFERES, M. A.; MAINARDES, J. **O Pacto Nacional pela Alfabetização na Idade Certa em ação: revisão de literatura**. Ensaio. Avaliação e Políticas Públicas em Educação. Rio de Janeiro, vol. 27, n. 102, jan/mar, 2019.
- ALVES, Maria Teresa Gonzaga; SOARES, José Francisco. **Contexto escolar e indicadores educacionais: condições desiguais para a efetivação de uma política de avaliação educacional**. Educação & Pesquisa, São Paulo, v. 39, n. 1, p. 177-194, jan./mar. 2013.
- AMARAL, Maria Ideuma Saraiva. **Gestão por resultado e avaliações externas: estudo de caso na EMEIEF Governador César Cals de Oliveira Filho (Maracanaú, 2013-2015)**. Fortaleza, 2017, 132p. Dissertação de Mestrado – Mestrado Profissional em Planejamento e Políticas Públicas, Universidade Estadual do Ceará.
- ANDRADE, Renato Júdice de; SOARES, José Francisco. **O efeito da escola básica brasileira**. Estudos em Avaliação Educacional, v. 19, n. 41, set./dez. 2008.
- ARAÚJO, K. H.; LEITE, R. H.; ADRIOLA, W. B. **Prêmios para escolas e professores com base no desempenho acadêmico discente: a experiência do estado do Ceará (Brasil)**. Revista Linhas. Florianópolis, v. 20, n. 42, p. 303-325, jan./abr. 2019.
- ARELARO, L. R. G.; JACOMINI, M. A.; KLEIN, S. B. **O ensino fundamental de nove anos e o direito à educação**. Educação e Pesquisa, São Paulo, v.37, n.1, 220p. 35-51, jan./abr. 2011.
- ARRETCHE, Marta Teresa da Silva. **Estado Federativo e Políticas Sociais: determinantes da descentralização**. Rio de Janeiro: Revan; São Paulo: FAPESP, 2000.
- BAKTHIN, Mikhail Mikhailovitch. **Estética da criação verbal**. 6ª edição. São Paulo: WMF Martins Fontes, 2011.
- BASSO, A. C. M. S.; RODRIGUES, C. F.; BROOKE, D. A. L.; VIGNOLI, D. A.; CANDIAN, J. F.; REZENDE, W. S. **Desigualdade de desempenho e raça: uma análise a partir do Paebs 2009**. Estudos em Avaliação Educacional, vol. 23, n. 51, 2012.
- BASTOS, Rayssa Lopes. A proposta curricular de língua portuguesa de Juiz de Fora/MG: reflexões acerca do processo de recontextualização na prática**. Petrópolis/RJ, 2017, 201p. Tese de Doutorado – Programa de Pós-Graduação em Educação, Universidade Católica de Petrópolis.
- BAUER, Adriana; ALAVARSE, O. M; OLIVEIRA, R. P. **Avaliações em larga escala: uma sistematização do debate**. Educação e Pesquisa. São Paulo, v. 41, p. 1367-1384, dez. 2015.
- BAUER, Adriana; GATTI, Bernardete A.; TAVARES, Marialva R. (Org.). **Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origens e pressupostos**. Vol.II. Florianópolis: Insular, 2013.

BENEDICT, Ruth. **O crisântemo e a espada: padrões da cultura japonesa**. São Paulo: Perspectiva, 2011.

**BEZERRA, Francisco Igor Magalhães Mapurunga Bezerra. O impacto do ensino noturno nos resultados do Spaece: o caso de duas escolas do Ceará. Juiz de Fora/MG, 2014, 107p. Dissertação de Mestrado – Mestrado Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.**

BOBBIO, Norberto. **A era dos direitos**. Rio de Janeiro: Campus, 1992.

BONAMINO, A. **Tempos de avaliação educacional: o SAEB, seus agentes, referências e tendências**. Rio de Janeiro: Quartet, 2002.

BONAMINO, Alicia; SOUSA, S. 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.

BONAVIDES, Paulo. **Ciência Política**. 10ª ed. São Paulo: Malheiros, 2000.

BORGES, Edna Martins. **Avaliações em larga escala no contexto escolar: percepção de diretores escolares da rede estadual de ensino de Minas Gerais**. Belo Horizonte, 2016, 327p. Tese de Doutorado – Programa de Pós-Graduação em Educação, Universidade Federal de Minas Gerais.

BOTO, Carlota. **A educação escolar como direito humano de três gerações: identidades e universalismos**. Educação e Sociedade, Campinas, vol. 26, n. 92, p. 777-798, Especial, outubro de 2005.

BOUDET, K. P.; CITY, E. A.; MURNANE, R. J. **Data wise: a step-by-step guide to using assessment results to improve teaching and learning**. Harvard Education Press, Cambridge, Massachusetts, 2013.

BOURDIEU, Pierre. **A distinção: crítica social do julgamento**. São Paulo: Unesp, 2003.

BOURDIEU, Pierre. **A economia das trocas simbólicas**. São Paulo: Perspectiva, 2011.

BOURDIEU, Pierre. **A miséria do mundo**. Petrópolis: Vozes, 2003(a).

BOURDIEU, Pierre. **Le champ scientifique**. In: Actes de la recherche en sciences sociales, vol. 2, nº 2-3, junho de 1976.

BOURDIEU, Pierre; PASSERON, Jean-Claude. **A reprodução: elementos para uma teoria do sistema de ensino**. 3ª ed. Rio de Janeiro: Francisco Alves, 1992.

BRANDÃO, Gildo Marçal. **Linhagens do pensamento político brasileiro**. DADOS – Revista de Ciências Sociais, Rio de Janeiro, Vol. 48, nº 2, 2005.

BRANDÃO, Júlia Barbosa. **O rateio de ICMS por desempenho de municípios no Ceará e seu impacto em indicadores do sistema de avaliação da educação**. Dissertação de Mestrado. 87p. Escola Brasileira de Administração Pública e de Empresas, Centro de Formação Acadêmica e Pesquisa. Fundação Getúlio Vargas. Rio de Janeiro, 2014.

BRASIL. **Avaliação nacional da alfabetização (ANA): documento básico**. – Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2013.

BROOKE, Nigel (org). **Marcos históricos na reforma da educação**. Belo Horizonte: Editora Fino Traço, 2012.

BROOKE, Nigel. **Responsabilização educacional no Brasil**. Revista Iberoamericana de Evaluación Educativa, vo. 1, n. 1, 2008.

- BROOKE, Nigel; ALVES, Maria Tereza Gonzaga; OLIVEIRA, Lina Kátia Mesquita de. **Avaliação da educação básica: a experiência brasileira**. Belo Horizonte: Fino Traço, 2015.
- BROOKE, Nigel; CUNHA, M. A. **A avaliação externa como instrumento da gestão educacional nos estados**. Estudos & Pesquisas Educacionais, São Paulo, v. 2, p. 17-79, 2011.
- BROOKE, Nigel; SOARES, José Francisco. **Pesquisa em eficácia escolar: origens e trajetórias**. Belo Horizonte: Editora UFMG, 2008.
- CÂMARA, Marcelo. **Análise dos resultados do pré-teste da Provinha Brasil de matemática**. Estudos em Avaliação Educacional, vol. 24, nº54, jan-abr, 2013.
- CALDERÓN, Adolfo Ignacio. *Usos e apropriações das avaliações em larga escala*. In: **Diálogos e proposições, Vol. II: Planos de Ação para a rede estadual de ensino de Minas Gerais**. Editora CAEd/UFJF: Juiz de Fora, 2017.
- CALDERÓN, A. I.; RAQUEL, B. M. G.; CABRAL, E. S. **O Prêmio Escola nota 10: meritocracia e cooperação para a melhoria do desempenho escolar**. Ensaio: Avaliação e Políticas Públicas Educacionais, Rio de Janeiro, v. 23, n. 87, p. 517-540, abr./jun. 2015.
- CATELLS, Manuel. **A sociedade em rede. A era da informação: sociedade, economia e cultura**. Vol. 1. Brasília: Paz e Terra, 2016.
- CASTRO, Cláudio de Moura. **The World Bank policies: damned if you do, damned if you don't**. Comparative Education. Vol. 38, n. 4, 2002.
- CASTRO, Maria H. G. **A consolidação da política de avaliação da educação básica no Brasil**. Meta: avaliação, Rio de Janeiro, v. 1, n. 3, p. 271-296, set./dez. 2009a.
- CASTRO, Maria H. G. **Sistemas de avaliação da educação no Brasil: avanços e novos desafios**. São Paulo em Perspectiva, São Paulo, v. 23, n. 1, p. 5-18, jan./jun. 2009b.
- CEARÁ, Assembleia Legislativa do Estado. **Educação de Qualidade começando pelo Começo. Relatório Final do Comitê Cearense para a Eliminação do Analfabetismo Escolar**, 2006.
- CEARÁ. Secretaria de Estado da Educação. **Boletim Pedagógico – Língua Portuguesa, Alfabetização**. 2014.
- CEARÁ. Secretaria de Estado da Educação. **Boletim Pedagógico – Língua Portuguesa, Alfabetização**. 2015.
- CEARÁ. Secretaria de Estado da Educação. **Boletim do Professor – Língua Portuguesa, Alfabetização**. 2017.
- CÔCO, Dilza. **Avaliação externa da alfabetização: o PAEBES-ALFA no Espírito Santo**. Vitória, 2014. 404p. Tese de Doutorado – Programa de Pós-Graduação em Educação, Universidade Federal do Espírito Santo.
- COELHO, V. S.; COSTA, M. I. S.; SCHALCH, L. **Arranjos contratuais para a provisão de serviços públicos de saúde e o fortalecimento da regionalização**. 41º Encontro Anual da Anpocs. Caxambu, 2017. Disponível em: <https://www.anpocs.com/index.php/papers-40-encontro-2/gt-30/gt25-18/11057-arranjos-contratuais-para-a-provisao-de-servicos-publicos-de-saude-e-o-fortalecimento-da-regionalizacao/file>
- COOK-GUMPERZ, J. *Introdução: a construção social da alfabetização*. In: COOK-GUMPERZ, J. (Org). **A construção social da alfabetização**. 2ª ed. Porto Alegre: Artmed, 2008.

COSTA, Helen Cristina Vieira. **Mão guiando mãos: o trabalho pedagógico com a produção textual de crianças do 2 ano do ensino fundamental no município de Fortaleza.** Fortaleza, 2018, 119p. Dissertação de Mestrado - Programa de Pós-Graduação em Educação, Universidade Federal do Ceará.

COSTA, P. B. M. **A contribuição do Fome Zero na consolidação da política brasileira de segurança alimentar.** 41º Encontro Anual da Anpocs. Caxambu, 2017. Disponível em: <https://www.anpocs.com/index.php/papers-40-encontro-2/gt-30/gt25-18/10840-a-contribuicao-do-fome-zero-na-consolidacao-da-politica-brasileira-de-seguranca-alimentar/file>

COUTO, José Cláudio Diniz. **Descontinuidade das políticas públicas em educação: ações políticas, financiamento e alternância de poder.** Tese de Doutorado. Programa de Pós-Graduação em Educação da UMESP. São Paulo, 2017.

CULLEN, J.B; REBACK, Randall. *Tinkering toward accolades: School gaming under a performance accountability system.* In: GRONBERG, T. J; JANSEN, D. W. (Org). **Advanced in applied microeconomics**, 14. Elsevier, 2006.

CUNHA, Luiz Antônio. *As políticas educacionais entre o presidencialismo imperial e o presidencialismo de coalizão.* In: **Crise da escola e políticas educativas.** 2ª ed. Belo Horizonte: Autêntica Editora, 2013.

CURY, Carlos Roberto Jamil. **Direto à educação: direito à igualdade, direito à diferença.** Cadernos de Pesquisa, n. 116, p. 245-262, julho de 2002.

DANTAS, Larissa Martins. **Repercussões da avaliação externa sobre a prática docente: o caso do Sistema Permanente de Avaliação da Educação Básica do Ceará (Spaece).** Fortaleza, 2015, 148p. Dissertação de Mestrado – Programa de Pós-Graduação em Educação, Universidade Federal do Ceará.

DEUBEL, André-Nöel Roht. **Políticas públicas: formulación, implementación y evaluación.** Bogotá: Ediciones Aurora, 2009.

DIAS, Ana Paula de Moura Ferreira. **Os efeitos do Programa de Intervenção Pedagógica no ciclo de alfabetização: o desafio para que todos os alunos leiam e escrevam até os oito anos de idade.** Juiz de Fora/MG, 2015, 136p. Dissertação de Mestrado – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.

DICKEL, Adriana. **A Avaliação Nacional da Alfabetização no contexto do Sistema de Avaliação da Educação Básica e do Pacto Nacional pela Alfabetização na Idade Certa: responsabilização e controle.** Cadernos Cedes, Campinas, v. 36, n. 99, mai/ago, 2016.

DICKEL, Adriana. **O impacto do PISA na Produção Acadêmica Brasileira: contribuições para a discussão do currículo escola.** Educação: Teoria e Prática, vol. 20, n. 35, jul/dez, 2010.

DINIZ, Adriana Moreira. **O caderno de boas práticas dos diretores de escola de Minas Gerais, sua relação com a prática docente e as repercussões deste instrumento nos resultados do Proalfa.** Juiz de Fora/MG, 2013, 122p. Dissertação de Mestrado – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.

DOURADO, Luiz Fernandes. **Políticas e gestão da educação básica no Brasil: limites e perspectivas.** Educação e Sociedade. Campinas, vol. 28, n. 100. Edição Especial, p. 921-946, out, 2007.

DUBET, François. **O que é uma escola justa?** Cadernos de Pesquisa, v. 34, n. 123, p. 539-555, set./dez. 2004.

EIRÓ, Flávio. **O Programa Bolsa Família e os pobres “não-mercedores”: poder discricionário e os limites da consolidação de direitos sociais.** 41º Encontro Anual da Anpocs. Caxambu, 2017. Disponível em: <https://www.anpocs.com/index.php/papers-40-encontro-2/gt-30/gt25-18/10834-o-programa-bolsa-familia-e-os-pobres-nao-mercedores-poder-discricionario-e-os-limites-da-consolidacao-de-direitos-sociais/file>

ESPÍRITO SANTO, Secretaria de Estado da Educação. **Revista Pedagógica – Língua Portuguesa e Matemática, 3º ano do ensino fundamental.** 2014.

ESPÍRITO SANTO, Secretaria de Estado da Educação. **Revista Pedagógica – Língua Portuguesa e Matemática, 3º ano do ensino fundamental.** 2015.

ESPÍRITO SANTO, Secretaria de Estado da Educação. **Revista Pedagógica – Língua Portuguesa e Matemática, 3º ano do ensino fundamental.** 2017.

ESTEBAN, M. T. **Considerações sobre a política de avaliação da alfabetização: pensando a partir do cotidiano escolar.** Revista Brasileira de Educação, v. 17, n. 51, set/dez, 2012.

FARIA, Carlos Aurélio Pimenta de. **A política de avaliação de políticas públicas.** Revista Brasileira de Ciências Sociais, vol. 20, nº 59, 97-109, outubro de 2005.

FELDFEBER, Myriam. *Internacionalização da educação, “Tratados de Livre Comércio” e políticas educativas na América Latina.* In: FERREIRA, E. B.; OLIVEIRA, D. A. (Org). **Crise da escola e políticas educativas.** Belo Horizonte: Autêntica, 2013.

FERNANDES, Reynaldo. **Índice de Desenvolvimento da Educação Básica (Ideb).** Brasília: MEC/INEP. Textos para Discussão, Série Documental, n. 26, 2007.

FERREIRA, Eliza Bartolozzi. *Políticas educacionais no Brasil no tempo de crise.* In: FERREIRA, Eliza Bartolozzi; OLIVEIRA, Dalila Andrade (org.). **Crise da escola e políticas educativas.** 2ª ed. Belo Horizonte: Autêntica Editora, 2013.

FERREIRA, Eliza Bartolozzi; OLIVEIRA, Dalila Andrade (org.). **Crise da escola e políticas educativas.** 2ª ed. Belo Horizonte: Autêntica Editora, 2013.

FERREIRA, E. B.; OLIVEIRA, D. A. (Org). **Crise da escola e políticas educativas.** Belo Horizonte: Autêntica, 2013.

FERREIRO, Emilia. **Com todas as letras.** 3ª ed. São Paulo: Cortez, 1993.

FERREIRO, Emilia. **Reflexões sobre alfabetização.** 26ª ed. São Paulo: Cortez, 2011.

FERREIRO, Emilia; TEBEROSKY, Ana. **Psicogênese da língua escrita.** Porto Alegre: Artmed, 2007.

FILGUEIRAS, K. (2013). **PROALFA (MG): Avaliação da alfabetização, em larga escala, no Brasil.** Acolhendo a Alfabetização Nos Países De Língua Portuguesa, 7(13), 38-64.

FILHO, A. C. C.; LÍCIO, E. C.; BARTHOLO, L.; MESQUITA, A. C. S.; PASSOS, L. **Lições sobre a articulação federativa para enfrentamento da miséria a partir da experiência dos programas estaduais de transferência de renda condicionada no Plano Brasil Sem Miséria.** 41º Encontro Anual da Anpocs. Caxambu, 2017. Disponível em: <https://www.anpocs.com/index.php/papers-40-encontro-2/gt-30/gt25-18/11033-licoes-sobre-a-articulacao-federativa-para-enfrentamento-da-miseria-a-partir-da-experiencia-dos-programas-estaduais-de-transferencia-de-renda-condicionada-no-plano-brasil-sem-miseria-1/file>

FONSECA, Andreia Serra Azul da. **Avaliação do Programa de Alfabetização na Idade Certa – Paic: reflexos no planejamento e prática escolar**. Fortaleza, 2013, 125p. Tese de Doutorado – Programa de Pós-Graduação em Educação, Universidade Federal do Ceará.

FOUCAULT, Michel. **Vigiar e punir**. Petrópolis: Vozes, 1987.

FRANCO, Karla Oliveira; CALDERÓN, Adolfo Ignacio. **O Simave à luz das três gerações de avaliação da educação básica**. Estudos em Avaliação Educacional, vol.28, nº67, jan-abr, 2017.

FRANCO, Creso; ALVES, Fátima; BONAMINO, Alicia. **Qualidade do ensino fundamental: políticas, suas possibilidades, seus limites**. Educação e Sociedade, Campinas, vol. 28, n. 100 - Especial, p. 989-1014, out. 2007.

FRANCO, K. O.; CALDERÓN, A. I. **O Simave à luz das três gerações de avaliação da educação básica**. Estudos em Avaliação Educacional. São Paulo, v. 28, n. 67, jan/abr, 2017.

FREIRE, Lilian Rose S. Carvalho. **Saresp 2005: as vicissitudes da avaliação em uma escola da rede estadual**. Dissertação de Mestrado. Faculdade de Educação, Universidade de São Paulo, São Paulo, 2008.

FREIRE, Paulo. **Educação como prática da liberdade**. São Paulo: Paz e Terra, 2009.

FREITAG, Raquel Meister Ko, ALMEIDA, Ayane Nazarela Santos; ROSÁRIO, Mônica Maria Soares. **Contribuições para o aprimoramento da Provinha Brasil enquanto instrumento diagnóstico do nível de alfabetização e letramento nas séries iniciais**. Revista Brasileira de Estudos Pedagógicos, Brasília, v. 94, n. 237, p. 390-416, maio/ago. 2013.

FREITAS, Dirce Nei Teixeira de. *Avaliação da Educação Básica no Brasil: características e pressupostos*. In: BAUER, Adriana; GATTI, Bernardete A.; TAVARES, Marialva R. (Org.). **Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origens e pressupostos**. Vol.I. Florianópolis: Insular, 2013c.

FREITAS, Luiz C. *Caminhos da avaliação de sistemas educacionais no Brasil: o embate entre a cultura da auditoria e a cultura da avaliação*. In: BAUER, Adriana; GATTI, Bernardete A.; TAVARES, Marialva R. (Org.). **Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origens e pressupostos**. Vol.II. Florianópolis: Insular, 2013.

FREITAS, Luiz C. **Os reformadores empresariais da educação: da desmoralização do magistério à destruição do sistema público de educação**. Educação e Sociedade, Campinas, v. 33, n. 119, p. 379-404, abr.-jun. 2012.

FREITAS, Luiz C. **Políticas de responsabilização: entre a falta de evidência e a ética**. Cadernos de Pesquisa, vol. 43, n. 148, p. 348-365, jan/abr, 2013b.

FREITAS, Rosângela Maia de. **O Prêmio Escola Nota Dez e o regime de cooperação estabelecido entre dois municípios cearenses**. Mossoró/RN, 2018, 184p. Dissertação de Mestrado – Programa de Pós-Graduação da Universidade Estadual do Rio Grande do Norte.

FREY, Klaus. **Políticas públicas: um debate conceitual e reflexões referentes à prática da análise de políticas públicas no Brasil**. Revista Planejamento e políticas públicas, Brasília, nº 21, p. 212-259, jun 2000.

FRIGOTTO, Gaudêncio. *Política e gestão educacional na contemporaneidade*. In: FERREIRA, E. B.; OLIVEIRA, D. A. (Org). **Crise da escola e políticas educativas**. Belo Horizonte: Autêntica, 2013.

- GATTI, B. A. *Possibilidades e fundamentos de avaliações em larga escala: primórdios e perspectivas contemporâneas*. In: BAUER, Adriana; GATTI, Bernardete A.; TAVARES, Marialva R. (Org.). **Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origens e pressupostos**. Vol.I. Florianópolis: Insular, 2013.
- GATTI, B. A.; BARRETO, E. S. S.; ANDRÉ, M. E. D. A. **Políticas docentes no Brasil: um estado da arte**. Brasília: UNESCO, 2011.
- GIDDENS, Anthony. **A constituição social da sociedade**. São Paulo: Martins Fontes, 2009.
- GONTIJO, C. M. M. **Alfabetização: políticas mundiais e movimentos nacionais**. Campinas: Autores Associados, 2014.
- GOULART, Cecília. *A avaliação da alfabetização em pesquisas, práticas e políticas públicas: debatendo posições teórico-metodológicas*. In: MORTATTI, M. R. L.; FRADE, I. C. A. S. (Org.). **Alfabetização e seus sentidos: o que sabemos, fazemos e queremos?** São Paulo: Unesp, 2014.
- HABERMAS, Jürgen. **Pensamento pós-metafísico**. Rio de Janeiro: Tempo Brasileiro, 1990.
- HABERMAS, Jürgen. **Teoria do agir comunicativo**. São Paulo: Martins Fontes, 2012.
- HAGOPIAN, Jesse (Ed.). **More than a score: the new uprising against high-stakes testing**. Chicago: Haymarket Books, 2014.
- HANUSHEK, E.; RAYMOND, M. **Does school accountability lead to improved student performance?** Journal of Policy Analysis and Management, New York, v. 24, n. 2, p. 297-327, 2005.
- IRFFI, G. D.; PETTERINI, F. C. **Avaliando o impacto da mudança da lei do ICMS do Ceará nos indicadores municipais de educação e saúde**. In: VII Encontro Economia do Ceará em Debate 2011, Fortaleza.
- JACOB, Brian. **Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago public schools**. Journal of Public Economics, v. 89, n.5-6, 761-796, 2005.
- JANUZZI, Paulo de Martino. **Monitoramento e avaliação de programas sociais: uma introdução aos conceitos e técnicas**. Campinas: Alínea Editora, 2016.
- JÚNIOR, Ruy Barboza de Oliveira. **As políticas educacionais do segundo mandato do governo Paulo Hartung e seus impactos no trabalho docente**. Vitória, 2013, 148p. Dissertação de Mestrado – Programa de Pós-Graduação em Educação, Universidade Federal do Espírito Santo.
- KLEIN, Ruben. **Utilização da Teoria da Resposta ao Item no Sistema Nacional de Avaliação da Educação Básica (Saeb)**. Revista Ensaio, n. 40, v. 11, 0. 283-296, jul/set, 2003.
- KLEIN, Ruben; FONTANIVE, Nilma Santos. **Avaliação em larga escala: uma proposta inovadora**. Em Aberto, Brasília, ano 15, n.66, abr./jun., 1995.
- KRAWCZIK, Nora. *A construção social das políticas educacionais no Brasil e na América Latina*. In: **O cenário educacional latino-americano no limiar do século XXI**. Campinas: Autores Associados, 2000.
- LEAL, Telma Ferraz; MORAIS, Artur Gomes de; PESSOA, Ana Cláudia Rodrigues Gonçalves; NASCIMENTO, Julliane Campelo do. **Habilidades de compreensão leitora: seu ensino e sua avaliação pela Provinha Brasil**. Revista Brasileira de Educação, vol. 22, nº 68, jan-mar, 2017.



