

UNIVERSIDADE DE LISBOA
Instituto de Educação



Avaliações Externas da Aprendizagem: Perspectivas de Especialistas, Decisores
Políticos e Docentes Brasileiros

Helen Vieira de Oliveira

Orientador: Prof. Doutor Domingos Manuel Barros Fernandes

Tese especialmente elaborada para obtenção do grau de Doutor em Educação,
especialidade de Avaliação em Educação

2020

UNIVERSIDADE DE LISBOA

Instituto de Educação



Avaliações Externas da Aprendizagem: Perspectivas de Especialistas, Decisores
Políticos e Docentes Brasileiros

Helen Vieira de Oliveira

Orientador: Prof. Doutor Domingos Manuel Barros Fernandes

Tese especialmente elaborada para obtenção do grau de Doutor em Educação,
especialidade de Avaliação em Educação

Júri:

Presidente: Doutora Maria Natália de Carvalho Alves, Professora Associada e membro
do Conselho Científico do Instituto de Educação da Universidade de Lisboa.

Vogais:

- Doutor Luis António Gomes Senna, Professor Titular da Faculdade de Educação da
Universidade do Estado do Rio de Janeiro, Brasil;
- Doutor Rui Eduardo Trindade Fernandes, Professor Auxiliar da Faculdade de
Psicologia e de Ciências da Educação da Universidade do Porto;
- Doutor Almerindo Janela Gonçalves Afonso, Professor Associado do Instituto de
Educação da Universidade do Minho;
- Doutor Domingos Manuel Barros Fernandes, Professor Catedrático do Instituto de
Educação da Universidade de Lisboa, orientador;
- Doutora Maria Leonor de Almeida Domingues dos Santos, Professora Associada com
Agregação Jubilada do Instituto de Educação da Universidade de Lisboa;
- Doutor Pedro Miguel Freire da Silva Rodrigues, Professor Auxiliar do Instituto de
Educação da Universidade de Lisboa.

Agradecimento

Primeiramente, agradeço a Deus, por guiar e dirigir minha vida de forma a conduzir-me pelo melhor caminho. Que o Seu caminho seja sempre o meu norte.

À minha família que esteve comigo e me apoiou em todos os momentos. Não tenho palavras para agradecer todo amor, parceria e paciência a mim dedicados. Em especial aos meus pais.

Ao meu professor e orientador Prof. Doutor Domingos Fernandes, por me ensinar a contar uma história e fazer dessa jornada uma caminhada com grandes construções. Obrigada pela forma significativa que conduziu esta etapa acadêmica.

Aos professores do Instituto de Educação da Universidade de Lisboa que participaram da minha formação, pelas discussões e ensinamentos apresentados ao longo do curso.

Ao meu orientador do mestrado Prof. Doutor Luiz Antônio Gomes Senna, por continuar no meu caminho e me apoiar nessa jornada.

À Prof.^a Doutora Zena Eisenberg, pela sua acolhida e incentivo no período final da tese.

À minha amiga Suammy, que iniciou o doutorado junto comigo e me apoio em todos os momentos. Obrigada pela parceria, pelos momentos de troca e pelo apoio constante.

A todas as pessoas que contribuíram para realização deste trabalho.

Minha Sincera Gratidão!

Resumo

A avaliação externa da aprendizagem dos alunos vem, ao longo de várias décadas, se desenvolvendo, ocupando assim, um lugar de destaque e sendo alvo de críticas por uns e de aprovação por outros. Um campo ainda cheio de controvérsias. No Brasil, essa avaliação está presente há, aproximadamente 30 anos, vindo a se expandir, significativamente, depois da implementação do SAEB – Sistema de Avaliação da Educação Básica. Nesse início, desenvolveu-se como uma ferramenta política, para dirigir o sistema e verificar a qualidade da educação. Com o passar dos anos, uma cultura avaliativa foi implementada, nos estados e municípios, principalmente, depois da implementação da Prova Brasil, começando a se desenvolver a função pedagógica. Entretanto, o desenvolvimento dessas funções, os múltiplos propósitos que a avaliação externa vem desempenhando, sua baixa credibilidade e utilização, seus efeitos controversos, evidenciam a necessidade de mais estudos nessa área. Essa pesquisa buscou discutir a avaliação externa da aprendizagem dos alunos, com base na apreciação de diferentes intervenientes no processo educativo. Assim, em um primeiro momento, realizou-se uma revisão teórica, com base na literatura nacional e internacional da área de avaliação da aprendizagem e avaliação externa. Em seguida, desenvolveu-se uma análise descritiva de documentos e relatórios técnicos das avaliações, desenvolvidas no Rio de Janeiro, desde o âmbito internacional, passando pelas avaliações nacionais, até chegar nas locais, que compreendem as avaliações desenvolvidas pelo estado e município. A pesquisa empírica foi desenvolvida em duas partes. Na parte qualitativa, foram realizadas 22 entrevistas ao todo, com especialistas, decisores políticos e professores da educação básica. A parte quantitativa foi empregada de forma a completar a pesquisa qualitativa. Dessa forma, foram aplicados 124 questionários a professores da rede municipal do Rio de Janeiro. Os resultados demonstraram que há uma maior referência da função política da avaliação externa, principalmente, no que diz respeito à característica somativa. Há pouco desenvolvimento da avaliação externa formativa, mesmo que seja considerada por todos os intervenientes como um instrumento relevante para o processo de ensino e aprendizagem. Em conclusão, esta pesquisa identificou alguns fatores que parecem estar a dificultar o desenvolvimento da avaliação externa, principalmente, no que tange à sua função pedagógica formativa.

Palavras-chave: Avaliação externa da aprendizagem dos alunos; Políticas educacionais; Qualidade da Educação

Abstract

The external evaluation of student learning has been developing for quite a few decades, thus occupying a prominent place and being criticized by some and approved by others. A field still full of controversy. In Brazil, this assessment has been present for approximately 30 years, and has expanded significantly after the implementation of the SAEB - Basic Education Assessment System. In the beginning, it developed as a political tool to run the system and to verify the quality of education. Over the years, an evaluative culture was implemented in the states and municipalities, especially after the implementation of Prova Brasil, beginning to develop the pedagogical function. However, the development of these functions, the multiple purposes that external evaluation has been performing, their low credibility and use, their controversial effects, highlight the need for further studies in this area. This research aimed to discuss the external assessment of students' learning, based on the appreciation of different actors in the educational process. Thus, at first, a theoretical review was conducted, based on the national and international literature in the area of learning assessment and external assessment. Then, a descriptive analysis of documents and technical reports of the evaluations developed in Rio de Janeiro, from the international level, through the national evaluations, until the local ones, comprising the assessments developed by the state and municipality. The empirical research was developed in two parts. In the qualitative part, 22 interviews were conducted, all in all, with specialists, policy makers and teachers of basic education. The quantitative part was used in order to complete the qualitative research. Thus, 124 questionnaires were applied to teachers of the city network of Rio de Janeiro. The results showed that there is a greater referencing of the political function of external evaluation, especially regarding the summative characteristic. There is little development of formative external assessment, even if it is considered by all stakeholders as a relevant tool for the teaching and learning process. In conclusion, this research has identified some factors that seem to be hindering the development of external evaluation, especially regarding its formative pedagogical function.

Keywords: External assessment of student learning; Educational policies; Quality of Education

ÍNDICE

Capítulo I Introdução.....	1
Capítulo II Avaliação Externa da Aprendizagem dos Alunos	6
Da Avaliação como Domínio do Conhecimento à Avaliação da Aprendizagem	7
Alguns Desenvolvimentos no Domínio da Avaliação	15
Linha temporal – Madaus e Stufflebeam.....	15
Métodos alternativos de avaliação – Stufflebeam.	17
Avaliações promissoras para o século XXI.....	19
Desenvolvimento da Avaliação Externa da Aprendizagem dos Alunos	24
Avaliação externa como diretriz das políticas educacionais.....	27
Avaliação externa e qualidade da educação: um assunto em pauta na agenda política.	32
Utilização da avaliação externa da aprendizagem.....	39
Relevância.	49
Credibilidade.	50
Comunicação dos resultados: a leitura e a apropriação dos resultados.	54
Efeitos da avaliação externa.....	55
Efeito da avaliação externa na sala de aula.	58
Efeito da avaliação externa no sistema educacional e nas políticas públicas.....	62
Capítulo III Metodologia	66
Delineamento da Pesquisa	66
Desenho da Pesquisa.....	71
Descrição dos Procedimentos Adotados	74
Participantes da Pesquisa	76
Coleta e Registro de Dados.....	80
Apresentação e Discussão dos Principais Resultados.....	85

Capítulo IV Do Internacional ao Local: Uma Discussão Acerca das Avaliações Externas Internacionais, Nacionais e do Estado do Rio de Janeiro.....	88
Avaliações Externas Internacionais	89
Programa Internacional de Avaliação de Estudantes (PISA).....	91
O PISA no Brasil.....	97
Avaliação Externa regional – América Latina.....	101
O LLECE no Brasil.....	103
Avaliações Externas Nacionais – Brasil	105
Sistema de Avaliação da Educação Básica (Saeb).....	106
Avaliações Externas Estaduais – Estado do Rio de Janeiro	111
Programa de Avaliação Externa – Saerj.....	113
Programa de Avaliação Diagnóstica do Desempenho Escolar.	116
Avaliações Externas Municipais – Município do Rio de Janeiro	120
Análise das Diferentes Avaliações Externas Presentes no Rio de Janeiro	123
Capítulo V A Avaliação Externa sob a Apreciação dos Diferentes Intervenientes: Descrição, Análise e Interpretação dos Dados	131
Narrativa dos Especialistas	132
Avaliação externa da aprendizagem.....	132
Qualidade da educação.....	157
Narrativa dos Decisores Políticos	172
Avaliação externa da aprendizagem.....	172
Qualidade da educação.....	191
Narrativa dos Professores da Educação Básica	203
Avaliação externa da aprendizagem.....	203
Qualidade da educação.....	221
Comparação entre os Grupos e Interpretação dos Resultados	233
Capítulo VI Conclusão	244
Limitações e Sugestões para Futuras da Pesquisas.....	252

Referências Bibliográficas	254
Anexos.....	269

Índice de tabelas

Tabela 1	Participação dos municípios nas diferentes avaliações externas.....	37
Tabela 2	Frequência observada dos principais objetivos da avaliação de alunos promovida pela rede municipal de ensino	38
Tabela 3	Tempo de atuação dos docentes participantes da pesquisa	82
Tabela 4	Perfil dos decisores políticos.....	85
Tabela 5	Natureza e propósitos	206
Tabela 6	Credibilidade	212
Tabela 7	Utilidade	215
Tabela 8	Utilização	219
Tabela 9	Satisfação dos diferentes intervenientes	221
Tabela 10	Avaliação externa e a qualidade do processo de ensino e aprendizagem	224
Tabela 11	Avaliação externa, prática de ensino e avaliação da aprendizagem	229
Tabela 12	Resultado da avaliação externa	233
Tabela 13	Comparação dos propósitos da avaliação externa entre os diferentes grupos	235
Tabela 14	Comparação das utilizações da avaliação externa entre os diferentes grupos	237

Índice de figuras

Figura 1	Atores educacionais e suas ações políticas	31
Figura 2	Utilização da avaliação externa segundo os usuários e o tipo de utilização.....	48
Figura 3	Fatores que influenciam o processo de aprendizagem	57
Figura 4	Matriz de investigação	70
Figura 5	Desenho da estratégia aninhada concomitante	73
Figura 6	Divisão das CREs no município do Rio de Janeiro	80
Figura 7	Desenvolvimento dos sistemas de avaliação presentes no Rio de Janeiro.....	126
Figura 8	Avaliações externas ao longo da história no Rio de Janeiro	128
Figura 19	Dimensão “Natureza e os propósitos”: especialistas	138
Figura 10	Dimensão “Relevância e importância”: especialistas	141
Figura 11	Dimensão “Credibilidade”: especialistas	145
Figura 12	Dimensão “Utilidade”: especialistas	148
Figura 13	Dimensão “Utilização”: especialistas	154
Figura 14	Dimensão “Satisfação dos diferentes intervenientes”: especialistas	156
Figura 15	Dimensão “Avaliação externa e qualidade do processo de ensino e aprendizagem”: especialistas.....	161
Figura 16	Dimensão “Avaliação externa, prática de ensino e avaliação da aprendizagem”: especialistas	167
Figura 17	Dimensão “Resultado da avaliação externa”: especialistas	171
Figura 18	Dimensão “Natureza e propósitos”: decisores políticos	175
Figura 19	Dimensão “Relevância e importância”: decisores políticos	178
Figura 20	Dimensão “Credibilidade”: decisores políticos	182
Figura 21	Dimensão “Utilidade”: decisores políticos	183
Figura 22	Dimensão “Utilização”: decisores políticos	187

Figura 23	Dimensão “Satisfação dos diferentes intervenientes”: decisores políticos	190
Figura 24	Dimensão “Avaliação externa e qualidade do processo de ensino e aprendizagem”: decisores políticos	192
Figura 25	Dimensão “Avaliação externa, prática de ensino e avaliação da aprendizagem”: decisores políticos	198
Figura 26	Dimensão “Resultado da avaliação externa”: decisores políticos...	202
Figura 27	Dimensão “Natureza e propósitos”: professores	205
Figura 28	Dimensão "Relevância e importância": professores	208
Figura 29	Dimensão "Credibilidade": professores	211
Figura 30	Dimensão "Utilidade": professores	214
Figura 31	Dimensão "Utilização": professores	218
Figura 32	Dimensão "Satisfação dos diferentes intervenientes": professores..	220
Figura 33	Dimensão “Avaliação externa e a qualidade do processo de ensino e aprendizagem”: professores	223
Figura 34	Dimensão: “Avaliação externa, prática de ensino e avaliação da aprendizagem”: professores	228
Figura 35	Dimensão: "Resultado da avaliação externa": professores.....	232
Figura 36	Continuum de desenvolvimento da dupla função da avaliação externa.....	242

Capítulo I

Introdução

A avaliação educacional é um vasto campo de análise, sendo a avaliação externa uma abordagem específica desse campo. Uma avaliação que, diferentemente da interna, é projetada e desenvolvida por pessoas de fora do espaço escolar, que devem apresentar mecanismos de processamento confiáveis e comunicação de resultados eficiente.

A avaliação externa apresenta uma multiplicidade de objetos de análise, como políticas públicas educacionais, sistemas educacionais, programas, escolas, professores, alunos etc. Por isso, faz-se necessário realizar uma delimitação maior do campo. Sendo assim, a temática dessa tese encontra-se mais precisamente no âmbito da avaliação externa da aprendizagem dos alunos.

Essa avaliação vem ocupando um lugar de destaque na maioria dos sistemas escolares, gradualmente, assumindo relevância para a geração de informações sobre a qualidade da educação; um instrumento para embasar tomada de decisões de gestores do sistema de ensino; uma contribuição no que se refere à melhoria da qualidade do ensino; um instrumento que possibilita gerar informações, tanto para avaliar e formular políticas públicas quanto informações sobre a escola, as turmas e os alunos; e uma contribuição para o processo de ensino e aprendizagem.

Porém, ao mesmo tempo em que essa avaliação externa vem ocupando um lugar de destaque, recebe também, muitas críticas, suscitando muitos debates. Sendo assim, uma área que ainda precisa de muita pesquisa e diálogo entre os diferentes atores do processo educacional.

A avaliação externa está presente, no Brasil, há aproximadamente 30 anos e veio, progressivamente aumentando, principalmente, depois da implementação do SAEB – Sistema de Avaliação da Educação Básica, desenvolvida pelo governo federal. Ao longo dos anos que se seguiram, outras avaliações externas foram ocorrendo no território brasileiro, desde avaliações internacionais até as locais, podendo citar o Programa Internacional de Avaliação – PISA, a Prova Brasil, o Programa Nova Escola – PNE, o Programa de Avaliação Externa – SAERJ, Prova Rio, dentre outras.

Esse protagonismo da avaliação externa advém de uma nova administração pública gerencial, que na educação é implantada no sentido de garantir o sucesso escolar,

com base na ênfase de resultados e indicadores de desempenho. Para atender a esse novo modelo administrativo, ocorre uma descentralização dos sistemas educacionais, a implementação de mecanismos de avaliação de resultados, a flexibilização do currículo e outras ações que apresentam como objetivo estabelecer uma cultura organizacional com base nos princípios de gestão estratégica e de controle de qualidade. Para isso, a descentralização passa a ter um papel importante para possibilitar o envolvimento dos atores educacionais na realização dos objetivos propostos (Castro, 2008).

Há, por conseguinte, uma tríade muito importante nesse diálogo, que interliga a avaliação externa, políticas públicas e qualidade do ensino. A partir da análise do contexto dessa avaliação e o diálogo estabelecido nessa triangulação, observa-se que a avaliação vem objetivando apresentar tanto uma função política quanto uma função pedagógica, de caráter somativo e formativo. Há explorações presentes da literatura no contexto brasileiro sobre as duas funções, entre as quais pode-se citar: Bauer, Alavarse e Oliveira (2015), Esteban e Fetzner (2015), Alavarse, Bravo e Machado (2012), Bonamino e Sousa (2012), Brooke e Cunha (2011), Horta Neto (2010), Freitas (2007), Oliveira e Araújo (2005), entre outros.

Na função pedagógica, a avaliação externa apresenta como objetivo proporcionar informações sobre o aluno e contribuir com o aporte de seu desenho e processo. Dessa forma, viria a complementar a avaliação realizada pelo professor (Nevo, 1997), servindo como um recurso para a sua reflexão. O que poderia promover um auxílio na aprendizagem de cada aluno ou de um grupo de estudantes.

Por sua vez, a função política apresenta diversos objetivos, desde o estudo comparativo e a avaliação do desempenho do sistema, como também, a influência das ações em sala de aula. Dessa forma, os resultados da avaliação são usados como indicadores de qualidade da educação, servindo também, como instrumentos de controle, regulação, pressão verticalizada e descentralização.

Porém, mesmo havendo pesquisas que visam discorrer sobre as avaliações externas e seus possíveis efeitos, na gestão escolar e na sala de aula, não há ainda, um entendimento sólido sobre esse fenômeno. Tampouco, uma análise comparativa aprofundada do desenvolvimento das funções apresentadas.

Além disso, há poucas discussões sobre outros pontos, como os mecanismos de processamento da avaliação externa e possíveis efeitos sobre as políticas públicas. Dado o exposto, sua contribuição como um instrumento para a melhoria da qualidade da

educação, ainda não está comprovada e permanece em debate. Dessa forma, esse campo necessita de mais pesquisas empíricas e uma solidificação do quadro conceitual.

Ao evidenciar essa lacuna na literatura, configurou-se como pertinente o estudo da avaliação externa da aprendizagem dos alunos, ainda que esse campo tenha se desenvolvido de forma considerável nas últimas décadas. Fica evidente, a necessidade de estudos que possam responder melhor sobre o diálogo dessa avaliação com as políticas públicas e a qualidade do ensino, a partir das perspectivas dos diferentes intervenientes do processo educacional. Por isso, este trabalho busca compreender a avaliação externa da aprendizagem dos alunos a partir da apreciação das palavras de protagonistas relevantes do campo da educação: professores da educação básica, decisores políticos e especialistas (e.g., professores universitários, pesquisadores).

Entende-se como necessária tal pesquisa, tendo em vista que o Rio de Janeiro adota as avaliações externas nacionais e municipais e, durante muito tempo, também teve as avaliações estaduais. Os professores, em sua maioria, com respeito às avaliações externas estaduais, no Rio de Janeiro, apresentavam movimento contrário, considerando-as como um instrumento de controle de suas práticas e de sua autonomia.

Por sua vez, os decisores políticos consideram-as como um instrumento para dar informações precisas aos docentes e aos alunos, de forma ágil e fiel, vindo por consequência, servir como uma ferramenta pedagógica, para além de ajudar a traçar políticas públicas para a melhoria da qualidade do ensino ofertado. Portanto, diferentes atores do processo educativo sustentam perspectivas distintas e mesmo antagônicas, relativamente, à utilização e aos efeitos das avaliações externas das aprendizagens.

Diante disso, esta pesquisa foi orientada pelas seguintes questões:

1. Como se caracterizam e qual a natureza das avaliações externas da aprendizagem dos alunos?
2. Que relações se podem estabelecer entre as políticas públicas de educação e as avaliações externas da aprendizagem?
3. Que relações se podem estabelecer entre as avaliações externas da aprendizagem e a qualidade do sistema de ensino?
4. Em que medida as avaliações externas são relevantes para melhorar a qualidade da aprendizagem e a qualidade dos processos de ensino?

Para responder a essas perguntas, com foco na tríade levantada, compreendeu-se que era necessário buscar as percepções de diferentes intervenientes do sistema

educacional, para que também houvesse uma visão mais profunda sobre a avaliação externa. Assim, três grupos foram considerados para a entrevista.

Os especialistas configuram o primeiro grupo sendo aqueles que apresentam uma visão teórica importante para a construção do objeto de pesquisa. O segundo grupo é formado pelos decisores políticos, visto que possuem autoridade no que diz respeito às políticas públicas educacionais, sendo aqueles que tomam decisões e produzem materiais de caráter legislativo ou normativo para a educação. Por fim, os professores da educação básica formam o terceiro grupo, uma vez que dão significados e sentidos à avaliação externa dentro do contexto escolar, entendendo que eles não apenas cumprem e executam ordens providas de órgãos superiores, mas também, constroem conhecimento e uma cultura própria que irão pautar suas ações em sala de aula. Com esses participantes, foram realizadas 22 entrevistas ao todo.

Com o intuito de ampliar a percepção dos professores e realizar uma compreensão mais ampla sobre o tema, decidiu-se por realizar também uma recolha de dados quantitativos como forma de complementar os dados qualitativos. As informações apresentadas foram provenientes de dados coletados por meio de questionário respondido por 124 professores que atuam em escolas públicas do município do Rio de Janeiro, nas disciplinas de Português e Matemática.

Para atender às questões de investigação, definiu-se dois objetos de pesquisa. O primeiro é a própria avaliação externa da aprendizagem, com o intuito de caracterizá-la. Para tal, as dimensões desse objeto buscam lançar luz sobre pontos importantes da avaliação externa, como: a) natureza e propósitos; b) relevância e importância; c) credibilidade; d) utilidade; e) utilização; e f) satisfação dos diferentes intervenientes.

O segundo objeto discorre sobre a qualidade da educação, buscando entender como ocorre a relação dele com a avaliação externa. Para tal, são analisadas as: a) relações das avaliações externas com os processos de ensino; b) relações das avaliações externas com os processos de aprendizagem; c) relações das avaliações externas com as práticas de ensino; d) relações das avaliações externas com as práticas de avaliação; e e) relações entre os resultados das avaliações externas e a qualidade da educação.

Para atender ao seu objetivo, a tese foi organizada em seis capítulos, sendo uma introdução, um de caráter teórico conceitual, outro metodológico e dois compreendendo descrição, análise e interpretação dos dados e uma conclusão.

O segundo capítulo versa sobre os conceitos teóricos, que são considerados fundamentais para uma compreensão da temática deste estudo, delineando a natureza, os

fundamentos e os propósitos da avaliação. Para além de realizar uma discussão da avaliação externa em três diferentes níveis: sala de aula (ensino-aprendizagem), gestão local (estabelecimento ou região) e sistema educacional e políticas educacionais. O que possibilitou a compreensão de emprego das diferentes funções da avaliação.

No terceiro capítulo, apresenta-se a metodologia da pesquisa. Neste momento, buscou-se discorrer como sucedeu o “olhar científico”, isto é, qual caminho decidiu-se seguir para estudar os objetos de pesquisa, evidenciando assim, os pressupostos metodológicos. Optou-se por realizar um estudo de natureza interpretativa, com base em Wolcott (1994), sendo dada atenção particular à interpretação dos significados que os participantes atribuem ao fenômeno estudado.

O quarto capítulo busca realizar uma análise descritiva das avaliações externas, no Rio de Janeiro, desde o âmbito internacional até o local, vindo a situar como a educação comparada, as organizações internacionais e os sistemas educacionais nacionais se harmonizam num contexto de avaliação externa. Primeiramente, discutiu-se as avaliações externas internacionais, regionais e nacionais numa perspectiva comparada. Em seguida, apresenta-se o sistema de avaliação do estado e do município do Rio de Janeiro, finalizando com uma comparação entre as avaliações externas locais e nacionais. A realização dessa análise descritiva e comparação entre as avaliações de diferentes âmbitos possibilitou a construção de um quadro das avaliações externas para fundamentar, juntamente com o referencial teórico, a interpretação dos dados.

Por fim, o quinto capítulo, que tem por objetivo realizar a descrição, análise e interpretação dos dados, com base no referencial teórico e na análise descritiva dos sistemas de avaliação. Apresentando assim, os resultados finais da investigação, delimitando a partir da percepção dos participantes dessa pesquisa, a avaliação da aprendizagem dos alunos e sua relação com a qualidade da educação.

Capítulo II

Avaliação Externa da Aprendizagem dos Alunos

Quando se fala em avaliação, deve-se levar em conta a diversidade de assuntos que podem ser abordados e observados amplamente por diferentes ângulos e perspectivas. Pode-se afirmar que, tradicionalmente, o aluno é o centro da avaliação educacional, sendo avaliado pelo sistema escolar durante todo o período letivo. No entanto, há outros campos que cada vez mais vêm surgindo como interesse de investigação, tais como a avaliação do desempenho do docente ou do sistema de ensino, da organização e funcionamento pedagógico do espaço escolar, de um dado programa, do currículo, dentre outros. Tal fato revela a variedade de avaliações no âmbito escolar. Assim como diferentes tipos de avaliação, temos diferentes abordagens, que podem estar embasadas em diferentes paradigmas, os quais apresentam pressupostos políticos, sociológicos, filosóficos e psicológicos distintos. Como afirma Cardinet (1993, p. 11), “Quanto mais penetramos no domínio da avaliação, mais tomamos consciência do carácter enciclopédico da nossa ignorância e cada vez mais pomos em causa as nossas certezas.”.

Devido ao grande número de áreas de investigação, em decorrência de um leque tão grande de interesses e possibilidades no campo de avaliação educacional, faz-se necessária, sendo até mesmo imprescindível, a delimitação e a definição da área de estudo. Este trabalho, como já mencionado, tem como interesse o campo da avaliação da aprendizagem dos alunos no que se refere à avaliação externa. Nesse sentido, este capítulo versa sobre estas duas áreas: avaliação da aprendizagem e avaliação externa da aprendizagem dos alunos, considerando também seu desdobramento nas dimensões políticas e de qualidade da educação.

Num primeiro momento, há uma análise sobre a natureza, os fundamentos e os propósitos da avaliação, como base para o desenvolvimento do tema. Essa parte, juntamente com o segundo momento, que aborda alguns domínios da avaliação, busca apresentar desde a natureza científica da avaliação até a sua construção teórica em diferentes abordagens. Em seguida, discute-se sobre a avaliação da aprendizagem dos alunos, enfocando os conceitos de avaliação formativa e somativa, conceitualizados por Scriven em 1967 no âmbito da avaliação de programas e trazidos para o meio escolar em 1971 através de Benjamim Bloom e seus colaboradores (Bloom, Hastings & Madaus,

1971). Depois, busca-se discorrer sobre o assunto nos tempos atuais, com autores anglo-saxões, francófonos e de língua portuguesa; como se trabalha a questão no contexto escolar e suas ligações com as avaliações interna e externa da aprendizagem.

No segundo ponto deste capítulo, o mesmo caminho será feito, porém, com a avaliação externa da aprendizagem. Realiza-se um delineamento histórico com base em Kellaghan e Madaus (2003), que apresentam sete períodos iniciados nos finais do século XIX até aos tempos mais atuais. O trabalho de Guba e Lincoln (1989), que discutiram quatro concepções de avaliação – como medida, como descrição, como juízo de valor e como negociação – contribuiu igualmente para enquadrar conceitualmente esta investigação.

A terceira parte versa sobre o desenvolvimento da avaliação externa da aprendizagem do aluno, evidenciando como a globalização e a educação comparada contribuíram para uma cultura de avaliação externa. Uma discussão sobre as diferentes funções, formatos e características também é realizada nessa parte.

No quarto e no quinto momentos, a avaliação externa é vista a partir do prisma das políticas públicas educacionais e qualidade da educação, com o objetivo de verificar o caminho que levou a avaliação externa a figurar como assunto do dia nas agendas políticas e também como o principal indicador para aferir a qualidade da educação.

Em seguida, realiza-se uma discussão sobre a utilização da avaliação externa em três diferentes níveis: sala de aula (ensino-aprendizagem), gestão local (estabelecimento ou região) e sistema educacional e políticas educacionais. Para além da distinção das diferentes utilizações possíveis da avaliação externa, ou seja, utilização do desenho e processo da avaliação; e do resultado da avaliação.

Em diálogo com a discussão sobre a utilização da avaliação externa, a última parte desse capítulo ocupa-se de abordar os efeitos dessa avaliação discutidos na literatura e observados em pesquisas brasileiras empíricas. Foi realizado um esforço no sentido de expor efeitos tanto positivos quanto negativos ou nulos.

Da Avaliação como Domínio do Conhecimento à Avaliação da Aprendizagem

É certo que a avaliação faz parte não só da esfera escolar, mas também das mais diversas áreas da vida humana. Como afirma Fernandes (2013, p. 12), a avaliação “é um domínio fundamental do conhecimento porque nos permite formular juízos acerca de todas as áreas de funcionamento da sociedade e tomar decisões fundamentadas, tendo em vista a sua melhoria”. No campo educacional, assim como nas demais áreas da vida

humana, a avaliação apresenta diversos objetos e funções, o que leva a diferentes formas de planejamento, organização, recolha de informações e divulgação. Para tal, é importante delinear a natureza, os fundamentos e os propósitos da avaliação, para que, dessa forma, venha aclarar o desenho do campo de análise deste estudo.

Como ponto de partida, julga-se importante abordar a questão da natureza científica da avaliação, na qual há uma discussão em torno de ser ou não uma ciência, devido a suas dimensões valorativas, éticas e políticas. Em consequência disso, algumas dualidades, no que diz respeito às questões da natureza da avaliação, são discutidas, com base em Fernandes (2007; 2010a; 2013), tais como: avaliação formal e avaliação informal; parcialidade e imparcialidade; objetividade e subjetividade. Para além disso, analisa-se a sua natureza transdisciplinar e os seus propósitos mais fundamentais.

Faz-se necessário diferenciar as avaliações que são realizadas diariamente por todas as pessoas, tais como a avaliação do tempo atmosférico, de um produto no mercado, do comportamento de outra pessoa; das avaliações mais sistemáticas, em que se necessita ir além das impressões pessoais, tal como a avaliação do desempenho de um aluno, do estado financeiro de uma empresa, da estrutura de um prédio. O primeiro tipo de avaliação é considerado informal ou tácito, um lugar-comum, e é usado como auxílio para tomadas de decisões mais assertivas no cotidiano, no que se refere a decisões pessoais e profissionais. No entanto, essa avaliação, devido à sua natureza parcial, ainda que possa ser útil e relevante num número de situações, pode revelar-se pouco confiável e credível porque está essencialmente confinada às experiências e saberes pessoais do “avaliador” não incluindo outras visões e perspetivas. Nessas condições, a avaliação formal é pensada de forma deliberada no sentido de avaliar qualquer objeto, reduzindo a parcialidade, já que vai para além de um critério pessoal, usando processos avaliativos que podem ser escrutinados coletivamente. Para tal, são definidos critérios, diversificando e mencionando as fontes de dados utilizadas, aclarando, com tais características, as evidências que estão contidas num juízo de valor.

É preciso considerar que tais formas de avaliação são diferentes ainda que devam ser consideradas como complementares. Segundo Fernandes (2007, p. 8), as relações entre a avaliação formal e a avaliação informal podem ser consideradas segundo as seguintes perspetivas:

1. a avaliação formal é vista como um desenvolvimento e uma melhoria no tipo de conhecimento que é gerado pela avaliação informal, porque o torna mais explícito e mais fundamentado em evidências diversas;
2. a avaliação formal e a avaliação informal geram diferentes tipos de conhecimento que, no mínimo, são considerados potencialmente complementares;
3. a avaliação formal e a avaliação informal desafiam-se e questionam-se mutuamente sem obedecer a qualquer hierarquia e geram formas interativas de conhecimento.

A avaliação baseada numa epistemologia objetivista, de acordo com Dias Sobrinho (2004), tem por propósito principal prestar informações objetivas, científicas, claras, incontestáveis, úteis para a orientação do mercado e do governo. Segue parâmetros que foram estabelecidos de forma prévia e objetiva, nos quais quase sempre se utilizam procedimentos de quantificação de produtos, voltados para o controle da qualidade dos serviços prestados e dos produtos oferecidos. No entanto, como aponta Horta Neto (2010), o fato de a avaliação utilizar técnicas objetivas leva apenas a uma suposta imparcialidade, pois além da influência de valores, visões de mundo e percepções dos avaliadores, há também as interferências dos governos.

Em compensação, a epistemologia subjetiva apresenta uma perspectiva diferente, no momento em que percebe a realidade como complexa, dinâmica, aberta e polissêmica, entendendo que a verdade é, em consequência, relativa e dependente das experiências humanas. Com isso, outra questão se abre – a dimensão ética da avaliação –, tendo em vista que, como pondera Dias Sobrinho (2004, p. 722):

O reconhecimento da avaliação como fenômeno plurifacetado e de responsabilidade social significa também admitir a sua dimensão ética, para além de sua complexidade epistemológica. Atribuir valor absoluto de verdade e objetividade aos números e seus efeitos de seleção e classificação é querer esconder e abafar o fato de que o campo social é penetrado de valores, interesses e conflitos. É esse caráter ético e, portanto, político que coloca a avaliação no centro das reformas e dos conflitos, pois o que está em jogo e em disputa é o modelo de sociedade.

Esta questão valorativa da natureza da avaliação é a base para a consideração, por parte de alguns setores do mundo acadêmico, de que a avaliação não é uma disciplina científica, pois a ciência não poderia emitir quaisquer juízos de valor. No entanto, há posição contrária a essa colocação, defendendo que “a produção científica, em qualquer domínio, está sempre associada a qualquer forma de avaliação” (Fernandes, 2013, p. 17). Com isso, a disciplina de Avaliação não seria apenas possível, mas estaria a serviço de outras disciplinas acadêmicas, conferindo-lhe, assim, uma natureza transdisciplinar tal como discutido e defendido por Scriven (2008).

Por fim, depois dessa breve discussão acerca da natureza e dos fundamentos da avaliação, tendo em vista obter uma percepção do âmago do campo em análise, cabe evidenciar alguns dos propósitos mais fundamentais da avaliação em geral que, de acordo com Fernandes (2013, p. 12), são:

apoiar a tomada de decisões; servir a prestação pública de contas; melhorar práticas e procedimentos; compreender problemas de natureza social, contribuindo para a identificação de soluções possíveis; compreender as experiências vividas por quem está envolvido numa dada prática social; e acreditar e reconhecer programas.

Ao considerar a avaliação como uma “prática complexa, para a qual concorrem fatores socioeconômicos, políticos, institucionais e pessoais” (Pacheco, 2012, p. 1), a função do professor nesse contexto é desempenhar diferentes papéis no domínio da avaliação – tanto no que diz respeito à utilização para melhorar a aprendizagem do aluno, quanto também para julgar ou classificar, como afirma Fernandes (2008). Isso corrobora com a natureza polissêmica e complexa da avaliação, que apresenta duas funções nesse contexto: formativa e somativa.

Scriven, no final dos anos 60 do século passado, foi o primeiro a diferenciar as duas principais funções da avaliação: a formativa, cujo objetivo é ajudar os desenvolvimentos de programas e outros objetos, sendo dessa forma realizada durante o programa; e a somativa, que ocorre no final de um período, e que ajuda a determinar o valor de um dado objeto ou se um determinado programa ou projeto deverá ou não continuar (Stufflebeam & Shinkfield, 2007). No entanto, apenas na década seguinte o conceito de avaliação formativa passa a ser utilizado no contexto da aprendizagem dos alunos, através dos escritos de Bloom et al. (1971), visto que começa a se preocupar com

“definir o grau de domínio de uma determinada tarefa de aprendizagem e identificar o estado de uma tarefa não dominada.” (p. 61, tradução nossa)

Atualmente, o que se compreende por avaliação formativa não é o mesmo que nos anos 60 e 70 (Fernandes, 2008). Para o autor, neste caso, apresenta-se uma visão mais restrita, muito “centrada em objetivos comportamentais e nos resultados obtidos pelos alunos, pouco interativa e, por isso, normalmente realizada após um dado período de ensino e de aprendizagem” (pp. 22-23). Já nos tempos atuais, observa-se uma avaliação mais complexa, na medida em que se configura de uma forma “interativa, centrada nos processos cognitivos dos alunos e associada aos processos de *feedback*, de regulação, de autoavaliação e de autorregulação das aprendizagens.” (p. 23).

Muitas são as definições apresentadas para a avaliação formativa nos dias de hoje, sejam elas de língua inglesa, francófona ou portuguesa. A seguir, serão apontadas algumas definições sobre ambas as funções. Antes, porém, cabe destacar que, como afirmam Harlen e Gardner (2010), a avaliação não é em si mesma formativa ou somativa; sua natureza se distingue pelo propósito empregado.

Black e Wiliam (1998; 2001) destacam a importância da interatividade do processo de ensino e aprendizagem, entendendo que os professores devem saber sobre os progressos e dificuldades, para que assim venham a adaptar o seu trabalho para satisfazer as necessidades dos alunos. A avaliação formativa seria quando “a evidência é na realidade usada para adaptar o trabalho de ensino para atender às necessidades dos alunos.” (2001, p. 2, tradução nossa), e sua prática produz ganhos significativos na aprendizagem, destacando as seguintes características: *feedback*; envolvimento dos alunos; resultados da avaliação sendo usados para ajustar o ensino; e autoavaliação. Citando estudos realizados no Reino Unido, os autores concluem que a avaliação formativa é uma prática “seriamente em necessidade de desenvolvimento” (2001, p. 5, tradução nossa).

Cardinet (1986) entende a avaliação formativa como orientadora no que tange ao aluno, em seu percurso no trabalho escolar, vindo a localizar dificuldades para que assim possa identificar processos para superá-las e prosseguir em sua aprendizagem. O autor também traz a questão do erro, em que o considera como “normal e característico de um determinado nível de desenvolvimento na aprendizagem” (Cardinet, 1986, p. 4).

A ideia de avaliação formativa como uma fonte de regulação é trazida por Perrenoud (1999), que entende essa avaliação como uma regulação intencional, dado que visa à determinação do caminho já percorrido e aquele que ainda falta percorrer, com o

objetivo de potencializar os processos de aprendizagem. O autor defende que, no lugar de avaliação, seria conveniente utilizar o termo “observação formativa”, tendo em vista que ao termo “avaliação” encontra-se muito associado à ideia de classificação, seleção, certificação. Já “observar” é “construir uma representação realista das aprendizagens, de suas condições, de suas modalidades, de seus mecanismos, de seus resultados” (Perrenoud, 1999, p. 104), sendo formativa quando vem a colaborar com a orientação e otimização das aprendizagens.

Para Ribeiro (1991, p. 84), a avaliação formativa pretende “determinar a posição do aluno ao longo de uma unidade de ensino, no sentido de identificar dificuldades e lhes dar solução”, sendo permanente ou contínua, uma vez que incide sobre “pequenos segmentos da matéria”. Realizar um levantamento dos objetivos atingidos e quais apresentaram dificuldades orienta a ação do professor, visto que este realizará esforços para “recuperar aprendizagens que não foram realizadas” (Ribeiro, 1991, p. 86). No entanto, a realidade de ensino oferece poucas condições para o desenvolvimento de uma avaliação continuada. A autora declara que também há práticas classificatórias na avaliação formativa, uma vez que primeiro classifica os alunos e depois busca que eles ultrapassem as dificuldades encontradas. Essa realidade reflete as ideias erradas sobre tal avaliação, visto que os professores entendem que, ao aplicar uma avaliação formativa, devem recolher, analisar e classificar.

Fernandes (2008, p. 26) propõe a Avaliação Formativa Alternativa (AFA) em contraponto à avaliação formativa behaviorista, mais restrita e pontual com vistas à verificação do alcance de objetivos comportamentais, ou uma avaliação de “contornos relativamente mal definidos”, que se pratica em muitas salas de aula. Para tanto, a AFA apresenta, como principal função, a regulação e melhoria dos processos de ensino e aprendizagem, ou seja, “deve permitir conhecer bem os saberes, as atitudes, as capacidades e o estágio de desenvolvimento dos alunos, ao mesmo tempo que deve proporcionar-lhes indicações claras acerca do que é necessário fazer a seguir.” (Fernandes, 2008, p. 30).

De acordo com Villas-Boas (2006, p. 3), nas escolas brasileiras há uma presença muito marcante de avaliações classificatórias, seletivas e excludentes, utilizadas com o intuito de atribuir notas e classificar como aprovado ou reprovado. Essa prática encontra-se presente na avaliação somativa. No que concerne à avaliação formativa, a autora afirma que o seu papel é valorizar “o aluno e sua aprendizagem e o torne parceiro de todo o processo, conduz à inclusão, e não à exclusão.”, vindo a promover “a aprendizagem do

aluno e do professor, e o desenvolvimento da escola, sendo, portanto, aliada de todos.”. A avaliação deve ocorrer não só de maneira frequente, mas também de forma interativa, tomando o aluno como referência para além dos critérios de avaliação. A autora considera que a avaliação formativa, em seu verdadeiro sentido, é um desafio a enfrentar.

Relativamente à estrutura da avaliação formativa, podem-se destacar dois elementos: autoavaliação e *feedback*. De acordo com Black e Wiliam (2001), a autoavaliação é um componente essencial da avaliação formativa, considerada como uma forma de melhorá-la. Para esses autores, os alunos são normalmente íntegros e confiáveis na avaliação de si mesmos e dos companheiros, chegando até mesmo a serem muito duros. Contudo, para que tal avaliação atinja o objetivo proposto, o aluno deve conhecer de forma clara as metas que a sua aprendizagem deve atingir. Esse ponto acaba por virar uma problemática, na medida em que muitos alunos não possuem essa clareza, o que vem a prejudicar o processo avaliativo. Com isso, os autores concluem afirmando que “Para a avaliação formativa ser produtiva, os alunos devem praticar regularmente a autoavaliação para que, dessa forma, possam compreender os principais propósitos da sua aprendizagem e, assim, entender o que precisam fazer para conseguir.” (Black & Wiliam, 2001, p. 7, tradução nossa).

Para além da autoavaliação, numa estrutura mais ampla, Perrenoud (1999, p. 97) versa sobre a autorregulação, que consiste em “reforçar as capacidades do sujeito para gerir ele próprio seus projetos, seus progressos, suas estratégias diante das tarefas e dos obstáculos”. Para tal, o autor discorre que é essencial que haja o desenvolvimento de práticas e instrumentos de autoavaliação e de apropriação de critérios avaliativos.

No que corresponde ao *feedback*, Sadler (1989) considera-o um elemento-chave da avaliação formativa, consistindo na distribuição de informação sobre como algo foi ou está sendo feito e abrangendo professores e alunos. Para Noizet e Caverni (1985) trata-se de uma dupla retroação, uma vez que se dirige aos alunos, para apoiá-los nas aprendizagens, e aos professores, como forma de auxiliar o desenvolvimento de seu programa pedagógico. Contudo, Fernandes (2008, p. 83), no que diz respeito à AFA, afirma que tal recurso “deve ser mais que uma simples mensagem”, pois a comunicação com os alunos deve ser efetiva no sentido de eles saberem o que fazer com a informação.

No que se refere à avaliação somativa, e principalmente sua relação com a avaliação formativa, Bloom et al. (1971, p. 61, tradução nossa) a definem como uma avaliação “muito geral, que serve como ponto de apoio para atribuir notas, classificar o aluno e transmitir os resultados em termos quantitativos, feita no final de um determinado

período”. Noizet e Caverni (1985) apresentam duas finalidades para essa avaliação: fornecimento de um balanço do estado do aluno e embasamento para a tomada de decisão. Ribeiro (1991), por sua vez, apresenta algumas vantagens, tais como: a possibilidade de verificar os resultados da aprendizagem; a realização de correções no processo de ensino; e a funcionalidade de classificação.

Quanto à relação entre essas duas modalidades de avaliação, há diferentes posicionamentos. Cardinet (1986) entende que elas se opõem; Noizet e Caverni (1985) abordam a questão da contaminação que ocorre na prática de uma avaliação na outra; Ribeiro (1991) defende que a avaliação somativa complementa a avaliação formativa; Fernandes (2008) aborda a coexistência de tais avaliações no sistema educativo e advoga a sua complementaridade.

Sobre essa relação, Pacheco (2012) desenvolve três pontos importantes para o contexto escolar: políticas tendencialmente formativas; práticas predominantemente somativas; e entre o desejado e o realizado. No que concerne às políticas tendencialmente formativas, o autor aborda a questão da regulação que provém de normativas nacionais, na qual há uma valorização do caráter somativo externo e uma contenção da avaliação somativa interna, na medida em que há uma sobrevalorização da avaliação formativa nesse contexto. No entanto, apesar das normativas nacionais, percebe-se uma predominância de práticas somativas, por fatores como: “caso da prescrição curricular em termos de um currículo nacional, da organização curricular em disciplinas e do peso que os testes mantêm na estrutura escolar e nas imagens sociais produzidas sobre a escola.” (Pacheco, 2012, p. 5), vindo, por conseguinte, a ter um discurso cujo objetivo é a avaliação formativa e a prática voltada mais para a somativa. Por fim, o autor refere-se à avaliação como um pêndulo que oscila entre a expectativa e o real, o ideal e o concreto, o referente e o referido; tendo como uma das tarefas mais exigentes, quiçá quase impossível, realizar um processo avaliativo integrado, no qual a melhoria e a certificação possam “caminhar lado a lado”.

A partir disso, observa-se a multiplicidade do campo da avaliação, um terreno no qual se necessita de discussões fundamentadas, tanto de seu caráter técnico quando de suas bases teóricas, buscando, assim, identificar quais são seus fundamentos e como, a partir deles, se desenvolve um desenho de avaliação e seus usos. Havendo a necessidade de entender que a avaliação não é um campo neutro, visto que é “produção de sentidos, prática social, portanto, intersubjetiva, relacional, aberta, polissêmica e carregada de

valores, que põe em questão os significados dos fenômenos.” (Dias Sobrinho, 2008, p. 193).