- LIBÂNEO, José Carlos. **Didática**. São Paulo: Cortez, 1994.
- LIMA, Alessio Costa. **Ciclo de avaliação da educação básica do Ceará: principais resultados**. Estudos em Avaliação Educacional, vol. 23, nº53, set-dez, 2012.
- LIMA, Alessio Costa; PEQUENO, Maria Iaci Calvancante; MELO, Maria Noraelena Rabelo. **Avaliação da alfabetização no Ceará: principais resultados da primeira edição do Spaece-Alfa**. Estudos em Avaliação Educacional, v. 19, n. 41, set/dez, 2008.
- LIMA, Diva. **O Sistema Permanente de Avaliação da Educação Básica do Ceará (Spaece) e sua influência sobre a gestão pedagógica de uma escola de ensino médio situada no município de Tauá – Ceará**. Fortaleza, 2015, 147p. Dissertação de Mestrado - Programa de Pós-Graduação em Educação, Universidade Federal do Ceará.
- LIMA, L. C.; AFONSO, A. J. **Reformas da educação pública: democratização, modernização, neoliberalismo**. Porto: Afrontamento, 2002.
- LIMA, Naira da Costa Muylaert. **Educação e desigualdades sociais: suas relações no estado do Espírito Santo**. Rio de Janeiro, 2016. 238p. Tese de Doutorado – Departamento de Educação, Pontifícia Universidade Católica do Rio de Janeiro.
- LÜCK, Heloísa. **Dimensões da gestão escolar e suas competências**. Curitiba: editora Positivo, 2009.
- LUCKESI, Cipriano. **Avaliação da aprendizagem escolar**. 21ª ed. São Paulo: Cortez, 2010.
- LUHMANN, Niklas. **Introdução à teoria dos sistemas**. Petrópolis: Vozes, 2010.
- MACHADO, Cristiane; ALAVARSE, Ocimar Munhoz. **Qualidade das escolas: tensões e potencialidades das avaliações externas**. Educação & Realidade, Porto Alegre, v. 39, n. 2, p. 413-436, abr./jun. 2014.
- MACHADO, C.; ALAVARSE, O. M.; ARCAS, P. H. **Sistemas estaduais de avaliação: interfaces com qualidade e gestão da educação**. RBPAE, vol. 31, n. 3, p.667-680, set/dez, 2015.
- MACIEL, Linéia Moreira. **O uso do Spaece pela secretaria municipal de educação de Quixadá: desafios e possibilidades**. Juiz de Fora/MG, 2013, 116p. Dissertação de Mestrado – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.
- MAIA, Cláudio Machado & GRANDO, Johnny Luiz. **Metodologia para definição de tipologias de políticas públicas, para entender a elaboração, implementação e avaliação**. Revista do Desenvolvimento Regional – Faccat – Taquara/RS – v.13, n.2, jul/dez. 2016.
- MAIA, Rita de Cássia Costa. **O desempenho de escolas da rede estadual no município de Patrocínio/MG no Proalfa: uma análise dos fatores afetos à gestão escolar de quatro casos comparados**. Juiz de Fora/MG, 2013, 112p. Dissertação de Mestrado – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.
- MAINARDES, Jefferson. **Abordagem do ciclo de políticas: uma contribuição para a análise de políticas educacionais**. Educ. Soc. Campinas, vol. 27, n. 94, p. 47-69, jan/abr. 2006.
- MARQUES, Cláudio de Albuquerque; RIBEIRO, Ana Paula de Medeiros; CIASCA, Maria Isabel Filgueiras Lima. **Paic: o pioneirismo no processo de avaliação municipal com autonomia**. Estudos em Avaliação Educacional, vol. 19, nº41, set-dez, 2008.

MARQUES, Cláudio de Albuquerque; AGUIAR, Rui Rodrigues; CAMPOS, Márcia Oliveira Cavalcante. **Programa Alfabetização na Idade Certa: concepções, primeiros resultados e perspectivas**. Estudos em Avaliação Educacional, vol. 20, nº43, mai-ago, 2009.

MARSHALL, Thomas. **Cidadania, classe social e status**. Rio de Janeiro: Zahar, 1967.

MARTINS, Ana Cláudia Osório. **Proalfa: avaliação e propostas pedagógicas**. Juiz de Fora/MG, 2013a, 316p. Dissertação de Mestrado – Programa de Pós-Graduação em Educação, Universidade Federal de Juiz de Fora.

MARTINS, A.M. **O campo das políticas públicas de educação: uma revisão de literatura**. Est. Aval. Educ. São Paulo, vol. 24, n. 56, p. 276 – 299, set/dez 2013.

MATOS, Ana Paula Pequeno. **Um estudo de caso sobre a gestão escolar no contexto dos resultados de matemática e português do Spaece**. Juiz de Fora/MG, 2017, 157p. Dissertação de Mestrado – Mestrado Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.

MAUÉS, Olgaíses Cabral. *O papel da escola na construção dos saberes e os limites da noção de competências*. In: FERREIRA, Eliza Bartolozzi; OLIVEIRA, Dalila Andrade (org.). **Crise da escola e políticas educativas**. 2ª ed. Belo Horizonte: Autêntica Editora, 2013.

MCNELLY, C. L. **Prescribing National Educational Policies: The Role of International Organizations**. University of Chicago. Comparative Educational Review. Vo. 39, n. 4, 1995.

MELLO, Guiomar Namó de. **Políticas públicas de educação**. Estudos Avançados, n. 5(13), 1991.

MINAS GERAIS, Secretaria de Estado da Educação. **Boletim Pedagógico – Proalfa 2009**. Disponível em : [http://www.simave.caedufjf.net/wp-content/uploads/2012/06/BOLETIM\\_PEDAGOGICO\\_PROALFA\\_2009.pdf](http://www.simave.caedufjf.net/wp-content/uploads/2012/06/BOLETIM_PEDAGOGICO_PROALFA_2009.pdf)

MINAS GERAIS, Secretaria de Estado da Educação. **Revista Pedagógica – Proalfa, 2012**.

MINAS GERAIS, Secretaria de Estado da Educação. **Revista Pedagógica – Proalfa, 2013**.

MINHOTO, Maria Angélica Pedra. **Impactos da “educação de resultados” no ofício docente**. In: XVI ENDIPE – Encontro Nacional de Didática e Práticas de Ensino. Unicamp, Campinas, 2012.

MORAES, Denise Alessandra Palhares Diniz. **Gestão escolar eficaz: o diferencial de uma escola em contexto de vulnerabilidade social**. Juiz de Fora/MG, 2014, 96p. Dissertação de Mestrado – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.

MORAIS, A. G. **Políticas de avaliação da alfabetização: discutindo a Provinha Brasil**. Revista Brasileira de Educação, v. 17, n. 51, set/dez, 2012.

MORAIS, A. G. *Precisamos de boas políticas públicas de avaliação da alfabetização: análise das razões de tal necessidade e de fatores que impedem que avancemos no cumprimento dessa republicana tarefa*. In: MORTATTI, M. R. L.; FRADE, I. C. A. S. (Org). **Alfabetização e seus sentidos: o que sabemos, fazemos e queremos?** São Paulo: Unesp. 2014.

MORAIS, A. G.; LEAL, T. F.; ALBUQUERQUE, E. B. C. **“Provinha Brasil”: monitoramento da aprendizagem e formulação de políticas educacionais**. Revista Brasileira de Política e Administração da Educação, Porto Alegre, vol. 25, n. 2, mai/ago, 2009.

MORTATTI, Maria do Rosário Longo (org). **Alfabetização no Brasil: uma história de sua história**. Marília: Cultura Acadêmica Editora, 2011.

MORTATTI, Maria do Rosário Longo. **Um balanço crítico da “década da alfabetização” no Brasil.** Caderno Cedes, Campinas, v. 33, n. 89, jan/abr, 2013.

MOTA, Maria Océlia. **Entre a meritocracia e a equidade: o Prêmio Escola Nota Dez na percepção e atuação de agentes implementadores.** Rio de Janeiro, 2018, 297p. Tese de Doutorado – Programa de Pós-Graduação em Educação, Pontifícia Universidade Católica do Rio de Janeiro.

NUNES, Graciele Aparecida. **Análise da implementação do Programa de Intervenção Pedagógica em uma superintendência regional de ensino (SRE/MG).** Juiz de Fora/MG, 2015, 106p. Dissertação de Mestrado – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.

OLIVEIRA, Ana Paula de Matos. **A Prova Brasil como política de regulação da rede pública do Distrito Federal.** Dissertação de Mestrado. Faculdade de Educação, Universidade de Brasília, Brasília/DF, 2011.

OLIVEIRA, Dalila Andrade. **Educação básica: gestão do trabalho e da pobreza.** Petrópolis: Vozes, 2000.

OLIVEIRA, Dalila Andrade. *Política educativa, crise da escola e a promoção de justiça social.* In: FERREIRA, E. B.; OLIVEIRA, D. A. (Org). **Crise da escola e políticas educativas.** Belo Horizonte: Autêntica, 2013.

OLIVEIRA, Dayane Louzada de. **Análise dos resultados do PAEBES de Ciências da Natureza do município de Marataízes-ES e seu reflexo na prática docente.** Alegre/ Espírito Santo, 2018. 111p. Dissertação de Mestrado – Centro de Ciências Exatas, Naturais e da Saúde, Universidade Federal do Espírito Santo.

OLIVEIRA, Diana Nara da Silva. **A educação do/no campo em escola de assentamento de reforma agrária em Jaguaruana/Ceará.** Limoeiro do Norte/CE, 2017a, 141p. Dissertação de Mestrado – Mestrado em Educação e Ensino, Universidade Estadual do Ceará.

OLIVEIRA, Nádia Ramos Grisson de. **Análise de ações escolares e resultados do Proalfa: um estudo de caso de uma escola de sucesso em Governador Valadares.** Juiz de Fora/Minas Gerais, 2017b, 116p. Dissertação de Mestrado – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.

OLIVEIRA, R. P. **IDEB e trabalho pedagógico da escola: uma articulação possível?** Revista Escola Pública, São Paulo, v. especial, p. 76-92, 2011b.

PARO, V. H. **Diretor escolar: educador ou gerente.** São Paulo: Cortez Editora, 2015.

PARO, V. H. **Gestão democrática da escola pública.** 4a. São Paulo: Cortez Editora, 2017.

PASQUALI, Luiz; PRIMI, Ricardo. **Fundamentos da Teoria da Resposta ao Item – TRI.** Avaliação psicológica. 2003, vol.2, n.2, pp. 99-110.

PAULA, Patrícia de Lima. **A influência da elaboração de itens na eficácia das avaliações em larga escala.** Juiz de Fora/Minas Gerais, 2018, 100p. Dissertação de Mestrado – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.

PERBONI, Edilene Mizaél de Carvalho. **Os sujeitos das políticas de avaliação sistêmica da alfabetização em Minas Gerais: uma analítica à luz da governamentalidade.** Itatiba/SP, 2015, 188p. Tese de Doutorado – Programa de Pós-Graduação em Educação, Universidade São Francisco.

PEREIRA, Flávio Alves. **A incidência das políticas de responsabilização do estado do Ceará nas ações de gestão pedagógica em âmbito municipal - o caso da Escola Maria Nair (IPU-CE)**. Juiz de Fora/MG, 2014a, 233p. Dissertação de Mestrado - Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora

PEREIRA, Selma Lúcia de Assis. **PAEBES: modos, formas e diálogos a partir dos usos dos resultados em língua portuguesa da avaliação externa estadual no município de Cariacica-ES**. Vitória, 2014b. 233p. Tese de Doutorado – Departamento de Educação, Universidade Federal do Espírito Santo.

PILATI, Orlando. **Sistema Nacional de Avaliação da Educação Básica (Saeb)**. Ensaio: Avaliação e Políticas Públicas em Educação, Rio de Janeiro, v. 2, n. 1, p. 11-30, out/dez, 1994.

POLON, Thelma Lucia Pinto. **Identificação dos Perfis de Liderança e Características relacionadas à Gestão Pedagógica Eficaz nas escolas participantes do Projeto Geres - Estudo Longitudinal Geração escolar 2005 - Pólo Rio de Janeiro**. Tese de Doutorado. Programa de Pós-Graduação em Educação. Pontifícia Universidade Católica do Rio de Janeiro. 2009.

PORTELA, Maíra Miranda. **Alfabetização matemática: um paralelo entre a avaliação nacional e cenários estaduais**. Juiz de Fora/ Minas Gerais, 2016, 118p. Dissertação de Mestrado – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.

PRADO, Maria Lúcia Viana do. **Análise dos cursos de capacitação para divulgação e apropriação de resultados do Simave/Proalfa e Proeb da Superintendência Regional de Teófilo Otoni/Minas Gerais**. Juiz de Fora/ Minas Gerais, 2014, 135p. Dissertação de Mestrado – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.

PRIMI, Ricardo. **Psicometria: fundamentos matemáticos da Teoria Clássica dos Testes**. Avaliação Psicológica, 11(2), pp. 297-307, 2012.

PRZEWORSKY, Adam. **Estado e economia no capitalismo**. Rio de Janeiro: Relume-Dumará, 1995.

REIS, Carla Cândida da Silva. **Da agenda à implementação: um olhar sobre as dinâmicas de interação dos agentes implementadores do Programa de Intervenção Pedagógica/Alfabetização na Idade Certa no estado de Minas Gerais**. Juiz de Fora/MG, 2013, 115p. Dissertação de Mestrado – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.

REIS, Thaize de Souza; SOUZA, Deisy das Graças de; ROSE, Júlio César de. **Avaliação de um programa para o ensino de leitura e escrita**. Estudos em Avaliação Educacional, vol. 20, nº44, set-dez, 2009.

REZENDE, Wagner Silveira. **A influência do contexto normativo do clima escolar no desempenho estudantil do Seape 2012**. Cadernos Cenpec, São Paulo, v.6, n.1, p.47-74, jan./jun. 2016.

REZENDE, Wagner Silveira. **Alguns problemas de interpretação sobre a natureza da avaliação em larga escala**. Pesquisa e Debate em Educação, v.4, n.2, 2014.

RIANI, Juliane de Lucena Ruas; SILVA, Vânia Cândida da; SOARES, Tufi Machado. **Análise da avaliação da alfabetização de Minas Gerais: evolução e desigualdade**. Estudos em Avaliação Educacional, vol. 23, nº53, set-dez, 2012.

RIBEIRO, Sérgio Costa. **A pedagogia da repetência**. Estudos Avançados. Vol. 5, nº 12, mai/ago, São Paulo, 1991.

RIBEIRO, Vera Masagão; FONSECA, Maria da Conceição Ferreira Reis. **Matriz de referência para a medição do alfabetismo nos domínios do letramento e do numeramento**. Estudos em Avaliação Educacional, vol. 21, nº45, jan-abr, 2010.

RIBEIRO, Vera Masagão; SOARES, Tufi Machado. **Construção de escala combinada para medição do alfabetismo em contexto não escolar**. Estudos em Avaliação Educacional, vol. 19, nº41, set-dez, 2008.

ROCHA, Gladys; MARTINS, Raquel Fontes. **Construção de um corpus de escrita infantil com itens de avaliações**. Estudos em Avaliação Educacional, vol. 23, nº51, jan-abr, 2012.

RODRIGUES, Suely da Silva. *Avaliação em larga escala e apropriação de resultados educacionais*. In: **Diálogos e proposições, Vol. II: Planos de Ação para a rede estadual de ensino de Minas Gerais**. Editora CAEd/UFJF: Juiz de Fora, 2017.

ROSA, Zely Mattos. **A gestão de resultados e a avaliação em larga escala no cotidiano de uma escola da rede municipal de Juiz de Fora**. Juiz de Fora/MG, 2013, 80p. Dissertação de Mestrado – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.

SALOMÃO, Rommy. **A formação continuada de professores alfabetizadores: do Pró-Letramento ao Pnaic**. 117 fl. Dissertação de Mestrado. Programa de Pós-Graduação em Educação, Universidade Estadual de Ponta Grossa, Ponta Grossa/PR, 2014.

SANTOS, D.; PRIMI R. **Desenvolvimento socioemocional e aprendizado escolar: Uma proposta de mensuração para apoiar políticas públicas**. Educação para o Século XXI. São Paulo, 2014. Disponível em: <http://educacaoosec21.org.br/wp-content/uploads/2013/07/desenvolvimento-socioemocional-e-aprendizado-escolar.pdf>

SARAMAGO, José. **Ensaio sobre a cegueira**. São Paulo: Companhia das Letras, 2006.

SARTES, L. M. A; SOUZA-FORMIGONI, M. L. O. **Avanços na Psicometria: da Teoria Clássica dos Testes à Teoria da Resposta ao Item**. Psicologia: Reflexão e Crítica, 26(2), 241-250, 2013.

SAUSSEZ, F.; LESSARD, C. *A educação baseada na prova. De que se trata? Quais são suas implicações?* In: FERREIRA, E. B.; OLIVEIRA, D. A. (Org). **Crise da escola e políticas educativas**. Belo Horizonte: Autêntica, 2013.

SAUSSURE, Ferdinand de. **Curso de linguística geral**. 27ª ed. São Paulo: Cultrix, 2006.

SAVIANI, D. **Escola e democracia**. Campinas: Editora Autores Associados, 2008.

SAVIANI, D. **O Plano de Desenvolvimento da Educação: análise do projeto do MEC**. Educação e Sociedade, Campinas, v. 28, n. 10, 2007.

SARAVIA, Enrique. **Introdução à teoria da política pública**. Brasília: ENAP, 2006.

SEGATTO, I. C. **Como ideias se transformam em reformas: um estudo comparativo das mudanças educacionais orientadas pelo desempenho nos estados brasileiros**. 2011. Dissertação (Mestrado em Administração Pública) – Escola de Administração de Empresas de São Paulo, São Paulo, 2011.

SILVA, Dogival Alencar de. **Nível do desenvolvimento do pensamento formal e desempenho em Português e Matemática no ensino médio**. Fortaleza, 2013a, 131p.

Dissertação de Mestrado – Programa de Pós-Graduação em Educação, Universidade Federal do Ceará.

SILVA, Josiane Toledo Ferreira. **A escrita na avaliação da alfabetização em larga escala**. Rio de Janeiro, 2013b, 136p. Tese de Doutorado – Programa de Pós-Graduação em Educação, Pontifícia Universidade Católica do Rio de Janeiro.

SILVA, Lucas Melgaço da. **Avaliações em larga escala na alfabetização: contextos no ensino público de um município do estado do Ceará**. Fortaleza, 2016a, 98p. Dissertação de Mestrado – Programa de Pós-Graduação em Educação, Universidade Federal do Ceará.

SILVA, Olívia Coelho da. **Formação do leitor: desafio à prática docente e à avaliação da aprendizagem**. Fortaleza, 2016b, 202p. Dissertação de Mestrado – Programa de Pós-Graduação em Educação, Universidade Federal do Ceará.

SILVA, Patrícia Aparecida da. **As influências do Sistema Mineiro de Avaliação na construção da qualidade da educação**. Uberlândia, 2016c, 237p. Dissertação de Mestrado – Programa de Pós-Graduação em Educação da UFU.

SILVA, Roberto Cláudio Bento. **Apropriação dos resultados do Spaece pelos gestores escolares: um estudo de caso envolvendo duas escolas do interior do Ceará**. Juiz de Fora/MG, 2014, 178 p. Dissertação de Mestrado – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.

SILVA, T. T. **Avaliação da alfabetização: um exame de diferentes edições da Provinha Brasil**. Dissertação de Mestrado. 128 fl. Programa de Pós-Graduação em Educação. Universidade Federal de Pernambuco. Recife, 2013.

SILVA, Vandrê Gomes da; CARVALHO, Cynthia Paes de. **Usos e efeitos das avaliações externas como objeto de pesquisa**. Estudos em Avaliação Educacional. Apresentação ao dossiê. São Paulo, v.15, n.59, p. 12-21, set/dez, 2014.

SIMÕES, Maria Inez Barroso. **O Programa de Intervenção Pedagógica do Estado de Minas Gerais – PIP**. Dissertação de Mestrado. 107p. Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública. Universidade Federal de Juiz de Fora, 2012.

SOARES, Erineuda do Amaral. **Sistema Permanente de Avaliação da Educação Básica do Ceará – Alfabetização: estratégias de acompanhamento de estudantes no contexto da prática**. Porto Alegre, 2016b, 136p. Dissertação de mestrado - Universidade do Vale do Rio dos Sinos.

SOARES, José Francisco. **A avaliação como instrumento de garantia do direito à educação** (entrevista). Cadernos Cenpec. São Paulo, v. 2, n.1, pp. 183-213, julho de 2012.

SOARES, José Francisco; ALVES, Maria Teresa Gonzaga. **Desigualdades raciais no sistema brasileiro de educação básica**. Educação e Pesquisa, São Paulo, v.29, n.1, p. 147-165, jan./jun. 2003.

SOARES, Magda. **Alfabetização: a questão dos métodos**. São Paulo: Contexto, 2016.

SOARES, Magda. **Alfabetização e letramento**. 5ª ed. São Paulo: Contexto, 2007.

SOARES, Magda. **Letramento e alfabetização: as muitas facetas**. Revista Brasileira de Educação. Rio de Janeiro, Apend; Autores Associados, Campinas, n. 25, 2004.

SOARES, Silvana Mendes Sabino. **O trabalho pedagógico com a leitura na educação infantil e no ensino fundamental: investigações sobre ensino e avaliação**. Fortaleza, 2017,

286p. Dissertação de Mestrado - Programa de Pós-Graduação em Educação, Universidade Federal do Ceará.

SOARES, Tufi Machado. **Modelo de três níveis hierárquicos para a proficiência dos alunos de 4a série avaliados no teste de língua portuguesa do SIMAVE/PROEB-2002**. Revista Brasileira de Educação, mai-ago, nº 29, 2005.

SOUSA, Francisco Jucivânio Félix de. **Reflexões sobre a apropriação de resultados de matemática do Spaece – o caso da escola do Jabuti**. Juiz de Fora/MG, 2016, 140p. Dissertação de Mestrado – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.

SOUSA, S. Z; OLIVEIRA, R. 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, Celina. **Políticas públicas: conceitos, tipologias e sub-áreas**. Disponível em <http://professor.pucgoias.edu.br/SiteDocente/admin/arquivosUpload/3843/material/001-%20A-%20POLITICAS%20PUBLICAS.pdf>. 2002.

SOUZA, Celina. **Políticas públicas: questões temáticas e de pesquisa**. Caderno CRH, Salvador, nº 39, p. 11-24, jul/dez 2003.

SOUZA, Celina. **Políticas públicas: uma revisão de literatura**. Sociologias, Porto Alegre, ano 8, nº16, p. 20-45, jul/dez 2006.

SOUZA, Maria Alba de. *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, Bernardete A.; TAVARES, Marialva R. (Org.). **Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: origens e pressupostos**. Vol. I. Florianópolis: Insular, 2013.

SPINK, Peter K. **Continuidade e descontinuidade em organizações públicas: um paradoxo democrático**. Cadernos Fundap, São Paulo, Ano 7, nº 13, p. 57-65, abril de 1987.

STOER, Stephen R; MAGALHÃES, António M. **The reconfiguration of the modern social contract: new forms of citizenship and education**. European Educational Research Journal, 1(4), p. 692-704, 2002.

TEDESCO, J. C. Prólogo. In: UNESCO. **Evaluar las evaluaciones: una mirada política acerca de las evaluaciones de la calidad educativa**. Buenos Aires: IPE/Unesco, p. 11-14, 2003.

TEIXEIRA, Renata Mota. **O Programa de Intervenção Pedagógica em duas escolas estaduais de Juiz de Fora: sucessos e entraves**. Juiz de Fora/MG, 2013, 104p. Dissertação de Mestrado – Programa de Pós-Graduação Profissional em Gestão e Avaliação da Educação Pública, Universidade Federal de Juiz de Fora.