No entanto, para além da essência da avaliação, julga-se necessário discorrer sobre o seu desenvolvimento ao longo do tempo, o que será feito na próxima seção. Entende-se que é necessária uma leitura para além de nossa época, tendo em vista que o que hoje está posto não diz respeito apenas ao momento, mas sim a toda uma construção histórica, travada por conflitos, rupturas e concepções.

Alguns Desenvolvimentos no Domínio da Avaliação

Esta seção tem como objetivo apresentar uma revisão das principais abordagens internacionais que desenharam a trajetória da avaliação educacional, apontando, assim, o objetivo e a forma como a avaliação se apresentou ao longo do tempo.

Uma análise do campo da avaliação sem levar em consideração sua construção histórica retrata uma visão limitada do contexto estudado, pois, como afirmam Guba e Lincoln (1989, p. 22, tradução nossa), a “avaliação como a conhecemos não aparece simplesmente num dia; ela é o resultado de um processo de desenvolvimento, de construção e reconstrução que envolve um número de influências que interagem”. Por isso, para que ocorra uma melhor compreensão do desenvolvimento e da construção teórica no domínio da avaliação em educação, julga-se necessária uma análise das diferentes abordagens de avaliação. Porém, não caberá realizar de forma detalhada a análise de cada perspectiva neste espaço, tendo em vista que já existe na literatura (cf. Fernandes, 2008; Guba & Lincoln, 1989; Kellaghan & Madaus, 2000; Stufflebeam & Shinkfield, 2007).

Cabe salientar que, nesta parte, optou-se por utilizar abordagem de avaliação no lugar de modelo de avaliação, em consonância com Stufflebeam (2000) e Fernandes (2010a), visto que aquela designação é de caráter mais amplo, o que possibilita englobar práticas que não tenham construção teórica sólida.

Linha temporal – Madaus e Stufflebeam.

Madaus e Stufflebeam (2000) apresentam uma divisão por períodos, chamadas por idades, que representam desde antes de 1900 até os dias atuais. O primeiro período, chamado de Idade da Reforma, tem como data de partida 1792. O século XIX é conhecido devido às grandes mudanças oriundas da Revolução Industrial, fato que provocou uma grande reforma no campo educacional e nos programas sociais, particularmente no Reino

Unido e nos Estados Unidos, o que os levou a realizar avaliações das instituições e dos programas.

O segundo período (1900-1930), conhecido por Idade da Eficiência e dos Testes, compreendeu uma fase na qual houve uma valorização da aplicação de testes e inquéritos para a verificação da eficiência de programas, escolas e professores.

A Idade Tyleriana, terceiro período, recebeu esse nome pela grande influência de Ralph Tyler, considerado o “pai da avaliação educacional”. Antes de suas pesquisas, a área da avaliação era centrada no aluno e na valorização dos êxitos alcançados por eles. Tyler passa a redirecionar tal visão para um campo mais amplo, no qual incluía outros objetos educacionais, como o currículo, além de entender que as decisões acerca de um programa deveriam ser a resposta da congruência entre os objetivos deste e seus resultados reais (Stufflebeam & Shinkfield, 2007).

No quarto período, conhecido como a Idade da Inocência ou Idade da Ignorância, foi necessário deixar a guerra e a depressão para trás e traçar um novo caminho, construir e expandir capacidades. Essa nova configuração influenciou na avaliação educacional. Trata-se de um período em que houve um grande desenvolvimento nos aspectos técnicos da avaliação, com a publicação de muitos testes padronizados nacionalmente.

De 1958 a 1972, ocorreu a Idade do Desenvolvimento, momento de mudanças que resultaram em uma expansão da avaliação como uma indústria e uma profissão. O lançamento do Sputnik, realizado pela Rússia em 1957, gerou a promulgação, no ano seguinte, da Lei de Educação de Defesa Nacional, o que levou a uma série de novos projetos nacionais de desenvolvimento curricular, especialmente nas áreas de Ciências e Matemática. A abordagem de Tyler foi utilizada, nesse momento, com a finalidade de ajudar a delinear objetivos para o novo currículo e, de forma posterior, verificar se estes foram alcançados. Houve também a criação dos testes padronizados nacionais, para verificação dos objetivos e dos conteúdos do novo currículo. Por fim, ocorreram esforços de muitos especialistas para avaliar o desenvolvimento curricular, tendo em vista o estabelecimento de uma nova estrutura.

Em princípios da década de 60, para os especialistas em educação, ficou claro que resultados das avaliações não eram úteis para os que estavam envolvidos com o desenvolvimento dos currículos e nem auxiliavam àqueles que desejavam saber sobre a eficácia dos programas. Tal aspecto negativo foi abordado primeiramente por Lee Cronbach, que criticou a falta de relevância e utilidade das avaliações propostas,

aconselhando os avaliadores a considerá-las como um processo de recolha e divulgação de informações que poderia vir a servir de guia para o desenvolvimento curricular.

A Idade da Profissionalização constitui o sexto período, que compreende os anos de 1973 a 1983, um momento de afirmação e de institucionalização da avaliação educacional. Autores se destacaram no desenvolvimento de novos modelos de avaliação, como Scriven, Stake e Stufflebeam.

Por fim, a última era, que abarca os períodos de 1983 a 2000, conhecida como Idade da Expansão e da Integração. Apesar do nome positivo escolhido por Madaus e Stufflebeam (2000) para esse período, reconhece-se que, no início, não houve investimento financeiro necessário, o que acarretou redução de custos. No entanto, em seguida, com o crescimento econômico, houve uma expansão no campo da avaliação, principalmente no que corresponde à área da avaliação profissional. Cabe destacar que, nessa época, tornou-se comum a *accountability* e as avaliações com foco em resultados nos Estados Unidos. Também foram desenvolvidas estruturas curriculares e testes com vistas à verificação do progresso no que tange aos padrões estabelecidos.

Para além da linha do tempo proposta por Madaus e Stufflebeam (2000), há outras formas de se estruturar as diferentes abordagens de avaliação, o que será visto a seguir com os métodos alternativos de avaliação de Stufflebeam.

Métodos alternativos de avaliação – Stufflebeam.

Outra proposta é a de Stufflebeam, que apresenta quatro categorias: pseudoavaliações; avaliações orientadas por questões ou métodos; avaliações orientadas pela melhoria/*accountability*; e agenda social/advocacia. Para poder agrupar as abordagens de avaliação nas categorias propostas, o autor apresenta o que entende por avaliação, que é “um estudo sistemático planejado, dirigido e realizado a fim de ajudar um grupo de clientes a julgar e/ou melhorar o mérito de um objeto.” (Stufflebeam & Shinkfield, 2007, p. 67, tradução nossa).

As pseudoavaliações são aquelas que apresentam problemas da divulgação de resultados, tendo em vista que, apesar de a recolha de informações ser realizada corretamente, a divulgação destas não preza pela verdade, sendo feita de forma seletiva, com informações falsas, não revelando as verdadeiras conclusões. Nessa categoria, enquadram-se as seguintes abordagens: avaliações controladas politicamente, que pretendem ter a credibilidade de uma avaliação, mas que possibilite um controle da parte do cliente, na medida em que possa criar uma falsa imagem do valor do objeto avaliado;

e a avaliação para a criação de uma imagem pública, que visa à criação de uma imagem positiva de uma instituição, de uma propaganda, de um sistema e outros, manipulando o cliente para que se possa garantir que a avaliação divulgue uma imagem que lhe interesse.

O segundo grupo é o das avaliações orientadas por questões ou métodos (quase-avaliações), tendo como finalidade verificar se os objetivos de um determinado programa foram atingidos. Com isso, há uma ênfase nos objetivos operacionais, que são especificados, o que leva essa avaliação a ser centrada nos métodos e nas questões, havendo uma menor atenção à avaliação dos objetos propriamente dita. Como exemplo desse grupo, há a avaliação baseada em objetivos, na qual se busca concluir se houve êxito no trabalho avaliado, se os objetivos foram alcançados, sendo Tyler o pioneiro desse tipo de avaliação; e a avaliação baseada no método experimental, que busca determinar ou demonstrar vínculos causais entre certas variáveis.

Por fim, as avaliações orientadas pela melhoria/*accountability*, cujo propósito é evitar avaliações tendenciosas, como apresentadas nas pseudoavaliações, e a recomendação de uma avaliação com embasamento em questões de valor e mérito, em contraste com as quase-avaliações. Tem como objetivo ajudar a melhorar o programa e julgar o mérito e a relevância dele. Destacam-se, como exemplos: a avaliação orientada para a decisão, que busca proporcionar conhecimentos e juízos de valor para embasar decisões, o que incentiva a utilização de uma avaliação mais contínua e sistemática; a avaliação orientada para consumidores, cujo objetivo é auxiliar às pessoas que estão relacionadas diretamente com o serviço a entender o funcionamento de suas atividades; a avaliação política, que tem o objetivo de descrever e emitir valor sobre as políticas aplicadas às instituições ou à sociedade; e a avaliação orientada para os consumidores, que busca auxiliar os clientes na escolha de serviços e bens de consumo.

Posteriormente, Stufflebeam (2000) acabou por dividir a última categoria em duas avaliações: advocacia, no qual o objetivo é avaliar o mérito e o valor de um determinado programa ou objeto; e agenda social, cujo objetivo é a contribuição com a melhoria e a transformação da sociedade, com um grau elevado de participação dos *stakeholders*. Nessa última, o autor destacou quatro avaliações das nove que considerou as mais promissoras para o século XXI, a saber: avaliação deliberativa e democrática; avaliação construtivista; avaliação focada na utilização e nos utilizadores; e avaliação receptiva. Tais abordagens serão analisadas mais adiante.

Como afirma Stufflebeam (2000), as abordagens das avaliações orientadas por uma agenda social possibilitam a utilização da avaliação de programas como forma de

capacitar os excluídos, na medida em que lhes possibilitam o acesso à informação. O avaliador tem um papel importante, visto que tem a intenção de servir os menos favorecidos e/ou marginalizados e corrigir injustiças educacionais e/ou sociais, tendo, por conseguinte, o compromisso de produzir resultados válidos.

Avaliações promissoras para o século XXI.

Na avaliação respondente, apresentada por Robert Stake, a atenção está estritamente voltada à avaliação de programas, não tendo o intuito de realizar um guia de planificação de uma avaliação, mas sim um suporte para aqueles que se lançam em tal tarefa. Stake (2003) considera essencial a capacidade de resposta a questões que são levantadas, principalmente, pelas pessoas envolvidas no programa, sendo sensível às partes interessadas, procurando ter em conta os seus interesses e preocupações. A avaliação ocorre de forma a possibilitar contínua adaptação das metas de avaliação e familiaridade do avaliador com o programa. Ao levar em consideração tal fato, o avaliador passa a ser visto como aquele que proporciona informação avaliativa útil às pessoas que estão envolvidas com o programa.

A abordagem de avaliação respondente configura-se em: a) descrição e juízo de um programa; b) distintas fontes de dados; c) análise da congruência e das contingências; d) identificação das normas pertinentes e as opostas; e) múltiplas utilizações das avaliações (Stufflebeam & Shinkfield, 2007).

No que se refere à estrutura funcional da avaliação, pode-se destacar, como ponto central, a observação, tendo em vista que é através dela que há coleta e registro dos dados. O valor de um programa não se encontra na mensuração dos dados, pois estes nem sempre permitem traduzir as realidades sociais e os significados que as pessoas lhes atribuem. Consequentemente, há uma preocupação em descrever detalhadamente os processos envolvidos no âmbito do que se está a avaliar. Sobre a devolutiva ao cliente, Stake (2003) salienta a necessidade de essa estar em formato e linguagem atraentes e adaptada às suas necessidades.

Outra forma de retratar o desenrolar histórico da avaliação educacional é apresentada por Guba e Lincoln (1989). Os autores expõem quatro gerações de avaliação, que são identificadas ao longo dos últimos anos. As três primeiras são apresentadas com base em outras abordagens; já a última é uma proposta dos autores como forma de ruptura com as anteriores, denominada avaliação construtiva. De forma sucinta, serão caracterizadas as três primeiras gerações.

O primeiro momento compreende a avaliação como medida, tendo em vista que apresentava como objetivo medir as aprendizagens escolares dos alunos. Esse período teve início na França, no momento em que o então ministro da Educação francês Alfred Binet, devido à pressão de professores para impedir o acesso de alunos com deficiências mentais, começou a realizar testes para medir a inteligência e aptidões dos alunos. Em 1912, surge então o coeficiente de inteligência, nada mais que o resultado do quociente da idade mental e da idade cronológica do sujeito.

A segunda geração emerge a partir da necessidade de ultrapassar algumas limitações presentes na primeira geração, tais como o fato de o conhecimento dos alunos ser considerado o único objeto de avaliação. Para superar essa situação, a segunda geração passou, então, a descrever padrões de pontos fortes e fracos com relação aos objetivos relacionados. Com isso, a avaliação não consistia apenas na medição, mas também na descrição de até que ponto os objetivos foram alcançados pelos alunos. Nesse contexto, o principal papel do avaliador é descrever o grau de consecução relativamente a cada um dos objetivos que previamente se definem.

A avaliação como julgamento ou formulação de um juízo acerca do valor e do mérito de um dado objeto é apresentada na terceira geração. Esse período é marcado pelo lançamento do Sputnik, em 1957, pela então União Soviética. Um grande movimento começou a ocorrer no sentido de se investir no campo da avaliação do currículo, dos projetos e das aprendizagens dos alunos. As avaliações passam, então, a dar ênfase à formulação de juízos acerca dos objetos. Isso posto, os avaliadores, para além do papel técnico e de descrição, passam a desempenhar a função de juízes. A partir desse período, começam a surgir ideias, tais como:

1. a avaliação deve induzir e/ou facilitar a tomada de decisões que regulem o ensino e as aprendizagens;
2. a recolha de informação deve ir para além dos resultados que os alunos obtêm nos testes;
3. a avaliação tem de envolver os professores, os pais, os alunos e outros intervenientes;
4. os contextos de ensino e de aprendizagem devem ser considerados no processo de avaliação;
5. a definição de critérios é essencial para que se possa apreciar o mérito e o valor de um dado objeto de avaliação. (Fernandes, 2008, p. 59)

A quarta e última geração representa uma proposta de Guba e Lincoln (1989) com base nas limitações presentes nas três primeiras gerações. Os autores apresentam uma abordagem de cunho filosófico, vindo a rejeitar os princípios do positivismo lógico, na medida em que depõem a existência de qualquer realidade objetiva, independente dos seres humanos, e valorizam a ideia de realidades subjetivas, construída pelas pessoas. Com isso, as avaliações não possuem mais como base as descrições, mas sim a construção proveniente de um processo interativo em que se inclui o avaliador e os *stakeholders* envolvidos no processo. Tais construções devem revelar valores subjacentes, saberes, crenças e atitudes, o que leva a um destaque aos instrumentos humanos da avaliação, dado que os avaliadores e partes interessadas são colocados no centro do processo.

A abordagem construtivista valoriza e promove o pluralismo no contexto avaliativo, estando a construção social no cerne da questão, visto que o entendimento do contexto avaliado está diretamente relacionado com questões sociais, culturais e psicológicas. Para tal, o avaliador deve mostrar sensibilidade no que se refere aos atores envolvidos, particularmente no que concerne às suas construções e aos seus valores. Seu papel passa de um caráter técnico de recolha de informações, como visto em outras abordagens, para um caminho de abertura e diálogo, que se constrói com negociação e mediação – negociação, no que tange à participação e ao envolvimento dos *stakeholders*, e mediação, no que diz respeito aos valores variáveis e conflitos que podem existir. Para além disso, há um traço de emancipação e capacitação dos envolvidos, porque isso é feito através de conscientização das partes interessadas de forma a provocar uma participação mais democrática (Stufflebeam & Shinkfield, 2000).

De acordo com Fernandes (2008, pp. 62-63), grande parte da avaliação de quarta geração tem por embasamento um conjunto de princípios, ideias e concepções, a saber:

1. os professores devem partilhar o poder da avaliação com os alunos e outros intervenientes e devem utilizar uma variedade de estratégias, técnicas e instrumentos de avaliação;
2. a avaliação deve estar integrada no processo de ensino e aprendizagem;
3. a avaliação formativa deve ser a modalidade privilegiada de avaliação, com a função principal de melhorar e de regular as aprendizagens;

4. o *feedback*, nas suas mais variadas formas, frequências e distribuições, é um processo indispensável para que a avaliação se integre plenamente no processo de ensino-aprendizagem;
5. a avaliação deve servir mais para ajudar as pessoas a desenvolverem as suas aprendizagens do que para julgá-las ou classificá-las numa escala;
6. a avaliação é uma construção social em que são tidos em conta os contextos, a negociação, o envolvimento dos participantes, a construção social do conhecimento e os processos cognitivos, sociais e culturais na sala de aula;
7. a avaliação deve utilizar métodos predominantemente qualitativos, não se pondo de parte a utilização de métodos quantitativos.

Stufflebeam e Shinkfield (2000) salientam as vantagens dessa abordagem, tais como: a) a completa divulgação dos processos de avaliação e suas conclusões; b) o envolvimento dos intervenientes; c) o reconhecimento dos participantes como um importante “instrumento” da avaliação, sendo consideradas as suas experiências, perspectivas e valores; e d) a relevância conferida aos dados de natureza qualitativa e à sua triangulação utilizando uma diversidade de processos.

De igual forma, os autores também asseveram sobre a limitação na aplicabilidade da abordagem e algumas desvantagens, como o planeamento e o orçamento, que acabam por ser prejudicados devido à abertura e à natureza exploratória e participativa. Nesse sentido, o avaliador postula que pode não haver, em última instância, conclusões corretas.

House e Howe (2000) conceitualizaram uma avaliação democrática e deliberativa inclusiva de natureza dialógica. Na ideia de deliberativa inclusiva entende-se que todo o interesse relevante deve ser representado na avaliação, com a devida expressão, sem desequilíbrio de poder. Por dialógica, a construção de diálogo entre as partes interessadas e os avaliadores, como também, em alguns momentos, com o público. O objetivo é dar voz a todas as partes interessadas, assegurar que interesses, ideias, opiniões serão representados sem deturpações. Por fim, as conclusões da avaliação devem, por sua vez, emergir da deliberação, raciocínio cuidadoso, reflexão e debate.

Os autores destacam as relações de poder existentes no grupo de interessados em uma avaliação, é certo que não há igual poder entre as partes, sendo, por conseguinte, estabelecida uma estrutura de diálogo não democrático. Tal fato não pode ser ocultado, sendo papel do avaliador estabelecer um contexto no qual seja possível que ocorram

deliberações livres e desobstruídas em todas as fases da avaliação. Com isso, afirmam House e Howe (2000), o avaliador

não é um espectador passivo, um facilitador inocente, ou um rei filósofo que toma decisões para os outros, mas sim um profissional consciencioso que adere a um conjunto de princípios cuidadosamente considerados defensáveis para reforçar a inclusão, diálogo e deliberação (pp. 410-411, tradução nossa).

A proposta de House e Howe (2003) não é um modelo de avaliação, mas sim uma teoria que sugere que os estudos têm de ser imparciais, ainda que se admita que o viés não possa ser totalmente eliminado.

Outro tipo de avaliação é a apresentada por Patton, a avaliação orientada para os utilizadores e para a utilização. Patton (2003) parte da premissa que a avaliação deveria ser julgada pela sua utilidade e seu uso efetivo, quer dizer a “Avaliação centrada na utilização está preocupada com a forma como as pessoas reais no mundo real aplicam os resultados da avaliação e experimentam o processo de avaliação. Portanto, o seu foco está no uso pretendido pelos usuários.” (Patton, 2003, p. 223, tradução nossa).

Ao avaliador, cabe o papel de facilitar o julgamento e tomada de decisão pelos usuários, não defendendo qualquer conteúdo particular de avaliação, modelo, método etc., mas sim uma relação de trabalho entre avaliador e usuários para ajudá-los a determinar o tipo de avaliação de que necessitam, com um uso planejado e facilitado. Tal fato faz com que haja uma grande variedade de processos de avaliação que pode ser usada. Cabe também ao avaliador prezar pela honestidade e integridade em todo o processo avaliativo, respeitando as pessoas envolvidas e afetadas pela avaliação, e ser sensível à diversidade de interesses e valores envolvidos. É de grande importância que se tenha a percepção de que a forma como se conduz a avaliação afeta a sua utilização por parte dos interessados (Patton, 2003).

Face ao exposto, fica evidente que muitas características apresentadas nas diferentes abordagens de avaliação ainda se fazem presentes nos sistemas educacionais. Como exemplo, pode-se citar a ideia de avaliação como medida, como afirma Luckesi (1998, p. 76): “a escola brasileira opera com a verificação e não com a avaliação da aprendizagem”, o que retoma a abordagem da avaliação como medida, apresentada por Guba e Lincoln (1989). Por isso, alguns autores atuais buscam clarificar a ideia de “pedagogia de exame” (Díaz-Barriga, 1999; Luckesi, 2014), com vistas a realizar uma diferenciação entre avaliar e medir (Hadji, 2001; Luckesi, 1998). E a definição de práticas

avaliativas libertadoras (Hoffmann, 2006), com vistas a desenraizar conceitos presentes nas décadas de 1980 e 90, mas que se encontram ainda hoje nos espaços escolares.

Desenvolvimento da Avaliação Externa da Aprendizagem dos Alunos

Apesar de ser um tema amplamente discutido na atualidade, a avaliação externa teve início na China, há cerca de 2.500 anos, com a finalidade de selecionar indivíduos para os serviços público e militar. Regras para a realização do exame foram estabelecidas, tais como: isolamento do local do exame com o exterior, um candidato por mesa, identificação do candidato feita apenas por número no papel da prova. Tal estrutura tinha como objetivo estabelecer maior rigor e fiabilidade ao procedimento empregado, uma preocupação que até hoje está relacionada à natureza da avaliação externa. O sistema de exame chinês ganhou então visibilidade e foi no século XVI, através dos jesuítas, que os exames chegaram à Europa, quando eles incorporaram os exames escritos em suas escolas. No entanto, com a finalidade de seleção para os serviços público e militar, em igualdade com a China, só foi estabelecido na Europa no século XVIII, na então Prússia (atual Alemanha), em 1748. Em seguida, em 1793, a França também adotou o mesmo sistema. Somente no século seguinte, o Reino Unido implementou o sistema de exames com o objetivo de maximizar atitudes, minimizar despesas e controlar o nepotismo (Kellaghan & Madaus, 2003).

No que tange às universidades, foi no século XIX que houve a incorporação para a seleção de alunos. O procedimento iniciou-se na Alemanha, o *Abitur*, em 1788. Em seguida, foi a vez da França, que implementou a avaliação em 1808. Por fim, ocorreu a utilização nas universidades de Londres, Cambridge, Oxford e Dublin, no Reino Unido. Em 1845, os Estados Unidos, mais precisamente em Boston, deram início então ao sistema de exames. Ao perceberem a dificuldade em se avaliar de forma oral um número elevado de alunos, decidem pela aplicação dos exames escritos. Para além da facilidade de se aplicar o exame a um grande número de alunos, reconheceu-se a questão de se poder aplicar as mesmas questões, em condições semelhantes em um tempo limitado, podendo haver comparação entre as pontuações (Kellaghan & Madaus, 2003).

De acordo com Kamens e McNeely (2009), houve um crescimento explosivo de avaliações educacionais de âmbito nacional e internacional em decorrência do impacto da globalização. Os autores consideram que a proliferação e aceitação internacional de testes padronizados é uma consequência das ações de forças ideológicas associadas ao processo de globalização acelerada que têm contribuído para que uma variedade de

organizações internacionais (e.g., OCDE, Banco Mundial, UNESCO) invistam os seus esforços nesse tipo de avaliação. Desse modo, procuram influenciar as estruturas econômicas e políticas dos países participantes.

Teodoro (2001) afirma que houve um forte impulso no que se refere à internacionalização das problemáticas educativas, devido à criação de um vasto sistema de organizações internacionais de natureza intergovernamental. Dessa forma, o desenvolvimento de políticas educativas de uma diversidade de países, principalmente de países subdesenvolvidos e emergentes, começou a depender de forma mais acentuada da legitimação e da assistência técnica das organizações internacionais.

Esse desenvolvimento veio a fundamentar-se na concepção de educação comparada, entendendo que apenas “quando comparamos o sistema educacional de nosso próprio país com o de outros, é que tomamos consciência de certos aspectos distintivos da educação nacional e, então, passamos a elaborar critérios que mais judiciosamente nos permitirão interpretar semelhanças e diferenças.” (Bonitatibus, 1989, p. 14). Plomp (1998) ressalta a relevância dos estudos comparativos internacionais para avaliar e controlar a qualidade dos sistemas educacionais. É nesse delineamento internacional, no qual se destacam a globalização e a educação comparada, que se estabelece uma cultura de avaliação externa.

Com o intuito de realizar um levantamento periódico de informações sobre o sistema educacional, a avaliação do sistema passa a realizar a verificação, com maior ênfase, do desempenho acadêmico do aluno, através de testes de rendimentos, e a análise dos questionários, para identificação de fatores que possam vir a influenciar esse desempenho, e de outros fatores escolares (Sousa & Oliveira, 2010). Segundo Clímaco (2005), frequentemente, esse tipo de avaliação é “referida como um modo de fundamentar as políticas de reforma e até certo ponto, é considerada como a expressão das preocupações com a qualidade e, genericamente, da democratização da educação.” (p. 55). A prática de avaliação de sistema, de acordo com Vasconcellos (2010), é uma tendência no momento, tanto no que diz respeito ao acompanhamento do sistema quanto à exigência de organismos internacionais.

Para que venha alcançar o seu objetivo frente a essas demandas e outras necessidades dos sistemas de educação, a avaliação externa assume um conjunto variado de funções, sendo as mais frequentes: certificação, que é utilizada para a comprovação de que um candidato atingiu um nível de aprendizagem ao fim de um período determinado; seleção, que visa ao controle de entrada e prosseguimento nos estudos; controle, que

possibilita ao governo assegurar que conteúdos básicos sejam ministrados nas escolas segundo o currículo nacional ou local; monitorização, que está relacionada à prestação de contas, tendo em vista que se utilizam os resultados da avaliação externa para pedir contas às escolas e aos professores; e orientação, que consiste em comunicar aos alunos avaliados os resultados, servindo como motivação para o empenho em tarefas para a melhoria da aprendizagem, auxiliando no seu progresso escolar (Fernandes, 2008; Kellaghan & Madaus, 2003; Mons, 2009; Natriello, 1987).

Para além de suas diferentes funções, há diversas características que se apresentam nas avaliações externas nos diferentes países. O período escolar em que os exames são aplicados, no início, no meio, no final, a cada bimestre, trimestre ou ciclo de estudo. O número de órgãos envolvidos na administração da avaliação externa também é uma característica que diferencia entre um país e outro. Pode haver uma única autoridade para o país ou diferentes sistemas (federal, estadual e municipal), com diferentes autoridades e agências de avaliação públicas ou privadas. Outra característica é o grau de externalidade da avaliação, podendo ser total, quando todo o processo é externo à escola, ou parcial, quando incorporam um elemento da escola, por exemplo, no momento em que o professor realiza a correção do exame. A variação no formato do exame também merece destaque, pois há questões como ensaios, respostas curtas, respostas objetivas etc. (Kellaghan & Madaus, 2003).

Apesar dessa diversidade no que tange à avaliação externa, com relação às funções e formatos, há um número de características comuns à maioria delas. É uma medida em larga escala, tendo em vista que abrange um contingente considerável de participantes, sendo, assim, em grande escala (país, região, grau, faixa etária etc.). As avaliações são desenvolvidas e/ou controladas por uma agência externa à escola, que é normalmente o governo, seja ele nacional, estadual ou municipal. No entanto, há ocorrências de agências de avaliação na administração do processo, assim, o governo fica apenas com a função de supervisão. Os exames apresentam como base os conteúdos das diferentes áreas que compõem o currículo escolar, ocorrendo, assim, uma maior ênfase no conteúdo, em detrimento das habilidades – geralmente, com ênfase em apenas duas disciplinas, leitura e matemática. Há uma padronização na administração, material e correção dos testes. A maioria dos exames apresenta várias funções, como certificar, selecionar, motivar e controlar. Em geral, os conteúdos, os critérios de correção e os resultados dos exames tornam-se públicos (Haney & Madaus, 1986; Kellaghan & Madaus, 2003; Stecher, 2002).

Por fim, o critério credibilidade é atribuído à avaliação externa como ponto de grande evidência. Determinados grupos sociais consideram-na uma ciência exata, objetiva, proporcionando informação precisa, sem ambiguidade e consistente. Tal fato garante a confiabilidade das informações apresentadas, para acima de qualquer suspeita, sobre o conhecimento dos alunos. Com isso, um dos processos que pode contribuir para melhorar a qualidade da educação é, para um número de autores, a avaliação dos sistemas educacionais, desde que a informação gerada seja credível e a sua utilização adequada à melhoria dos processos e produtos envolvidos (Fernandes, 2014).

No entanto, de acordo com Afonso (2014), é indispensável uma análise crítica à excessiva centralidade da avaliação externa da aprendizagem dos alunos, pois, apesar de ser apenas uma dimensão do vasto campo da avaliação educacional, estamos num momento em que a avaliação institucional das escolas, a avaliação do desempenho docente e a avaliação de políticas subordinam-se ou referenciam-se a ela.

É evidente que a avaliação externa ocupa, atualmente, um espaço central na discussão nos espaços escolares, nas discussões sobre o sistema educacional e nas agendas de políticas públicas locais, nacionais e internacionais. Desde discursos que lhe dão *status* de salvadora da pátria até aqueles que apresentam movimentos absolutamente contrários. Entender quais são as bases dos diferentes posicionamentos é entender também a diversidade do assunto, que se apresenta de forma multifacetada, sendo necessário realizar um diálogo a partir de diferentes perspectivas, atores envolvidos e fundamentos. Por isso, as próximas linhas serão direcionadas para a dimensão política da avaliação, com o intuito de alargar para além dos aspectos técnicos.

Avaliação externa como diretriz das políticas educacionais.

Na opinião de Ozga (2000, p. 22), “todas as investigações em educação se relacionam, em maior ou menor grau, com política”. Em consonância com essa ideia, entende-se igualmente que o campo de investigação na área de avaliação externa da aprendizagem dos alunos está inevitavelmente relacionado com as políticas públicas de educação. Por isso, faz-se necessária a coadunação dos dois campos, para que dessa forma, possa haver uma melhor compreensão da avaliação. Tal desenho possibilita uma visão mais completa do objeto de pesquisa, entendendo que há uma relação direta entre as duas discussões.

Não há um consenso na literatura acerca da definição de políticas públicas, devido à sua inserção recente como domínio do conhecimento. Em geral, pode-se afirmar que se

compreende como decisões e ações públicas, com o intuito de realizar a alocação imperativa de valores envolvendo bens públicos, tendo como propósito manter o equilíbrio social ou introduzir desequilíbrios para modificar uma dada realidade. Nesse contexto, há divergências de conceitualização entre os atores para uma determinação de linhas de conduta coletivas, dentro de um quadro de cooperação-integração reciprocamente reconhecido (Rua, 2012; Saraiva, 2006; Souza, 2006).

Com base nessa definição, observa-se que a política pública está relacionada com as ações formuladas pelo governo e com as relações entre os atores políticos. Dessa forma, um trabalho no campo das políticas públicas deve compreender esses atores, para além de realizar uma análise do que o governo pretende fazer e o que efetivamente faz. No desenvolvimento da política pública, o primeiro momento consiste na formulação, uma preparação da decisão política; em seguida, ocorre a implementação, que consiste na tomada de decisão por um político e sua formalização; por fim, a mensuração, uma análise dos efeitos produzidos pelas políticas públicas, sendo assim uma avaliação *a posteriori* (Saraiva, 2006). No entanto, cabe acrescentar que tais etapas são permeadas por diferentes atores, com diferentes papéis e com diferentes organizações, o que torna imprescindível realizar uma análise a partir da percepção desses sujeitos.

As pessoas envolvidas no cenário das ações das políticas públicas são chamadas de atores políticos, sendo os que formulam as políticas públicas e também aqueles cujos interesses podem vir a ser afetados, positiva ou negativamente, pelo caminho tomado por uma determinada política. Esse segundo grupo possui também o poder de influenciar a formulação e a implementação de políticas, tendo um poder não apenas de receptor, mas também um papel ativo nas decisões políticas.

Ao considerar a definição de Rua (2012, p. 17) de que as políticas públicas são o “conjunto das decisões e ações relativas à alocação imperativa de valores envolvendo bens públicos.”, pode-se considerar que políticas públicas educacionais dizem respeito às ações e decisões que se referem à educação escolar. É importante, por isso, destacar seu caráter processual (Ozga, 2000), visto que há todo um processo de negociação, de contestação ou mesmo de luta entre aqueles que estão envolvidos diretamente com as decisões e ações e também os diversos atores políticos.

Segundo Plank (2001), a relação de dependência que se estabelece entre o sistema educacional e as autoridades políticas gera três consequências principais: a) o controle sobre políticas educacionais está nas mãos de pessoas que estão fora do sistema educacional; b) os responsáveis dentro do sistema educacional geralmente não são

profissionais da área de educação; c) a continuidade de políticas educacionais e a acumulação de conhecimento especializado ficam prejudicadas, visto que ocorre rotatividade a cada eleição ou mudança ministerial. Segundo o autor, ao tomar posse um novo Ministro da Educação, ocorrem, aproximadamente, 300 nomeações políticas no Brasil, no que diz respeito aos níveis mais elevados do sistema, movimentando, dessa forma, um número considerável de atores políticos burocráticos.

A constituição e a trajetória histórica das políticas educacionais no Brasil, em especial os processos de organização e gestão da educação básica nacional, 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, pp. 925-926)

Com isso há de se considerar que existe uma grande descontinuidade das políticas públicas no Brasil, proveniente de mudanças no governo, que, em grande parte, podem ocorrer para neutralizar as marcas do antecessor (Spink, 1987) ou por pressão de diferentes atores políticos. Assim, é importante considerar e entender a trajetória das políticas públicas. Tal fato propicia uma melhor compreensão do que foi e como as ações presentes estão relacionadas com ações e decisões que vão para além do espaço e do tempo imediatos, ou seja, o presente.

Após a II Guerra Mundial, a influência da administração de empresas começa a manifestar-se na administração pública, as ideias de descentralização e de flexibilização administrativa ganham espaço em todos os governos. Assim, começa a vigorar a administração gerencial (Bresser-Pereira, 1996; Secchi, 2009).

De acordo com Bresser-Pereira (1996, pp. 11-12), aos poucos os contornos da administração pública gerencial foram se delineando, sendo eles:

(1) descentralização do ponto de vista político, transferindo recursos e atribuições para os níveis políticos regionais e locais; (2) descentralização administrativa, através da delegação de autoridade para os administradores públicos transformados em gerentes crescentemente autônomos; (3) organizações com poucos níveis hierárquicos ao invés de piramidal; (4) pressuposto da confiança limitada e não da desconfiança total; (5) controle por resultados, *a posteriori*, ao

invés do controle rígido, passo a passo, dos processos administrativos; e (6) administração voltada para o atendimento do cidadão, ao invés de auto-referida.

Dessa forma, a governança pública ganha um novo delineamento. Segundo Kooiman (1993), o Estado adota um movimento de horizontalidade entre atores públicos e privados no processo de elaboração de políticas públicas, no qual há um compartilhamento de tarefas e responsabilidades, o que é considerado como uma gestão cooperativa. Porém, Richards e Smith (2002) consideram que o Estado não perdeu seu protagonismo em uma administração com base em cooperação, mas sim mantém um papel de liderança e cria centros múltiplos de elaboração de políticas públicas, passando a ter um maior papel no que tange à política de coordenação e controle.

Na educação, a gestão gerencial vem sendo implantada no sentido de garantir o sucesso escolar, com base na ênfase de resultados e indicadores de desempenho. Para tal, o novo modelo de gestão educacional visa à descentralização dos sistemas educacionais, à implementação de mecanismos de avaliação de resultados, à flexibilização do currículo, às políticas compensatórias, aos programas focalizados e ao financiamento da educação. Apresenta como objetivo a busca por estabelecer uma cultura organizacional com base nos princípios de gestão estratégica e de controle de qualidade. Para tal, a descentralização assume um papel importante no que tange a possibilitar o envolvimento dos atores educacionais na consecução dos objetivos propostos (Castro, 2008).

Os atores políticos educacionais, a partir dessa nova gestão administrativa na política educacional, recebem um novo papel, tendo em vista que ocorre um compartilhamento de tarefas e responsabilidades ou a criação de centros múltiplos de elaboração de políticas públicas, porém o governo permanece em seu papel de liderança e de controle (Richards & Smith, 2002). Sendo, por conseguinte, importante realizar uma análise das posições e ações desses atores, pois é a partir de suas interações, disputas, composições e dimensões de atuação que resulta a política pública educacional. A Figura 1 apresenta os atores educacionais que possuem ações diretas nas políticas públicas de avaliação em diferentes níveis no contexto brasileiro.

ATORES EDUCACIONAIS	AÇÕES
Instituições Internacionais (Banco Mundial, ONU, Unesco, BID, OCDE, Unicef, PNUD e outros)	Realizam diagnósticos, análises e propostas voltadas para a solução de problemas de âmbito econômico e educacional, vindo a influenciar na formulação de políticas para a educação de vários países.
Ministério da Educação (MEC)	Tem como área de competência a política nacional de educação; a avaliação, a informação e a pesquisa educacionais; a pesquisa e a extensão universitárias; o magistério e a assistência financeira.
Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep)	Subsídia a formulação de políticas educacionais dos diferentes níveis de governo com intuito de contribuir para o desenvolvimento econômico e social do país. Operacionaliza o sistema de avaliação nacional.
Secretarias de Educação (estado e município)	Elaboram as políticas educacionais locais, coordenam a sua implantação e avaliam os seus resultados.
Agências de Avaliação (CAEd, Fundação Cesgranrio, Cespe)	Operacionalizam programas estaduais e municipais destinados a mensurar o rendimento de estudantes das escolas públicas e algumas da iniciativa privada.
Entidades Científicas do campo educacional (Anped, Anpae, Anfope, FCC, SBPC, Fineduca, Cedes, Abave entre outros)	Promovem a participação das comunidades acadêmica e científica na formulação e desenvolvimento da política educacional do país. Apresentam-se como um espaço de debate das questões científicas e políticas, e de relevante produção acadêmica.
Institutos e fundações privadas empresariais (Fundação Lemann, Instituto Ayrton Senna, Fundação Victor Civita, Itaú Social, entre outros)	Realizam ações voltadas para a educação básica, como produção de conhecimento e desenvolvimento de ações inovadoras, no sentido de garantir o direito à educação de qualidade.
Sindicatos (Sinpro-Rio, Sepe)	Realizam ações de representação, negociação e assistência aos profissionais de determinada categoria, defendendo seus direitos profissionais, sociais e políticos.

Figura 1

Atores educacionais e suas ações políticas

Apesar da disposição apresentada na Figura 1 e de acordo com Zanten (2004), citado em Barroso (2006, p. 12), deve sublinhar-se que a

heterogeneidade dos interesses em presença, a complexidade das estratégias dos atores envolvidos, os meandros da tomada de decisão política, as reinterpretações quando da sua concretização, os efeitos de retorno e os reajustamentos regulares das disposições iniciais tornarem impossível qualquer raciocínio linear e casual.

A nova gestão educacional também influencia a avaliação externa, tendo em vista que, a partir dela, se presume que as funções e os efeitos esperados dessa avaliação sejam múltiplos. Com isso, ela deve assegurar uma avaliação do desempenho do aluno; servir como um indicador da qualidade do serviço educacional oferecido; servir como ponte entre os atores responsáveis pela operacionalização desse serviço (como professores, escolas, autoridades locais) e os governantes, geralmente nacionais, responsáveis pelas

decisões políticas; servir com ferramenta norteadora, vindo a orientar as atividades dos atores locais e informar o mais alto nível hierárquico sobre o desempenho desses agentes; e, por fim, servir como ferramenta de informação e controle para atores externos à escola, sendo parte de um sistema de prestação de contas (Mons, 2009).

Essa visão de ação pública sequencial e racional, na qual o desenvolvimento e a implementação de políticas públicas são alcançados caso os meios certos sejam colocados diante de objetivos claramente definidos, apresenta uma visão de cima para baixo da ação pública, ou seja, há um controle descendente, na medida em que as autoridades públicas realizam um controle organizacional, no qual as pessoas do topo (gestores do sistema educacional) desenvolvem metas para as pessoas da base (atores escolares) perseguirem, sendo uma pressão verticalizada. A avaliação externa, nesse contexto, é uma ferramenta importante da etapa final, tendo em vista que permite monitorar o desempenho final e pressionar para que as metas sejam alcançadas (Freitas, 2007; Moe, 2000; Mons, 2009).

No Brasil, há um sistema hierárquico do tipo piramidal, que compreende desde a instância nacional, passando pelas secretarias locais, até chegar à escola, para além dos outros atores políticos contemplados na Figura 1. A escola, como base da pirâmide, apresenta diferentes atores, como a equipe de gestão escolar, os professores, os funcionários, os alunos, os pais e as associações (grêmio estudantil e associação de pais e mestres). A relação entre os diversos atores políticos educacionais, nos diferentes níveis do sistema hierárquico, é construída com base em diálogos, disputas e imposições, que são permeados por ideologias. Cabe destacar, assim, que a construção de um quadro mais completo sobre a educação consiste em uma análise não só do espaço escolar, mas do contexto no qual ela está inserida, tendo em vista que ela se constrói não apenas a partir das atividades escolares, mas também com base em ações políticas e socioculturais, que influenciam diretamente uma diversidade de dinâmicas escolares (e.g., organização e funcionamento pedagógico e administrativo, regulamentos internos, políticas de ensino e de avaliação pedagógica, sistemas de apoio às aprendizagens dos alunos).

Avaliação externa e qualidade da educação: um assunto em pauta na agenda política.

Há quase duas décadas, Enguita (2001) discorreu sobre o termo qualidade no campo da educação, entendendo que nessa área, essa palavra estaria na “moda”. E assim se faz até os dias atuais, pois a qualidade da educação ainda hoje figura nas pautas de debates das diversas instituições de ensino, dos diferentes atores políticos e com grande peso na mídia.

No campo da educação, ao longo dos últimos anos, “qualquer proposição relativa a conservar, melhorar ou mudar isto ou aquilo, não importa o que seja, deve explicar-se em termos de qualidade.”, afirma Enguita (2001, p.25). Com isso, a qualidade parece ter se transformado na palavra-chave para todos os problemas da educação brasileira. No entanto, cabe ressaltar que, apesar da importância e do peso a ela atribuídos, há um movimento que acaba por torná-la vazia, na medida em que as discussões que lançam mão de tal base, na maioria das vezes, a entendem como um conceito unívoco e consensual, quando, na verdade, estamos perante um conceito complexo, de natureza polissêmica. De fato, trata-se de um conceito que não pode deixar de estar associado a uma diversidade de perspectivas filosóficas, políticas e ideológicas, mas também educacionais e pedagógicas, sustentadas pelos autores que têm procurado caracterizar e definir aquele conceito. Devido a essa polissemia e uma postura recorrente em produzir discursos vagos e retóricos sobre a qualidade da educação brasileira, considerou-se importante discorrer sobre o assunto, pois se entende que

não se trata simplesmente de examinar qual seria o ‘verdadeiro conceito de qualidade de educação’ . . . , mas de confrontar as alternativas em função de suas implicações morais e das práticas educacionais que implicam ou que delas derivam e para as quais se pede a adesão (Carvalho, 2014, p. 329).

De acordo com Risopatrón (2007), o conceito de qualidade é um significante e não um significado. Bondioli (2004) afirma que a qualidade não é um dado de fato, não é um valor absoluto, mas sim negociável, na medida em que a qualidade apresenta um caráter de transação, pois ocorre no debate entre os indivíduos e grupos que possuem interesse e responsabilidade em relação à educação. De igual forma, Silva (2008) considera que a qualidade da educação não é um objeto de investigação empírica, mas um ajuizamento de valor a partir da concepção que se tenha de qualidade, uma análise a partir de certas condições históricas, territoriais, culturais, de classe ou de grupo social. Por fim, porém, longe de esgotar o assunto, para Horta Neto (2010), a qualidade da educação estaria ligada à existência de um padrão, dependendo, por conseguinte, de uma referência a partir da qual seja possível comparar os processos educacionais que se deseja medir a esses padrões.

Como se pode ver, o termo qualidade tem sido objeto de diferentes significados, tendo em conta os valores, as experiências e, em geral, as concepções acerca de um alargado leque de questões filosóficas, ideológicas, políticas, culturais e outras,

sustentadas por quem o define. Nesses termos e tal como afirmam Dourado e Oliveira (2009), um primeiro ponto a ser destacado é que a “qualidade é um conceito histórico” e situado no tempo e no espaço, sendo, por isso, sujeito a alterações ao longo dos anos.

O conceito de qualidade da educação foi levado ao patamar de “palavra de ordem” a partir do movimento que começou na década de 1980, de acordo com Virgínio Sá (2008), mais precisamente em 1983, com o relatório *A Nation at Risk*, elaborado pela Comissão Nacional para a Excelência na Educação do governo norte-americano. O relatório discorreu sobre a relação direta entre a perda da competitividade relativa da economia americana com uma suposta queda no nível de desempenho das escolas, principalmente das públicas, levando, por conseguinte, a uma urgente reforma educacional nacional.

A América Latina, também em 1980, inicia suas discussões em torno do campo da qualidade da educação, muito em decorrência dos esforços empregados para o aumento do número de matrícula nas escolas. Nas décadas anteriores, a pauta vigente na educação trazia como questão a igualdade de oportunidades, ou seja, a democratização da educação no sentido de ampliação do acesso à escola. Assim, a preocupação, nesse momento, girou em torno da ampliação quantitativa da educação. Por conseguinte, a demanda pela oferta de vaga era mais forte do que pensar sobre a forma do processo educativo e as condições necessárias para um ensino de qualidade (Oliveira & Araújo, 2005).

No Brasil, os indicadores de qualidade foram sendo construídos ao longo do tempo, formando três momentos distintos de qualidade da educação, estabelecidos simbolicamente e concretamente na sociedade. A qualidade, então, é estabelecida, num primeiro momento, com um indicador de oferta limitada, sendo a educação voltada para atender o interesse de uma minoria privilegiada; passando pelo indicador relacionado à ideia de fluxo, quando a definição de qualidade está no número de alunos que progredem ou não no sistema de ensino; por fim, o atual momento, é a ideia de qualidade relacionada com a medição de desempenho acadêmico dos alunos (Oliveira & Araújo, 2005).