TRINDADE, Iole Maria Faviero. *Caminhos e descaminhos investigativos na área de alfabetização*. In: MORTATTI, Maria do Rosário Longo (org). **Alfabetização no Brasil: uma história de sua história**. Marília: Cultura Acadêmica Editora, 2011.

UNESCO. **Educação para todos: o compromisso de Dakar**. Brasília: Unesco, Consed, Ação Educativa, 2001.

UNESCO; MINISTÉRIO DA EDUCAÇÃO. **Alfabetização como liberdade**. Brasília: MEC, 2003.

VIANNA, Heraldo Marelím. **Avaliações nacionais em larga escala: análises e propostas**. Estudos em Avaliação Educacional. São Paulo, n. 27, p. 41-76, jan/jul, 2003.

VIEIRA, Evaldo A. **As políticas sociais e os direitos sociais no Brasil: avanços e retrocessos.** Revista Serviço Social & Sociedade, 18 (53), p. 67 – 73, 1997.

VILLAS BOAS, Benigna Maria de Freitas. **Avaliação no Bloco Inicial de Alfabetização no Distrito Federal.** Estudos em Avaliação Educacional, vol. 18, nº 36, jan-abr, 2007.

WEBER, Max. *A objetividade do conhecimento nas ciências sociais.* In: COHN, Gabriel (Org.). FERNANDES, Florestan (Coord.). **Weber – Sociologia.** Coleção Grandes Cientistas Sociais, 13. São Paulo: Ática, 1999, p. 79-127.

ZANCHETTA JUNIOR, J. **Práticas de leitura literária e a contribuição do Pnaic.** Revista Brasileira de Educação, vol. 22, n. 68, jan/mar, 2017.



## **APÊNDICE A - Sobre três pontos relativos às pesquisas conduzidas no âmbito da *policy science*: conflito, descentralização e análise de programas específicos**

Em relação ao conflito, vale notar que, em todas as fases da política, de sua elaboração à sua execução, o registro é o de que as decisões tomadas não se baseiam em um processo linear de raciocínio e reflexão individuais, sem atritos e contratempos. O enfoque no conflito afirma que diferentes grupos de interesse disputam espaço para fazer com que seus interesses sejam atendidos através das políticas públicas a serem implementadas. As políticas públicas, assim, seriam objetos perpassados por disputas de poder.

Essa perspectiva tem o mérito de perceber a realidade social de forma muito mais realista e menos idealizada do que a leitura proposta pelo racionalismo estrito, que caracterizou os anos iniciais da *policy science*. Ela chama a atenção para a insuficiência da explicação do modelo estritamente racionalista, para o qual as decisões sobre políticas públicas deveriam ser analisadas tendo como foco o processo racional de escolha por parte dos tomadores de decisão. Nesse registro, o que importa como fator de explicação para a decisão são as limitações de informação impostas ao decisor, sem que outros atores sejam considerados como parte importante desse processo.

Não há dúvidas, portanto, acerca da pertinência e da importância que a dimensão do conflito tem para a análise das políticas públicas. De fato, a *policy arena* é um contexto conflituoso, no qual diferentes grupos estabelecem relações de disputa e apoio eventual, concorrendo por recursos escassos (particularmente, pelo direito de dizer a agenda). Mesmo as alianças partidárias podem ser momentaneamente suspensas para que alianças no âmbito da *policy arena* sejam estabelecidas. Ao lançar luz sobre os conflitos, essa leitura das políticas públicas descobre a influência que diferentes atores possuem nesse processo, tirando o enfoque limitado apenas nos tomadores de decisão. Mesmo as expectativas da sociedade exercem algum grau de constrangimento aos decisores, visto que sua quebra tem custos políticos.

Se sua importância é reconhecida para a análise das políticas, o excesso de atenção que recebe pode levar a problemas de interpretação, comprometendo a pesquisa. Nesse cenário, os conflitos seriam elementos capazes de explicar tudo o que diz respeito às políticas públicas, sendo alçados ao status de premissas, que, portanto, não precisam de explicações. Com isso, passa-se da necessidade de explicar como os conflitos operam para utilizar os conflitos como explicação de todo o resto. Isso faz obscurecer, por exemplo, o papel dos consensos na definição das políticas. Mesmo num cenário pautado por disputas, há um grau de consenso necessário

para que uma política seja definida e implementada. Uma análise completa do processo de definição das políticas deve levar em consideração tanto os conflitos quanto a produção de consensos.

Além disso, uma concentração analítica excessiva nas disputas pode levar a leituras enviesadas e maniqueístas sobre o papel que os diferentes atores desempenham no processo de definição e implementação das políticas. De acordo com Arretche (2000), políticas públicas são elaboradas através de um processo político, com grupos sociais exercendo pressões sobre os formuladores da política, buscando o atendimento de seus interesses próprios, principalmente no que diz respeito à formulação da agenda. Muitas vezes, essa perspectiva pode levar a interpretações radicalizadas e polarizadas, que transformam os tomadores de decisão em marionetes nas mãos de grupos econômicos privilegiados. Nesse sentido, o enfoque no conflito pode levar a uma leitura semelhante, em termos de polarização, ao racionalismo que critica. Se o racionalismo atribuía quase que exclusivamente aos agentes governamentais o papel de tomada de decisão, o enfoque excessivo nos conflitos pode negar a eles qualquer importância. Leituras desse tipo tendem a ver o Estado como um mero reflexo das relações de produção, sendo um conjunto de estruturas instrumentalizadas pelos grupos econômicos mais poderosos.

O Estado não é um mero reflexo das pressões das elites sociais, como pode fazer sugerir uma versão vulgar do marxismo (Souza fala em estruturalismo e funcionalismo, 2002, p. 6). Mesmo que não se considere o tema das classes sociais, pode-se supor que o Estado está a serviço de uma pluralidade de interesses difusos, que disputam espaços de influência nas decisões do governo (o pluralismo), leitura que também desconsidera o papel do próprio governo. Contudo, o Estado possui, também, seus próprios interesses. Da mesma forma, não é possível negar que tais grupos possuem capacidade de influência, levando à conclusão de que o governo não toma decisões isoladamente, sem considerar as intenções desses grupos, conforme podem fazer sugerir versões mais radicais do que Souza (idem) chama de elitismo. O Estado possui autonomia própria, que é relativa, e é, simultaneamente, influenciado por diferentes grupos de interesse (conforme lição de PRZEWORSKI, 1995).

Atrelada a esse tema, encontra-se a discussão sobre a crise da capacidade de governar do Estado, acompanhada do debate em torno da descentralização. Reconhecer a influência de determinados grupos não governamentais em relação à definição das políticas públicas não significa reconhecer que a capacidade dos governos de formular e implementar políticas públicas, e mesmo de governar, tenha diminuído.

Embora as políticas públicas possam incidir sobre a esfera privada, elas não são privadas. Mesmo quando as entidades privadas participam efetivamente

na sua elaboração e implementação, a participação está amparada em uma decisão pública, tomada por agentes governamentais embasados na autoridade do Estado. Conceitualmente, uma política pública demanda uma decisão política. Na prática, a maioria das decisões políticas não se transforma em uma política pública. (MAIA & GRANDO, 2016, p. 175).

Ou seja, mesmo quando a execução de um determinado programa é realizada por alguma entidade privada, a natureza de sua ação permanece pública, pois oriunda de uma decisão pública inicial (de enfrentar determinado problema através de uma política pública e de transferir a ação para a uma entidade privada). De todo modo, o território da descentralização no âmbito das políticas públicas está longe de ser consensual. As críticas são destinadas, particularmente, à baixa capacidade que as políticas descentralizadas têm de lidar com o tema da redução das desigualdades. A essa crítica juntam-se outras, dirigidas aos modelos sociais e políticos com base no neoliberalismo. A perspectiva trazida pelo gerencialismo é vista como uma de suas vertentes, fazendo com que suas propostas sejam vistas também com desconfiança. É o caso, por exemplo, das consultorias e assessorias contratadas pelos governos sob a defesa da busca pela efetividade.

A análise das políticas, com base no gerencialismo, é envolvida pelo tema da otimização, com foco na relação custo/benefício, monitorando os benefícios experimentados pelos usuários e buscando, quando a situação assim o exigir, orientar a tomada de decisão no que tange à necessidade de ajustes. A crítica, nesse caso, se destina ao estreitamento do entendimento das políticas públicas que tal enfoque revela. A análise das políticas deveria levar em consideração outros elementos:

Entretanto, não considerar elementos referentes aos arranjos institucionais (conflitos e consensos que pautaram aquela determinada política), aos atores envolvidos (associações, fóruns, conferências, conselhos, órgãos de governo que representam nichos de interesses) e à estabilidade institucional de políticas setoriais que envolvem percentuais orçamentários previstos em lei, pode comprometer os resultados, pois a redução não explica, por si só, fracassos, sucessos, limites, potencialidades, fragilidades, enfim, os motivos pelos quais uma medida/programa/projeto pode ter alcançado melhores resultados ou não. (MARTINS, 2013, p. 285).

Nesse cenário, as consultorias encomendadas pelos governos são tomadas como instrumentos para fins propagandísticos. Assim, estaríamos diante de uma análise enviesada, com objetivos diferentes daqueles que deveriam ser perseguidos por uma análise comprometida das políticas públicas.

Faria (2005), ao tratar do tema do desenvolvimento das avaliações de políticas públicas, afirma que seu destaque nos últimos anos foi possível em virtude do florescimento do gerencialismo, cujos fundamentos são vistos com desconfiança pelo autor. Apesar disso, Faria reconhece que são os adeptos do gerencialismo aqueles que estão mais dispostos a analisar questões mais abrangentes das políticas públicas, como a geração de *accountability* derivado da avaliação e o papel da avaliação como controle social do Estado (FARIA, 2005, p. 106).

Por fim, outro aspecto da pesquisa sobre políticas públicas que merece ser mencionado, e que possui consequências diretas para a presente tese, se relaciona com a recusa, oriunda de parte da literatura especializada, de pesquisas com enfoque em programas específicos, tendo em vista a defesa de produções com conclusões mais generalizáveis e resultados menos pontuais e de maior alcance.

Abordando o tema sobre qual deveria ser o enfoque das pesquisas de políticas públicas educacionais, Souza afirma que:

(...) o foco analítico da política pública reside na identificação do problema que ela procura solucionar, na forma como esse problema e concretiza no sistema jurídico e político (*politics*) e na sociedade política (*polity*). Esse pressuposto – assim como a definição de tipologias – dificilmente aparece nos estudos em políticas educacionais, o que nos leva a inferir que a área nem sempre estabelece diálogos com as teorias e modelos construídos no campo mais amplo da análise das políticas públicas.

Nesse sentido, os estudos em avaliação de políticas educacionais tendem – sobretudo mais recentemente – a centrar seu foco na avaliação de programas e projetos, lançando mão de outras perspectivas teóricas, não localizadas exatamente no campo das políticas públicas (MARTINS, 2013, p. 289).

A visão de Martins apresenta alguns problemas. Primeiro, afirma que o problema da concentração em análises de programas específicos caracteriza os estudos em políticas educacionais, quando, na verdade, essa crítica é direcionada, por outros autores (como Frey e Souza) a todo o campo de pesquisa das políticas públicas. Além disso, como vimos, desde suas origens, a *policy science* é definida (de modo positivo, diga-se) como sendo constituída por conhecimentos oriundos de diferentes áreas do saber. Na passagem de Martins (2013), essa característica é vista como um problema para a análise de políticas educacionais.

A defesa de uma leitura holística das políticas públicas, envolvendo do contexto internacional às influências de grupos de interesse locais sobre a decisão dos governos, não é antagônica à análise de programas específicos. As dimensões micro e macro podem caminhar

juntas. A ausência de uma delas, por si só, não é fator que desqualifica a pesquisa, dependendo dos objetivos perseguidos pelo pesquisador. Para que a análise de um programa tenha lugar no campo é forçoso que todos os elementos macropolíticos sejam abarcados na análise? Sem eles, a análise não encerra sentido? Dificilmente, seria possível sustentar uma perspectiva como essa. Antes de tudo, é tarefa árdua ter um mapeamento completo de tudo o que afeta uma política pública, atividade de pesquisa sempre aberta à percepção de novos fatores. Além disso, mesmo mapeados os fatores que influenciam a política, sejam do âmbito macro ou do micro, não é simples levar a cabo sua completa investigação.

A defesa da visão macrossocial para o entendimento das políticas públicas é legítima e pertinente, mas não pode se tornar o fundamento último da *policy science*. Em seu nome, muitas vezes, os efeitos do macrocontexto sobre a política são apresentados de maneira protocolar, pronta, sem uma análise sequer, apenas com a referência, por exemplo, à existência de um sistema financeiro internacional que condiciona as decisões locais. Não há poder explicativo em afirmações como essa. No âmbito das políticas públicas educacionais, um movimento foi iniciado com vistas a valorizar também o microcontexto das políticas:

Contudo, parte da literatura passou a questionar os desenhos metodológicos utilizados na área das políticas educacionais, antes mais voltados à análise de macroestruturas, iniciando um movimento de estruturação de investigações dedicadas a investigar aspectos da micropolítica (BALL, 1989, 2006). Assim, estudos voltados a analisar as reformas implementadas desde o início dos anos 1980 – centradas na responsabilização das escolas e de seus profissionais – passaram a questionar as motivações políticas que fundamentam esse processo, adotando abordagens que buscam apreender a visão/opinião/percepção de diretores, professores, coordenadores pedagógicos, pais e alunos sobre as medidas, com vistas a construir percursos mais flexíveis de investigação na área (MARTINS, 2013, p. 290).

Mesmo com tal interesse despertado, o enfoque mais valorizado pelas pesquisas sobre políticas públicas é aquele que presta contas ao nível do macrocontexto. Frey (2000), por exemplo, vê as análises de programas específicos como uma limitação da pesquisa no campo da *policy science*. Segundo ele, os estudos dão ênfase a casos empíricos, com validade apenas situacional (2000, p. 243). Com isso, corre-se o risco de tratar como padrão algo circunstancial. Mas, cabe o questionamento: o que se busca com a análise de políticas públicas são padrões? Essa afirmação de Frey não se encaixa com a definição dada para o principal objetivo do campo, qual seja, o de compreender as ações do governo. Se cada governo é potencialmente diferente

do anterior, embora deva, necessariamente, lidar com sua herança, como pretender apenas investigar os padrões dessa ação? Soma-se a isso o elemento temporal, que faz com que as circunstâncias se alterem. A análise de uma política pública deve se basear em arcabouço teórico validado, mas este não é uniforme, visto que precisa se ajustar à circunstancialidade de cada política. Conforme o próprio Frey (2000, p. 243) reconhece, a análise de políticas públicas não possui uma teoria uniforme e um modelo unívoco. A visão de Frey parece oferecer escolhas polarizadas ao pesquisador. Ou as políticas são estudadas através dos programas que as encarnam, levando a conclusões limitadas e a estudos sem grande validade para o campo, ou as políticas são estudadas a partir de uma análise do macrocontexto, levando em consideração todos os elementos que podem explicá-la. Não nos parece adequado colocar o problema nesses termos. Como investigar, por exemplo, políticas educacionais municipais no Brasil, tendo em vista a heterogeneidade que esse cenário oferece, sem prestar contextos à microanálise (e sem desconsiderar os elementos do macrocontexto que podem contribuir para a explicação)? O autor mesmo parece reconhecer os desafios impostos à *policy science* em casos como esse:

Tomando como exemplo a política municipal no Brasil, vemo-nos defrontados com vários problemas peculiares. Primeiro, temos que levar em conta que o conhecimento científico no tocante ao quadro de referência, ou seja no tocante à configuração dos arranjos institucionais e das características dos processos políticos municipais, é bastante limitado. Existem, de forma geral, relativamente poucos estudos científicos sobre a política municipal, o que vale tanto para as condições institucionais quanto para os processos políticos de decisão e de planejamento (FREY, 2000, p. 243).

De acordo com a leitura de Frey, compartilhada por Souza (2002) a análise do programa não pode prescindir da análise do contexto político no qual está inserido. A dimensão da *policy* exige sempre o entendimento da *politics*. Analisar essas duas dimensões envolve, no entendimento de Martins, a superação do que a literatura sobre políticas públicas tem produzido no contexto nacional: “Examinar essas duas dimensões extrapola a denominada discussão de literatura ou discussão de contexto – item largamente encontrado na justificativa/introdução de teses, artigos, relatórios de pesquisa – pois exigiria metodologia específica, vinculando as dimensões anteriormente referidas, que concatenam as políticas públicas” (MARTINS, 2013, p. 289).

Frey (2000), como adepto da perspectiva do neoinstitucionalismo para a análise de políticas públicas, defende uma proposta de pesquisa baseada na análise da inter-relação entre instituições políticas, o processo político e os conteúdos da política, não apenas no aumento do

conhecimento sobre programas e projetos desenvolvidos pelo governo. Para a compreensão das políticas, o entendimento do papel das instituições é central, deslocando o interesse dos programas em si mesmos: “faz-se necessário certo grau de deslocamento do foco dos estudos da parte dos programas, planos e resultados materiais para os pontos de interseção entre estes por um lado e os processos e instituições por outro, dependendo a necessária intensidade do grau de deslocamento do próprio grau de fluidez que caracteriza esses pontos de interseção” (FREY, 2000, p. 247). O autor é realista ao defender a necessidade de alguns levantamentos primários sobre a *politics* para a interpretação da *policy*. Entretanto, atribui ao processo político uma importância maior do que a própria política pública. Dependendo do enfoque dado à pesquisa, as diferentes dimensões que envolvem as políticas públicas podem receber pesos e atenções diversas (*polity*, *politics*, *policy*). O peso a cada uma delas variará de acordo com os objetivos da pesquisa.

A perspectiva defendida por Frey (2000), compartilhada com outros autores, ao colocar a política como objeto prioritário em relação à política pública em si mesma, termina por instrumentalizar as investigações sobre a *policy*, transformando-a em espelho do processo político, um instrumento para compreendê-lo. O interesse próprio na política pública é reduzido e ela só interessa como reflexo das opções do governo. Nossa proposta é romper com essa perspectiva. Não se trata de seguir a ideia, metodologicamente equivocada, da teoria da curvatura da vara, que tanto eco encontrou na seara educacional (ver SAVIANI, 2008). A crítica não é à defesa de observar as ações do governo através das políticas públicas, dimensão, sem dúvida, importante para sua análise, mas, antes, de defender relativa autonomia, como objeto de pesquisa, dos programas que encarnam as políticas. Dizer que as políticas públicas podem ser pensadas como reflexo do governo não é o mesmo que afirmar que apenas isso interessa à *policy science*. As políticas, implementadas através de programas e conjuntos de ações, despertam interesse próprio, que pode envolver ou não a compreensão mais profunda do aspecto macrossocial. Esse é um ponto importante. Sem reduzir a importância da leitura do contexto que a envolve, o estudo de uma política pública pode prescindir de um estudo mais aprofundado sobre o referido contexto, dependendo dos interesses de pesquisa do analista e do que está disponível para ser pesquisado.

Propomos, portanto, uma inversão: a *policy science* não se interessa apenas pelas relações entre as instituições, o processo e os conteúdos da política, tema caro à ciência política, mas se debruça também sobre a análise dos programas através dos quais uma política pública é implementada. Assim, a luz descola-se da política para a política pública em si mesma. As

políticas públicas não são apenas um objeto através do qual o governo é analisado. Elas têm importância relativamente autônoma em relação ao cenário político. O interesse pela *politics* é que serve à investigação da *policy*.

Independentemente do contexto político, interessa saber se uma política está sendo executada como planejada (monitoramento e avaliação de execução) e se alcança os objetivos previstos (avaliação de resultados). Analisar o programa tem valor mesmo que não se possa estudar todos os elementos importantes para explicá-lo. Comparar e descrever são ações importantes, sob pena de não produzirmos informação alguma sobre determinadas políticas públicas. O interesse pelo contexto, que é sempre muito variado e exige uma seleção dos aspectos a serem investigados, terá seu alcance definido a partir das características do programa que se deseja analisar.

As exigências para que uma política pública seja analisada são tantas que, seguindo a cartilha de Frey (2000) e outros autores, a *policy science* corre o risco de ser, entre nós, uma ciência sem objeto. Toda análise é, em maior ou menor grau, parcial e fragmentária. A busca pode ser holística, mas, na prática, há muitas limitações para a pesquisa das políticas públicas. Enquanto discursa sobre a necessidade de produzir estudos mais atentos ao macrocontexto de produção das políticas públicas, lançando luz sobre a forma como a agenda foi definida e mapeando os grupos de interesse que exerceram influência sobre essa definição, ao passo que critica os trabalhos descritivos sobre determinadas políticas, por serem específicos em demasia e não contribuírem de modo significativo para o avanço do campo com conclusões mais generalistas, a pesquisa em políticas públicas em educação no Brasil segue sem produzir estudos significativos sobre programas e ações em específico e sem pesquisar, com metodologia e referencial teórico mais robustos, aquilo que diz ser o mais importante.



## APÊNDICE B - Uma reflexão sobre o conceito de políticas públicas

O objetivo deste anexo não é fazer um apanhado geral das concepções sobre políticas públicas, de modo a mapeá-las. O estado da arte foi realizado por outros estudiosos, como Souza (2002, 2003, 2006), Martins (2013) e Frey (2000). Um dos objetivos é, partindo desse mapeamento, identificar as fragilidades conceituais do campo de pesquisa da *policy science* para o estudo da avaliação da alfabetização como política pública no cenário brasileiro, reconhecendo a existência de quatro problemas centrais, que precisam ser enfrentados. Outro é fornecer uma contribuição original para o campo, apresentando um conceito de política pública que consideramos mais ajustado do que os diversos conceitos oferecidos pelos pesquisadores da área até agora. Evidentemente, não se trata de apresentar um conceito absolutamente fechado, mas, antes, de oferecer uma definição mais precisa e menos genérica do termo, submetendo-o à crítica.

Um primeiro problema de conceituação que se apresenta é a confusão entre o campo de pesquisa e seu objeto de interesse. A mesma expressão é usada para se referir a esses dois elementos. De acordo com a definição que Martins (2013) apresenta, por exemplo, as políticas públicas são um campo do conhecimento, uma disciplina científica que, embora recebendo contribuições de diferentes campos do saber (sociologia, economia, administração, estatística etc), pode ser definida como uma subárea da ciência política. Mesmo que se considere que a posição de subárea pertence à história de nascimento e consolidação da disciplina, e que hoje a *policy science* responda como área própria de conhecimento, dotada de objeto, epistemologia e metodologia próprios, o problema permanece de pé. A definição das políticas públicas apenas como área do conhecimento obscurece a polissemia da expressão. Seu significado, mais do que disciplina acadêmica, envolve seu objeto. Políticas públicas são aquilo que um determinado campo de pesquisa, denominado, também, muitas vezes, de políticas públicas, se propõe a investigar. De um lado, um campo de investigação sobre as ações do Estado. De outro, as próprias ações do Estado. Tratar as duas coisas com o mesmo signo pode levar a uma série de problemas de interpretação para o campo.

Nas compilações sobre o estado da arte em língua portuguesa, o que ganha destaque, em termos de conceituação, é a disciplina acadêmica, o campo de pesquisa. É o que se vê, por exemplo, nos autores supracitados. O enfoque dado à disciplina é uma possível causa das dificuldades em se conceituar o que são as políticas públicas. Em função disso, utilizaremos,

ao longo do texto, a expressão *policy science* para se referir ao campo<sup>209</sup> de pesquisa. A expressão políticas públicas será reservada ao objeto de estudo desse campo. E como objeto, é preciso defini-lo.

Afinal, se a disciplina se dedica a entender e explicar as ações do Estado, é preciso definir, com clareza, que ações são essas. Na definição apresentada por Martins (2013), conforme exposto em linhas anteriores, o interesse dos estudos sobre políticas públicas é por tudo aquilo que os governos fazem. Diante disso, todo ato do governo é objeto de interesse desse campo de pesquisa? Dificilmente, uma definição como essa faria sentido. Nem toda ação ou omissão do governo é uma política pública. As políticas públicas não abrangem todos os atos dos governos, mas determinadas ações específicas, com características próprias, que as singularizam em relação ao todo. Voltaremos a esse ponto a seguir.