Esse terceiro momento, que a educação brasileira vem vivenciando, é a utilização dos índices dos testes padronizados da capacidade cognitiva dos estudantes para aferir a qualidade da educação, como o Sistema de Avaliação da Educação Básica (Saeb), a Avaliação Nacional da Alfabetização (ANA), o Exame Nacional do Ensino Médio (Enem), dentre outros. Segundo Gatti (2013, p. 30), “essas avaliações se tornaram o carro-chefe das ações políticas em educação, muito especialmente em nível federal”, relativamente novas no Brasil, porém difundidas há mais tempo em outros países.

De acordo com Fernandes (2013), esse indicador de qualidade está voltado para um pressuposto epistemológico de qualidade como medida, tendo em vista que é determinada por critérios e *standards*. Essa visão vem sendo muito difundida nos sistemas educacionais e nas sociedades em geral, apresentando, como lógica da avaliação, as seguintes ações: definição de critérios que caracterizem o mérito; definição de *standards*; medição do desempenho que é comparado com os *standards*; e análise e interpretação dos resultados de forma que se produza um juízo acerca do mérito, valor ou significado do objeto. Alavarse, Bravo e Machado (2012) em pesquisa realizada com base no índice de Desenvolvimento da Educação Básica – Ideb, identificaram uma tendência atual nas políticas educacionais com relação à implementação de avaliação externa com o objetivo de incrementar a qualidade da educação. No entanto, os autores asseveram a necessidade de mais investigação sobre essa tendência.

Nesse contexto de avaliação externa da aprendizagem do aluno, deve-se destacar a diferenciação apresentada por Harvey (2006) entre o conceito de qualidade e o conceito de processos externos de garantia de qualidade. Este é sobre a medida da qualidade de um processo, considerada no seu resultado, tendo como objetivo controle, responsabilização e melhoria. A qualidade fornece a base conceitual para o processo de garantia de qualidade. A avaliação externa não representa a qualidade da educação, mas sim um dos processos externos à escola que são utilizados para a garantia da qualidade. Ou seja, a avaliação seria utilizada como forma de controle, regulação, responsabilização para a melhoria da qualidade da educação. Assim, o texto de Harvey (2006) lança luz na diferenciação entre a natureza do processo e o resultado. A avaliação seria uma das formas para, a partir dos resultados, poder planejar, elaborar ações que possam levar a uma melhoria do processo. Como afirmam Alavarse et al. (2012, p. 18), “se, de um lado, qualidade não se confunde com desempenho em leitura e resolução de problemas, de outro, esses tópicos não são, de forma nenhuma, estranhos ao processo escolar que se pretenda de qualidade”.

É a partir da utilização da avaliação externa como um processo para controle, regulação, responsabilização e melhoria da qualidade da educação, que se pode categorizar essa avaliação em duas representações diferentes de usar o teste a serviço da política: testes *low-stakes*, responsabilização branda, e testes *high-stakes*, responsabilização forte.

Os testes de responsabilização branda não têm consequências significativas, tangíveis ou diretas anexas aos resultados, pois se entende que apenas a informação é

suficiente como incentivo para as pessoas agirem. Em contrapartida, testes com responsabilização forte consideram que somente os resultados não são suficientes para motivar a ação, assim, supõe-se que deve haver a promessa de recompensas ou a ameaça de sanções para que haja mudança (Haney & Madaus, 1986; Heubert & Hauser, 1999; McDonnell, 1994). Com isso, a avaliação externa ganha um papel político de instrumento central para a regulação dos sistemas educacionais e responsabilização de seus atores, responsabilização essa realizada com base no desempenho do aluno. Dessa forma, há uma dualidade político-pedagógica, que se constrói do macro (política) para o micro (pedagógica).

Há uma falta de convergência nos estudos sobre a efetividade da utilização de testes de responsabilização forte. Alguns estudos apontam para que estudantes em estados com testes de responsabilização forte tiveram, em média, mais ganhos significativos nos testes do que alunos em estados com pouca ou nenhuma responsabilização (Carnoy & Loeb, 2002; Linn, 2000). Outros estudos revelam que tais políticas pareciam ter efeitos prejudiciais no currículo, didática do professor, na permanência dos alunos na escola, mostrando, por conseguinte, que não apresentam verdadeiros benefícios (Hong & Youngs, 2008; Shepard, 2000). No entanto, faz-se importante pontuar, assim como afirmam Koretz, McCaffrey e Hamilton (2001), que poucos esforços vêm sendo empregados no sentido de avaliar diretamente os ganhos obtidos em contextos de avaliação de responsabilização forte, assim como os estudos existentes carecem de instrumentos com ferramentas convencionais adequadas de validação para a realização da pesquisa.

No cenário brasileiro, com base em Bonamino e Sousa (2012, p. 377), há três gerações de avaliação. As avaliações de primeira geração seriam as que apresentam o objetivo de acompanhar a evolução da qualidade da educação. Para tal, divulgam seus resultados na Internet, para consulta pública, ou pela mídia, sem que as escolas recebam de forma direta esses resultados. O Saeb é um exemplo desse tipo de avaliação, tendo em vista que apresenta baixa interferência na escola, mostrando-se mais adequado para diagnosticar e monitorar a qualidade da educação básica de forma geral, sem apresentar índices individuais de desempenho por escola ou aluno.

A segunda geração, para além de realizar a divulgação pública, faz a devolução dos resultados para a escola, sem estabelecer consequências materiais, mas sim simbólicas. Com isso, o pressuposto desse tipo de avaliação é que apenas o conhecimento dos resultados favorece a realização de ações por parte das equipes escolares para a

melhoria da educação. Um exemplo de avaliação de segunda geração é a Prova Brasil, que produz informações com o objetivo de ajudar os governantes nas decisões políticas e os atores escolares, visando à melhoria da qualidade da educação.

Por fim, a terceira geração de avaliação educacional, que apresenta políticas de responsabilização forte, ou *high-stakes*, tendo em vista que contempla sanções ou recompensas em decorrência dos resultados dos alunos. Com o desenvolvimento dos sistemas próprios de avaliação dos estados e dos municípios, há mais incidências desse tipo de avaliação. Como exemplo, há o Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (Saresp), o Sistema de Avaliação Educacional de Pernambuco (Saepe), o Sistema de Avaliação da Educação do Estado do Rio de Janeiro (Saerj), dentre outros.

Para ter uma visão mais clara dos sistemas municipais de avaliação, Bauer, Sousa, Horta Neto, Valle e Pimenta (2017) realizaram um mapeamento das iniciativas da avaliação no ensino fundamental. Uma pesquisa por meio de questionário, que contemplou três dimensões: interação do município com avaliações em larga escala propostas pelo governo federal e pelos governos estaduais; avaliações propostas pela rede municipal; usos dos resultados das avaliações implantadas nos municípios. Dos 5.568 municípios existentes no Brasil, a pesquisa conseguiu a resposta de 4.309, o que representa 77,4% do total de municípios. Alguns dados são interessantes para o escopo desta tese.

Tabela 1

Participação dos municípios nas diferentes avaliações externas

Participação dos Municípios	Porcentagem
Adesão às avaliações do governo federal (Provinha Brasil, Prova Brasil e ANA)	≥ 90%
Adesão à avaliação estadual	67%
Avaliação própria municipal	37%

Fonte: Bauer et al. (2017)

A partir desses dados, fica evidente a coexistência das avaliações administradas por diferentes instâncias governamentais (Tabela 1), evidenciando a multiplicidade de avaliações. Na Tabela 2, as respostas evidenciam a crença de que a avaliação externa pode induzir à melhoria da qualidade, pois as alternativas que obtiveram maior escolha seguem

nesse sentido: melhorar a aprendizagem dos alunos (85,3%) e melhorar o índice de desenvolvimento da educação básica (40,2%).

Tabela 2

Frequência observada dos principais objetivos da avaliação de alunos promovida pela rede municipal de ensino

Principais objetivos da avaliação de alunos		n	%
Acompanhamento e monitoramento da rede	Melhorar a aprendizagem dos alunos	1.119	85,3
	Melhorar o índice de desenvolvimento da educação básica	524	40,2
	Reduzir a taxa de repetência e/ou abandono	427	32,8
Subsidiar o planejamento educacional	Definir prioridades para a gestão da educação municipal	520	39,9
	Obter informações para a formação continuada	450	34,6
	Rever propostas curriculares do município	442	33,9
	Aprimorar os processos de gestão da rede municipal	237	18,2
	Obter informações para a premiação/bonificação de escolas, gestores, professores e/ou alunos	45	3,5
Divulgar os resultados das escolas para os pais e a comunidade		72	5,5
Outros objetivos		23	1,8

Fonte: Bauer et al. (2017, p.7)

Com relação ao uso dos resultados relacionados à implantação de incentivos a escolas, alunos e professores, a pesquisa conclui que “tais práticas não são tão recorrentes como se poderia esperar pelo que vem sendo divulgado na literatura acadêmica” (Bauer et al., 2017, p. 16), tendo em vista que apenas 13,3% afirmaram premiar alunos (computador, viagem etc.); 9,5% afirmaram dar prêmios às escolas (computadores, recursos financeiros etc.); e 6,3% oferecem bônus salarial aos profissionais das escolas.

A pesquisa apresentada evidencia que, atualmente, os índices dos testes padronizados da capacidade cognitiva dos estudantes são utilizados para aferir a qualidade da educação, tendo em vista que a maioria dos municípios entrevistados considera que avaliação externa pode induzir à melhoria da qualidade. Com isso, a concepção que se apresenta com base nessa relação, avaliação externa e qualidade da

educação, diz respeito ao processo de ensino e aprendizagem, na forma de ensinar e na garantia de se aprender, vindo a proporcionar o domínio eficaz dos conteúdos prescritos em um determinado currículo mínimo. A atenção maior está no monitoramento do desempenho dos alunos, para que se possa, se necessário, mudar a condução do que é ensinado e como é ensinado, do que é aprendido e como é aprendido (Madaus & Russell, 2010). Esse é o recorte que se faz ao se falar de qualidade da educação nos tempos atuais, pouco se fala em acesso, permanência, número de concluintes e insumos.

Utilização da avaliação externa da aprendizagem.

A avaliação externa da aprendizagem dos alunos vem sendo utilizada pelas políticas públicas educacionais como instrumento para o alcance de diferentes propósitos. Nas últimas décadas, seu uso se desenvolveu significativamente, cumprindo funções em diferentes níveis governamentais. De acordo com Mottier Lopez (2009, citado por Yerly, 2014), há três níveis interdependentes de utilização: nível micro – sala de aula (ensino-aprendizagem) –, nível meso – gestão local (estabelecimento ou região) – e nível macro – sistema educacional e políticas educacionais.

A avaliação externa do desempenho do aluno tornou-se uma ferramenta de política educacional, desde o nível internacional até o local, vindo a apresentar diferentes objetivos, que vão desde estudos comparativos, responsabilização de instituições, avaliação do desempenho do sistema como um todo até a influência de ações em sala de aula. Os resultados da avaliação são usados como indicadores de qualidade da educação, servindo também como instrumentos de controle, regulação, pressão verticalizada e descentralização (Freitas, 2007; Mons, 2009).

Os programas de avaliação vêm usando a perspectiva epistemológica de qualidade como medida, produzindo, por conseguinte, dados essencialmente quantitativos e indicadores observáveis. Nessa perspectiva, uma vez que não é possível medir diretamente a qualidade, são normalmente utilizados indicadores que são, ou deveriam ser, socialmente construídos e que permitem reconhecer ou identificar a qualidade (Horta Neto, 2010; Oliveira & Araújo, 2005). Nesses termos, um indicador é “uma variável operativa usada para mensurar, indiretamente, um conceito que não admite mensuração direta” (Fonseca, 2010, p.10). No campo da educação, há uma grande diversidade de indicadores orientados para diferentes aspectos, em consequência das diferentes agendas políticas e do grande número de dados por elas produzidos. No entanto, apesar dessa variedade, os indicadores devem ter por objetivo fornecer uma ideia tão aproximada

quanto possível sobre a qualidade da educação para os diferentes atores envolvidos com o sistema escolar, sobretudo para aqueles que são responsáveis por dirigir as suas instituições.

A construção de um sistema de indicadores é entendida com base em algumas funções, como instrumento para descrever a situação educacional, fazer comparações, formular metas, fornecer explicações, diagnosticar problemas, tomar decisões, realizar avaliação e monitoramento. Para isso, necessita apresentar algumas características, tais como: validade, confiabilidade, fidedignidade, utilidade, tempo de disponibilidade, viabilidade de custo, relevância voltada para as políticas públicas, informações de fácil entendimento para o público em geral (Darling-Hammond, 1991; Fonseca, 2010).

Contudo, alguns autores afirmam que os indicadores são medidas que podem apenas retratar o que está sendo analisado, porém, devido a suas limitações, não avalia de fato a qualidade da educação, isto é, eles indicam um fato, porém apresentam limitações para explicar os porquês de esse fato estar ocorrendo (Bottani, 1998; Horta Neto, 2010; Vitelli, Fritsch & Corsetti, 2018). Uma das grandes limitações apontadas pela literatura é o uso de um indicador de forma isolada, pois representa apenas um aspecto de uma realidade complexa. Por isso, deve-se considerar a importância de se trabalhar um grupo de indicadores, o que possibilita retratar várias visões da realidade. Para além disso, levar em consideração que um dos maiores perigos dos indicadores é em decorrência da facilidade com que pode haver uma incompreensão, por serem usados ou interpretados de maneiras inválidas, dando falsas impressões. Tal fato se agrava quando os indicadores são cada vez mais usados para avaliar professores e julgar escolas. Em tais condições, os indicadores podem ser vistos como prejudiciais (Darling-Hammond, 1991).

Outra questão apontada pela literatura é o grande número de dados gerados pelos indicadores, tendo em vista que perdem sentido pela quantidade de informações sem análise. Bottani (1998) e Souza e Oliveira (2010) não acreditam ser na abundância de informações e de medidas que se produzirá um melhor conhecimento, não vendo, dessa forma, sentido na produção de centenas deles.

Atualmente, a avaliação externa da aprendizagem do aluno vem sendo utilizada como um indicador da qualidade da educação. Com a aplicação do teste, pressupõe-se que seja possível avaliar o que o aluno aprendeu ou não dos conteúdos propostos. Essa forma de verificação da qualidade ainda encontra resistência entre os profissionais da educação, considerando-se que “não faz parte da nossa história e da nossa cultura educacional pensar a qualidade enquanto medida, apesar da importância socialmente

atribuída aos exames vestibulares na difusão de certo padrão de qualidade de ensino” (Oliveira & Araújo, 2005, p. 12).

Porém, cabe ressaltar que a utilização da avaliação como medida não é algo novo, vindo a ser apresentada por Guba e Lincoln (1989), na década de 1980, como a primeira geração de avaliação, tendo em vista que tinha como objetivo medir as aprendizagens escolares dos alunos. E, desde então, vem sendo discutida na literatura (cf. Hadji, 2001; Hoffmann, 2006; Luckesi, 1998, 2014), chegando-se a afirmar que “a escola brasileira opera com a verificação e não com a avaliação da aprendizagem” (Luckesi, 1998, p. 76). Com isso, apesar de ser novo no campo da qualidade da educação, está presente no campo da avaliação há muito tempo.

Diante de tal situação, é pertinente afirmar que não é unânime a utilização de exames padronizados como indicadores para nortear políticas e programas educacionais ou verificar a qualidade da educação, cabendo uma análise mais profunda sobre as bases dos diferentes posicionamentos. A aceitação de tais avaliações entre a comunidade acadêmica e educacional brasileira está longe de ser um consenso. De acordo com a literatura (cf. Alavarse, 2013; Bauer et al., 2015; Oliveira & Araújo, 2005), podem-se destacar três grupos que apresentam distintos posicionamentos sobre a questão. O primeiro grupo é composto pelos que acreditam na capacidade de se revelar a verdade por meio das avaliações externas, dando-lhes um caráter de garantia absoluta. O segundo grupo desconsidera as eventuais contribuições que podem decorrer das medidas educacionais em larga escala, uma vez que elas não fazem parte da história e da cultura educacional brasileira. Há também um terceiro grupo que reconhece as contribuições, porém se contrapõe à forma como elas têm se difundido no território nacional, bem como ao uso que vem sendo feito, como, por exemplo, ser considerado como indicador único e principal da qualidade da educação.

A avaliação educacional vem sendo considerada cada vez mais indispensável para descrever, compreender e agir sobre uma grande variedade de problemas que afetam os sistemas educativos e formativos. Trata-se de um processo social complexo que envolve pessoas que funcionam em determinados contextos, com os seus valores, as suas práticas e políticas próprias e envolve também a natureza do que está a ser avaliado que, por sua vez, também tem as suas finalidades, lógicas e políticas próprias. Nessas condições, não se pode propriamente esperar que os resultados de um número de avaliações sejam consensualmente aceites por

todos aqueles a quem, de algum modo, a avaliação possa interessar. (Fernandes, 2013, p. 13)

O uso da avaliação externa da aprendizagem dos alunos se desenvolveu significativamente nos últimos anos. Hoje cumpre múltiplas funções, em diferentes níveis dos sistemas, sendo utilizada e impactando diferentes sujeitos do processo educacional. Assim, é necessário ter em mente as pessoas que estão envolvidas no processo avaliativo, os seus contextos e qual lugar que elas ocupam, para poder ter uma visão mais completa do quadro apresentado.

A função política da avaliação externa estaria no nível macro, nas avaliações externas internacionais e nacionais, e, no nível meso, nas avaliações estaduais e municipais. No nível macro, as avaliações internacionais e nacionais estariam voltadas mais para os decisores políticos, com o objetivo de uma educação comparada, tendo em vista que, devido à sua natureza muito abrangente, seria difícil afetar os professores em sala de aula. Em menor instância, há as avaliações subnacionais, que buscam lançar luz em experiências locais, analisando-as como um sistema educacional e realizando um estudo comparado com as avaliações nacionais (Alavarse et al., 2012; Sousa & Oliveira, 2010). Nos Estados Unidos, iniciou-se a avaliação subnacional com sistema de responsabilização escolar nos anos de 1980, realizando desenvolvimento de padrões e conteúdo para a aprendizagem em cada série. Nos anos 90, passou-se a realizar testes de forma individual por escola e a realização de classificação por desempenho. A federação exige que todos os estados realizem procedimentos de prestação de contas, caso contrário eles não receberiam financiamento federal para os programas educacionais (Hanushek & Raymond, 2005; Linn, 2000).

A República Federativa do Brasil compreende, no que contempla à sua estrutura político-administrativa, a União, os Estados, o Distrito Federal e os Municípios, considerados entes federativos¹. Assim, ao longo dos anos, pode-se observar o fortalecimento de iniciativas governamentais das diferentes instâncias de implementação de avaliação externa da educação básica. Efetivamente, a partir de 1990, houve uma necessidade de obter informações sobre o sistema educacional brasileiro, tendo em vista que os diferentes governos procuraram seguir a tendência internacional, ou seja, busca da qualidade da educação e, para tal, a necessidade de conhecer os desempenhos dos alunos

¹ De acordo com a Constituição de 1988, são considerados como autônomos no que se refere à tomada de decisões e responsabilidade.

através de uma diversidade de avaliações externas. Seguindo essa tendência, a avaliação acabou por ganhar relevância nas políticas públicas de educação, tal como vem acontecendo em grande parte do mundo. Assim, como iniciativa do governo federal, surgiu o Sistema de Avaliação da Educação Básica (Saeb), que veio atender e iniciar o cumprimento da Emenda Constitucional nº 53 e o art. 208, IV, da Constituição Federal de 1988. Tal movimento abriu espaço para que estados e municípios também colocassem em vigor a geração de informações concretas da população escolar da educação básica.

confere à União, aos Estados, ao Distrito Federal e aos Municípios o financiamento de todos os níveis da educação básica, a melhoria da qualidade de ensino, de forma a garantir padrão mínimo definido nacionalmente, estabelece a responsabilidade do estado em assegurar padrões mínimos de acesso, permanência e desempenho escolar dos alunos. Para cumprir essa tarefa complexa, que envolve grandes desafios, entre eles as desigualdades extra e intraescolares: a pobreza e a violência; as novas formas de estrutura familiar; as particularidades de cada local, de cada escola e do desenvolvimento cognitivo de cada aluno é necessário reunir informações concretas sobre a população atendida e o ensino ofertado para, deste modo, programar ações que visem atingir o objetivo traçado (Brasil, 1988).

Como objetivos principais do sistema de avaliação em seu primeiro ano de aplicação, o Saeb apresentou, dentre outros, o desenvolvimento e/ou aprofundamento das capacidades avaliativas das unidades gestoras do sistema educacional (o próprio Ministério, as Secretarias Estaduais e Órgãos Municipais de Educação); e a regionalização da operação do processo avaliativo, com o intuito de possibilitar a gestão desse processo diretamente por instâncias locais.

Dessa forma, há uma profusão de iniciativas subnacionais geridas pelos estados a partir dos anos de 1990. Contudo, apenas nos anos 2000, há uma evidência maior na quantidade dessas avaliações (Ver Anexo 1). No âmbito municipal, observa-se que o movimento foi mais recente; de acordo com pesquisas, verifica-se que os municípios vêm realizando avaliações externas próprias para que possam ancorar suas políticas educacionais em seus resultados, tornando-as diretrizes de suas propostas e ações para a educação pública, podendo realizar comparações com as avaliações nacionais (cf. Alavarse et al., 2012; Alavarse, Machado & Arcas, 2017; Bauer et al., 2017; Bonamino

& Sousa, 2012; Brooke, 2006; Brooke & Cunha, 201; Cerdeira & Almeida, 2012; Gatti, 2009, 2013; Sousa & Oliveira, 2010; Sousa, Pimenta & Machado, 2012).

Na função pedagógica, a avaliação externa teria como objetivo contribuir com informação sobre o aluno e com o aporte de seu processo e desenho. Com isso, seria uma complementação da avaliação interna e um recurso para reflexão do professor sobre o sistema de avaliação, promovendo um auxílio na aprendizagem de cada aluno ou de um grupo de estudantes. Ela não precisa fornecer esquemas explicativos para o sucesso escolar, mas sim fazer com que os atores do sistema escolar reflitam os resultados de suas ações, para que possam ter uma imagem de suas práticas e possam melhorá-las, caso os resultados não sejam os esperados (Mons, 2009; Nevo, 1997). Com relação à função pedagógica da avaliação externa, Pacheco (1994) apresenta quatro dimensões, a saber: dimensão pessoal, que visa à estimulação do sucesso dos alunos; dimensão didática, que objetiva o diagnóstico, melhoramento e verificação dos resultados; dimensão curricular, que possibilita o melhoramento de currículo frente às necessidades dos alunos; e dimensão educativa, correspondendo à avaliação da qualidade da educação.

Com relação ao nível micro de utilização da avaliação externa, há uma percepção por parte dos decisores políticos de que não podem regular diretamente a instrução em sala de aula, mas que podem influenciar a instrução de forma indireta, ao atribuir recompensas ou sanções aos resultados da avaliação externa. Com isso, esse tipo de avaliação é vista como um sistema de monitoramento do desempenho dos alunos, visando mudar a condução do que é ensinado e aprendido e como é ensinado e aprendido (Madaus & Russell, 2010).

Quanto aos professores, podem-se destacar três posicionamentos. O primeiro é um movimento de falta de credibilidade, no qual os professores não consideram que a avaliação externa esteja ajudando as escolas em seus objetivos, não oferece *feedback* útil e não está avaliando de forma adequada a aprendizagem dos alunos. Assim, esses professores não acreditam que a avaliação externa esteja ajudando a melhorar a qualidade da educação, tendo um efeito nulo ou negativo sobre as ações em sala de aula. (Herman & Golan, 1993; Mons, 2009; Stecher, 2002). Outro grupo seria o dos professores que consideram que a avaliação externa é relevante em sua prática, tendo em vista que contribui para a organização de suas ações pedagógicas ao fornecer dados que auxiliam em suas decisões e planejamentos, o que contribui com pistas sobre o que ensinar, quando ensinar os alunos. Trata-se, portanto, de um processo que contribui diretamente com o trabalho desenvolvido em sala de aula (Evers & Walberg, 2002; Madaus, Russell &

Higgins, 2009; Nevo, 1997). O terceiro grupo considera que há relevância dessa avaliação nas ações em sala de aula, no entanto desconfiam da forma como está sendo posta em prática (Marchesi, 2002; Mons, 2009).

Como exemplo do terceiro grupo, pode-se destacar a pesquisa realizada nos municípios do Rio de Janeiro e Duque de Caxias, ambos no estado do Rio de Janeiro, em que Almeida (2013) indica que a opinião dos professores sobre as avaliações não é totalmente desfavorável, mesmo que ocorram preocupações com relação às limitações e incorreções dos instrumentos de avaliação ou críticas à sua implementação.

No entanto, Mons (2009) afirma que a avaliação externa não é uma política que leva a um posicionamento (positivo ou negativo) dos diferentes atores *a priori*, mas sim que as reações desses atores estariam fortemente condicionadas pelo contexto em que elas se desenvolvem. A avaliação externa, afirma o autor, estaria mais para uma ferramenta neutra no início, permanecendo, assim, quando as consequências associadas a ele permanecem mínimas. Contudo, os testes com responsabilização forte criam reações negativas nos atores educacionais, tendo em vista que levam a um desenvolvimento de efeitos perversos que afetam negativamente os processos de aprendizagem dos alunos.

Dessa forma, é importante destacar que esse contexto expressa o projeto educacional e social que se tem por norte, tendo em vista que a

avaliação educacional mantém intrínseca relação com uma dada concepção de qualidade da educação. Seja qual for o objeto de avaliação – alunos, currículo, profissionais, instituições, planos, políticas, entre outros – o delineamento adotado em sua implantação e o uso que se fizer de seus resultados expressam o projeto educacional e social que se tem por norte. (Sousa, 2014, p. 408)

Com isso, faz-se necessário entender qual é a utilidade da avaliação para os diferentes níveis, nas diferentes dimensões, no sentido de aprofundar o estudo para verificar, além dos usos que estão sendo dados, se a avaliação está sendo útil, tanto no sentido pedagógico quanto no de políticas educacionais.

De acordo com Patton (2005, p. 72, tradução nossa),

as avaliações devem ser julgadas por sua utilidade e uso real; portanto, os avaliadores devem facilitar o processo de avaliação e projetar qualquer avaliação com consideração cuidadosa de como tudo o que é feito, do começo ao fim, afetará o uso. O uso diz respeito a como pessoas reais no mundo real aplicam os resultados da avaliação e experimentam o processo de avaliação.

A percepção dos atores educacionais sobre a utilidade da avaliação externa é de extrema importância, tendo em vista que, ao realizar uma avaliação, deve haver a preocupação de que seja realmente útil para todos os que têm interesse nessa avaliação e/ou para todos aqueles que, de algum modo, possam por ela ser afetados. Assim como afirma Dantas (2009, p. 81), cabe ressaltar que a percepção da utilidade da avaliação externa está também atrelada a posições ideológicas do indivíduo (ou seu grupo), o que poderia vir a discordar da abordagem ou dos objetivos da política de avaliação, “nesse caso, a expectativa que se tem da avaliação pode ser de uma ferramenta a serviço de uma ideologia outra que não a sua, o que a torna não útil e, por vezes, deletéria”.

De acordo com Reboloso et al. (2002), a utilidade envolve quatro componentes objetivamente analisáveis: a) a relevância dos dados em relação à necessidade de informação; b) o tempo relacionado à entrega dos resultados da avaliação; c) a apresentação do relatório (entendida como a organização lógica da informação, no sentido de ser coerente e de fácil compreensão); e d) o impacto real do relatório (em termos de uso demonstrável ou sua influência na legislação ou na tomada de decisões). Os autores afirmam que as medidas de utilidade se destinam a reforçar a execução de negociações específicas com o cliente e o monitoramento da satisfação e do uso.

Patton (2003) afirma que a avaliação deveria ser julgada pela sua utilidade e seu uso² efetivo, ou seja, a “avaliação centrada na utilização está preocupada com a forma como as pessoas reais, no mundo real, aplicam os resultados da avaliação e experimentam o processo de avaliação. Portanto, o seu foco está no uso pretendido pelos usuários.” (p. 223, tradução nossa).

Atualmente, há duas discussões com relação à utilização do indicador de avaliação externa da aprendizagem dos alunos para pensar na melhoria da qualidade da educação. A primeira estaria voltada para a lógica interna dos processos de avaliação, uma discussão direcionada para a análise de princípios, procedimentos, limites e potencialidades. Assim, a utilização estaria atrelada ao processo ou desenho da avaliação, isto é, a avaliação tanto propicia dados para um processo de pensamento dos utilizadores como também contribui com características ligadas ao seu desenho, como descritores das matrizes de referência e formatos de itens. A segunda discussão, e a mais frequente, compreende a utilização de resultados, proveniente das descobertas da avaliação (Alkin & Taut, 2003; Patton, 1997, 2003; Sousa & Oliveira, 2010; Silva, Gimenes, Moriconi & Louzano, 2013).

² Cabe ressaltar que utilização e uso, neste trabalho, são tratados como sinônimos, seguindo a linha teórica da área.

Para além da utilização da avaliação com base no resultado ou no processo, Leviton e Hughes (1981) apresentam uma divisão a partir de três níveis, a saber: instrumental, conceitual ou simbólico. O uso instrumental diz respeito à utilização da avaliação de forma direta, usada para a tomada de decisão ou solução de problemas, uma função de apoio. O uso conceitual refere-se a casos em que nenhuma ação direta foi tomada, no entanto, o entendimento das pessoas foi afetado após a avaliação, ou seja, visando influenciar o pensamento de um formulador de política sobre um determinado problema, mas sem dar informações para qualquer uso específico, o que constituiria uma função educativa. O uso simbólico é quando uma pessoa usa a mera existência da avaliação, ao invés de qualquer aspecto de seus resultados, para persuadir ou convencer outras pessoas a apoiar uma posição, uma função política. De acordo com os autores, há um consenso de que a utilização da avaliação não é um conceito unitário, mas sim existem diferentes tipos de utilização com base nos propósitos a que servem.

Shulha e Cousins (1997) afirmam que, em esforço para aprimorar os usos da avaliação nas dimensões apresentadas por Leviton e Hughes (1981), as pesquisas subsequentes vieram a desenvolver uma “lista de compras” que poderiam influenciar a utilização da avaliação externa, tais como: a) relevância; b) credibilidade; c) envolvimento dos usuários; d) eficácia da comunicação; e) potencial para o processamento de informações; f) necessidade dos clientes em obter informações; g) grau antecipado de mudança de programa; h) valor percebido da avaliação como ferramenta de gestão; i) qualidade da execução da avaliação; e j) características contextuais da decisão ou definição de políticas.

No entanto, Cousins e Leithwood (1986, citado por Shulha & Cousins, 1997) realizam um esforço para ir além da descrição desses fatores que influenciam a utilização, desenvolvendo um método meta-analítico para avaliar o peso relativo dos fatores em sua capacidade de prever o uso. Os autores concluíram que a qualidade, a sofisticação e a intensidade dos métodos de avaliação apresentavam maior capacidade de influenciar o uso dos resultados da avaliação externa.

No Brasil, de acordo com Brooke e Cunha (2011, p. 20), “os mesmos instrumentos estão sendo usados com objetivos diversos, de modo que a finalidade precípua ou original do instrumento pouco esclarece sobre seus desdobramentos posteriores”.

No que tange aos usuários da avaliação, a utilização pode ser dividida em dois grupos: as instâncias de gestão educacional, voltadas para a elaboração e implementação de políticas públicas, como o Ministério da Educação, a Secretaria de Educação do estado

e do município; e as instâncias locais, que são as ações por parte da escola, como os gestores escolares, os professores, os alunos e os responsáveis pelos alunos.

A partir da análise de duas pesquisas voltadas para o levantamento da utilização da avaliação (Brooke & Cunha, 2011; Silva et al., 2013), foi elaborada a Figura 2, com a finalidade de diferenciar a utilização, segundo os usuários da avaliação e o tipo de utilização.

	Instância de Gestão Educacional – MEC e Secretarias (Macro e Meso)	Instância Local – Gestão Escolar e Professor (Micro)
Processo ou Desenho da Avaliação	Revisão ou criação de orientações curriculares para a Rede de Ensino.	Realização de provas simuladas; Alterações na concepção/forma de avaliação da aprendizagem dos professores; Planejamento de atividades de acordo com os descritores ou com a matriz de referência da avaliação externa.
Resultados da Avaliação	Avaliar e orientar a Política Educacional; Informação para as escolas sobre a aprendizagem dos alunos; Informação para a sociedade; Estabelecimento de padrões e metas; Criação de indicadores de qualidade e equidade educacionais; Composição dos critérios para pagamento de bônus a profissionais e/ou escolas; Monitoramento pedagógico; Formação continuada para profissionais; Composição de critérios para alocação de profissionais e recursos; Utilização na política de avaliação docente; Produção de materiais pedagógicos; Certificação de alunos e escolas.	Análise dos resultados da avaliação; Uso dos resultados como critérios para a formação de grupos de alunos; Diversificação e/ou intensificação de atividades pedagógicas; Comunicação dos resultados aos alunos e pais; Elaboração de plano de ação escolar; Cobrança/reconhecimento interno às escolas a partir dos resultados; Formação continuada de professores; Monitoramento da qualidade de ensino; Subsídio para avaliação institucional; Foco nos alunos com dificuldades; Ações de reforço escolar; Planejamento das ações de sala de aula; Acompanhamento individual do desempenho escolar pelos alunos e seus familiares.

Fonte: Brooke e Cunha, 2011 e Silva et al., 2013.

Figura 2

Utilização da avaliação externa segundo os usuários e o tipo de utilização

Bauer et al. (2015) afirmam que o papel que as avaliações verdadeiramente irão assumir no conjunto de política educacional é determinado pelo uso dado aos resultados. Para os autores, as avaliações apresentam potencial para produzir avanços no conhecimento do universo educacional e podem estabelecer pontos de apoio para políticas de melhoria de sistema educacionais, se garantidas determinadas condições e adequada utilização. Porém, destacam que, em geral, o uso dos resultados não é definido *a priori* ou nem sequer é considerado na adoção de avaliações em larga escala, vindo, por conseguinte, a prejudicar sua utilização.

Há um grande número de pesquisas voltadas para a utilização dos resultados na esfera da avaliação da aprendizagem, seja no âmbito da utilização na escola, pelo professor e por seus gestores, seja na formulação e avaliação de políticas públicas, em diferentes instâncias governamentais. Porém, o campo carece de pesquisas voltadas para dimensões mais específicas da avaliação externa.

Com isso, é importante destacar algumas dimensões para um aprofundamento no tema de utilização, como relevância, credibilidade, comunicação dos resultados da avaliação externa. É importante considerar que essas dimensões não esgotam o assunto, necessitando de mais estudos, e também que as dimensões apresentadas são interligadas umas com as outras, porém foram apresentadas separadamente para ajudar na construção e percepção das múltiplas variáveis presentes na avaliação externa.

Relevância.

A avaliação externa da aprendizagem no Brasil, ao longo dos quase 30 anos, veio progressivamente ganhando espaço nas agendas políticas, mas não necessariamente nas escolas. Muitas discussões surgem no sentido da relevância da avaliação externa para diferentes esferas, desde ações políticas, passando pela gestão da unidade escolar, até chegar à ação do professor em sala de aula.

De acordo com Libânio (2013, pp. 86-87), o processo de ensino compreende um conjunto de ações organizadas do professor e dos alunos, visando alcançar determinados resultados. O autor entende que o ensino e a aprendizagem são duas facetas de um mesmo processo, ou seja, “o professor planeja, dirige e controla o processo de ensino, tendo em vista estimular e suscitar a atividade própria dos alunos para a aprendizagem”.

A avaliação externa seria um instrumento nesse processo, tendo em vista que pode servir como guia para o planejamento do professor, contribuindo com pistas sobre o que e quando ensinar os alunos, evitando, assim, perda de tempo didático (Evers & Walberg,

2002; Madaus et al., 2009). A avaliação seria, portanto, um meio de desenvolvimento e controle da aprendizagem do aluno. No entanto, alguns autores consideram que as práticas escolares, em um contexto de avaliação externa, viriam a reproduzir características de tais avaliações, para preparação dos alunos com o objetivo de alcançar os índices estabelecidos, o que Díaz-Barriga (1999) veio a considerar como “pedagogia do exame” e que não tem nenhuma relevância no processo de ensino e aprendizagem.

Para além da relevância no sentido de contribuir para as decisões e o planejamento do professor em sala de aula, há também que se considerar a relevância social da avaliação. Sua relevância estaria na contribuição como política pública para o aprofundamento dos valores democráticos, levando a uma valorização de políticas de acesso e condições de permanência, equidade, construção e socialização dos conhecimentos como bem público, associação entre as funções cognitiva e formativa integral do profissional social (Dias Sobrinho, 2008, p. 824).

Relativamente aos exames, Fernandes (2008) afirma que se todos os alunos tivessem acesso a uma educação e a uma formação semelhantes ou iguais, princípio da equidade, por consequência, apresentariam resultados semelhantes nas avaliações. Por isso, “a igualdade de resultados acaba por constituir a base mais sólida para a igualdade de oportunidades para todos os grupos de alunos que saem das escolas” (Fernandes, 2008, p. 111). Porém, Kellaghan e Madaus (2003) consideram difícil determinar se a diferença de resultado é reflexo das realizações dos grupos, ou seja, *status* socioeconômicos, localização, etnia, gênero; ou se tais diferenças ocorrem devido aos exames.

No entanto, as mesmas avaliações podem gerar efeito contrário, ou seja, o de aumentar a desigualdade, na medida em que realizam ações que buscam apenas alcançar os índices governamentais estabelecidos, deixando para trás os alunos com baixo desempenho, promovendo o aumento da segregação e não na construção de aprendizagens (Freitas, 2016; Ristoff, 1995).

Credibilidade.

De acordo com Pasquali (2009, p. 993), a “psicometria representa a teoria e a técnica de medida dos processos mentais, especialmente aplicada na área da Psicologia e da Educação”, sendo a validade e a fiabilidade os dois parâmetros mais importantes para uma medida ou um exame, considerados como aspectos técnicos para a qualidade e a credibilidade da avaliação externa.

Para Messick (1995, p. 5, tradução nossa),

A validade é um julgamento avaliativo geral do grau em que a evidência empírica e as justificativas teóricas apoiam a adequação e a apropriação de interpretações e ações baseadas em escores de testes ou outros modos de avaliação (Messick, 1989). A validade não é uma propriedade do teste ou avaliação como tal, mas sim do significado dos resultados dos testes ou de outras formas de avaliação. Esses escores são uma função não apenas dos itens ou condições de estímulo que proporcionam, mas também das pessoas que respondem, bem como do contexto da avaliação.

Os estudos de validade são divididos em quatro categorias: validade preditiva, validade concorrente, validade de conteúdo e validade de construto. As duas primeiras categorias podem ser analisadas como validade orientada por critérios, tendo em vista que o investigador tem interesse em algum critério que se deseja prever. Se o critério for obtido em um tempo após o teste, representa uma validade preditiva. Se a pontuação do teste e a pontuação do critério forem determinadas essencialmente ao mesmo tempo, corresponde à validade concorrente. Assim, a validade de critério é relativa ao grau de eficácia que o teste permite prever do desempenho específico de um sujeito. A validade de conteúdo corresponde à congruência entre o que se pretende avaliar e a natureza das questões que integram o exame ou a prova que se construiu, é necessário verificar se o teste apresenta uma amostra significativa de um universo de comportamentos (conteúdos). Não há validade se o exame não avalia o que supostamente pretendia ser avaliado. Por fim, a validade de construto refere-se a em que medida o teste constitui uma representação legítima do construto, adequada e consistente (Cronbach & Meehl, 1955).

Messick (1995) afirma que uma investigação de qualquer aspecto de validade pode contribuir para uma compreensão da validade de construto, pois essa validade abrange quase todos os aspectos de validade. A validade de construto é muito útil para explicar “a natureza dos instrumentos que medem traços para os quais não se possuem critérios externos. Assim sendo, é necessário partir de uma variável logicamente definida.” (Vianna, 1983, p. 56).

A validade de construto pode ser dividida em seis aspectos, a saber: conteúdo, substantivo, estrutural, generalização, externo e consequente (Messick, 1995). No aspecto do conteúdo, é realizada uma análise entre o conteúdo de um teste e o construto que ele pretende medir. De acordo com o autor, estudos de avaliação do conteúdo são feitos a partir de análise do conteúdo das tarefas, do currículo e da teoria do domínio. Com relação

ao aspecto substantivo, a análise é realizada no processo usado pelos avaliados quando respondem às questões da avaliação e na consistência desses processos com o construto projetado pela avaliação. O aspecto estrutural da validade de um construto examina a adequação de um sistema de pontuação no que se refere ao domínio de construção, tendo em vista que a análise de uma estrutura interna de um teste pode indicar em qual grau os itens e os componentes deste estão de acordo com o construto. A generalização enfoca na consistência dos resultados da avaliação, a fim de compreender os limites do construto. No que diz respeito ao aspecto externo, examina-se a consistência da avaliação com relação a outras medidas e construtos. Por fim, considera-se o aspecto consequencial da validade de construto, que tem como foco verificar as consequências intencionais e não intencionais do uso do teste e seu impacto (Miller & Linn, 2000).

Em crítica realizada por Vianna (1983), evidencia-se que a validade de muitos instrumentos de medida é apenas inferida, o que não condiz com uma metodologia científica. É necessário que haja provas evidentes de que um teste construído para um determinado fim é efetivamente válido.

Por fim, há uma discussão no sentido de que a validade não está apenas relacionada ao teste em si ou aos resultados que dele se derivam, mas também à interpretação e à utilização que se faz tanto dos resultados quanto das consequências das decisões advindas com base nesses resultados, ou seja, uma validade para além dos parâmetros psicométricos (Messick, 1995).

Por fiabilidade, entende-se a característica de medir, sem erros, consistência em seus resultados, na medida em que estes se mantêm mesmo em tempo ou ocasiões diferentes. De acordo com Gipps (1994, citado por Fernandes, 2008), há quatro técnicas clássicas para avaliar a fiabilidade de um teste, a saber: realizar a aplicação do mesmo teste com um intervalo de alguns dias e comparar os desempenhos; administrar versões comparáveis do mesmo texto a amostras semelhantes e comparar os resultados; realizar uma divisão aleatória do teste em duas partes para poder comparar os resultados dos alunos; e determinar um coeficiente de consistência interna, com base nas correlações que se calculam a partir de todas as possíveis divisões do teste. É necessária, por conseguinte, uma sequência de atividades para que assim possa aferir a fiabilidade de um teste.

Sousa e Oliveira (2010) destacam problemas que podem comprometer a fiabilidade dos resultados. Um deles está relacionado a algumas práticas de aplicação dos instrumentos, tendo em vista que ainda é comum a utilização de professores da escola para aplicação, juntamente com políticas em que os resultados da avaliação impactariam

salários, bonificação ou recursos para a escola. Outra questão levantada está relacionada ao interesse político de secretários e governantes, pois utilizariam os resultados positivos da avaliação como uma forma de aprovação das suas iniciativas de política educacional.

Outro fator que é importante destacar é um certo grau de desconfiança por parte dos atores escolares com relação aos avaliadores externos. No Brasil, esses avaliadores, na sua grande maioria, estão ligados às universidades, que, apesar de apresentarem grande competência acadêmica e prestigiadas pesquisas, estão distanciadas do espaço escolar, não proporcionam um diálogo de seu trabalho com a escola. As avaliações são geralmente utilizadas apenas como campo de pesquisa, sem retorno direto dos dados pesquisados, e o professor, considerado “um boneco de ventríloquo”, sendo visto como aquele que ou aplica saberes “produzidos por peritos que detêm a verdade a respeito de seu trabalho ou é o brinquedo inconsciente no jogo das forças sociais que determinam o seu agir, forças que somente os pesquisadores das ciências sociais podem realmente conhecer” (Tardif, 2002, p. 230).

Essa desconfiança, por conseguinte, não está no caráter técnico da avaliação, mas sim na percepção dos atores escolares de que os avaliadores externos, por não terem relação com o contexto escolar, não seriam capacitados para avaliá-lo, pois desconhecem a realidade de uma sala de aula. Com isso, essa descrença acaba por influenciar a percepção acerca da credibilidade da avaliação externa por parte dos atores escolares.

Para Tardif (2002, p. 230), é de suma importância levar em consideração a subjetividade dos atores escolares, ou seja, a subjetividade dos professores, para que se possa compreender a natureza do ensino, tendo em vista que

um professor de profissão não é somente alguém que aplica conhecimentos produzidos por outros, não é somente um agente determinado por mecanismos sociais: é um ator no sentido forte do termo, isto é, um sujeito que assume sua prática a partir dos significados que ele mesmo lhe dá, um sujeito que possui conhecimento e um saber-fazer provenientes de sua própria atividade e a partir dos quais ele a estrutura e a orienta.

Dessa forma, há necessidade de considerar a relação entre a credibilidade e a confiança docente que, em geral, não passa apenas pelos aspectos técnicos, mas também pela relação entre os atores escolares e aqueles que elaboram e analisam os resultados das avaliações. É fundamental estabelecer uma relação de confiança no sentido de que as evidências fornecidas pela avaliação são convincentes ou confiáveis para o uso daqueles

que têm interesse nela, corroborando com a concepção de Patton (1997) de que a credibilidade é um conceito complexo e que envolve a percepção de precisão, justiça e confiança da avaliação.

Comunicação dos resultados: a leitura e a apropriação dos resultados.

A comunicação do resultado da avaliação externa não se restringe apenas à divulgação, mas também à apropriação desse resultado pelos diferentes atores educacionais, com vistas à efetiva utilização dos resultados. Pesquisas evidenciam que é um desafio incorporar adequadamente os resultados da avaliação à prática docente. Os gestores reconhecem essa dificuldade e realizam movimentos para traduzir os resultados em uma linguagem compreensível, realizando seminários, jornadas de capacitação e produção de materiais resumidos para as escolas. Contudo, os resultados continuam a não provocar os impactos esperados na qualidade da educação, seja por que as tentativas de simplificação dessas informações complexas não têm sido efetivas ou por que não faz parte da dinâmica escolar pautar seu planejamento em resultados de avaliação (Fontanive, 2013; Fontanive, Elliot & Klein, 2007; Sousa & Oliveira, 2010).