Bonavides (2000) também define política pública como um campo de pesquisa dedicado a compreender a ação do governo no que diz respeito a determinados temas sociais e econômicos de interesse público, ou seja, interessado nas formas de governar e nas instituições políticas. Esse conceito tão pouco nos serve. Primeiro, define política pública como *policy science*, confundindo o campo de pesquisa com seu objeto, conforme ressaltamos. É o tipo de definição que mais confunde do que ajuda a esclarecer. Além disso, foca na ação do governo, no ato de governar, e não nos programas e ações em si mesmos. Vale ressaltar que, ao afirmarmos a pertinência da análise dos programas e ações não queremos eliminar o interesse pelo contexto e pelas explicações macrosociológicas. Estas, na verdade, dominam o campo. O intuito é chamar a atenção para a relativa autonomia, como objetos de pesquisa, dos programas em si, tema que também será retomado a seguir.

Souza segue o caminho de conceituar políticas públicas como campo científico e as define como:

(...) o campo do conhecimento que busca, ao mesmo tempo, ‘colocar o governo em ação’ e/ou analisar essa ação (variável independente) e, quando necessário, propor mudanças no rumo ou curso dessas ações (variável dependente). A formulação de políticas públicas constitui-se no estágio em que os governos democráticos traduzem seus propósitos e plataformas eleitorais em programas e ações que produzirão resultados em mudanças no mundo real (2006, p. 26).

---

<sup>209</sup> Campo científico no sentido que Bourdieu (1976) usa o termo, ou seja, um conjunto de regras, normas e relações, mantidas ou alteradas por indivíduos investidos em determinados papéis valorizados no campo, que é caracterizado por arranjos mais ou menos estáveis, cuja legitimidade é dada por disputas internas de poder.

Seguindo Saravia (2006), Souza acrescenta, ainda, que há quatro elementos que possibilitam a existência de políticas: i) institucional – a decisão da política pública é da competência de autoridades formais, constituídas legalmente, possuindo, assim, um caráter vinculante frente à coletividade; ii) decisório – a política pública é formada por uma sequência de decisões, em que problemas são identificados, alçados ao status de problemas públicos e recebendo a atenção dos decisores através de ações específicas; iii) comportamental – uma política é um curso de ação, que não se esgota em uma decisão singular e específica, implicando um conjunto de ações ou um comportamento inativo; iv) causal – as políticas possuem produtos de suas ações com pretensões de exercer efeitos no mundo social (SOUZA, 2006). Essas características, não há dúvida, são próprias das políticas públicas, entretanto, não se chega a um conceito para a expressão que a diferencie do campo científico do qual é objeto de interesse e que conceda aos programas a importância que eles possuem.

Em outro texto (2002), Souza reconhece que políticas públicas é uma expressão polissêmica, que recebeu diferentes conceituações de acordo com a perspectiva de diversos autores:

Não existe uma única, nem melhor, definição sobre o que seja política pública. Mead (1995) a define como um campo dentro do estudo da política que analisa o governo à luz de grandes questões públicas. Lynn (1980) a define como um conjunto específico de ações do governo que irão produzir efeitos específicos. Peters (1986) segue o mesmo veio: política pública é a soma das atividades dos governos, que agem diretamente ou através de delegação, e que influenciam a vida dos cidadãos. Dye (1984) sintetiza a definição de política pública como “o que o governo escolhe fazer ou não fazer”. A definição mais conhecida continua sendo a de Laswell, ou seja, decisões e análises sobre política pública implicam em responder às seguintes questões: quem ganha o quê, por que e que diferença faz. (SOUZA, 2002, p. 4).

Todas essas definições apresentam elementos importantes para o conceito de políticas públicas, mas são fragmentárias, genéricas e abstratas. A rigor, não se apresentam como definições, pois o conceito, ao fim, não é definido com clareza. Não há, nelas, diferenciação entre os diferentes atos do governo, por exemplo. Assim, tudo que o governo faz é política pública. Não parece promissor seguir esse caminho. Nem tudo o que o governo faz é política pública, conforme já ressaltamos. Além disso, continua presente a indiferenciação entre o campo de estudo (*policy science*) e o objeto de estudo (a política pública em si). Em nenhuma das definições dos autores clássicos, o aspecto conflituoso aparece, o que faz com que muitos

críticas sejam direcionadas a essa concentração na racionalidade da ação governamental, ignorando o papel dos demais atores e grupos. É importante notar que o problema aqui é sequer considerar o tema do conflito. Ele se faz presente na análise de políticas, embora não deva ser o único foco de interesse. O que há de comum em todas as definições, contudo, é a presença do governo, que receberá maior ou menor atenção de acordo com o enfoque dado. Soma-se a isso a visão holística que as caracteriza.

Neste ponto, entra em cena outro problema de conceituação. Ao definir o objeto da *policy science* como as ações do governo, é preciso diferenciar políticas públicas dos programas e conjunto de ações que as encarnam. Os autores do campo prestam pouco atenção a esse ponto. As políticas públicas envolvem decisões políticas, rearranjos, negociações, disputas de poder, definição de agenda, implementação etc. É um processo amplo e complexo. Elas representam o posicionamento político e ideológico do governo diante de temas sociais importantes e são encarnadas, para sua realização, em programas e ações específicas. Este é um ponto fundamental para a análise de políticas públicas. Os programas e ações sistematizadas, que visam à execução de uma política, são os objetos, por excelência, da *policy science*. Contudo, a produção científica do campo tende a obscurecer sua importância, criticando os estudos que prestam reverência aos programas e chamando atenção para a preeminência do interesse pelo contexto político, macrosociológico, que envolve a produção da política, como vimos no capítulo 1.

Os principais teóricos sobre políticas públicas, incluindo aqueles que foram fundamentais para a constituição do campo, não dão ênfase a como elas se encarnam em programas e ações, mas, particularmente, na forma como o governo decide, nas influências sobre as decisões. Assim, conforme ressaltam Maia & Grandó (2016) e seguindo o exposto no tópico anterior, Laswell busca conciliar o conhecimento científico com a forma de agir do governo, ao passo que Simon se interessa pelas limitações oferecidas à racionalidade dos decisores, tendo em vista a indisponibilidade ou incompletude de informações. Lindblom, por sua vez, chama atenção para as relações de poder que caracterizam as diferentes fases do processo de tomada de decisão, enquanto que Easton propõe a ideia de política pública como sistema em que implementação, formulação e resultados são influenciados pelos grupos de interesse e pela mídia. Nenhum deles dá destaque ao tema da análise das políticas públicas através dos programas que as colocam em prática.

A teoria aponta para três dimensões das políticas públicas que precisam ser consideradas (DEUBEL, 2009; SOUZA, 2002; FREY 2000): *polity*, *politics* e *policy*. Analiticamente, essas

três dimensões se diferenciam, levando o analista a se concentrar em elementos diferentes da ação política. *Polity* é a política como âmbito do governo, as instituições políticas, a ordem do sistema político, sua estrutura institucional. O termo *politics* se refere às atividades de organização do poder, aos processos políticos, abarcando a luta por ele. *Policy*, por sua vez, diz respeito aos programas e ações desenvolvidos pelas autoridades políticas, aos conteúdos concretos das políticas, à configuração dos programas políticos. Nosso interesse se direciona, precipuamente, à dimensão da *policy*, com *polity* e *politics* se configurando como dimensões explicativas importantes para a *policy*. Historicamente, as políticas públicas foram tratadas como variáveis dependentes da atividade política, como resultados diretos das ações governamentais (a *policy* seria dependente da *politics*). As três dimensões se influenciam mutuamente, sendo difícil estabelecer uma causalidade clara e direta de como operam essas influências. A análise da *policy* tende a ser criticada, por tratar *polity* e *politics* como variáveis independentes, ou seja, como se a estrutura e a atividade políticas fossem fixas e imutáveis. Na verdade, nenhuma dessas dimensões pode ser tratada, absolutamente, como variável dependente. Isso não impede, porém, de concentrarmos nosso interesse no estudo da *policy*.

Portanto, é importante deixar claro, mais uma vez, que o foco da presente pesquisa não é olhar o governo como agente, refazendo, passo a passo, a tomada de decisão no que tange às políticas públicas, mas analisar programas que encarnam determinadas políticas públicas, recorrendo ao entendimento de como o governo decide caso seja esse um fato importante para a interpretação dos programas em questão e passível de ser plenamente pesquisado.

Um terceiro problema que a conceituação das políticas públicas apresenta é a consideração do não fazer do governo (sua omissão) como elemento integrante das definições. É o que se vê, por exemplo, em Maia & Grando, quando definem políticas públicas como a “(...) soma das atividades do governo, que escolhe o que fazer e o que não fazer, resultando em arranjos institucionais fomentados por organizações de interesses privados de determinados segmentos econômicos e agenda públicas (...)” (2016, p. 165). Não fazer nada em relação a um problema, portanto, também é considerado como uma forma de política pública.

Não entendemos “não fazer nada” como uma forma de política pública<sup>210</sup>. Não agir diante de determinadas circunstância revela a intenção do governo, suas prioridades, suas

---

<sup>210</sup> É possível que a autora não tenha notado, visto que compreende o não agir como parte da política pública adotada pelo governo, mas Souza, em uma passagem, afirma que só quando a política é posta em ação a análise dos pesquisadores tem lugar: “após desenhadas e formuladas, se desdobram em planos, programas, projetos, bases de dados ou sistemas de informação e pesquisas. Quando postas em ação, são implementadas, ficando, então, submetidas a sistemas de acompanhamento e avaliação e à análise dos pesquisadores” (SOUZA, 2003, p. 14).

escolhas. Contudo, esse comportamento passivo diante de determinados problemas não é, a rigor, política pública. É uma decisão, justamente, de não produzir uma política pública para lidar com determinados problemas. A política pública precisa ser encarnada através de ações e programas. Ela é necessariamente ativa. Considerar não fazer como política pública oferece, ainda, um problema metodológico. Como ter como objeto de estudo algo que não existe? Como veremos, a implementação, por exemplo, não é uma fase necessária da política pública? Como implementar o que não existe? Na perspectiva do ciclo de políticas, essa definição não encerra sentido. Só pode ser política pública alguma ação específica do governo, e, ainda, não qualquer ação. Uma ação qualificada, planejada, sistematizada e encarnada em um ou mais programas com objetivos e público-alvo especificados, para enfrentar problemas selecionados como fundamentais pela agenda governamental. Um conceito de política pública que aceita a ausência de ação como política não tem objeto definido e não pode, por isso, ser analisado e investigado por algum ramo da ciência. O não agir do governo interessa à política, mas não à análise de políticas públicas propriamente dita.

Isso nos leva à conclusão de que o elemento central para a análise de políticas públicas é sua implementação. Não há política pública sem implementação. Se só há intenções, não há objeto de estudo para a *policy science*. Não quer dizer que as intenções e negociações que antecedem a efetivação de uma política pública não sejam importantes para compreendê-la. No entanto, se não há implementação, as intenções, a definição da agenda e todo o jogo de poder que as envolve são objeto de estudo para a ciência política e para a sociologia política. A *policy science*, é o que defendemos, tem como objeto políticas públicas encarnadas em conjuntos de ações e programas. Por isso, a concentração na definição da agenda não pode ser tomada como um objeto autônomo para a *policy science*. Só faz sentido estudar a agenda, para o campo, se a política, de fato, foi implementada. Discussões sobre a agenda de políticas que não chegaram a acontecer caracterizam estudos de intenções, mas não de ações do governo. Do ponto de vista da abordagem do ciclo de políticas, a agenda é apenas a primeira fase de análise. Ela pode esclarecer problemas relacionados à efetivação de políticas, por exemplo, e seu estudo é fundamental para a compreensão mais profunda de uma política, desde que exista uma política implementada para ser analisada.

Essa perspectiva difere, substancialmente, do que é defendido pela *policy science*. A agenda, as intenções, o intuito inicial do governo, mesmo que não sejam efetivadas através de ações concretas, são tratadas como objetos de interesse do campo, sendo consideradas, em muitos casos, como o aspecto mais importante da análise de políticas públicas. Reconhecemos

a importância dessa fase, desde que a política seja implementada, momento a partir do qual o campo ganha seu objeto de estudo.

Um quarto problema de conceituação se refere à consideração de que as políticas públicas são fenômenos utilizados para compreender o funcionamento do Estado, se apresentando como a análise das consequências que determinadas escolhas do governo têm para a solução de problemas específicos e se interessando pelos resultados que um determinado sistema político apresenta. Nesse caso, estamos diante de um uso instrumental da análise de políticas públicas, qual seja, eles são um espelho para a compreensão do Estado, de como os governos agem e se comportam. O interesse final é compreender a ação governamental. As políticas são apenas os objetos através dos quais isso se torna possível. A política pública torna-se, assim, objeto da ciência política. Não é esse o intuito da pesquisa. Compreender o papel das políticas públicas em cenários eleitorais, o uso das políticas públicas como parte do comportamento político ou como reflexo de determinadas concepções ideológicas, todos eles interesses da ciência política, só nos interessam se apresentarem algum elemento heurístico para as políticas públicas analisadas.

Mais uma vez, é preciso esclarecer a crítica que está sendo feita. O estudo das políticas públicas pode levar a conclusões sobre como o governo age e como o Estado se comporta. Entretanto, as políticas públicas, elas mesmas, são objeto de interesse para uma ciência dedicada à sua compreensão, a *policy science*. Afirmar que analisar políticas públicas serve apenas ao propósito de compreender a ação governamental reduz a importância do estudo das políticas em si. Novamente, não se trata de advogar que o macrocontexto e o entendimento das ações do governo não são importantes. Trata-se de conceder algum grau de autonomia para as políticas e seus programas, reconhecendo que eles são o objeto, por excelência da *policy science*. É o que buscamos com esta tese. Ao nos dedicarmos ao estudo de políticas de avaliação da alfabetização, através de programas específicos, estamos interessados nos programas eles mesmos, em suas ações, suas consequências, suas implicações para o combate ao analfabetismo, seus resultados, os problemas de implementação que enfrentaram e uma série de outros aspectos. Não pretendemos usar as políticas de avaliação da alfabetização para compreender o governo como um todo. Isso não significa ignorar o contexto político em que essas políticas foram criadas e implementadas. Ele nos interessa na medida em que nos ajuda a compreender melhor os programas.

Aqui, se apresenta um problema de caráter epistemológico. Mesmo que desejássemos investigar o comportamento e as ações do governo através da análise de programas de avaliação

da alfabetização, isso não seria possível. Para que fosse, seria necessário que as ações do governo, nas mais diversas áreas de atuação (educação, saúde, segurança pública, meio ambiente, finanças públicas etc), fossem concatenadas, organizadas a partir de uma lógica e de uma ideologia coerentes, homogêneas, perpassando todas as políticas públicas do governo, ligando-as de modo que formassem um todo único. Não é o caso. Políticas do mesmo governo em diferentes áreas não se apresentam de modo tão unívoco, ao contrário, podem ser percebidas, muitas vezes, como dotadas de lógicas e interesses contraditórios. Para analisar o comportamento geral do governo, é preciso olhar para todas as políticas, o que, eventualmente, levará à conclusão de que a racionalidade do governo não opera da mesma forma em todas as áreas. Outra saída seria analisar como o governo age em cada uma das suas áreas de atuação, usando as políticas públicas como instrumento para tanto. Mesmo nesse caso, defendemos nosso intuito de não seguir esse caminho. Nossa proposta segue direção distinta: objetivamos compreender as políticas públicas e os programas que as colocam em prática, observando, entre outras coisas, o contexto político que as cerca. Nesse caso, o que está em primeiro plano é a política pública em si, com o contexto político se apresentado como uma variável de explicação.

Um elemento importante para a definição de políticas públicas é seu caráter processual. Partindo da noção de políticas públicas anunciada por Martins (2013), com base na ideia de que elas se definem como aquilo que o governo faz, aumenta a dificuldade em se estabelecer um conceito unívoco para a expressão, na medida em que o que os governos fazem é algo muito complexo. A complexidade da ação do governo é tributária do fato do mesmo não ser linear, ou não apresentar sempre o mesmo comportamento diante de situações semelhantes. Isso se deve, em grande parte, aos conflitos que caracterizam o ato de governar. As decisões do governo não são tomadas apenas com base em ideologias bem fundamentadas e lógicas. Tais decisões são influenciadas por uma série de fatores, entre eles, os interesses de grupos sociais em conflito, que disputam a influência da agenda.

Além disso, cumpre ressaltar que as decisões não são atos isolados e estanques. Em geral, elas são frutos de um processo de negociações e, como sequência, iniciam novos processos. Essa dimensão processual das políticas públicas, mais do que justificar a necessidade de se atentar para o conflito como fator característico de sua constituição, é fundamental para a explicação das mudanças e constantes ajustes pelos quais as políticas passam. Trata-se de um aspecto inerente às políticas públicas. São definidas, implementadas e avaliadas ao longo de um processo, em geral, não dotado de linearidade. Seguindo a definição de hipercomplexidade do mundo social (LUHMANN, 2010; HABERMAS, 1990, 2012), as políticas públicas são



objetos em constante transformação. Há, pelo menos, duas formas de entender o caráter processual das políticas públicas e as transformações pelas quais elas passam: i) positivamente, quando a transformação é atribuída a uma necessária adaptação ao contexto social; e ii) negativamente, quando a transformação é atribuída à fragilidade da proposta da política, propícia a ser alterada ao sabor das mudanças de governo. Na prática, essas duas perspectivas, antes de se antagonizarem, podem ser observadas de maneira complementar, dependendo do contexto. Entretanto, sua separação, para fins analíticos, pode ser útil. No primeiro caso, as transformações têm fundamento sociológico. O mundo social se apresenta de maneira tão complexa que as condições que deram origem à política pública, sejam elas econômicas, culturais, políticas ou institucionais, podem ser alteradas em um curto espaço de tempo. Como a política depende de uma série de fatores para ver seus efeitos pretendidos aflorarem, a alteração em um ou mais desses fatores pode comprometer o alcance de seus objetivos, exigindo mudanças em seu desenho (em casos mais drásticos, uma mudança da própria política).

Soma-se a isso o fato das políticas públicas serem ancoradas em valores. Uma política destinada ao combate ao analfabetismo, por exemplo, se baseia no valor dado à alfabetização nas sociedades letradas, bem como em suas consequências nos âmbitos individuais e sociais. Caso haja modificação nos valores que dão base à política pública, as consequências dessa mudança se fazem sentir em seu desenho ou em sua concepção. Pode-se argumentar que a modificação de valores não costuma operar de forma tão drástica em curtos espaços de tempo. Apesar disso, não está descartado, numa sociedade que opera a partir de redes tão heterogêneas e de amplo alcance (CASTELLS, p. 2016), que substantivas modificações nos valores atribuídos a determinados fenômenos tenham lugar. As intenções que motivaram a escolha de uma política não são perenes. Podem ser alteradas de acordo com mudanças significativas no cenário político e social. Outra forma de mudança que pode ser considerada positiva é a que se relaciona com a capacidade de avaliação da política pública. A fase de avaliação da política, conduzida de forma séria e compromissada com a idoneidade de análise dos resultados e com o rigor metodológico, pode levar à conclusão de que a política precisa sofrer ajustes para que alcance seus objetivos. Nesse caso, a mudança teria lugar em virtude de elementos técnicos.

De uma forma ou de outra, quando as modificações na política ocorrem por esses motivos, elas podem ser atribuídas, positivamente, à capacidade de adaptação das políticas públicas, à sua flexibilidade. O valor positivo atribuído a essa característica encontra respaldo na ideia de que uma política que não se atenta às condições sociais sobre as quais pretende produzir efeitos é uma política engessada e pouco sensível ao mundo que a entorna.

Por outro lado, as mudanças podem ser percebidas a partir de uma leitura negativa. As alterações constantes nas políticas não seriam tributárias de sua adaptação consciente a mudanças no cenário social, a partir de uma acurada percepção sociológica de seus decisores e implementadores, mas, antes, de sua grande fragilidade. Nesse caso, as modificações sofridas pelas políticas se devem à sua instrumentalização política, no pior sentido do termo, com as políticas sendo utilizadas pelos governos para o atendimento de interesses próprios ou de grupos influentes. De forma complementar, as políticas poderiam ser percebidas como instrumentos propagandísticos, que, mesmo atingindo seus efeitos, não seriam mantidas ou renovadas caso um novo governo assumisse o poder. Tendo como base esse tipo de interpretação, as políticas nunca seriam dotadas de estabilidade e continuidade, independentemente do alcance de seus efeitos

Ainda, a mudança da política pública pode ser atribuída a fatores econômicos. Um dos pontos mais sensíveis, e menos enfrentados pela pesquisa, em relação às políticas públicas se relaciona com seu financiamento. Esse é um ponto crucial para todas as fases do ciclo das políticas. A decisão sobre a implementação ou não de uma política passa, necessariamente, pelos seus custos. Esse fator é fundamental para a escolha do tipo de política que será implementada. Num cenário perpassado por conflitos de interesses, os diferentes grupos podem usar esse argumento para fazer oposição a determinada escolha. Além disso, na fase de avaliação, esse pode ser o fator decisivo para definir a continuidade ou não de uma política, com base na relação custo-benefício. Quando a mudança em uma política ocorre em virtude de fatores econômicos, ela pode ser percebida de forma positiva ou negativa, a depender das circunstâncias. Se a política era excessivamente custosa e com poucos resultados, o ajuste será bem visto. Caso contrário, sendo uma política com grandes benefícios, os ajustes em virtude de cortes no orçamento podem ser vistos de forma negativa.

Diante do exposto, como conceituar políticas públicas de modo que os problemas aqui apresentados possam ser contornados? Com a consciência de que o conceito a seguir apresentará, também, limitações, propomos uma definição de políticas públicas que, ao mesmo tempo, estabelece uma diferença entre campo e objeto de estudo, que chame a atenção para os programas, que não considere a omissão como objeto de pesquisa da *policy science* e que não seja uma instrumentalização do conteúdo para o simples entendimento da ação do governo. Assim, políticas públicas são: *decisões do governo para o enfrentamento de problemas sociais alçados ao status de problemas públicos, perpassadas por consensos e dissensos, efetivadas através de programas e conjuntos de ações organizados de forma sistemática, com objetivos e*

*público-alvo definidos, e com um período de duração determinado, no qual espera-se que os resultados previstos sejam alcançados, que, ao serem implementados, tornam-se objeto de pesquisa, por excelência, da policy science.*

Através dessa conceituação, temos: i) o governo é ponto central para a definição, pois é o agente, por excelência, das políticas públicas (ponto comum às diferentes definições que assumidas pela expressão), a despeito da eventual participação de atores privados; ii) as decisões são influenciadas por diferentes atores e grupos, sendo o produto de dissensos e consensos até que um arranjo final seja definido; iii) as políticas públicas são encarnadas através de programas e ações, sendo que a omissão do governo, embora seja uma tomada de posição, não se caracteriza como política pública; iv) não é qualquer ação do governo que interessa à análise de políticas públicas, mas sim aquelas organizadas e sistematizadas através de programas e ações que possuem duração ao longo de um tempo definido, além de objetivos e público-alvo próprios; v) os programas possuem autonomia em termos de interesse para a pesquisa, se apresentam como o objeto, por definição, da *policy science*, diferenciando, assim, campo de pesquisa e objeto de pesquisa. Acreditamos, ao levar em consideração esses elementos, oferecer uma conceituação satisfatória de políticas públicas como suporte para a pesquisa das políticas de avaliação da alfabetização.

## APÊNDICE C - Modelos de formulação e análise de políticas públicas

Antes de passar à apresentação dos modelos de formulação e análise de políticas produzidos pela literatura especializada, é preciso estabelecer uma distinção entre duas ações: de um lado, o desenvolvimento de modelos para formular políticas; de outro, o desenvolvimento de modelos para analisá-las. A distinção não é trivial e envolve uma tomada de posição em relação ao que se espera da *policy science* como campo do conhecimento. As consequências dessa tomada de posição se fazem sentir, entre outras coisas, na metodologia adotada para o estudo, bem como na seleção do objeto de pesquisa.