Rodrigues (2011), ao analisar o uso e as repercussões de resultados do Saesp, conclui que a maioria dos professores não apresenta resistência quanto à avaliação externa, aceitando, de forma passiva, a sistemática de aplicação. No entanto, há uma dificuldade com relação à leitura dos resultados e ignoram a função da avaliação de definir políticas públicas com o objetivo de minimizar as dificuldades encontradas pela escola no desenvolvimento de uma educação de qualidade.

Em pesquisa realizada por Souza (2009) sobre os efeitos da Prova Brasil nas escolas de Brasília, a autora afirma que há uma distância entre a avaliação externa e a apropriação dos resultados. Identificou-se que os professores têm acesso aos resultados, no entanto, não souberam indicar como utilizá-los. Outro fator de desconhecimento por parte dos docentes foi a construção do processo de avaliação. Com isso, fica evidente a falta de apropriação do resultado e do desenho da avaliação, um campo, afirma a autora, que deve ser estudado, repensado e reformulado pelos diferentes atores envolvidos. A pesquisa conclui que os resultados da avaliação externa são pouco aproveitados pela maioria dos professores no ambiente escolar.

A comunicação dos resultados, muitas vezes, é a expressão concreta de denúncia da baixa qualidade do ensino, vindo a ser um anúncio negativo do trabalho desenvolvido nas escolas brasileiras. Tal fato acaba por gerar uma resistência por parte dos atores

escolares na apropriação desses resultados, visto que os professores não se motivam a usar resultados que lhes chegam de forma coercitiva, sem um diálogo, apenas apontando erros e pressionando por mudanças. Com isso, a avaliação deve abrir espaço para diálogo, não para “descrições assertivas unilaterais e de julgamentos autoritários.” (Nevo, 1997, p. 94).

A ênfase na dificuldade de leitura dos resultados pelos atores escolares acaba por ser considerada um dos maiores empecilhos para a indução da melhoria da qualidade da educação com base na avaliação externa. No entanto, pouco se discute como se dá o processo de comunicação dos resultados. O que se vê é um discurso de incapacidade dos professores em ler esses dados ou uma resistência, fazendo com que os governantes acabem por fomentar mais formação de professores para determinado fim. Por isso, a formação de professores, nos últimos anos, vem direcionando também sua estrutura com base na avaliação externa. Nas linhas que se seguem, buscou-se delinear como o desejo pela utilização da avaliação externa redefiniu a formação dos professores.

Efeitos da avaliação externa.

As características dos testes empregados nas avaliações externas são, em grande medida, reproduzidos em muitos momentos em sala de aula, seja no preparo do aluno para a realização de tais exames, seja na utilização de questões dos testes, seja na ênfase em determinados conteúdos curriculares etc. Esses são exemplos de possíveis impactos dos exames sobre os acontecimentos em sala de aula, os processos educativos e seus participantes (Scaramucci, 2004), o que vem sendo identificado na literatura como efeito retroativo.

Para Scaramucci (2004), efeito retroativo e impacto são tratados como sinônimos, o que não ocorre na obra de Alderson e Wall (1993), que fazem uma diferenciação entre efeito retroativo, impacto e influência, na medida em que *washback* (termo utilizado pelos autores) diz respeito aos efeitos dos testes no processo de ensino-aprendizagem; impacto corresponde aos efeitos dos testes nas pessoas, nas escolas, nas políticas, no sistema educacional e na sociedade; e, por fim, as influências como sendo por atitudes de professores e alunos com relação aos testes. Os autores também destacam a predominância na literatura em referir-se aos efeitos negativos, porém eles exploram a potencial relação tanto dos efeitos negativos quanto dos positivos entre os testes e o processo de ensino e aprendizagem.

Alderson e Wall (1993, pp. 8-9) esquematizam um conjunto de hipóteses que ajuda a esclarecer um pouco o conceito. Um teste irá influenciar o ensino, a aprendizagem, como os professores ensinam, o que eles ensinam, o que os alunos aprendem, como os alunos aprendem, o ritmo e a sequência da aprendizagem, o ritmo e a sequência do ensino, o grau de profundidade da aprendizagem, o grau de profundidade do ensino, as atitudes com relação ao conteúdo, ao método de ensino e de aprendizagem etc. Testes que apresentam consequências importantes terão efeito retroativo e testes que não apresentam efeitos importantes não terão efeito retroativo. Testes terão efeitos retroativos em todos os alunos e professores, e testes terão efeitos retroativos somente em alguns alunos e alguns professores.

Para este trabalho, utilizaremos efeito como sinônimo de impacto, com base em Scaramucci (2004), não adotando, por conseguinte, a diferenciação trazida por Alderson e Wall (1993).

Em estudo realizado sobre as consequências dos testes de alto risco no sistema educacional, Stecher (2002) apresenta efeitos positivos, neutros/negativos (ver Figura 3), focando sua discussão nos efeitos dos testes sobre os alunos, os professores e os diretores. O autor destaca duas razões para se preocupar com as consequências dos testes em sala de aula. A primeira razão diz respeito aos defensores dos testes que esperam que seus resultados incentivem os professores a adotarem práticas mais eficazes e motivem os alunos a trabalharem mais, levando assim a uma mudança na prática educacional. Com isso, deve haver uma análise para verificar se esses testes atendem a tais objetivos. O segundo ponto estaria relacionado à melhora de pontuação sem haver melhora no desempenho, tendo em vista que esses testes podem provocar mudanças em comportamentos somente com o objetivo de impactar a pontuação, o que pode afetar a validade dos resultados do teste.

De acordo com Stecher (2002), muito dos efeitos são difíceis de medir, como verificar a influência da avaliação externa na elaboração de políticas públicas. Outros efeitos são passíveis de medição quantitativa, mas de difícil medição qualitativa, como verificar quantas horas foram gastas para o ensino de um conteúdo (quantitativo), analisar a qualidade desse ensino. Por isso, tal campo carece de estudos consistentes e de validade científica.

Efeito Positivo	Efeito Negativo
Efeito nos Estudantes	
Fornecer aos alunos mais informações sobre seus próprios conhecimentos e habilidades;	Frustra os alunos e os desencoraja de tentar;
Motiva os alunos a trabalhar mais na sua aprendizagem;	Torna os alunos mais competitivos;
Enviar sinais claros aos alunos sobre o que estudar;	Faz com que os alunos desvalorizem notas e avaliações escolares.
Ajudar os alunos a associar o esforço pessoal às recompensas.	
Efeito nos Professores	
Apoiar o melhor diagnóstico das necessidades individuais dos alunos;	Incentiva os professores a se concentrarem mais em conteúdo de teste específico do que em padrões curriculares;
Ajudar os professores a identificar áreas de força e fraqueza em seu currículo;	Conduz os professores a se envolverem em preparação de teste inadequado;
Ajudar os professores a identificar o conteúdo não dominado pelos alunos e a redirecionar as instruções;	Desvaloriza o senso de valor profissional dos professores;
Motivar os professores a trabalhar mais e de maneira mais inteligente;	Atrai professores a enganar ao preparar ou administrar testes.
Incentivar os professores a participar no desenvolvimento profissional para melhorar a instrução.	
Efeito nos Administradores	
Faz com que os administradores examinem as políticas da escola relacionadas ao currículo e à instrução;	Leva os administradores a promulgar políticas para aumentar os resultados dos testes, mas não necessariamente aumentar o aprendizado;
Ajudar os administradores a avaliar a qualidade de seus programas;	Faz com que os administradores realoquem recursos para os assuntos testados à custa de outros assuntos;
Leva os administradores a mudar as políticas da escola para melhorar o currículo ou a instrução;	Leva os administradores a desperdiçar recursos na preparação de testes;
Ajudar os administradores a tomar melhores decisões de alocação de recursos, por exemplo, fornecer desenvolvimento profissional.	Distrai os administradores de outras necessidades e problemas da escola.
Efeito nos Políticos Públicos	
Ajudar os formuladores de políticas a avaliar a eficácia das políticas educacionais;	Fornecer informações enganosas que levam os formuladores de políticas a tomar decisões pouco otimizadas;
Melhora a capacidade dos formuladores de políticas de monitorar o desempenho do sistema escolar;	Promove o espírito de "culpar as vítimas" entre os formuladores de políticas;
Promove melhor alocação de recursos educacionais do estado.	Incentiva uma visão simplista da educação e seus objetivos.

Fonte: Stecher, 2002, pp.86-87

Figura 3

Fatores que influenciam o processo de aprendizagem

Ao longo dos últimos anos, percebe-se um elevado aumento na quantidade de avaliações externas no contexto escolar. Cada vez mais, instituições governamentais e não governamentais vêm empreendendo esforços na busca por melhorias em seus sistemas de avaliação. No centro dessa política, “estão presentes questões ligadas à qualidade, às medidas em educação, à responsabilização por resultados e à prestação de contas” (Horta Neto, 2010, p. 85). Por isso, Alavarse (2013) afirma que essa discussão vem sendo permeada de recusas e entusiasmos, podendo afirmar que são argumentos que se colocam a favor e contra à avaliação externa, muito em decorrência de seus efeitos.

Dessa forma, faz-se necessária uma análise, com base em pesquisas empíricas brasileiras, sobre os efeitos positivos e negativos nos diferentes níveis, desde a sala de aula até as políticas públicas.

Efeito da avaliação externa na sala de aula.

Alguns estudos apontam que a avaliação externa tem pouco ou nenhum efeito sobre o que os professores fazem em sala de aula. Cimbricz (2002) considera que, apesar de os testes influenciarem a prática docente, essa influência depende de como os professores interpretam o teste e o utilizam para guiar as suas ações. Firestone, Mayrowetz e Fairman (1998) afirmam que os testes podem promover certas mudanças, como o alinhamento de assuntos ensinados no teste, no entanto, não apresenta sucesso na mudança nas estratégias de instruções básicas, ou seja, na prática docente, “a avaliação baseada no desempenho pode alterar comportamento e procedimentos específicos na sala de aula, mais facilmente do que paradigmas gerais para ensinar um assunto” (Firestone et al., 1998, p. 111, tradução nossa). Os autores afirmam que, para que houvesse mudança de prática, os professores deveriam ter a oportunidade de refletir sobre os modelos e resultados da avaliação externa e lançar-se em ações que pudessem aprofundar o entendimento sobre o que ensinam e como ensinam. Acabam, assim, por questionar se a avaliação externa estaria promovendo oportunidades de reflexão. Concluem que avaliação com responsabilidade forte pode encorajar alguns professores a mudar a sua prática, no entanto, seus efeitos motivacionais são limitados e não conduzem a tais movimentos em larga escala. Por conseguinte, consideram que os efeitos da avaliação externa sobre o ensino podem estar sendo superestimados.

Para que haja uma efetiva utilização da avaliação externa por parte do professor, de forma que venha a contribuir em sua prática, ele deve receber ajuda para refletir sobre suas próprias crenças, assim como as dos alunos, colegas, pais e gestores escolares (Shepard, 2000).

Por outro lado, há evidências de que a avaliação externa influencia profissionais a terem maior comprometimento devido à política de responsabilização de professores e escolas pelos resultados obtidos; impulsionam mudanças em currículos inadequados; ajuda a concentrar a instrução nas habilidades (por exemplo, habilidades básicas); fornece aos professores informações diagnósticas para melhorar instrução; identifica áreas curriculares que necessitam de melhorias (cf. Ahn & Vigdor, 2010; Atkinson et al., 2009;

Carnoy & Loeb, 2002; Koretz et al., 2001; Madaus et al., 2009; Mons, 2009; Stecher, 2002; Winters & Sundararaman, 2008).

Castro (2009b), em pesquisa realizada no programa de incentivo às escolas públicas estaduais de São Paulo, destaca que as políticas educacionais com base na avaliação externa tiveram êxito, tendo em vista que teve um bom resultado nas escolas de ensino médio, onde 84% das unidades se desenvolveram positivamente, sendo que 77% atingiram ou superaram as metas previstas.

Em contrapartida, há outra corrente que acredita que os efeitos da avaliação externa afetam os professores, suas práticas e seu processo de ensino de forma negativa – efeitos esses esperados ou não pelas autoridades que administram as avaliações. Há estudos que afirmam que esse tipo de avaliação pode gerar o afunilamento do currículo, induzindo os professores a se concentrarem apenas nos conteúdos e habilidades abordados na avaliação externa. Enfatizam as habilidades escolares, principalmente leitura e matemática, muitas vezes dando pouca atenção ao conhecimento e às habilidades que seriam úteis no cotidiano dos alunos. Com isso, há uma redução no currículo proposto, evidenciando apenas os conteúdos avaliados pela avaliação externa, seja por matrizes de referências locais ou nacionais. Assim a avaliação acaba por definir o que é ensinado e aprendido (cf. Au, 2007; Chamberlin, Wragg, Haynes & Wragg, 2002; Fletcher, 1995; Kellaghan & Madaus, 2003; Madaus & Russell, 2010; Madaus et al., 2009; Ravitch, 2009; Stecher, 2002; Shepard, 1990).

Au (2007) constatou que há uma relação significativa entre implementar uma avaliação de responsabilização forte e produzir mudanças no currículo. Isso acontece por haver um controle do conteúdo, no momento em que se ensina para o teste, havendo estreitamento do currículo; controle formal, quando ocorrem mudanças na forma como o conhecimento é abordado, aumentando a fragmentação do conhecimento, principalmente de forma a ensinar em relação direta com o teste; e, por fim, o controle pedagógico, no qual há a pedagogia em resposta ao teste, com aulas mais expositivas. O autor conclui que a avaliação com responsabilização forte estimula o alinhamento curricular à avaliação.

A partir da análise das iniciativas do governo federal, como a implantação do Saeb, o Exame Nacional do Ensino Médio (Enem) e o Exame Nacional de Cursos (ENC), Sousa (2003) observou o potencial dessas avaliações de condicionarem os currículos e, desse modo, intensificarem desigualdades escolares e sociais. Esteban e Fetzner (2015) abordam os reflexos da avaliação externa nacional brasileira na organização do trabalho

docente e nas práticas curriculares. Concluem que ações escolares em que predominam exercícios de treinamento para as provas e propostas curriculares com objetivos padronizados e fragmentados levam a um “aprisionamento curricular”.

Em pesquisa realizada com professores do estado do Rio de Janeiro, Cerdeira e Almeida (2012) observaram no discurso dos professores que a avaliação externa estadual causa efeitos perversos no currículo, tendo em vista que realiza ênfase nas disciplinas de Língua Portuguesa e Matemática, acarretando também práticas docentes que focalizam em conteúdos estabelecidos pelas provas, deixando de lado outros tipos de habilidade.

Menegão (2015), em pesquisa acerca do impacto da avaliação externa sobre o currículo escolar afirma que esta tem contribuído para apagar aspectos curriculares, dimensões, disciplinas e conteúdos, que não podem ser medidos. A autora conclui que a utilização da matriz de referência como um currículo escolar é tida como um consenso, um artifício do qual nem professores, nem gestores e nem mesmo alunos escapam.

Outro efeito levantado pelos estudos é o impacto na metodologia utilizada pelos professores em sala de aula, ou seja, os professores restringem a utilização de métodos com o intuito de preparar os alunos para o exame. Tal fato leva a uma simplificação do trabalho do professor, reduzindo sua capacidade de criatividade, reflexão, ensino centrado no aluno (cf. Alderson & Wall, 1993; Darling-Hammond, 1994; Herman & Golan, 1993; Kellaghan & Madaus, 2003; Looney, 2009; Madaus & Russell, 2010; Smith, 1991; Stecher, 2002;). Cabe destacar que há evidências de que a didática do professor seria pouco influenciada pela avaliação externa – a forma como ele ensina e sua relação com o aluno –, mas o tipo de atividades e a ênfase nos conteúdos trabalhados na avaliação externa sofreriam maior influência, dando mais ênfase a simulados, conteúdos abordados nos testes etc. (Firestone et al., 1998; Luxia, 2007).

Sousa e Arcas (2010) constatam a preocupação dos professores no estado de São Paulo em trabalhar com ênfase nas habilidades, competências e conteúdos das matrizes de referência da avaliação estadual, com vistas à preparação dos alunos para que, assim, possam ter um bom desempenho, o que evidencia a matriz de referência da avaliação externa como um currículo para os professores.

Sobre o sentimento de professores e alunos com relação à avaliação externa, pesquisas apontam que esta induz sentimentos negativos nos professores, de vergonha, constrangimento, culpa, ansiedade, acusação e raiva. Os professores das disciplinas que são avaliadas acabam por ser acusados pelos professores de outras disciplinas, principalmente em escolas com baixo rendimento; pedem para mudar de escola; sofrem

com a evidenciação de escolas por meio da publicização de *ranking* (cf. Herman & Golan, 1993; Madaus & Russell, 2010; Madaus et al., 2009; Mons, 2009; Smith, 1991). Ball (2003), ao se referir à performatividade de professores em ambientes em que se trabalha em resposta a metas, indicadores e avaliações, afirma que há uma esquizofrenia estrutural e individual de valores e propósitos, e uma inautenticidade e falta de sentido.

Com relação aos estudantes, a avaliação externa pode gerar ansiedade, preocupação e medo, podendo acarretar um sentimento negativo com relação à escola e afetar o processo de aprendizagem. Como também alienação ao exame ou ações de rebelião (Kellaghan & Madaus, 2003; Madaus et al., 2009; Mons, 2009).

Em pesquisa realizada sobre gerencialismo e responsabilização em escolas do município de Campinas – SP, Rodrigues (2018) afirma que as políticas de mérito, tanto o bônus quanto a prova, principalmente em escolas que não oferecem condições de trabalho, promovem nos professores desânimo, frustração e desestímulo com a carreira. Com o objetivo de realizar uma análise de como a avaliação externa se relaciona com a organização no ambiente escolar, Polato (2014) afirma que políticas voltadas para a melhoria de índices das avaliações fazem com que escolas apliquem ações que geram desmotivação nos alunos e não garantem a aprendizagem, vindo a promover abandono escolar.

Por fim, verifica-se que ocorre uma desvalorização dos julgamentos dos professores sobre o desenvolvimento de seus alunos. A desconfiança nos julgamentos qualitativos dos professores tem sido um valor oculto da padronização de resultado da avaliação externa (Madaus & Russell, 2010; Mons, 2009), o que acaba por evidenciar a avaliação externa como única forma confiável de avaliar o desempenho do aluno. Isso cria uma tensão entre os dois níveis de avaliação (interno e externo), pois sugere uma hierarquia entre eles, tendo em vista que a avaliação externa passa a ser considerada de melhor qualidade. Há, assim, um desafio para encontrar equilíbrio entre as duas avaliações (Depresbiteris, 2001).

Esse quadro é ratificado por pesquisas que concluem a falta de formação adequada dos professores para o desenvolvimento da avaliação interna, tanto uma formação inicial quanto a continuada (cf. Alavarse, 2013; Rojas, 2007; Villas-Boas, Freitas & Soares, 2016). No entanto, Nevo (1997) afirma que a avaliação interna é uma pré-condição para que a avaliação externa venha a ter utilidade, devendo haver um diálogo entre as duas.

Ao analisar as avaliações da aprendizagem realizadas nas salas de aula francesas, Bercier-Larivière e Forgette-Giroux (1999) apresentam uma crítica à ausência de um

modelo de qualidade dos resultados das avaliações internas. Evidenciam que o conceito de precisão deve ser considerado nesse tipo de avaliação, como um critério de qualidade para os resultados, dada a relevância da tarefa de avaliação em relação às habilidades visadas, sua consistência com a atividade pedagógica, a necessidade de transparência para com o aluno e a ausência de qualquer desvantagem circunstancial. Os autores destacam que os professores, muitas vezes, levam em consideração elementos que foram ignorados durante o processo de aprendizagem ou não têm nada a ver com a habilidade avaliada, como, até mesmo, comportamento e esforço.

Em estudo realizado com base na Prova São Paulo, Alavarse et al. (2017) pontuam a incongruência entre o desempenho na avaliação interna e externa dos alunos, ou seja, alunos que foram reprovados na escola, ficaram com o conceito Adequado na avaliação externa e vários alunos que foram aprovados não passaram do nível Abaixo do Básico na Prova São Paulo. Com base nesses dados, os pesquisadores discutiram com os professores sobre os instrumentos e critérios utilizados para a avaliação interna. O que se evidenciou no estudo foi que, principalmente no caso dos alunos em que se mostrou discrepância no desempenho, lançavam mão de critérios alheios ao que se estaria avaliando, como disciplina, participação e realização de tarefas, critérios esses que não deveriam ser objetos de avaliação utilizados, para fins de processos de aprendizagem e do próprio fluxo escolar.

Efeito da avaliação externa no sistema educacional e nas políticas públicas.

Os efeitos da avaliação externa do desempenho do aluno podem afetar o sistema educacional. Alguns fatores positivos giram em torno da melhoria da qualidade do sistema educacional e a redução das desigualdades educacionais.

A partir de 1990, a avaliação ganha papel central como elemento para a promoção da qualidade da educação no Brasil. Desde a educação básica até o ensino superior, o discurso de a avaliação poder produzir um ensino de melhor qualidade vem ganhando espaço nas agendas políticas. As possíveis contribuições da avaliação para a melhoria do ensino estão fundamentadas em dois argumentos: o primeiro diz respeito à lógica interna dos processos avaliativos, vindo a ser referência para a discussão das potencialidades e dos limites das práticas adotadas; o segundo leva em consideração a utilização dos resultados na condução das políticas educacionais (Alkin & Taut, 2003; Patton, 1997, 2003; Silva et al., 2013; Sousa & Oliveira, 2010).

Como um efeito positivo, pode-se considerar na redução das desigualdades educacionais, que é pautada na não diferenciação entre os estudantes, evitando, por conseguinte, o favorecimento de estratos sociais mais altos. A objetividade e a imparcialidade da avaliação obrigam os professores a respeitar o currículo nas escolas com uma população desfavorecida, levando ao melhoramento dos resultados dos estudantes, diminuindo também a proporção de alunos evadidos (Eberts, Hollenbeck & Stone, 2000; Mons, 2009).

No entanto, a literatura também apresenta argumentos contrários, alguns autores afirmam que a avaliação externa pode induzir o aumento da desigualdade, tendo em vista que se investe mais nos melhores alunos, levando a uma discriminação dos considerados menos capazes (Haney & Madaus, 1986; Stecher, 2002). Dados de pesquisa também evidenciam que não há efeitos positivos das políticas de testes de responsabilidade forte no desempenho do aluno (Lee, 2008).

Os atores do sistema educacional passam a realizar ações com base na pressão para a melhoria de resultados, realizando uma pressão verticalizada, através de ações como formação de professores, gerenciamento de tempo de conteúdos a serem abordados, alteração na forma de ingresso de alunos na unidade escolar etc. (Freitas, 2007; Haney & Madaus, 1986; Madaus et al., 2009; Stecher, 2002). Para além disso, promovem injustiças com relação à bonificação de professores e premiação das escolas, desencadeando, assim, um sentimento negativo nos professores com relação às avaliações e ao sistema educacional.