Como vimos, historicamente, o campo das políticas públicas se preocupou não apenas em descrever e analisar as políticas, mas em orientar o governo na tomada de decisão. Com enfoque pragmático, a *policy science* teria como produto a orientação para os governos. A preocupação com a racionalidade das decisões governamentais estimulou a ideia de que a ciência deveria ser capaz de fornecer os fundamentos de uma decisão esclarecida e bem fundamentada por parte dos governos. Diante disso, a descrição e a análise das políticas gozariam de um status menos valorizado quando comparadas à orientação para a tomada de decisões. Vale a pena problematizar essa ideia: por mais que os achados da ciência possam ser utilizados como critérios para legitimar a tomada de decisões no âmbito político, a *policy science* não pode exigir que suas descobertas e conclusões sejam, necessariamente, utilizadas pelos governos. Assim, cabe aos governos, operando em um campo político próprio, com regras próprias (ver BOURDIEU, 2011), decidir sobre o uso ou não das descobertas científicas. A ciência pode ser apropriada pela política, mas isso depende da discricionariedade dos atores políticos. Em outras palavras, mesmo que considere seus achados como informações qualificadas, a *policy science* não pode obrigar os governos a agirem de acordo com suas conclusões. Assim, tal ciência não pode se definir a partir de uma função que não lhe cabe diretamente. Mesmo que os governos não façam uso dos seus achados, a *policy science* permanece tendo uma razão de ser, qual seja, produzir conhecimento qualificado, sistematizado e válido sobre políticas públicas. Se o governo fará uso ou não desse conhecimento extrapola o âmbito de ação do campo científico. Seu objetivo principal é a produção de informações qualificadas sobre seu objeto de interesse.

Essa perspectiva coloca em cena um questionamento sobre o papel secundário e pouco valorizado atribuído à descrição e à análise de políticas públicas. Uma coisa é dizer como o governo deve agir (um tempo anterior à decisão). Outra é dizer como o governo agiu e quais

foram as consequências dessa ação (um tempo posterior à decisão). Mesmo que o governo não siga as orientações da *policy science*, seu objeto de pesquisa permanece de pé: os programas e ações adotados pelo governo para a execução de políticas públicas. Por isso, nos interessa mais os modelos de análise e descrição de políticas públicas do que aqueles que buscam oferecer normas para sua formulação. Nesse último caso, o perigo de uma espécie de receitismo para a ação do governo é real.

Souza (2002, p. 17), por exemplo, apresenta um modelo descritivo/operacional para a formulação de políticas públicas. Sua ideia é usar cada um dos elementos elencados e descritos como modelo para a formulação de políticas. Conforme já ressaltamos, trata-se de questão controversa a ideia de desenvolver modelos que funcionem como orientação para os governos formularem políticas. Controversa ou não, tal concepção não faz parte da proposta da presente tese. Nosso objetivo não é propor ao governo uma maneira adequada de desenhar políticas públicas, mas, antes, analisar políticas públicas governamentais já implementadas no âmbito da avaliação da alfabetização. De todo modo, o quadro descritivo de Souza compartilha uma série de elementos utilizados como critérios para compor nossa proposta de tipologia para as políticas públicas, tais como: os objetos da política, a temporalidade prevista para o alcance dos resultados, a escala, os custos, as políticas prévias, a *policy community*, os sistemas de incentivo, controle e regras.

Passando para os modelos de formulação e análise de políticas públicas, a literatura sobre o tema apresenta um conjunto variado de propostas. Começamos com o incrementalismo. De acordo com Lindbloom (SOUZA, 2002), os recursos governamentais destinados a determinadas políticas e programas não partem do zero, mas são decisões incrementais, se baseando na existência de políticas, programas e decisões anteriores. Para Lindbloom, isso significa que as decisões do governo não são substantivas, por não inovarem o cenário, sendo incrementais ou marginais. Em contextos de maior estabilidade política e social, o incrementalismo apresenta maior poder explicativo. Ele não se aplica de modo ajustado a períodos de grande transformação, em que rupturas com o passado são promovidas em prol de grandes reformas. De todo modo, a ideia proposta por Lindbloom tem força explicativa, despindo sua concepção do tom crítico atribuído a decisões que se baseiam em decisões anteriores.

Na realidade, torna-se difícil imaginar decisões relacionadas a políticas públicas que não se baseiam no contexto político anterior. Mesmo quando estamos diante de governos com concepções ideológicas diferentes, é improvável que as mudanças entre um e outro

pressuporiam um total e completo desmanche do que já havia sido realizado anteriormente em determinado contexto social. Aliás, como vimos no capítulo 2 da tese, este é um dos pontos de crítica em relação às políticas públicas em educação no Brasil, a descontinuidade que as caracteriza. A troca de governo, se implicasse sempre a falência das políticas vigentes em benefício de novas políticas, cujos efeitos propagandísticos se manifestariam na criação da “cara” do novo governo, levaria a um ciclo sempre inacabado de políticas públicas, tratadas, inevitavelmente, como ineficazes, visto que seus efeitos jamais seriam observados. Nesse sentido, o incrementalismo chama atenção para um ponto importante. As decisões dos governos anteriores se apresentam como limites, constrangimentos mais ou menos incisivos para decisões futuras. Se, de um lado, isso pode ser percebido como um problema para que vejamos realizadas mudanças mais profundas e de grande monta, de outro, dá uma sobrevida limitada a políticas já existentes, tendo em vista os custos econômicos, sociais e políticos que uma mudança completa nas políticas pode exigir.

Ao descrever a abordagem do ciclo de políticas, Souza (2002 p. 8) aponta os seguintes estágios da política: definição de agenda, identificação de alternativas, avaliação das opções, seleção das opções, implementação e avaliação. Segundo a autora, a abordagem do ciclo se concentra demasiadamente na definição da agenda, questionando o motivo pelo qual algumas questões são trazidas à cena do debate das políticas públicas, ao passo que outras são ignoradas. Para responder a essa questão, três respostas são apresentadas: i) a agenda é definida com base na natureza dos próprios problemas, assumindo que sua gravidade exige uma ação governamental para seu enfrentamento; ii) a agenda é construída com base na formação de uma consciência coletiva em torno do problema; e iii) a agenda é definida de acordo com a capacidade dos participantes, ou seja, grupos sociais são mais ou menos capazes de influenciar a agenda conforme seus próprios interesses.

De saída, é preciso reconhecer que essas respostas não são excludentes e podem se manifestar de maneira conjunta. Como não há uma explicação na abordagem do ciclo sobre como se forma a consciência coletiva de um problema, podemos pressupor que a própria gravidade do problema é parte responsável pela formação de tal consciência. Indo além, poderíamos nos questionar sobre a ideia de uma natureza dos problemas. A gravidade dos problemas públicos é ontológica, ou seja, parte de uma espécie de essência dos problemas, ou a própria ideia de que determinado fenômeno é um problema social é produto da sociedade e não da ontologia? Se algo é tratado como um problema é por ter sido pensado como um problema em determinado contexto social. A concepção ontológica de problemas sociais é

difícil de ser sustentada. A formação da consciência coletiva, por sua vez, precisa ser explicada. Ela sugere um processo coletivo de definição de prioridades com base na identificação de problemas, pendendo para uma interpretação de democracia participativa, onde a sociedade em geral deve definir o que é um problema social e o que não é.

Longe de desconsiderar a importância da sociedade na definição de problemas sociais, é necessário, contudo, relativizar tal participação, em virtude dos diferentes pesos que governo, grupos sociais mais influentes e o restante da população têm na definição da agenda de políticas públicas, o que nos leva à terceira resposta apresentada. Se, de um lado, é preciso mitigar a participação da sociedade na definição de políticas, sem anulá-la, contudo, de outro, faz-se necessário, como o polo oposto da explicação, relativizar o peso, quase determinístico, que pode ser atribuído aos grupos de interesse. Em leituras mais radicais, os grupos sociais mais influentes são vistos quase como os únicos capazes de definir a agenda, ignorando, simultaneamente, os interesses do próprio governo e a parcela de influência da sociedade em geral. A abordagem do ciclo de políticas apresenta o caráter processual das políticas, que teriam um ciclo de vida dividido em diferentes etapas.

Em seu mapeamento da literatura sobre políticas públicas, Souza (2002) apresenta, além da abordagem do ciclo de políticas, outros modelos para formulação e análise das políticas. A seguir, apresentamos as características gerais de cada modelo, ressaltando o que há, segundo nosso interesse de pesquisa, de mais interessante para a análise de políticas públicas e o que merece ressalvas e críticas.

O modelo *garbage can* data dos anos 1970 e parte do princípio de que há muitos problemas a serem enfrentados e poucas soluções disponíveis para enfrentá-los (SOUZA, 2002, p. 9). Assim, há uma inversão: os problemas são selecionados de acordo com as soluções disponíveis. Isso ocorre devido à limitação de soluções e da própria compreensão dos tomadores de decisão acerca dos problemas e das soluções. Como vimos no capítulo 2, a avaliação educacional em larga escala busca, justamente, lidar com o problema da escassez de informações para a tomada de decisões. No modelo *gargabe can*, as soluções são aquelas disponíveis em cada momento e isso limita a ação dos decisores, que agem por tentativa e erro. Como se vê, o modelo tem como foco a racionalidade dos decisores, sem levar em consideração a participação de outros atores. Contudo, ele tem o mérito de reconhecer dois pontos importantes: muitos problemas reconhecidos não apresentam soluções claras, algo que deve ser considerado pelos formuladores de políticas públicas, e os decisores possuem limitações para a tomada de decisão. O primeiro ponto coloca em cena o fato de que, nem sempre, a seleção de

um problema se dá por influência de grupos de interesse ou por decisão única dos governantes. A ausência de soluções claras para um problema pode, em determinado momento, ser o empecilho para que ele seja enfrentado através de uma política pública.

O modelo de coalizão de defesa, datado de início dos anos 1990, parte da crítica direcionada tanto ao ciclo de políticas quanto ao modelo *garbage can*, que afirma que ambos os modelos não têm poder explicativo em relação às mudanças que ocorrem nas políticas públicas (SOUZA, 2002, p. 9). O modelo propõe que as políticas públicas são formadas por subsistemas de valores, crenças e ideias, que impactam a formulação das políticas, servindo como constrangimento ou recurso para elas, algo ignorado pelos dois modelos aos quais dirige a crítica. Este modelo tem a virtude de reconhecer que valores, crenças e ideias devem ser considerados para o entendimento mais profundo de uma política pública (na verdade, poderíamos pensar que ele oferece suporte para que esses elementos seja selecionados como variáveis de interesse para a análise da política), contudo, não oferece uma explicação abalizada de como operam tais subsistemas, nem de como eles são constituídos.

O modelo das arenas sociais, por sua vez, propõe que as políticas públicas são iniciadas, fundamentalmente, pelos chamados empreendedores de políticas públicas, ou *policy community* (SOUZA, 2002, p. 10), uma comunidade de atores, especializados ou não, dispostos a investir em políticas públicas específicas, de modo que obtenham retorno para suas próprias demandas. Esses empreendedores seriam os grandes responsáveis por transformar um problema em um problema público, integrante da agenda governamental. De acordo com esse modelo, o governo passa a tratar uma questão como digna de receber uma política pública a partir da ação desses atores. Para tanto, a *policy community* chama a atenção dos tomadores de decisão dos governos lançando mão de, pelo menos, três mecanismos fundamentais: i) o uso indicadores e informações qualificadas que apresentam a dimensão do problema; ii) a exposição das consequências da manutenção do problema, que se repete sem que uma ação seja apresentada para enfrentá-lo de modo eficaz; e iii) o uso dos resultados da política pública em questão, mostrando suas falhas e sua ineficácia para enfrentar o problema. Mais uma vez, tais mecanismos podem ser utilizados de modo conjunto.

A avaliação educacional em larga escala serve aos propósitos desses mecanismos, na medida em que pode exercer o papel de desnudar um problema, ajudando a identificá-lo ou a sua dimensão. Essencialmente, o que seus resultados podem oferecer é um retrato da aprendizagem nas redes de ensino, ajudando a identificar problemas crônicos (dimensão temporal) e gerais (dimensão da amplitude). Ao avaliar sistematicamente a aprendizagem, um



sistema de avaliação influencia a agenda e a definição das políticas públicas. A título de exemplo, quando os resultados das redes de ensino para o ciclo de alfabetização são divulgados, mostrando que um percentual bastante significativo dos alunos matriculados nos anos iniciais não está plenamente alfabetizado ao fim do 3º ano do ensino fundamental, dificilmente, devido à publicidade de tais resultados, o governo pode ignorar sumariamente essa informação. Ele será cobrado por isso. Aliás, esse é um dos pontos para os quais a crítica em relação à avaliação em larga escala é dirigida com mais força. Como visto no corpo da tese, os sistemas de avaliação educacional no Brasil foram iniciativa dos governos, do Estado, e não das próprias redes de ensino e demais atores educacionais. Sendo assim, os resultados da avaliação são desqualificados pelos críticos como falseados, apresentando um cenário que não condiz com a realidade, servindo à propaganda do governo e escondendo o tamanho do problema. Entretanto, os resultados apresentam um cenário demasiadamente preocupante para que sejam usados como propaganda de governo.

A análise das políticas públicas proposta pelo modelo das arenas sociais tem a virtude de chamar a atenção para o fato de que as políticas não são fruto apenas da vontade dos tomadores de decisão do governo, envolvendo outros atores, dotados de interesses próprios. Além disso, dá ênfase a um elemento importante: a influência não direta e baseada apenas na manifestação volitiva dos grupos de interesse. Para que algo seja capaz de influenciar a agenda governamental, é necessário que algum respaldo seja dado. Esse suporte vem através de evidências e a avaliação educacional em larga escala ganha importância para a definição da agenda. Por outro lado, o modelo dá a entender que é apenas através da manifestação da *policy community* que uma agenda de políticas públicas é construída, ignorando que o próprio Estado, como ator interessado, possa ser ativo na definição da agenda, e não apenas receptor das evidências oferecidas pelos empreendedores das políticas públicas, além de não fazer menção ao papel de outros atores sociais, como professores e diretores, por exemplo, no caso das políticas educacionais.

O modelo do equilíbrio interrompido, com origem nos anos 1990, apresenta premissas muito parecidas com as do incrementalismo. De acordo com esse modelo, uma política pública é, em regra, caracterizada por períodos longos de estabilidade, interrompidos por períodos breves de instabilidade, nos quais mudanças mais significativas são realizadas nas políticas anteriores (SOUZA, 2002, p. 11). Assim, as experiências anteriores são as bases sobre as quais uma nova política é construída, a partir de um processo de implementação. Esse modelo se baseia na noção de sistema e apresenta grande afinidade com a concepção da Teoria dos

Sistemas de Luhmann (2010), influenciada, por sua vez, por noções da biologia e da informática. O modelo tem a virtude de reconhecer que as políticas públicas possuem uma base anterior com a qual dialogam antes de sua implementação, o que quer dizer que as políticas públicas não surgem do zero. Contudo, trata-se de um modelo que busca complexificar sua explicação, através do recurso ao conceito de sistema, sem que se produza uma concepção útil e original para a análise de políticas públicas. O incrementalismo chegou à mesma conclusão sem recorrer à controversa concepção de sistema para o entendimento do mundo social.

O modelo do novo gerencialismo público aponta que a agenda dos governos é influenciada, mais do que por grupos específicos de interesse (como proposto pelo modelo das arenas sociais), pela política fiscal, advogando eficiência e gastos mais enxutos por parte do Estado. A eficiência, inclusive, passa ser o principal objetivo da política pública (SOUZA, 2002, p. 12). Ela é um dos fatores que dá credibilidade à política, preconizando regras preestabelecidas em prejuízo da grande discricionariedade experimentada pelos governos na condução das políticas. Além disso, a busca pela eficiência seria o grande fundamento para a delegação de políticas públicas para entidades externas ao governo, aliada à capacidade técnica especializada de tais entidades e à sua independência do ciclo eleitoral. Souza afirma que, concorrendo com o modelo do novo gerencialismo público, há um movimento de fazer com que as políticas públicas, desde sua definição, assumam um caráter mais participativo. A ideia de concorrência apresentada por Souza (*idem*), nesse caso, não encerra sentido. O modelo do novo gerencialismo público é definido por características fiscais, mas não pela forma como a participação dos demais atores acontece. Isso quer dizer que tal modelo pode ser mais ou menos participativo, mantendo as características, principalmente fiscais, do novo gerencialismo. Esse modelo tem a virtude de chamar a atenção para as limitações fiscais impostas aos governos e para a consequente delegação de determinadas políticas para entidades privadas. Contudo, é um modelo com foco na definição da política, não trazendo nenhuma contribuição clara para a compreensão da implementação de um programa ou política. Ainda, centra-se no problema da delegação por parte do governo, vendo a necessidade de afirmar que o governo permanece tomando decisões acerca das políticas, mesmo com os processos de delegação.

Podemos observar algumas características comuns a esses modelos, apesar da diversidade de propostas. Souza (2002) ressalta que todos os modelos deixam claro que i) (o estudo da) política pública procura distinguir entre o que o governo pretende fazer e o que ele faz, de fato; ii) a análise pode recair sobre os diferentes níveis de governo e também sobre atores externos ao governo; iii) a política pública não é limitada apenas por regras; iv) a política é uma

ação intencional, tendo objetivos previstos (essa é uma questão importante para a conceituação de política pública); v) os impactos da política pública são de longo prazo (Souza tem razão em relação a isso, no entanto, chega a essa conclusão com base em outras fontes, visto que os modelos apresentados não tocam no tema do tempo necessário para a observância dos efeitos); e vi) a política pública vai além de sua decisão, envolvendo também sua implementação, execução e avaliação (o ciclo de políticas deixa isso claro, mas os demais modelos não abrangem tal distinção com tanta convicção e clareza).

Além dos modelos apresentados por Souza, vale a pena acrescentar, nesse rol exemplificativo, dois outros modelos apresentados por Frey (2000): o modelo do neoinstitucionalismo e o modelo de análise de estilos políticos. O neoinstitucionalismo tem foco no papel das instituições, entendidas como padrões regularizados de papéis e interações. O modelo salienta a importância das instituições para explicar a tomada de decisão no âmbito das políticas. As decisões sobre políticas públicas não são o resultado apenas de barganhas negociadas entre indivíduos perseguindo interesses próprios. Processos institucionais de socialização também as afetam. Os decisores agem de acordo com regras e práticas socialmente construídas (eles passaram por processos compartilhados de socialização). O tema das instituições coloca em cena um elemento extra àqueles que integram a análise de políticas públicas: as instituições fazem com que determinado problema chegue às instâncias decisórias, ao sistema político, afinal. A política em si, encarnada através de um conjunto de ações ou programas, é formulada com base na identificação de um problema e chega à sociedade política (*polity*) através das instituições.

Esse modelo dirige críticas ao paradigma da escolha racional, apresentando suas limitações, principalmente no que tange ao reducionismo do processo decisório a escolhas baseadas em poucas, desqualificadas ou excessivas informações. Há um ceticismo em relação à capacidade dos governos de tomarem decisões ajustadas no que tange às políticas públicas, em virtude de seus próprios interesses, de sua racionalidade limitada, de informações incompletas e da influências de atores e grupos privados.

As instituições trazem um elemento ético, de dever, de cumprimento com as expectativas projetadas para determinado papel. Frey (2000) chama atenção para o fato de que o neoinstitucionalismo tem maior poder explicativo em contextos de democracias consolidadas. Há uma contraposição entre esse modelo e a *policy science* tradicional, centrada na escolha racional. A virtude do modelo é chamar a atenção para o fato de que não existem apenas interesses individuais e egóicos, pois ações individuais estão embebidas em crenças compartilhadas, mantidas e

disseminadas por instituições. Este é um tema clássico da teoria social (Giddens, 2009, e Bourdieu, 2003, são exemplos). Contudo, é preciso não considerar as instituições, por mais importantes que sejam, como os únicos fatores explicativos para uma análise de políticas públicas.

Frey (2000) destaca, ainda, no âmbito do neoinstitucionalismo, a *policy network*, um conjunto de interações entre diferentes instituições, grupos e sociedade na gênese e implementação da política, com importância para a compreensão dos processos decisórios. Elas são menos formais do que as relações sociais institucionalizadas, mas possuem tal consistência e periodicidade que uma distribuição de papéis especificada é definida. As redes não se constroem em torno da política como um todo, mas, em regra, acerca de questões específicas de cada política, existindo concorrência interna entre os membros da rede e entre as redes. Essa é uma análise de grupos de interesse.

A análise de estilos políticos também parte da crítica à *policy science* tradicional, rechaçando a ideia de que processos políticos são determinados pelos conteúdos da política (máxima de Lowi). Instituições consolidadas são importantes para o êxito de políticas públicas e compreender fatores culturais, padrões de comportamento político e atitudes de atores políticos singulares é fundamental para avaliar a qualidade dos programas elaborados.

A comparação entre culturas políticas é um elemento central para essa linha de pesquisa. Isso é um problema, pois não há uma homogeneização da cultura política, principalmente hodiernamente, em contextos cada vez mais multifacetados e heterogêneos, em virtude dos processos de globalização e criação de redes internacionais (novamente, ver CASTELLS, 2016). São dois os objetivos principais da análise de estilos políticos: i) investigar quais são os fatores que podem levar à consolidação de um certo estilo político; nesse caso, o estilo é tratado como variável dependente das estruturas sociais e econômicas (por conta disso, recebe todas as críticas direcionadas a versões ortodoxas do marxismo); ii) investigar como os programas podem ser influenciados por certo estilo político; nesse caso, o estilo político é tratado como variável independente (FREY, 2000, p. 239).

Os tipos ideais de cultura política<sup>211</sup> elencados por Frey podem ter alguma aplicação em contextos e circunstâncias específicas, dependendo dos interesses do pesquisador, mas, dificilmente, podem ser tratados como instrumentos analíticos genéricos, aplicados a todo e

---

<sup>211</sup> A cultura paroquial é aquela na qual não existe uma visão política sistêmica; a cultura de súditos é por uma população interessada apenas nos resultados da política, mas desempenhando um papel passivo; e a cultura da participação é definida pela participação ativa da população no acontecimentos políticos (FREY, 2000, p. 237). Há um elemento hierarquizante e evolutivo nessa classificação, que não pode escapar das críticas: os tipos não são característicos de um país, podendo conviver ao mesmo tempo em contexto nacionais heterogêneos, caso de países como o Brasil; além disso, o modelo sugere que o tipo mais evoluído é o da participação, se sujeitando, dessa forma, a todas as críticas dirigidas a concepções evolucionistas no âmbito social.

qualquer contexto. A cultura política não define a estrutura política e, podendo ir contra ela, não é fator definitivo do conteúdo das políticas. Além disso, em particular num território como o Brasil, há diferenças significativas de cultura política (basta pensarmos na variedade oferecida pelos governos municipais e mesmo estaduais). A essa concepção podem ser dirigidas todas as críticas destinadas ao culturalismo antropológico<sup>212</sup>. Frey defende que o neoinstitucionalismo e análise de estilos políticos devem ser combinados, oferecendo ao analista uma melhor instrumentação para a pesquisa.

Por fim, é preciso ressaltar que cada um dos modelos aqui apresentados cobre uma parte significativa da análise das políticas públicas, sem dar conta do todo. Como vimos, a *policy science* é definida por Souza (2002) por seu holismo, por tentar dar conta da explicação do todo. No entanto, todos os modelos aqui apresentados são parciais. Eles lançam luzes sobre determinados aspectos, mantendo outros nas sombras. Antes de pensarmos tal característica dos modelos como um demérito, vale a pena chamar a atenção para o fato de que esses modelos são mais ou menos úteis para o analista de acordo com seus objetivos de pesquisa e tendo em vista a política ou programa que se quer analisar. Diante disso, o modelo das arenas sociais, por exemplo, é mais ajustado para uma pesquisa destinada a conhecer como determinado problema se tornou importante para o governo, sendo que, em contextos anteriores, ele não havia chamado a atenção dos tomadores de decisão. Nesse caso, a investigação sobre os grupos de interesse que participaram da definição da agenda é crucial para a compreensão de como o problema se tornou um problema público, observando como seus argumentos se tornaram os argumentos do governo para justificar a criação da política. Por outro lado, se o interesse recai sobre como uma política é implementada através de um programa específico, o modelo do ciclo de políticas pode ser mais elucidativo, visto que confere importância a essa etapa da política. Nesse caso, a investigação se debruçará sobre os implementadores e aqueles que fazem parte do que se costuma chamar de ponta do processo.

Além disso, cumpre notar o fato de que os modelos não são excludentes. As diferentes propostas de análise de cada um dos modelos podem ser organizadas em conjunto, de modo a servir a diferentes frentes de investigação por parte do analista de políticas públicas. Apesar da fragmentação que os caracteriza, um dos modelos parece permitir que todos os outros se encaixem em sua proposta. Trata-se da abordagem do ciclo de políticas. Por isso, ela foi selecionada como o modelo adotado para a análise empreendida nesta tese.

---

<sup>212</sup> A obra de Ruth Benedict, *O crisântemo e a espada*, é um esforço paradigmático para o estabelecimento de uma cultura nacional como categoria de investigação antropológica. A perspectiva do culturalismo, representada por ela, foi alvo de muitas críticas, em virtude, principalmente, da premissa de uma homogeneidade de contextos nacionais que não existia na realidade, apenas como categoria abstrata e estereotipada de análise.

## **APÊNDICE D - Limites das tipologias de Lowi e Deubel para o estudo de programas de avaliação da alfabetização e proposta do critério como gerador do tipo**

Conforme apresentado no capítulo 1 desta tese, a tipologia proposta por Lowi é uma das mais conhecidas na pesquisa em políticas públicas (MAIA & GRANDO, 2016). Como vimos, ele propõe uma tipologia com base na forma como as consequências da política produzem efeitos na população-alvo e como movimentam a máquina burocrática. Em específico, seu interesse tem um viés socioeconômico: a política tem seu tipo definido com base na distribuição (ou não) e na medida da distribuição dos benefícios gerados por ela, além de levar em consideração o aparato burocrático. Por isso, os quatro tipos de política que apresenta. Embora sua definição tipológica seja, ainda hoje, muito utilizada por pesquisadores da área para caracterizar políticas públicas, é preciso reconhecer os limites de sua tipologia para pesquisas destinadas a políticas públicas específicas, como é o caso aqui apresentado.

Lowi desenvolveu uma tipologia de caráter generalista, sem pensar em políticas específicas. A rigor, qualquer política pública pode ser caracterizada a partir dos tipos definidos por ele. Esse é, simultaneamente, o ponto forte e o ponto fraco de sua proposta tipológica. Do ponto de vista teórico, de modo a produzir explicações holísticas, sua tipologia pode ser útil a qualquer pesquisador, não importando a política que deseja investigar. Isso amplia, a princípio de modo ilimitado, as possibilidades de aplicação e uso de sua teoria. Por outro lado, ao serem tão genéricos, os tipos criados por ele não cobrem todas as características de interesse de uma política pública. Seu uso permanece adequado quando o pesquisador deseja dar ênfase à forma como os benefícios da política chegam ao público de destino. Se os interesses de pesquisa passam a ser outros, a aplicação da tipologia de Lowi perde em força explicativa e em capacidade de diferenciação das políticas. Sua tipologia se presta a diferenciar políticas públicas de diferentes áreas, mas não é tão útil quando se deseja diferenciar políticas que têm os mesmos objetivos e formas de distribuição de seus benefícios, sendo diferenciadas por outros critérios. Uma tipologia que não diferencia políticas públicas não atinge seu objetivo primordial. A diferenciação é a razão de ser da criação de tipologias, não importa a área de aplicação. Com as políticas públicas não é diferente: uma proposta tipológica deve ser capaz de diferenciar as políticas a partir de critérios que interessam ao pesquisador.

Nesse ponto, é preciso chamar a atenção para os problemas que uma tipologia generalista, como a de Lowi, pode gerar. Se ela é ampla, não leva em consideração as especificidades de cada política. Tais especificidades, contudo, podem ser exatamente aquilo

que o pesquisador deseja investigar, comparar e usar como critério de diferenciação. Assim, quando pensamos em avaliações da alfabetização como políticas públicas, estamos justamente diante de um cenário de especificidades, de características que fazem com que essas políticas sejam, até certo ponto, diferentes de todas as outras políticas públicas. Evidentemente, há características comuns a diversas políticas, motivo pelo qual elas são definidas como tais. Para definir a avaliação como política pública, é preciso reconhecer quais são as características que ela possui que a colocam como uma política pública, o que há em comum com outras políticas. O ciclo de políticas de Ball (conforme apresentado por Martins, 2013, e Mainardes, 2006) pode ser utilizado como um dos critérios para tanto (como outras políticas, a avaliação passa por uma definição da agenda, por processos de implementação, pela execução das ações previstas, pela avaliação de seus resultados etc). Entretanto, quando a decisão do pesquisador é por investigar a partir de uma análise mais fina e minuciosa, o que está em jogo é a identificação de especificidades e singularidades de determinada política.

Esse é o caso da investigação aqui proposta, acerca de três sistemas estaduais de avaliação da alfabetização no Brasil. Interessa menos, para os fins desta pesquisa, se as políticas de avaliação da alfabetização dos estados aqui analisados possuem características comuns com políticas de outros setores, como saúde e segurança pública, por exemplo, para fins de investigação da coerência do poder público no que tange às suas diferentes ações. Interessa mais a análise comparada da política pública em três cenários diferentes, encontrando, com objetivo heurístico, similaridades e diferenças entre elas. Tais semelhanças e diferenças são encontradas a partir da definição de critérios através dos quais a comparação pode ser feita. A escolha dos critérios, vale notar, é arbitrada pelo pesquisador, de acordo com os objetivos da pesquisa e com o que será definido como prioritário e mais importante. Essa definição será, em regra, uma tarefa difícil e ingrata. De um modo geral, por mais minuciosos que sejam os pesquisadores, os critérios a partir dos quais as políticas podem ser comparadas são muito amplos. Assim, a definição dos critérios é, antes, uma escolha: os critérios são selecionados, dentre muitos possíveis, de acordo com os objetivos da pesquisa e com a capacidade do pesquisador de investigar todos eles, tendo em vista a emergência de produzir resultados e informações.

O fato dos critérios serem definidos de modo arbitrário não é, de saída, um problema. Trata-se de um reconhecimento, não de uma crítica (pelo menos, não necessariamente). Além disso, ser arbitrário não significa ausência de justificção. A seleção depende do pesquisador e do tipo de pesquisa que deseja levar a cabo, mas isso não significa que cada decisão não deva

ser justificada e explicada. Aliás, é justamente por se tratar de uma escolha que os critérios devem ser explicados. Eles podem fazer sentido quando o interesse de pesquisa é um, ao passo que serão descartados se o interesse for outro. A justificativa para a escolha de um conjunto de critérios e não de outro passa, também, pela literatura sobre o tema, que ajudará a definir o que tem sido tratado como importante em determinada área de estudo.

Diante disso, cumpre reconhecer que a tipologia de Lowi não se revela útil para a análise aqui empreendida. Seu uso faria com que a investigação tivesse pouco poder explicativo, permanecendo em um nível muito genérico de análise, sem adentrar nas especificidades das políticas de avaliação da alfabetização, que é justamente o que interessa à pesquisa. Mais uma vez, isso não significa que sua tipologia não permaneça sendo útil para outros tipos de investigação. Contudo, para análise das políticas de avaliação da alfabetização, seu poder de diferenciação entre as políticas é muito baixo.

Tudo o que foi dito em relação à tipologia de Lowi se aplica aos tipos de políticas públicas propostos por Deubel (2009), como também vimos no capítulo 1 da tese. Seus quatro tipos de política são destinados a orientar a formulação e a implementação das políticas, mas não necessariamente a analisá-las (embora uma coisa não exclua a outra). Assim como Lowi, Deubel permanece interessado em uma perspectiva teórica holística, que dê conta de toda e qualquer política pública. Sua aplicação, a princípio, parece permitir análises mais específicas no âmbito educacional do que a tipologia de Lowi, mas, mesmo assim, não leva em consideração as especificidades da avaliação como política pública.

Como, então, selecionar os critérios para a análise comparada de políticas de avaliação da alfabetização? Tendo em vista que o interesse da pesquisa é a comparação entre três sistemas de avaliação da alfabetização, o ponto de partida é selecionar o que importa para a análise da avaliação educacional. É preciso chamar atenção para um ponto. Definir as características que importam para a investigação da avaliação educacional, o que já se configura como um nível de especificação de análise das políticas públicas de maneira geral, poderia levar a um segundo nível de especificação, qual seja, a análise das características da avaliação da alfabetização. Nesse caso, teríamos três níveis de análise das políticas públicas: i) políticas públicas em geral (nível mais abstrato); ii) políticas de avaliação educacional (nível intermediário); e iii) políticas de avaliação da alfabetização (nível mais específico). Esses níveis poderiam chegar a cinco, caso fossem acrescidos da análise das políticas educacionais e das políticas educacionais em alfabetização.



Para os fins desta pesquisa, a decisão foi por construir uma tipologia das políticas de avaliação da educação, aplicando-a às avaliações da alfabetização. Essa decisão não impede sua revisão. Por ora, no entanto, os critérios selecionados e apresentados no capítulo 4 dizem respeito, em sua maioria, a políticas de avaliação da educação como um todo, não apenas às avaliações da alfabetização (exceção feita ao critério *Concepção de alfabetização que dá base à avaliação*). Como o que se aplica ao gênero se aplica também à espécie, as referências no texto sempre falam em avaliações da alfabetização.

A seleção apresentada a seguir se baseia na ideia de que os critérios cobrem os pontos mais importantes para a análise de uma política de avaliação da alfabetização no contexto dos estados brasileiros. Com base neles, pode-se oferecer uma visão detalhada das políticas de avaliação, permitindo a comparação entre os diferentes sistemas aqui analisados. Reforço, novamente, que outros critérios poderiam ter sido selecionados, mas os que aqui se encontram foram tratados como suficientes para construir a análise comparativa desejada.

É preciso ressaltar alguns pontos importantes sobre a tipologia aqui proposta. O primeiro deles se relaciona com a própria ideia de tipologia, que difere significativamente das tipologias de Lowi e Deubel, usadas aqui como contraposição. No caso das tipologias propostas por esses autores, cada tipo de política envolve um conjunto de características que compõem esse tipo. Os tipos, dessa forma, são compostos e não unitários. Diferentes critérios são selecionados e reunidos para formar um tipo. No caso de Deubel, isso fica ainda mais claro, visto que ele reúne diferentes características das políticas para definir seus quatro tipos (conforme já foi dito, Lowi se concentra na forma como os benefícios da política chegam a seus beneficiários e no aparato burocrático para a construção de sua tipologia). Essa mesma lógica poderia ser aplicada aqui para a construção de uma tipologia das políticas de avaliação da alfabetização de outra natureza e grau de complexidade. A título de exemplo, um tipo 1 de política de avaliação da alfabetização, digamos, poderia ser formado por uma política composta por um federalismo de adesão (critério hipotético 1), por um vínculo frágil com uma política de formação de professores na alfabetização (critério hipotético 2) e por ações brandas de responsabilização atreladas aos resultados da avaliação (critério hipotético 3). Nesse caso, o tipo é definido pela combinação de três critérios (evidentemente, outros critérios poderiam ser usados para compor o tipo).

Contudo, a tipologia aqui proposta tem uma natureza diferente. Ela não reúne diferentes critérios para a formação de um tipo de política. Antes, o que ela faz é tratar cada critério como elemento gerador de um tipo. Diante disso, no lugar da reunião de critérios para a formação do

tipo, o que vemos é cada critério definir um tipo com base nele. É preciso reconhecer que há vantagens e desvantagens em seguir qualquer um desses caminhos. A tipologia por reunião de critérios apresenta tipos mais complexos de políticas, pois diferentes características são levadas em consideração para sua definição. Entretanto, esse caminho não pode ser trilhado com a observância de muitos critérios, sob pena de ficar impossível criar tipos com tantos critérios diferentes reunidos (as combinações seriam excessivas).

Com isso, agir dessa forma aumenta a complexidade do tipo diminuindo o número de características observadas. Por outro lado, a tipologia do critério como tipo não oferece a mesma complexidade de um tipo que reúne diferentes critérios, mas permite a observação e análise de diferentes características. Não parece haver, de saída, uma prevalência teórica de uma forma de construção de tipologia sobre a outra. Mais uma vez, trata-se de uma escolha do pesquisador. A seleção por um caminho de construção dos tipos trará, ao mesmo tempo, perdas e vantagens. No âmbito desta tese, a escolha foi pelo critério como gerador do tipo, tendo em vista que o interesse é abarcar o maior número possível de características de uma política de avaliação da alfabetização. Como não estamos diante de uma área de pesquisa com produção massiva e consolidada sobre o tema (qual seja, tipos de políticas de avaliação da alfabetização), parece ser mais convincente explorar tanto quanto possível a análise com base em muitas características da política. Isso não impede, porém, que, em esforços vindouros, uma reunião dos diferentes critérios aqui selecionados seja realizada para a construção de tipos mais complexos de políticas de avaliação da alfabetização (o que significaria a redução do número de critérios selecionados).

O segundo ponto, relacionado ao primeiro, diz respeito ao fato de que, se cada critério é um tipo, as políticas serão de tantos tipos diferentes quanto forem os critérios. Assim, os critérios não são, a princípio, excludentes. Há um conjunto de critérios selecionados e, dessa forma, as políticas serão definidas com base em cada um deles. As divisões entre eles também são arbitrárias, o que significa dizer que, em alguns casos, dois ou mais critérios poderiam ser reunidos em um único critério, ou, caso contrário, o que está sob o rótulo de um único critério poderia ter sido dividido em dois ou mais critérios. A decisão por cada um deles foi justificada ao longo do capítulo 4. Ainda, o fato de serem tratados como critérios independentes, para efeitos de análise, não significa que eles não influenciem uns aos outros. A título de exemplo, isso quer dizer que, se uma política de avaliação da alfabetização tem um determinado alcance, isso afetará as possibilidades de seu atrelamento a políticas de responsabilização.

O terceiro ponto acerca da tipologia apresentada é o fato de ela não excluir, também, as tipologias construídas por outros autores, como Lowi e Deubel. Aqui, a proposta é específica, ao passo que naqueles autores é genérica. Isso quer dizer que uma política pode ser classificada, por exemplo, como distributiva (em Lowi) e de tipo 2 (em Deubel), e, ainda sim, ser classificada em cada um dos critérios aqui propostos.

O quarto ponto diz respeito à pretensão dessa tipologia. O objetivo não é fazer uma proposta definitiva e hermética para a análise das políticas de avaliação da alfabetização, mas, antes, propor uma tipologia útil para a análise comparativa nessa seara. Como não existe literatura consolidada sobre o tema, o caminho é exploratório, conforme ressaltado anteriormente. Os critérios, portanto, não representam um rol taxativo. Eles estão abertos a ajustes, acréscimos e exclusões. De todo modo, o que se espera é que essa tipologia seja um ponto de partida para a definição das características que importam para a análise de políticas públicas de avaliação da alfabetização. Conforme afirmado no capítulo 4 e na conclusão, essa pode ser uma contribuição significativa da tese: explorar uma temática que recebeu pouca atenção específica dos pesquisadores e lançar luz sobre os critérios que ajudam a definir uma tipologia para as políticas de avaliação da alfabetização.

## ANEXO A - Matrizes de Referência do Paebes Alfa – 2009 a 2017

## Matriz de Referência do Paebes Alfa – 1º, 2º e 3º EF Língua Portuguesa 2009

Tópico	Descritores
T1. Aquisição do sistema da escrita	D1. Reconhecer especificidades da linguagem escrita
	D2. Diferenciar letras de outros sinais gráficos
	D3. Identificar letras do alfabeto
	D4. Distinguir, como leitor, diferentes tipos de letras
	D5. Reconhecer as direções e o alinhamento da escrita
	D6. Reconhecer a palavra como unidade gráfica
	D7. Identificar sons de sílabas em palavra(s)
	D8. Identificar relações fonema/grafema (som/letra)
	D9. Ler palavras silenciosamente
	D10. Demonstrar conhecimentos sobre os usos sociais da ordem alfabética.
T2. Leitura/compreensão	D10. Demonstrar conhecimentos sobre os usos sociais da ordem alfabética.
	D11. Identificar gêneros textuais diverso
	D12. Identificar finalidade de textos de diferentes gêneros e suportes
	D13. Localizar informação explícita em frases/ textos
	D14. Identificar elementos que constroem a narrativa
	D15. Formular hipóteses
	D16. Identificar assunto de um texto
	D17. Inferir o sentido de uma palavra ou expressão
	D18. Inferir informação implícita em textos diversos
	D19. Identificar efeitos de humor ou ironia em textos diversos
	D20. Identificar marcas linguísticas que evidenciam os interlocutores de um texto
	D21. Estabelecer relações entre partes de um texto que contribuem para a sua continuidade
	D22. Estabelecer relações lógico-discursivas presentes no texto, em situação de uso da língua
D23. Identificar efeito de sentido decorrente do uso de sinais de pontuação ou recursos gráficos	
T3. Escrita	D24. Escrever palavras
	D25. Escrever frases/textos

Fonte: Paebes Alfa 2009. Tabulação do autor.

**Matriz de Referência do Paebes Alfa – 1º, 2º e 3º EF Língua Portuguesa – 2010 e 2011**

Tópico	Competências	Descritores
T1. Aquisição do sistema da escrita	C1. Identificação de letras do alfabeto	D1 - Reconhecer especificidades da linguagem escrita
		D2 - Diferenciar letras de outros sinais gráficos.
		D3 - Identificar letras do alfabeto.
		D4 - Distinguir, como leitor, diferentes tipos de letras.
	C2. Uso adequado da página	D5 - Reconhecer as direções e o alinhamento da escrita da língua portuguesa.
T2. Apropriação do sistema alfabético.	C3. Reconhecimento da palavra como unidade gráfica	D6 - Compreender a função da segmentação de espaços em branco, na delimitação de palavras em textos escritos (consciência de palavras).
	C4. Aquisição da consciência fonológica.	D7 – Identificar, ao ouvir uma palavra, o número de sílabas que ela contém (consciência silábica).
		D8 – Identificar sílabas e sons (consciência silábica e consciência fonêmica).
		D9 – Identificar relações fonema/grafema, som/letra (consciência fonêmica).
	C5. Leitura de palavras, frases e pequenos textos.	D10 – Ler palavras silenciosamente.
		D11 – Ler frases e pequenos textos, localizando informações explícitas contidas neles.
T3 – Usos sociais da leitura e da escrita.	C6. Implicações do suporte e do gênero na compreensão de textos.	D12 – Reconhecer os usos sociais da ordem alfabética.
		D13 – Identificar suportes e gêneros textuais diversos.
		D14 – Reconhecer a finalidade de gêneros diversos.
T4- Leitura: compreensão, análise e avaliação.	C7. Localização de informações explícitas em textos.	D15 – Localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados.
		D16 – Identificar elementos que constroem a narrativa.
	C8. Interpretação de informações implícitas em textos.	D17 – Inferir informações implícitas em textos.
		D18 – Identificar assunto de frases e textos.
		D19 – Formular hipóteses.
		D20 – Identificar efeitos de humor ou ironia em textos diversos.
	C9. Coerência e coesão no processamento de textos	D21- Estabelecer relações lógico-discursivas presentes no texto.
		D22 – Estabelecer relações de continuidade temática, a partir da recuperação de elementos da cadeia referencial do texto.
		D23 – Identificar o efeito de sentido decorrente do uso de recursos gráficos, da pontuação, da seleção lexical e repetições.
		D24 – Identificar marcas linguísticas que evidenciam o enunciador no discurso direto ou indireto.
C10. Avaliação do leitor em relação aos textos lidos.	D25 – Distinguir fato de opinião sobre o fato.	
	D26 – Identificar tese e argumentos.	
	D27 – Avaliar a adequação da linguagem usada à situação, sobretudo a eficiência de um texto ao seu objetivo ou finalidade.	
T5 – Produção escrita	C11. Escrita de palavras e frases.	D28 – Escrever palavras.
		D29 - Escrever frases.
	C12. Produção de textos.	D30 - Produzir textos.

Fonte: Paebes Alfa 2010. Tabulação do autor.

**Matriz de Referência do Paebes Alfa – 1ºEF Língua Portuguesa – 2014/2017**

<b>T1- Reconhecimento de convenções do sistema alfabético.</b>	
<b>C1. Identificação de letras do alfabeto.</b>	
<b>D01</b>	Reconhecer especificidades da linguagem escrita.
<b>D02</b>	Identificar letras do alfabeto.
<b>D03</b>	Diferenciar letras de outros sinais gráficos, como os números, sinais de pontuação ou de outros sistemas de representação.
<b>D04</b>	Distinguir, como leitor, diferentes tipos de letras.
<b>C2. Uso adequado da página.</b>	
<b>D05</b>	Reconhecer as direções e o alinhamento da escrita da língua portuguesa.
<b>T2. Apropriação do sistema alfabético.</b>	
<b>C3. Reconhecimento da palavra como unidade gráfica.</b>	
<b>D06</b>	Compreender a função da segmentação de espaços em branco, na delimitação de palavras em textos escritos (consciência de palavras).
<b>C4. Aquisição da consciência fonológica.</b>	
<b>D07</b>	Identificar o número de sílabas de uma palavra (consciência silábica).
<b>D08</b>	Identificar sílabas e sons (consciência silábica e consciência fonêmica).
<b>D09</b>	Identificar relações fonema/grafema, som/letra (consciência fonêmica).
<b>C5. Leitura de palavras, frases e pequenos textos.</b>	
<b>D10</b>	Ler palavras silenciosamente.
<b>D11</b>	Ler frases e pequenos textos, localizando informações explícitas contidas neles.
<b>T3 – Usos sociais da leitura e da escrita.</b>	
<b>C6. Implicações do suporte e do gênero na compreensão de textos.</b>	
<b>D12</b>	Reconhecer o local de inserção de determinada palavra numa sequência em ordem alfabética.
<b>D13</b>	Identificar gêneros textuais diversos.
<b>D14</b>	Reconhecer a finalidade de gêneros diversos.
<b>T4- Leitura: compreensão, análise e avaliação.</b>	
<b>C7. Localização de informações explícitas em textos.</b>	
<b>D15</b>	Localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados.
<b>D16</b>	Identificar elementos que constroem a narrativa.
<b>C8. Interpretação de informações implícitas em textos.</b>	
<b>D17</b>	Inferir informações implícitas em textos.
<b>D18</b>	Identificar assunto de textos.
<b>T5 – Produção escrita.</b>	
<b>C11. Escrita de palavras e frases.</b>	
<b>D28</b>	Escrever palavras.
<b>D29</b>	Escrever frases.
<b>C12. Produção de textos.</b>	
<b>D30</b>	Produzir textos.

Fonte: Paebes Alfa 2016. Tabulação do autor.

### Matriz de Referência do Paebes Alfa – 2ºEF Língua Portuguesa – 2014/2017

<b>T1- Reconhecimento de convenções do sistema alfabético.</b>	
<b>C1. Identificação de letras do alfabeto.</b>	
<b>D02</b>	Identificar letras do alfabeto.
<b>D03</b>	Diferenciar letras de outros sinais gráficos, como os números, sinais de pontuação ou de outros sistemas de representação.
<b>D04</b>	Distinguir, como leitor, diferentes tipos de letras.
<b>C2. Uso adequado da página.</b>	
<b>D05</b>	Reconhecer as direções e o alinhamento da escrita da língua portuguesa.
<b>T2. Apropriação do sistema alfabético.</b>	
<b>C3. Reconhecimento da palavra como unidade gráfica.</b>	
<b>D06</b>	Compreender a função da segmentação de espaços em branco, na delimitação de palavras em textos escritos (consciência de palavras).
<b>C4. Aquisição da consciência fonológica.</b>	
<b>D07</b>	Identificar o número de sílabas de uma palavra (consciência silábica).
<b>D08</b>	Identificar sílabas e sons (consciência silábica e consciência fonêmica).
<b>D09</b>	Identificar relações fonema/grafema, som/letra (consciência fonêmica).
<b>C5. Leitura de palavras, frases e pequenos textos.</b>	
<b>D10</b>	Ler palavras silenciosamente.
<b>D11</b>	Ler frases e pequenos textos, localizando informações explícitas contidas neles.
<b>T3 – Usos sociais da leitura e da escrita.</b>	
<b>C6. Implicações do suporte e do gênero na compreensão de textos.</b>	
<b>D12</b>	Reconhecer o local de inserção de determinada palavra numa sequência em ordem alfabética.
<b>D13</b>	Identificar gêneros textuais diversos.
<b>D14</b>	Reconhecer a finalidade de gêneros diversos.
<b>T4- Leitura: compreensão, análise e avaliação.</b>	
<b>C7. Localização de informações explícitas em textos.</b>	
<b>D15</b>	Localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados.
<b>D16</b>	Identificar elementos que constroem a narrativa.
<b>C8. Interpretação de informações implícitas em textos.</b>	
<b>D17</b>	Inferir informações implícitas em textos.
<b>D18</b>	Identificar assunto de textos.
<b>D20</b>	Identificar efeitos de humor em textos diversos.
<b>C9. Coerência e coesão no processamento de textos</b>	
<b>D21</b>	Estabelecer relações lógico-discursivas presentes no texto.
<b>D22</b>	Estabelecer relações de continuidade temática, a partir da recuperação de elementos da cadeia referencial do texto.
<b>D23</b>	Identificar o efeito de sentido decorrente do uso de recursos gráficos, da pontuação, da seleção lexical e repetições.
<b>T5 – Produção escrita.</b>	
<b>C11. Escrita de palavras e frases.</b>	
<b>D28</b>	Escrever palavras.
<b>D29</b>	Escrever frases.
<b>C12. Produção de textos.</b>	
<b>D30</b>	Produzir textos.

Fonte: Paebes Alfa 2016. Tabulação do autor.

### Matriz de Referência do Paebes Alfa – 3ºEF Língua Portuguesa – 2014/2017

<b>T1- Reconhecimento de convenções do sistema alfabético.</b>	
<b>C1. Identificação de letras do alfabeto.</b>	
<b>D02</b>	Identificar letras do alfabeto.
<b>D03</b>	Diferenciar letras de outros sinais gráficos, como os números, sinais de pontuação ou de outros sistemas de representação.
<b>D04</b>	Distinguir, como leitor, diferentes tipos de letras.
<b>C2. Uso adequado da página.</b>	
<b>D05</b>	Reconhecer as direções e o alinhamento da escrita da língua portuguesa.
<b>T2. Apropriação do sistema alfabético.</b>	
<b>C3. Reconhecimento da palavra como unidade gráfica.</b>	
<b>D06</b>	Compreender a função da segmentação de espaços em branco, na delimitação de palavras em textos escritos (consciência de palavras).
<b>C4. Aquisição da consciência fonológica.</b>	
<b>D07</b>	Identificar o número de sílabas de uma palavra (consciência silábica).
<b>D08</b>	Identificar sílabas e sons (consciência silábica e consciência fonêmica).
<b>D09</b>	Identificar relações fonema/grafema, som/letra (consciência fonêmica).
<b>C5. Leitura de palavras, frases e pequenos textos.</b>	
<b>D10</b>	Ler palavras silenciosamente.
<b>D11</b>	Ler frases e pequenos textos, localizando informações explícitas contidas neles.
<b>T3 – Usos sociais da leitura e da escrita.</b>	
<b>C6. Implicações do suporte e do gênero na compreensão de textos.</b>	
<b>D12</b>	Reconhecer o local de inserção de determinada palavra numa sequência em ordem alfabética.
<b>D13</b>	Identificar gêneros textuais diversos.
<b>D14</b>	Reconhecer a finalidade de gêneros diversos.
<b>T4- Leitura: compreensão, análise e avaliação.</b>	
<b>C7. Localização de informações explícitas em textos.</b>	
<b>D15</b>	Localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados.
<b>D16</b>	Identificar elementos que constroem a narrativa.
<b>C8. Interpretação de informações implícitas em textos.</b>	
<b>D17</b>	Inferir informações implícitas em textos.
<b>D18</b>	Identificar assunto de textos.
<b>D20</b>	Identificar efeitos de humor em textos diversos.
<b>C9. Coerência e coesão no processamento de textos</b>	
<b>D21</b>	Estabelecer relações lógico-discursivas presentes no texto.
<b>D22</b>	Estabelecer relações de continuidade temática, a partir da recuperação de elementos da cadeia referencial do texto.
<b>D23</b>	Identificar o efeito de sentido decorrente do uso de recursos gráficos, da pontuação, da seleção lexical e repetições.
<b>D24</b>	Identificar marcas linguísticas que evidenciam o enunciador no discurso direto ou indireto.
<b>C10. Avaliação do leitor em relação aos textos lidos.</b>	
<b>D25</b>	Distinguir fato de opinião sobre o fato.
<b>T5 – Produção escrita.</b>	
<b>C11. Escrita de palavras e frases.</b>	
<b>D28</b>	Escrever palavras.
<b>D29</b>	Escrever frases.
<b>C12. Produção de textos.</b>	
<b>D30</b>	Produzir textos.

Fonte: Paebes Alfa 2016. Tabulação do autor.



### Matriz de Referência do Paebes Alfa – 1º a 3ºEF Matemática – 2016/2017

T1. Reconhecimento de números e operações	
C1. Mobilizar ideias, conceitos e estruturas relacionadas à construção do significado dos números e suas representações	
D01	Associar quantidades de objetos/pessoas/animais à sua representação numérica.
D02	Associar um número natural à sua escrita por extenso.
D03	Comparar ou ordenar quantidades de objetos/pessoas/animais.
D04	Comparar ou ordenar números naturais.
D05	Reconhecer números ordinais ou indicadores de posição.
C2. Mobilizar conceitos e propriedades numéricas, para resolver problemas	
D06	Resolver problemas com números naturais, envolvendo diferentes significados da adição ou subtração.
D07	Resolver problemas com números naturais, envolvendo diferentes significados da multiplicação ou divisão.
C3. Mobilizar conceitos e propriedades numéricas para efetuar operações	
D08	Efetuar a adição ou subtração de números naturais.
T2. Noções de espaço e forma	
C4. Reconhecer figuras geométricas planas ou espaciais	
D09	Identificar a representação de figuras bidimensionais.
D10	Identificar a representação de figuras tridimensionais.
C5. Localizar objetos em representações do espaço	
D11	Identificar a localização ou movimentação de pessoas, objetos ou pontos em representação plana do espaço.
T3. Noções de grandezas e medidas	
C6. Mobilizar conceitos e propriedades relacionadas a grandezas e medidas para comparar, identificar ou efetuar medições	
D12	Comparar ou ordenar comprimento, altura e espessura.
D13	Identificar ou relacionar cédulas e moedas do Sistema Monetário Brasileiro.
D14	Estabelecer relações entre unidades de medidas de tempo.
C7. Reconhecer grandezas e suas diferentes unidades de medida	
D15	Identificar diferentes maneiras de medir uma grandeza.
D16	Ler horas em relógios digitais e/ou analógicos.
T4. Tratamento da informação	
C8. Reconhecer informações e dados apresentados em gráficos, tabelas ou gêneros textuais	
D17	Identificar informações apresentadas em quadros ou tabelas.
D18	Identificar informações apresentadas em gráficos de colunas.
D19	Identificar informações apresentadas em diferentes gêneros textuais.

Fonte: Paebes Alfa 2017. Tabulação do autor.

## ANEXO B - Matrizes de Referência do Proalfa – 2009 a 2017

### Matriz de Referência do Proalfa – Língua Portuguesa - 2008/2009

Conhecimentos	Competências	Descritores
Características da tecnologia da escrita	C1. Domínio de conhecimentos e capacidades que concorrem para a apropriação da tecnologia de escrita	D1. Identificar letras do alfabeto.
		D2. Conhecer as direções e o alinhamento da escrita da língua portuguesa.
		D3. Diferenciar letras de outros sinais gráficos, como os números, sinais de pontuação ou de outros sistemas de representação.
Conhecimentos	Competências	Descritores
Decifração e fluência	C2. Decifração com maior ou menor fluência	D4. Distinguir, como leitor, diferentes tipos de letras.
		D5. Identificar, ao ouvir uma palavra, o número de sílabas (consciência silábica).
		D6. Identificar sons de sílabas (consciência fonológica e consciência fonêmica).
		D7. Identificar o conceito de palavra (consciência de palavra).
		D8. Ler palavras em voz alta.
Conhecimentos	Competências	Descritores
	C3. Recuperação de informações no contexto de práticas sociais de leitura	D9. Ler, em voz alta, uma frase/ ou um texto.
		D10. Ler palavras silenciosamente.
		D11. Localizar informação em uma frase/texto.
Conhecimentos	Competências	Descritores
Compreensão	C3. Recuperação de informações no contexto de práticas sociais de leitura	D12. Identificar elementos que constroem a narrativa
		D13. Inferir uma informação
		D14. Identificar assunto
		D15. Estabelecer relações lógico-discursivas
Conhecimentos	Competências	Descritores
Usos sociais da leitura e da escrita	C4. Implicações do suporte e do gênero na compreensão de textos	D16. Estabelecer relações de continuidade temática a partir da recuperação de elementos da cadeia referencial do texto
		D17. Reconhecer os usos sociais da ordem alfabética
		D18. Identificar gêneros textuais diversos e sua finalidade.
		D19. Formular hipóteses
Conhecimentos	Competências	Descritores
Avaliação e posicionamento do leitor em relação aos textos	C5. Julgamento e crítica	D20. Distinguir fato de opinião sobre o fato
		D21. Identificar tese e argumentos
		D22. Avaliar adequação da linguagem usada à situação, sobretudo, a eficiência de um texto ao seu objetivo ou finalidade.
		D23. Determinar o ponto de vista do enunciador ou de personagens sobre fatos, apresentados explícita e implicitamente no texto.
		D24. Identificar efeito de sentido decorrente de recursos gráficos, seleção lexical e repetição.
Conhecimentos	Competências	Descritores
Escrita	C6. Escrita de palavras (Codificação)	D25. Escrever palavras
	C7. Produção escrita	D26. Escrever frases/ textos

Fonte: Proalfa 2009. Tabulação do autor.

## Matriz de Referência do Proalfa – Língua Portuguesa - 2010

Tópicos	Competências	Descritores
T1- Reconhecimento de convenções do sistema alfabético	C1. Identificação de letras do alfabeto	D1. Identificar letras do alfabeto
		D2. Diferenciar letras de outros sinais gráficos, como os números, sinais de pontuação ou de outros sistemas de representação
		D3. Distinguir, como leitor, diferentes tipos de letras
	C2. Uso adequado da página	D4. Conhecer as direções e o alinhamento da escrita da língua portuguesa
T2- Apropriação do sistema alfabético	C3. Aquisição de consciência fonológica	D5. Identificar, ao ouvir uma palavra, o número de sílabas (consciência silábica)
		D6. Identificar sons de sílabas (consciência fonológica e consciência fonêmica)
	C4. Reconhecimento da palavra como unidade gráfica	D7. Compreender a função de segmentação de espaços em branco na delimitação de palavras em textos escritos
	C5. Leitura de palavras e pequenos textos	D8. Ler palavras
D9. Ler pequenos textos		
T3 - Leitura: compreensão, análise e avaliação	C6. Localização de informações explícitas em textos	D10. Localizar informação explícita em textos de maior extensão e de gêneros e temas menos familiares
		D11. Identificar elementos que constroem a narrativa
	C7. Interpretação de informações implícitas em textos	D12. Inferir informações em textos
		D13. Identificar assunto de texto
		D14. Formular hipóteses
	C8. Coerência e coesão no processamento de textos	D15. Estabelecer relações lógico-discursivas presentes no texto
		D16. Estabelecer relações de continuidade temática a partir da recuperação de elementos da cadeia referencial do texto
		D17. Identificar efeito de sentido decorrente de recursos gráficos, seleção lexical e repetição
	C9. Avaliação do leitor em relação aos textos	D18. Identificar marcas linguísticas que evidenciam o enunciador no discurso direto ou indireto
D19. Distinguir fato de opinião sobre o fato		
D20. Identificar tese e argumentos		
D21. Avaliar a adequação da linguagem usada à situação, sobretudo, a eficiência de um texto ao seu objetivo ou finalidade		

<b>T4- Usos sociais da leitura e da escrita</b>	C10. Implicações do gênero e do suporte na compreensão de textos	<b>D22. Reconhecer os usos sociais da ordem alfabética</b>
		<b>D23. Identificar gêneros textuais diversos</b>
		<b>D24. Reconhecer finalidade de gêneros textuais diversos</b>
<b>T5 - Produção escrita</b>	C11. Escrita de palavras	<b>D25. Escrever palavras</b>
	C12. Escrita de frases/textos	<b>D26. Escrever frases/textos</b>

Fonte: Proalfa 2010. Tabulação do autor.

## Matriz de Referência do Proalfa – Língua Portuguesa - 2011 e 2012

Tópicos	Competências	Habilidades
T1- Reconhecimento de convenções do sistema alfabético.	C1. Identificação de letras do alfabeto.	H1- Reconhecer especificidades da linguagem escrita.
		H2- Identificar letras do alfabeto.
	C2. Uso adequado da página.	H3- Diferenciar letras de outros sinais gráficos, como os números, sinais de pontuação ou de outros sistemas de representação.
		H4- Distinguir, como leitor, diferentes tipos de letras.
T2. Apropriação do sistema alfabético.	C3. Reconhecimento da palavra como unidade gráfica.	H5- Reconhecer as direções e o alinhamento da escrita da língua portuguesa.
	C4. Aquisição da consciência fonológica.	H6- Compreender a função da segmentação de espaços em branco, na delimitação de palavras em textos escritos (consciência de palavras).
		H7- Identificar, ao ouvir uma palavra, o número de sílabas que ela contém (consciência silábica).
		H8- Identificar sílabas e sons (consciência silábica e consciência fonêmica).
	C5. Leitura de palavras, frases e pequenos textos.	H9- Identificar relações fonema/grafema, som/letra (consciência fonêmica).
T3 – Usos sociais da leitura e da escrita.	C6. Implicações do suporte e do gênero na compreensão de textos.	H10- Ler palavras silenciosamente.
		H11- Ler frases e pequenos textos, localizando informações explícitas contidas neles.
		H12- Reconhecer os usos sociais da ordem alfabética.
T4- Leitura: compreensão, análise e avaliação.	C7. Localização de informações explícitas em textos.	H13- Identificar suportes e gêneros textuais diversos.
		H14- Reconhecer a finalidade de gêneros diversos.
	C8. Interpretação de informações implícitas em textos.	H15- Localizar informações explícitas em textos de maior extensão ou em textos que apresentam dados.
		H16- Identificar elementos que constroem a narrativa.
		H17- Inferir informações implícitas em textos.
		H18- Identificar assunto de frases e textos.
	C9. Coerência e coesão no processamento de textos	H19- Formular hipóteses.
		H20- Identificar efeitos de humor ou ironia em textos diversos.
		H21- Estabelecer relações lógico-discursivas presentes no texto.
		H22- Estabelecer relações de continuidade temática, a partir da recuperação de elementos da cadeia referencial do texto.
C10. Avaliação do leitor em relação aos textos lidos.	H23- Identificar o efeito de sentido decorrente do uso de recursos gráficos, da pontuação, da seleção lexical e repetições.	
	H24- Identificar marcas linguísticas que evidenciam o enunciador no discurso direto ou indireto.	
	H25- Distinguir fato de opinião sobre o fato.	
T5 – Produção escrita.	C11. Escrita de palavras e frases.	H26- Identificar tese e argumentos.
		H27- Avaliar a adequação da linguagem usada à situação, sobretudo a eficiência de um texto ao seu objetivo ou finalidade.
	C12. Produção de textos.	H28- Escrever palavras.
		H29- Escrever frases.
		H30- Produzir textos.

Fonte: Proalfa 2012. Tabulação do autor.

**Matriz de Referência do Proalfa – Língua Portuguesa – 2013 e 2014**

<b>Competências</b>	<b>Descritores</b>
C1. Identificação de letras do alfabeto	H1 – Identificar letras do alfabeto
	H2 – Diferenciar letras de outros sinais gráficos, como os números, sinais de pontuação ou de outros sistemas de representação.
	H3 – Distinguir, como leitor, diferentes tipos de letras.
C2. Uso adequado da página	H4 – Conhecer as direções e o alinhamento da escrita da Língua Portuguesa
C3. Aquisição de consciência fonológica	H5 - Identificar, ao ouvir uma palavra, o número de sílabas (consciência silábica)
	H6 – Identificar sons de sílabas (consciência fonológica e consciência fonêmica)
C4. Reconhecimento da palavra como unidade gráfica	H7 – Compreender a função de segmentação de espaços em branco na delimitação de palavras em textos escritos.
C5. Leitura de palavras e pequenos textos	H8 – Ler palavras.
	H9 – Ler pequenos textos.
C6. Localização de informações explícitas em textos	H10 – Localizar informação explícita em textos de maior extensão e de gêneros e temas menos familiares.
	H11 – Identificar elementos que constroem a narrativa.
C7. Interpretação de informações implícitas em texto	H12 – Inferir informações em textos.
	H13 – Identificar assunto de texto.
	H14 – Formular hipóteses.
C8. Coerência e coesão no processamento de texto	H15 – Estabelecer relações lógico-discursivas presentes no texto
	H16 – Estabelecer relações de continuidade temática a partir da recuperação de elementos da cadeia referencial do texto.
	H17 – Identificar efeito de sentido decorrente de recursos gráficos, seleção lexical e repetição.
	H18 – Identificar marcas linguísticas que evidenciam o enunciador no discurso direto ou indireto.
C9. Avaliação do leitor em relação aos textos.	H19 – Distinguir fato de opinião sobre o fato.
	H20 – Identificar tese e argumentos.
	H21 – Avaliar a adequação da linguagem usada à situação, sobretudo, a eficiência de um texto ao seu objetivo ou finalidade.
C10. Implicações do gênero e do suporte na compreensão de textos	H22 – Reconhecer os usos sociais da ordem alfabética.
	H23 – Identificar gêneros textuais diversos.
	H24 – Reconhecer finalidade de gêneros textuais diversos.
C11. Escrita de palavras	Escrever palavras.
C12. Escrita de frases/textos	Escrever frases/textos

Fonte: Proalfa 2013. Tabulação do autor.

## Matriz de Referência do Proalfa – Língua Portuguesa – 3º ano EF – 2015 e 2016

Tópicos	Competências	Descritores	Habilidade
<b>T1 - Apropriação do sistema alfabético.</b>	<b>C1. Aquisição da consciência fonológica.</b>	D01	Identificar sílabas de uma palavra.
	<b>C2. Leitura de palavras e frases.</b>	D02	Ler palavras formadas por sílabas canônicas.
		D03	Ler palavras formadas por sílabas não canônicas.
		D04	Ler frases.
<b>T2 - Leitura: compreensão, análise e avaliação.</b>	<b>C3. Localização de informação explícita em textos.</b>	D05	Localizar informação explícita.
		D06	Reconhecer os elementos que compõem uma narrativa e o conflito gerador.
	<b>C4. Interpretação de informações implícitas em textos.</b>	D07	Inferir informações em textos.
		D08	Inferir o sentido de palavra ou expressão a partir do contexto.
		D09	Reconhecer o assunto de um texto lido.
	<b>C5. Coerência e coesão no processamento de textos.</b>	D10	Reconhecer o sentido das relações lógico-discursivas marcados por advérbios e adjuntos adverbiais, etc.
		D11	Reconhecer as relações entre partes de um texto, identificando os recursos coesivos que contribuem para a sua continuidade.
		D12	Identificar efeito de sentido decorrente de recursos gráficos, seleção lexical e repetição.
		D13	Identificar marcas linguísticas que evidenciam o uso da linguagem e o interlocutor de um texto.
	<b>C6. Avaliação do leitor em relação aos textos.</b>	D14	Distinguir um fato de opinião.
	<b>T3 - Usos sociais da leitura e da escrita.</b>	<b>C7. Implicações do gênero e do suporte na compreensão de textos.</b>	D15
D16			Reconhecer o gênero discursivo.
D17			Identificar a finalidade de textos de diferentes gêneros.

Fonte: Proalfa 2016. Tabulação do autor.

## Matriz de Referência do Proalfa – Língua Portuguesa – 3º ano EF – 2017

Tópicos	Competências	Descritores	Habilidade
T1 - Apropriação do sistema alfabético.	C1. Aquisição da consciência fonológica.	D01	Identificar sílabas de uma palavra.
	C2. Leitura de palavras e frases.	D02	Ler palavras formadas por sílabas canônicas.
		D03	Ler palavras formadas por sílabas não canônicas.
		D04	Ler frases.
T2 - Leitura: compreensão, análise e avaliação.	C3. Localização de informação explícita em textos.	D05	Localizar informação explícita.
		D06	Reconhecer os elementos que compõem uma narrativa e o conflito gerador.
	C4. Interpretação de informações implícitas em textos.	D07	Inferir informações em textos.
		D08	Inferir o sentido de palavra ou expressão a partir do contexto.
		D09	Reconhecer o assunto de um texto lido.
	C5. Coerência e coesão no processamento de textos.	D10	Reconhecer o sentido das relações lógico-discursivas marcados por advérbios e adjuntos adverbiais, etc.
		D11	Reconhecer as relações entre partes de um texto, identificando os recursos coesivos que contribuem para a sua continuidade.
		D12	Identificar efeito de sentido decorrente de recursos gráficos, seleção lexical e repetição.
		D13	Identificar marcas linguísticas que evidenciam o uso da linguagem e o interlocutor de um texto.
	C6. Avaliação do leitor em relação aos textos.	D14	Distinguir um fato de opinião.
	T3 - Usos sociais da leitura e da escrita.	C7. Implicações do gênero e do suporte na compreensão de textos.	D15
D16			Reconhecer o gênero discursivo.
D17			Identificar a finalidade de textos de diferentes gêneros.
D18			Escrever palavras a partir de imagem.
D19			Produzir texto.

Fonte: Proalfa 2017. Tabulação do autor.



**Matriz de Referência do Proalfa – Matemática – 3º ano EF – 2017**

<b>I. ESPAÇO E FORMA</b>	
<b>D01</b>	Identificar a localização ou movimentação de pessoas, objetos ou pontos em representação plana do espaço.
<b>D02</b>	Identificar representações de figuras bidimensionais.
<b>D03</b>	Identificar representações de figuras tridimensionais.
<b>II. GRANDEZAS E MEDIDAS</b>	
<b>D04</b>	Comparar ou ordenar comprimento, altura e espessura.
<b>D05</b>	Utilizar conversão entre unidades de medidas de tempo na resolução de problema.
<b>D06</b>	Reconhecer horas em relógios digitais e/ou analógicos.
<b>D07</b>	Corresponder cédulas e/ou moedas.
<b>III. NÚMEROS E OPERAÇÕES / ÁLGEBRA E FUNÇÕES</b>	
<b>D08</b>	Identificar composições ou decomposições de números naturais.
<b>D09</b>	Corresponder números reais a pontos da reta numérica.
<b>D10</b>	Associar quantidades de objetos à sua representação numérica.
<b>D11</b>	Comparar ou ordenar quantidades pela contagem.
<b>D12</b>	Corresponder números naturais à sua escrita por extenso.
<b>D13</b>	Identificar números naturais segundo critérios de ordem.
<b>D14</b>	Reconhecer números ordinais ou indicadores de posição.
<b>D15</b>	Executar adição ou subtração com números naturais.
<b>D16</b>	Executar multiplicação ou divisão com números naturais.
<b>D17</b>	Utilizar números naturais, envolvendo diferentes significados da adição ou da subtração, na resolução de problemas.
<b>D18</b>	Utilizar números naturais, envolvendo diferentes significados da multiplicação ou da divisão, na resolução de problemas.
<b>IV. TRATAMENTO DA INFORMAÇÃO</b>	
<b>D19</b>	Identificar informações a partir de dados dispostos em tabelas.
<b>D20</b>	Identificar dados apresentados por meio de gráficos.

Fonte: Proalfa 2017. Tabulação do autor.

## ANEXO C - Matrizes de Referência do Spaae Alfa – 2009 a 2017

## Matriz de Referência – Leitura – 2º ano EF – Spaae Alfa – 2009 – 1ª parte

MATRIZ DE REFERÊNCIA PARA AVALIAÇÃO DA ALFABETIZAÇÃO				
Sistema Permanente de Avaliação da Educação Ceará – SPAECE-Alfa				
Eixo 1: Apropriação do sistema de escrita				
Tópicos	Descritor	Detalhamento	1º ano	2º ano
1 - Distinção entre letras e outras formas gráficas.	D 1 - Identificar letras entre rabiscos, desenhos, números e outros símbolos gráficos.	Avalia se o aluno identifica letras dentre várias formas gráficas, tais como rabiscos, desenhos, números e outros símbolos gráficos, apresentadas em diferentes sequências.	x	x
	D 2 - Reconhecer as letras do alfabeto.	Avalia se o aluno reconhece uma determinada letra, ou uma sequência de letras.	x	x
2 - Domínio das convenções gráficas.	D 3 - Identificar as direções da escrita.	Avalia se o aluno identifica a direção correta da escrita (de cima para baixo, da esquerda para a direita), identificando a localização do início e término da escrita em uma página de caderno ou em um texto.	x	x
	D 4 - Identificar o espaçamento entre palavras na segmentação da escrita.	Avalia se o aluno identifica o espaçamento entre palavras na segmentação da escrita, contando as palavras de uma frase ou os espaçamentos entre elas.	x	x
	D 5 - Reconhecer as diferentes formas de grafar uma mesma letra.	Avalia se o aluno reconhece uma mesma letra escrita em maiúscula ou minúscula, na forma cursiva ou de imprensa.		x
3 - Desenvolvimento da consciência fonológica.	D 6 - Identificar rimas.	Avalia se o aluno identifica os sons semelhantes (no final da palavra).	x	x
	D 7 - Contar as sílabas de uma palavra.	Avalia se o aluno conta sílabas (os "pedacinhos") de uma palavra.	x	x
	D 8 - Identificar sílabas (consoante/vogal) no início de palavras.	Avalia se o aluno identifica o som da sílaba inicial, formada pelo padrão consoante/vogal, de uma palavra.	x	x
	D 9 - Identificar sílabas (consoante/vogal) no meio e fim de palavras.	Avalia se o aluno identifica o som da sílaba medial ou final, formada pelo padrão consoante/vogal, de uma palavra.	x	x

## Matriz de Referência – Leitura – 2º ano EF – Spaece Alfa – 2009 – 2ª parte

MATRIZ DE REFERÊNCIA PARA AVALIAÇÃO DA ALFABETIZAÇÃO				
Sistema Permanente de Avaliação da Educação Ceará – SPAECE-Alfa				
Eixo 2: Leitura				
Tópicos	Descritor	Detalhamento	1º ano	2º ano
4 - Decodificação e compreensão de palavras.	D 10 - Decodificar palavras no padrão consoante/vogal.	Avalia se o aluno decodifica palavras formadas por sílabas no padrão canônico: consoante/vogal (ex: sí-la-ba).	x	x
	D 11 - Decodificar palavras nos padrões: vogal, consoante/vogal/consoante, consoante/consoante/vogal.	Avalia se o aluno decodifica palavras formadas por sílabas nos padrões não canônicos: vogal (ex: a-ba-ca-te); consoante/vogal/consoante (ex: tex-to, ve-ri-fi-car); consoante/consoante/vogal (ex: pa-la-vra), dentre outros.		x
	D 12 - Compreender palavras no padrão consoante/vogal.	Avalia se o aluno lê com compreensão palavras formadas por sílabas no padrão canônico: consoante/vogal (ex: sí-la-ba).	x	x
	D 13 - Compreender palavras nos padrões: vogal, consoante/vogal/consoante, consoante/consoante/vogal.	Avalia se o aluno lê com compreensão palavras formadas por sílabas nos padrões não canônicos: vogal (ex: a-ba-ca-te); consoante/vogal/consoante (ex: tex-to, ve-ri-fi-car); consoante/consoante/vogal (ex: pa-la-vra) dentre outros.		x
5 - Decodificação e compreensão de textos.	D 14 - Compreender frases.	Avalia se o aluno lê com compreensão frases com estrutura sintática simples, na ordem direta (voz ativa).	x	x
	D 15 - Localizar informação em textos.	Avalia se o aluno localiza informação explícita em diferentes gêneros textuais.		x
	D 16 - Reconhecer o assunto de um texto.	Avalia se o aluno identifica o assunto principal de um texto.		x
	D 17 - Identificar finalidade de textos de diferentes gêneros.	Avalia se o aluno identifica a finalidade, ou “para quê”, de textos de diferentes gêneros.		x

Fonte: Spaece Alfa 2009.

**Matriz de Referência – Leitura – 2º ano EF – Spaece Alfa – 2010 e 2011**

<b>Eixo 1: Apropriação do sistema de escrita - habilidades relacionadas à identificação e ao reconhecimento de aspectos relacionados à tecnologia da escrita.</b>	
<b>1.1 - Quanto ao reconhecimento de letras.</b>	
<b>D1</b>	Identificar letras entre desenhos, números e outros símbolos gráficos.
<b>D2</b>	Reconhecer as letras do alfabeto.
<b>1.2 - Quanto ao domínio das convenções gráficas.</b>	
<b>D3</b>	Identificar as direções da escrita.
<b>D4</b>	Identificar o espaçamento entre palavras na segmentação da escrita.
<b>D5</b>	Reconhecer as diferentes formas de grafar uma mesma letra ou palavra.
<b>1.3 - Quanto ao desenvolvimento da consciência fonológica.</b>	
<b>D6</b>	Identificar rimas.
<b>D7</b>	Identificar o número de sílabas de uma palavra.
<b>D8</b>	Identificar sílabas canônicas (consoante / vogal) em uma palavra.
<b>D9</b>	Identificar sílabas não canônicas (vogal, consoante / vogal / consoante, consoante / consoante / vogal etc.) em uma palavra.
<b>Eixo 2: Leitura - habilidades relacionadas à leitura de palavras, de frases e de textos.</b>	
<b>2.1 - Quanto à leitura de palavras.</b>	
<b>D10</b>	Ler palavras no padrão canônico (consoante / vogal).
<b>D11</b>	Ler palavras nos padrões não canônicos (vogal, consoante / vogal / consoante, consoante / consoante / vogal etc.).
<b>2.2 - Quanto à leitura de frases.</b>	
<b>D12</b>	Ler frases.
<b>2.3 - Quanto à leitura de textos.</b>	
<b>D13</b>	Localizar informação explícita em textos.
<b>D14</b>	Inferir informação em texto verbal.
<b>D16</b>	Interpretar textos não verbais e textos que articulam elementos verbais e não verbais.
<b>D17</b>	Reconhecer o tema ou assunto de um texto ouvido.
<b>D18</b>	Reconhecer o tema ou assunto de um texto lido.
<b>D22</b>	Identificar o propósito comunicativo em diferentes gêneros.

Fonte: Spaece Alfa 2010. Tabulação do autor.

**Matriz de Referência – Leitura – 2º ano EF – Spaece Alfa – 2012 a 2017**

<b>Eixo 1: Apropriação do sistema de escrita - habilidades relacionadas à identificação e ao reconhecimento de aspectos relacionados à tecnologia da escrita.</b>	
<b>1.1 - Quanto ao reconhecimento de letras.</b>	
<b>D1</b>	Identificar letras entre desenhos, números e outros símbolos gráficos.
<b>D2</b>	Reconhecer as letras do alfabeto.
<b>1.2 - Quanto ao domínio das convenções gráficas.</b>	
<b>D3</b>	Identificar as direções da escrita.
<b>D4</b>	Identificar o espaçamento entre palavras na segmentação da escrita.
<b>D5</b>	Reconhecer as diferentes formas de grafar uma mesma letra ou palavra.
<b>1.3 - Quanto ao desenvolvimento da consciência fonológica.</b>	
<b>D6</b>	Identificar rimas.
<b>D7</b>	Identificar o número de sílabas de uma palavra.
<b>D8</b>	Identificar sílabas canônicas (consoante / vogal) em uma palavra.
<b>D9</b>	Identificar sílabas não canônicas (vogal, consoante / vogal / consoante, consoante / consoante / vogal etc.) em uma palavra.
<b>Eixo 2: Leitura - habilidades relacionadas à leitura de palavras, de frases e de textos.</b>	
<b>2.1 - Quanto à leitura de palavras.</b>	
<b>D10</b>	Ler palavras no padrão canônico (consoante / vogal).
<b>D11</b>	Ler palavras nos padrões não canônicos (vogal, consoante / vogal / consoante, consoante / consoante / vogal etc.).
<b>2.2 - Quanto à leitura de frases.</b>	
<b>D12</b>	Ler frases.
<b>2.3 - Quanto à leitura de textos.</b>	
<b>D13</b>	Localizar informação explícita em textos.
<b>D14</b>	Inferir informação em texto verbal.
<b>D16</b>	Interpretar textos não verbais e textos que articulam elementos verbais e não verbais.
<b>D17</b>	Reconhecer o tema ou assunto de um texto ouvido.
<b>D18</b>	Reconhecer o tema ou assunto de um texto lido.
<b>D21</b>	Reconhecer o gênero discursivo.
<b>D22</b>	Identificar o propósito comunicativo em diferentes gêneros.

Fonte: Spaece Alfa 2017. Tabulação do autor.

## ANEXO D – Roteiros das entrevistas semi-estruturadas

### Roteiro de entrevista individual com os atores que fizeram parte das secretarias estaduais de educação

UNIVERSIDADE FEDERAL DE JUIZ DE FORA – UFJF PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO - PPGE
Roteiro de entrevista semi-estruturada (individual)

Prezado(a) respondente,

A seguir, é apresentado um conjunto de perguntas sobre o PROGRAMA. Elas fazem parte do instrumentário de pesquisa de doutorado conduzida por mim, Wagner Silveira Rezende, sob a orientação da professora Hilda Micarello, no âmbito do Programa de Pós-Graduação em Educação da Universidade Federal de Juiz de Fora. A pesquisa busca descrever o programa de avaliação da alfabetização do estado entre os anos de 2009 e 2017. Sua visão é fundamental para o desenvolvimento da pesquisa, motivo pelo qual agradeço, desde já, sua participação.

<b>Tendo em vista o (PROGRAMA) entre os anos de 2009 e 2017:</b>
--

1. Como é o programa em sua visão? Por gentileza, faça uma breve descrição de como você entende o programa no âmbito das ações da secretaria de educação e da rede de ensino.
2. O programa contribuiu para a melhoria da qualidade da alfabetização no estado no referido período? Se sim, a que você atribui essa melhoria?
3. A que você atribui a longevidade do programa de avaliação?

**Roteiro de entrevista individual com a especialista do CAEd/UFJF responsável pela equipe de aplicação de testes**

UNIVERSIDADE FEDERAL DE JUIZ DE FORA – UFJF PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO - PPGE
Roteiro de entrevista semi-estruturada (individual)

Prezada respondente,

A seguir, é apresentado um conjunto de perguntas sobre o Paebes Alfa, o Proalfa e o Spaece Alfa. Elas fazem parte do instrumentário de pesquisa de doutorado conduzida por mim, Wagner Silveira Rezende, sob a orientação da professora Hilda Micarello, no âmbito do Programa de Pós-Graduação em Educação da Universidade Federal de Juiz de Fora. A pesquisa busca descrever o programa de avaliação da alfabetização do estado entre os anos de 2009 e 2017. Sua visão é fundamental para o desenvolvimento da pesquisa, motivo pelo qual agradeço, desde já, sua participação.

**I. Sobre o tipo de aplicação**

- a. Quais foram as principais mudanças na aplicação dos testes de alfabetização desses três sistemas ao longo do tempo (incluindo a inclusão da avaliação da escrita)?
- b. A mudança da série avaliada leva a mudanças no tipo de aplicação?
- c. A aplicação tem sido de mesmo tipo nos três sistemas? Se não, quais são as principais diferenças entre eles?
- d. Houve alguma mudança em relação ao perfil do aplicador para esses programas ao longo do referido período?
- e. No caso de mudanças terem acontecido, você acredita que elas melhoraram a qualidade da aplicação do teste?

**II. Sobre a capacitação dos aplicadores**

- a. Como é o processo de capacitação dos aplicadores para os testes de alfabetização nos três sistemas? Houve mudança nesse processo ao longo do tempo? Há diferenças na capacitação entre os sistemas pesquisados?
- b. Em caso de mudanças terem acontecido, você acredita que elas contribuíram para a melhoria do processo de aplicação? Se sim, de que maneira?

c. Quem, em regra, são os aplicadores desses testes? Professores da rede? Profissionais da secretaria? Aplicadores externos? Houve mudança nesse perfil ao longo do recorte temporal da pesquisa (2009-2017)? Há diferenças entre os sistemas no que tange ao perfil do aplicador?

d. No caso de aplicadores externos, como eles são recrutados?

e. Os aplicadores são remunerados? Se sim, pela própria secretaria ou pelo Caed?

### **III. Sobre a participação das secretarias**

a. Como as secretarias de educação participam do processo de seleção dos aplicadores?

b. Há participação das secretarias desses três sistemas no processo de capacitação dos aplicadores? Se sim, o que é definido pelas secretarias e o que é de responsabilidade de definição do CAEd?

c. A decisão sobre datas de aplicação é de responsabilidade das secretarias? Quais os principais critérios que interferem na escolha dessa data ao longo do tempo e nos três sistemas?

d. O formato e o tipo da aplicação são definidos pelas secretarias ou pelo CAEd (por exemplo, se a aplicação será em um único dia, se terá apenas uma folha de teste, como será o tipo de cartão resposta, se o teste de escrita virá em separado ou acoplado ao teste de leitura etc)?

### **IV. Sobre os resultados**

a. Como as mudanças na aplicação (de seu formato ao tipo de aplicador) podem interferir nos resultados dos testes?

b. Você acredita que, nesses três sistemas e ao longo do tempo pesquisado, as mudanças identificadas na aplicação exerceram efeitos nos resultados?

**V.** Alguma informação importante que você gostaria de acrescentar?



**Roteiro de entrevista individual com a especialista do CAEd/UFJF responsável pela equipe de divulgação dos resultados**

UNIVERSIDADE FEDERAL DE JUIZ DE FORA – UFJF PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO - PPGE
Roteiro de entrevista semi-estruturada (individual)

Prezada respondente,

A seguir, é apresentado um conjunto de perguntas sobre o Paebes Alfa, o Proalfa e o Spaece Alfa. Elas fazem parte do instrumentário de pesquisa de doutorado conduzida por mim, Wagner Silveira Rezende, sob a orientação da professora Hilda Micarello, no âmbito do Programa de Pós-Graduação em Educação da Universidade Federal de Juiz de Fora. A pesquisa busca descrever o programa de avaliação da alfabetização do estado entre os anos de 2009 e 2017. Sua visão é fundamental para o desenvolvimento da pesquisa, motivo pelo qual agradeço, desde já, sua participação.

1. Quais são os canais utilizados para realizar a divulgação de resultados nos três programas de avaliação?
2. Quais são os destinatários dos materiais de divulgação de resultados?
3. Como é processo de produção desses materiais? Há participação das equipes das secretarias responsáveis pela avaliação?
4. Houve alguma mudança significativa, ao longo do período proposto para análise, nesses materiais?
5. Há algum material desenvolvido especificamente para os alunos?

**Roteiro de entrevista coletiva com as especialistas do CAEd/UFJF responsáveis pela elaboração de itens para a alfabetização e construção de testes** (foram realizadas duas entrevistas coletivas, tendo como base este mesmo roteiro: uma com foco em leitura e outra com foco em matemática)

UNIVERSIDADE FEDERAL DE JUIZ DE FORA – UFJF PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO - PPGE
Roteiro de entrevista semi-estruturada (coletiva)

Prezadas respondentes,

A seguir, é apresentado um conjunto de perguntas sobre o Paebes Alfa, o Proalfa e o Spaece Alfa. Elas fazem parte do instrumentário de pesquisa de doutorado conduzida por mim, Wagner Silveira Rezende, sob a orientação da professora Hilda Micarello, no âmbito do Programa de Pós-Graduação em Educação da Universidade Federal de Juiz de Fora. A pesquisa busca descrever o programa de avaliação da alfabetização do estado entre os anos de 2009 e 2017. Sua visão é fundamental para o desenvolvimento da pesquisa, motivo pelo qual agradeço, desde já, sua participação.

**Sobre Matrizes de Referência, itens e testes**

**i)** as matrizes de referência foram alteradas, no âmbito de cada programa, ao longo do período? Se sim, que mudanças foram essas? Como elas influenciaram o sistema de avaliação (houve alteração no tipo de teste por causa disso ou na forma de aplicação?); o que a comparação das matrizes dos três programas nos diz? Elas são semelhantes, iguais, possuem diferenças significativas? Se sim, quais são elas?

**ii)** os testes foram alterados ao longo do tempo? Se sim, que tipo de alteração foi feita e com base em qual justificativa? Os cadernos foram alterados? O número de itens mudou?

**iii)** os itens e sua forma de elaboração foram alterados? Se sim, por que motivo? O que a análise de itens tem a nos dizer sobre isso? Os itens estiveram plenamente ajustados à população nos três programas?

**iv)** qual a concepção de alfabetização por trás de cada uma das matrizes de referência? Tal concepção foi alterada ao longo do tempo? A concepção é a mesma para os três programas?

**v)** há relação direta entre o padrão de desempenho adequado (ou nome equivalente) e a condição de alfabetizado?

**Roteiro de entrevista coletiva com os especialistas do CAEd/UFJF responsáveis pela equipe de produção escrita**

UNIVERSIDADE FEDERAL DE JUIZ DE FORA – UFJF PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO - PPGE
Roteiro de entrevista semi-estruturada (coletiva)

Prezadas respondentes,

A seguir, é apresentado um conjunto de perguntas sobre o Paebes Alfa, o Proalfa e o Spaece Alfa. Elas fazem parte do instrumentário de pesquisa de doutorado conduzida por mim, Wagner Silveira Rezende, sob a orientação da professora Hilda Micarello, no âmbito do Programa de Pós-Graduação em Educação da Universidade Federal de Juiz de Fora. A pesquisa busca descrever o programa de avaliação da alfabetização do estado entre os anos de 2009 e 2017. Sua visão é fundamental para o desenvolvimento da pesquisa, motivo pelo qual agradeço, desde já, sua participação.

**Sobre a avaliação da escrita na alfabetização**

- i) Ela ocorreu para os três programas analisados? Desde que edição?
- ii) Qual é a estrutura dos testes de escrita? São compostos, em geral, por quantos itens?
- iii) Como é o processo de elaboração dos itens?
- iv) O que é avaliado em um teste de escrita?
- v) As matrizes de escrita mudaram ao longo do tempo?
- vi) Como é o processo de correção da escrita?

**ANEXO E – Termo de cooperação com os municípios – Paebes Alfa e Paebes**

**TERMO DE COOPERAÇÃO TÉCNICA Nº.....**

**PROCESSO SEDU/Nº.....**

**TERMO DE ADESÃO E COOPERAÇÃO TÉCNICA Nº... , QUE  
ENTRE SI CELEBRAM A SECRETARIA DE ESTADO DA  
EDUCAÇÃO – SEDU E O.....,  
NA FORMA ABAIXO.**

O Estado do Espírito Santo, pessoa jurídica de direito público interno, inscrito no CNPJ sob o Nº 27.080.530/0001-43, por intermédio da SECRETARIA DE ESTADO DA EDUCAÇÃO - SEDU, com sede na Avenida César Hilal, Nº 1.111, Santa Lúcia, Vitória, Capital do Estado do Espírito Santo, inscrita no CNPJ sob o Nº 27.080.530/0001-93, neste ato representado pelo SR....., portador da CI Nº ..... e CPF Nº ....., residente neste Estado e o Município.....inscrito no CNPJ/MP sob o nº ....., com sede na ..... CEP: ..... – Município de ..... – Espírito Santo, neste ato representado pelo Secretário Municipal de Educação, SR....., brasileiro, portador da CI N.º ....., expedida pelo órgão....., CPF N.º....., em conformidade com os autos do processo Nº..... resolvem celebrar o presente Termo de Cooperação Técnica, mediante as cláusulas e condições seguintes:

**CLÁUSULA PRIMEIRA - DO OBJETO**

O presente termo tem por objeto a adesão da rede municipal ao Programa de Avaliação da Educação Básica – PAEBES desta Secretaria de Educação, com vistas à realização da avaliação dos alunos do ensino fundamental, cujos resultados fornecerão indicadores de eficiência ao sistema educacional do Estado do Espírito Santo.

**CLÁUSULA SEGUNDA - DAS ATRIBUIÇÕES DOS PARTICIPES**

I- Caberá à Secretaria de Estado da Educação:

- a) Assegurar a participação de todas as escolas do município nas avaliações do Programa;
- b) Oferecer ao Município informações técnicas sobre o Programa de Avaliação da Educação Básica do Espírito Santo – PAEBES, quando solicitada;
- c) Capacitar o Coordenador Municipal sobre os procedimentos necessários a aplicação dos instrumentos da avaliação em conformidade os manuais de orientação;
- d) Fornecer os instrumentos para a avaliação compreendendo os testes e questionários e instrumentos de apoio;
- e) Fornecer os resultados da avaliação e assessorar sua análise junto a equipe técnica da Secretaria Municipal de Educação;
- f) Oferecer oportunidades de formação continuada na área de avaliação educacional aos técnicos da secretaria municipal de educação;

I- Caberá ao Município

- a) Organizar reuniões com profissionais da rede municipal de ensino – equipe técnica da secretaria municipal de educação, representantes de pais no conselho de escola e corpo docente da escola, com vistas à sensibilização da avaliação educacional;
- b) Designar um técnico para ser o interlocutor do Município junto a SEDU;
- c) Designar um ou mais Técnicos para responderem pela Coordenação Municipal de todas as atividades referentes a avaliação do Paebes Alfa e do Paebes no município, mediante direcionamento técnico da Sedu;
- d) Alocar profissionais para atuarem como aplicadores nas escolas, em conformidade com os critérios estabelecidos pela Sedu;
- e) Garantir a participação do Coordenador Municipal na capacitação realizada pela SEDU;
- f) Capacitar os aplicadores para atuarem na aplicação da avaliação nas escolas da rede municipal;
- g) Promover as condições adequadas ao envolvimento da sua rede escolar e da sua equipe da Secretaria Municipal de educação no processo de análise e disseminação dos resultados;
- h) Utilizar os resultados gerados pela avaliação na definição e na implementação de políticas públicas, com vistas à melhoria da qualidade do ensino.

### CLÁUSULA TERCEIRA - DOS RECURSOS FINANCEIROS

Fica estabelecido que, não haverá repasse de recursos financeiros por nenhuma das instituições parceiras para o desenvolvimento do Programa.

### CLÁUSULA QUARTA - DO ACOMPANHAMENTO

A SEDU, por meio da equipe da Gerência de Informação e Avaliação Educacional - GEIA, exercerá função gerencial de supervisão durante o período regulamentar da execução deste instrumento, ficando assegurada a seus agentes a reorientação das ações, quando necessário.

Parágrafo único: O Município franqueará livre acesso aos servidores da supervisão desta SEDU, a qualquer tempo e lugar, a todos os atos e fatos praticados, relacionados direta ou indiretamente a este Termo de Adesão e Cooperação Técnica, quando em missão de supervisão.

### CLÁUSULA QUINTA - DA EXECUÇÃO

I - As avaliações referentes ao ensino fundamental ocorrerão da seguinte forma:

- a) Primeiro, segundo, terceiro e quinto anos do ensino fundamental com a finalidade de diagnosticar estágios dos alunos em relação à alfabetização, com provas de leitura, escrita e matemática.
- b) Nonos anos do ensino fundamental e terceira série do ensino médio com a finalidade de aferir o desempenho dos alunos em Língua Portuguesa (leitura) e Matemática, ao final do ano letivo. E em anos alternados Ciências Humanas ou Ciências da Natureza.

### CLÁUSULA SEXTA – DA VIGÊNCIA

O presente Termo terá vigência a partir do dia subsequente ao da data de publicação no Diário Oficial do Estado, com duração até 01/06/2025, podendo ser prorrogado por interesse de ambas as partes.

### CLÁUSULA SÉTIMA – DA RECISÃO

O presente Termo será rescindido, independentemente do instrumento de sua formalização pela inexecução das obrigações estipuladas, por mútuo consenso ou mediante denúncia da parte interessada, com antecedência mínima de 30 (trinta) dias.