Com base em pesquisa realizada no município do Rio de Janeiro, Carvalho, Oliveira e Lima (2014) salientam a mudança da gestão escolar ao longo do tempo em resposta à pressão por melhores resultados nas avaliações externas. Rosistolato e Prado (2014) também identificaram mudanças no contexto escolar trazidas pelas avaliações externas, principalmente relacionadas à distribuição do conhecimento escolar e à orientação das atividades escolares para as provas, nos municípios do Rio de Janeiro e de Duque de Caxias.

Em pesquisa realizada com professores do estado de São Paulo, Novaes (2014) afirma a necessidade de maior atenção para a relação existente entre a avaliação externa, a política de responsabilização e a política de bonificação. Os professores consideram injusta a hierarquização e a classificação de escolas, mediante a adoção de *rankings* e a premiação/castigo de professores com base na avaliação externa.

Como afirma Cerdeira e Almeida (2012), em pesquisa realizada sobre a avaliação Nova Escola do município do Rio de Janeiro, mesmo que visem à melhoria da qualidade da educação, as políticas de responsabilização não vêm tendo boas repercussões nas escolas brasileiras. O programa recebeu muitas críticas, por parte dos professores, representados pelo Sindicato Estadual dos Profissionais da Educação do Rio de Janeiro (Sepe/RJ). Em agosto de 2005, o sindicato moveu três ações contra o governo do estado do Rio do de Janeiro com relação ao programa, todos devido a alguma discordância em relação à gratificação dos professores.

De acordo com Brooke (2013b), a realização de pesquisas sobre a política de incentivo monetário no contexto brasileiro ainda é muito incipiente, falta metapesquisa com fonte confiável, critérios rigorosos, para que assim possa chegar a uma conclusão sobre os efeitos dos intensivos monetários.

No nível macro, o das políticas públicas, deve-se realizar uma análise de quais são os efeitos da avaliação, no entanto, são poucos os estudos voltados para esse nível (Stecher, 2002). Um dos efeitos positivos é a construção de uma cultura de avaliação dos serviços prestados pelos órgãos públicos e a transparência sobre os seus processos e resultados, vindo também a promover maior planejamento para alocação de recursos do estado (Afonso, 2012; Bauer et al., 2015; Stecher, 2002).

Entretanto, movimentos como pressão verticalizada, visão simplista da realidade educacional – tendo em vista que desconsidera outras dimensões para além dos resultados obtidos na avaliação externa – e o trabalho com base em informações enganosas, forjadas para obter resultados positivos para divulgar na mídia e promover governos, são alguns efeitos negativos provenientes da avaliação externa nas políticas públicas (Freitas, 2007; Stecher, 2002).

Ao considerar os resultados da avaliação externa como indicadores de qualidade da educação, ações políticas e práticas pedagógicas, deve-se ter claro como esse indicador está sendo utilizado e quais são os efeitos provenientes de seu uso, sejam eles positivos, negativos ou nulos. Esses efeitos podem ser esperados ou não, podendo impactar o sistema como um todo ou cada um dos intervenientes que compõem os diferentes níveis.

Os efeitos têm grande referência na literatura, principalmente no que diz respeito aos efeitos negativos na sala de aula. Contudo, sua categorização não é um consenso, visto que são considerados de formas diferentes, isto é, alguns efeitos podem ser avaliados como positivos para uns e não necessariamente positivos para outros. Assim, fica evidente que a avaliação apresenta diferentes abordagens, as quais podem estar embasadas em

diferentes paradigmas, que apresentam pressupostos políticos, sociológicos, filosóficos e psicológicos distintos.

O objetivo deste capítulo foi lançar um pouco mais de luz sobre o objeto de estudo, buscando diferentes posicionamentos, embasando-os em pesquisas empíricas, com o intuito de realizar uma ligação entre a teoria e o “chão” da escola. Assim, foi realizada uma análise da avaliação externa a partir de sua construção histórica e teórica, buscando embasar as diferentes discussões e posicionamentos sobre o tema. Cabe destacar que esse esforço é indispensável para qualquer investigador e avaliador que deseje descrever, analisar e interpretar rigorosamente uma dada realidade (Fernandes, 2013). Muito se fala da realidade, pouco se conjuga essa realidade com o instrumento que se usa para avaliá-la. Esse caminho é perigoso, porque minimiza a descrição, a análise e a interpretação da realidade, deixando-a de certa forma opaca, pois não se conhece o instrumento, suas dimensões, características, alcances e limitações.

Não foi objetivo deste capítulo defender um lado ou uma verdade sobre a avaliação externa. Por isso, buscou-se uma diversidade de pontos de vista, de pesquisas e teorias, para poder desenhar o quadro atual dessa avaliação, que longe está de ser um processo acabado, mas sim um campo que, como referido ao longo deste capítulo, carece de muita pesquisa e reflexão.

Capítulo III

Metodologia

Na literatura, há diversas definições sobre o que é uma pesquisa. Gil (2008) afirma que é um processo racional e sistemático através do qual se busca proporcionar respostas para problemas propostos, com base em procedimentos científicos, sendo desenvolvida a partir de utilização cuidadosa de métodos, técnicas e outros procedimentos. No entanto, Goldenberg (2011) salienta a necessidade de não reduzir a pesquisa a procedimentos metodológicos, mas sim desenvolver um “olhar científico”. Com isso, este capítulo busca discorrer sobre o enquadramento da pesquisa, dando ao leitor subsídios para a compreensão e entendimento do trabalho através da discussão das respectivas opções. Considerou-se particularmente importante evidenciar os propósitos da pesquisa e como se teve em conta a ideia do “olhar científico”, na medida em que se problematizaram as escolhas consideradas mais assertivas em detrimento de outras, evidenciando, assim, os pressupostos metodológicos.

Delineamento da Pesquisa

Quão fortes são as palavras, que poder há em formar e transformar. Cecília Meireles, de maneira brilhante, traduz em poesia o poder inquestionável do dizer, que pode destruir “Reis, impérios, povos, tempos”, como também, proporcionar a “liberdade das almas”. Portanto, cabe-nos refletir sobre elas ou apenas sentenciar “Ai, palavras”³.

É inegável o poder da palavra na escola, seja ela proferida dentro de seus muros ou fora deles. Não se fala aqui da palavra apenas como contributo gráfico/sonoro na comunicação humana, mas sim no que tange à sua textualidade, isto é, “de um discurso que a sustenta, que a provê de realidade significativa” (Orlandi, 1995, p. 111).

Com base na “estranha potência das palavras”, o problema que motivou esta investigação reside no fato de se considerar relevante compreender de forma tão profunda quanto possível “as palavras” de diferentes intervenientes do campo da educação sobre a avaliação externa da aprendizagem dos alunos. Dito de outra forma, compreender a avaliação externa da aprendizagem dos alunos a partir da apreciação das palavras de protagonistas relevantes do campo da educação: professores da educação básica,

³ Meireles, Cecília (1967). Ai, palavras. In: *Obra Poética* (2ª ed.). Rio de Janeiro: José Aguilar Editora, pp. 560-561.

decisores políticos e especialistas (e.g., professores universitários, pesquisadores). A escolha destes intervenientes se deu tendo em vista que as referidas avaliações externas encontram-se na discussão de muitas pautas políticas e podem ocupar um papel importante na avaliação de cada aluno (Kellaghan & Madaus, 2003). Dessa forma, consideramos que as palavras proferidas tornam-se um instrumento revelador de “condições estruturais, de sistemas de valores, normas e símbolos (sendo ela mesma um deles), e, ao mesmo tempo, possui a magia de transmitir . . . representações de grupos determinados em condições históricas, socioeconômicas e culturais específicas.” (Minayo & Sanches, 1993, p. 245).

Assim como caracterizam Bonamino e Souza (2012), a avaliação externa da aprendizagem dos alunos começou a ganhar força na década de 1990 no Brasil, com a implantação do Sistema de Avaliação da Educação Básica (Saeb). O objetivo era realizar o diagnóstico e o monitoramento da qualidade da educação básica, sendo considerada como uma avaliação de primeira geração. Em 2005 surge a Prova Brasil que apresenta, juntamente com o Saeb, uma política de responsabilização branda (*low stakes*), pois gera uma consequência simbólica, uma vez que se limita a traçar metas e a divulgar resultados dos alunos por escola e rede de ensino, sem fazer relação com premiação ou sanções a esses resultados. Compõe, assim, uma avaliação de segunda geração, que “tem como pressuposto que o conhecimento dos resultados favorece a mobilização das equipes escolares para a melhoria da educação, bem como a pressão dos pais e da comunidade sobre a escola” (Bonamino & Sousa, 2012, p. 375). Em outro movimento, surge a terceira geração de avaliação, muito em resposta às avaliações de âmbito nacional e à criação do Índice de Desenvolvimento da Educação Básica⁴ (Ideb), surgem os sistemas próprios de avaliação estaduais e municipais. Em sua maioria, apresentam responsabilização forte (*high stakes*), considerando que têm impactos diretos nos profissionais de educação através da utilização de benefícios salariais e premiações.

Ao longo dos anos, observou-se um grande aumento das avaliações geridas por estados e municípios. Em 2015, havia 18 estados com sistemas próprios de avaliação, tendo mais 2 estados com planos de implantação em 2016, totalizando, por conseguinte, 20 estados das 27 unidades federativas (CAEd, 2015). Isso indica que o objetivo de implantar uma determinada política de avaliação, presente desde o primeiro ciclo do Saeb,

⁴ O indicador é calculado a partir dos dados sobre aprovação escolar, obtidos no Censo Escolar, e das médias de desempenho nas avaliações do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), o Sistema de Avaliação da Educação Básica (Saeb) e a Prova Brasil.

foi alcançado. Com relação a esse aumento no número de avaliações externas de sistemas próprios de avaliação, a academia apresenta opiniões diversas, a saber: uma postura mais crítica, entendendo que tal movimento decorre de uma política de mercado (Freitas, 2007); a ideia de que esse fato provém de uma pressão política, tendo em vista o baixo desempenho educacional (Brooke, 2013a); por fim, a necessidade de um diagnóstico mais preciso, com o objetivo de promover a melhoria da qualidade educacional (Silva et al., 2013).

Outro ponto de importante relevância é a variedade de características, objetivos, utilização e divulgação de resultados nas avaliações externas da aprendizagem dos alunos no Brasil. Essa variedade pode ocorrer devido a diferentes fatores, como os aprimoramentos ocorridos ao longo dos anos; uma política de autonomia dos entes federados, ficando a critério de cada estado e municípios a criação e a administração de um sistema próprio de avaliação; as constantes discontinuidades de políticas públicas, levando a mudanças frequentes nos desenhos das avaliações externas.

Assim, toma-se importante realizar uma análise das avaliações externas a partir da apreciação de *stakeholders*, visto que são sujeitos ou do contexto escolar ou de estruturas que se relacionam de forma direta com a educação. A investigação considera como participantes privilegiados para procederem a uma análise das avaliações externas os professores da educação básica, uma vez que são eles que estão no espaço escolar atuando de forma direta com os alunos e com todos os processos inerentes ao desenvolvimento das avaliações externas nas escolas; os decisores políticos, que tomam as decisões e constroem os discursos oficiais que são utilizados para nortear e influenciar o processo de ensino-aprendizagem; e os especialistas, tais como os professores universitários e os pesquisadores, que estudam e investigam os assuntos da educação.

Convém observar que não foi encontrado nenhum estudo que buscasse compreender a avaliação externa a partir das percepções de diferentes intervenientes do processo educacional. O que se encontrou foi uma análise a partir da percepção de professores, porém também não há muitos estudos com esse escopo. A maior parte das pesquisas realizadas nos últimos anos nessa área tem como foco a análise da avaliação externa no âmbito nacional, principalmente no que se refere ao Saeb, ao Ideb e à Prova Brasil, com poucos movimentos no que se refere aos níveis estaduais e municipais. Os temas mais abordados encontram-se no estudo das relações entre a avaliação externa e a gestão escolar, o currículo, a qualidade da educação e a apropriação de resultados.

Nessas condições, conclui-se que não se configura ainda uma pesquisa com o escopo que se pretende alcançar com esta investigação e é nessa lacuna que este trabalho foi desenvolvido. Nesse sentido, parece ser necessário conhecer mais profundamente as perspectivas de uma variedade de intervenientes acerca de um conjunto de aspectos relacionados com a avaliação externa.

Com isso, esta pesquisa buscou responder às seguintes questões, todas tendo em conta as perspectivas veiculadas pelos principais intervenientes do processo educativo – professores da educação básica, decisores políticos e especialistas:

1. Como se caracterizam e qual a natureza das avaliações externas da aprendizagem dos alunos?
2. Que relações se podem estabelecer entre as políticas públicas de educação e as avaliações externas da aprendizagem?
3. Que relações se podem estabelecer entre as avaliações externas da aprendizagem e a qualidade do sistema de ensino?
4. Em que medida as avaliações externas são relevantes para melhorar a qualidade da aprendizagem dos alunos e a qualidade dos processos de ensino?

A problemática aqui apresentada reflete-se num conjunto de objetos e dimensões de pesquisa que se organizaram na matriz de investigação que se apresenta na Figura 4. Cabe salientar que coaduna com a ideia de Fernandes, Rodrigues e Nunes. (2012). Mesmo entendendo que a estruturação de uma matriz é de certa forma arbitrária, considera-se ser um instrumento que contribui para garantir a focagem da pesquisa nas questões essenciais e para ajudar a manter bons níveis de consistência na recolha e na análise dos dados.

Para atender às questões de investigação, pensou-se em dois objetos. O primeiro é a própria avaliação externa da aprendizagem, com o intuito de caracterizá-la. Para tal, as dimensões desse objeto buscam lançar luz sobre pontos importantes da avaliação externa. A separação em diferentes dimensões contribui para maior clareza da análise do objeto, no entanto, cabe ressaltar que as dimensões dialogam entre si.

Para o objeto de avaliação externa da aprendizagem, foram consideradas relevantes as seguintes dimensões propostas para esta pesquisa: a) natureza e propósitos; b) relevância e importância; c) credibilidade; d) utilidade; e) utilização; e f) satisfação dos diferentes intervenientes.

O segundo objeto versa sobre a qualidade da educação, buscando entender como ocorre a relação dele com a avaliação externa. Para tal, foram escolhidas dimensões como: a) relações das avaliações externas com os processos de ensino; b) relações das avaliações externas com os processos de aprendizagem; c) relações das avaliações externas com as práticas de ensino; d) relações das avaliações externas com as práticas de avaliação; e e) relações entre os resultados das avaliações externas e a qualidade da educação.

Objetos	Dimensões
Avaliação Externa da Aprendizagem	Natureza e propósitos
	Relevância e importância
	Credibilidade
	Utilidade
	Utilização
	Satisfação dos diferentes intervenientes
Qualidade da Educação	Relações das avaliações externas com os processos de ensino
	Relações das avaliações externas com os processos de aprendizagem
	Relações das avaliações externas com a prática de ensino
	Relações das avaliações externas com as práticas de avaliação
	Relações entre os resultados das avaliações externas e a qualidade da educação

Figura 4

Matriz de investigação

Entende-se que a análise do objeto de pesquisa é realizada a partir de um ponto de vista, fato que está estritamente ligado à subjetividade do pesquisador e de suas escolhas teóricas. Assim como elucida Löwy (2013, p. 56), “como descobrir os elementos objetivamente essenciais que permitem compreender a estrutura social se a escolha entre o essencial e o acessório é, como o demonstrou de forma marcante Weber, inevitavelmente determinada por nossos valores subjetivos?”.

Por isso, é importante considerar que o processo de construção do objeto de investigação desta pesquisa deu-se no momento em que passei a trabalhar para uma agência de avaliação, com elaboração e revisão de itens das avaliações externas para

vários estados e municípios brasileiros; e também ministrando oficinas de apropriação de resultados para professores e gestores da rede pública de ensino.

Nesta experiência, principalmente nas oficinas, tive contato com as falas dos professores com relação à avaliação externa. Primeiramente, ficava clara a falta de conhecimento dos atores escolares sobre essas avaliações, desde a sua elaboração, a forma de correção pela Teoria da Resposta ao Item (TRI)⁵, até a leitura das escalas de proficiência e dos resultados. Os docentes destacavam que a falta de conhecimento gerava a não utilização do resultado que voltava para a escola. É importante esclarecer que essa oficina de apropriação é apresentada em poucos municípios brasileiros e não é uma ação constante. Com isso, seu potencial de alcance ainda parece ínfimo com relação ao número de estados e municípios que possuem sistema próprio de avaliação. Outro ponto que cabe evidenciar é com relação à resistência que os professores apresentavam no que concerne às avaliações externas, que era claramente perceptível nas concepções trazidas em suas falas.

Desse modo, a partir do trânsito entre diferentes instâncias, com diferentes intervenientes deste processo, surgiu a inquietação em investigar mais as avaliações externas levando em consideração esses sujeitos.

Por fim, é oportuno frisar que o projeto que deu origem a esta investigação foi submetido à Comissão de Ética do Instituto de Educação da Universidade de Lisboa tendo sido objeto de parecer favorável (Anexo 2).

Desenho da Pesquisa

Uma das primeiras questões com que um pesquisador se defronta é a que se refere à escolha da abordagem de investigação que irá utilizar para que, dessa forma, possa construir um fio condutor lógico e consistente ao longo do estudo. Assim, é a partir do processo da pesquisa que se qualificam as técnicas e os procedimentos necessários para responder ao que foi proposto (Goldenberg, 2011). Não há um método melhor ou pior, por isso deve-se procurar uma melhor adequação entre método, objetivo e condições em que a pesquisa será desenvolvida. Um bom método é aquele que permite uma construção correta dos dados, ajudando a refletir sobre a dinâmica da teoria, devendo ser operacionalmente exequível (Minayo & Sanches, 1993).

⁵ A Teoria da Resposta ao Item (TRI) considera um item uma unidade básica de análise, sendo qualificado a partir de três parâmetros: o poder de discriminação, o grau de dificuldade e a possibilidade de acerto ao acaso.

Com isso, para a consecução dos objetivos desta investigação, pareceu adequado utilizar uma metodologia de investigação por método misto. Neste trabalho, adotei a definição proposta por Johnson e Onwuebuze (2004), que compreendem métodos mistos como “pesquisa em que o pesquisador mistura ou combina técnicas de pesquisa quantitativa e qualitativa, métodos, abordagens, conceitos ou linguagem em um único estudo” (p.17, tradução nossa).

Como o objetivo do trabalho é aprofundar o conhecimento da avaliação externa da aprendizagem a partir da apreciação de diferentes intervenientes do processo educativo, optou-se por utilizar a triangulação de dados, tendo em vista que busca compreender e explicar um determinado objeto em suas múltiplas dimensões, ou seja, olhar para o mesmo fenômeno a partir de mais de uma fonte de dados, tendo em vista que informações advindas de diferentes ângulos contribuem no que diz respeito à corroboração, à elaboração ou à iluminação do problema de pesquisa (e.g., Denzin, 1978; Denzin & Lincoln, 2006; Minayo, Assis & Souza, 2005). Sendo assim, a abordagem metodológica foi escolhida com o intuito de contribuir para o aumento do conhecimento sobre o assunto, ou seja, possibilitar uma compreensão em profundidade do fenômeno estudado.

De acordo com Creswell (2010), há algumas estratégias para a realização da coleta de dados. Na perspectiva sequencial, há uma diferença temporal na coleta de dados quantitativos e qualitativos, sua vantagem estaria na capacidade de resolver questões específicas que surgem no processo de coleta de dados. Na abordagem concorrente ou concomitante, os dados são coletados de forma simultânea, podendo ocorrer devido à impraticabilidade de projetos sequenciais ou quando é irrelevante sequenciar dados. Para a condução desta pesquisa, tendo em vista que as dimensões para análise foram definidas antes de se iniciar a coleta de dados, a partir da literatura da área e de análise documental, considerou-se que a melhor abordagem seria a concomitante, pois seria irrelevante adotar a abordagem sequencial. Dessa forma, foi utilizada a estratégia aninhada concomitante, pois os dados qualitativos e quantitativos foram coletados de forma simultânea. Além disso, o método qualitativo teve predominância sobre o método quantitativo, visto que o objetivo geral do estudo é um aprofundamento da compreensão de uma determinada realidade, como afirma Minayo (1994, p.23), através do trabalho com “o universo dos significados, dos motivos, das aspirações, das crenças, dos valores e das atitudes” dos participantes. A utilização de dados quantitativos no estudo justificou-se para proporcionar uma análise mais alargada das respostas dos professores da educação básica,

tendo em vista a dimensão dessa população. Dessa forma, os dados de natureza quantitativa complementaram e iluminaram os dados de natureza qualitativa e as inferências que a partir deles se puderam realizar. Num certo sentido, assumindo os dados de natureza qualitativa uma clara predominância nesta pesquisa, pode dizer-se que, tal como sugerido por Creswell (2010), a sua parte quantitativa acabou por ser embutida na parte qualitativa, conforme se ilustra na Figura 5.

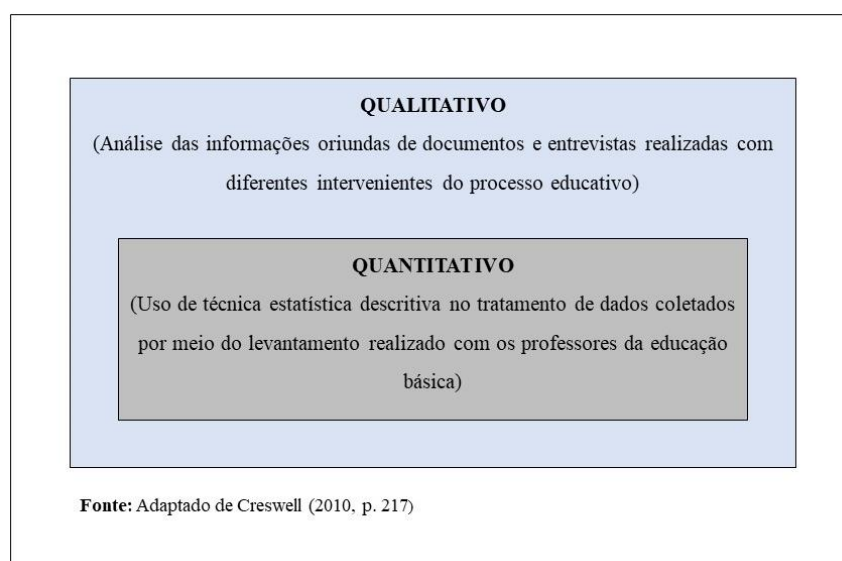


Figura 5

Desenho da estratégia aninhada concomitante

Tendo em conta o exposto, o presente estudo é de natureza interpretativa ainda que tanto a descrição quanto a análise sejam considerados processos fundamentais para a transformação e a organização dos dados que, como se afirmou, são majoritariamente de natureza qualitativa. Nesse contexto, é dada particular atenção à interpretação dos significados que os participantes atribuem aos fenômenos sob investigação que, naturalmente, envolvem questões de poder, ideologia, história e subjetividade (Merriam, 1998; Moita Lopes, 1994; Schwandt, 2006). Será com base nessa perspectiva, fortemente baseada nas percepções dos participantes, que se procurará aprofundar o conhecimento acerca das avaliações externas.

Descrição dos Procedimentos Adotados

Para entender melhor o contexto da pesquisa, optou-se, em um primeiro momento, por realizar descrição e análise de diferentes sistemas de avaliação, desde o âmbito internacional até o local, utilizados no Brasil. Para tal, foram realizadas buscas em documentos oficiais produzidos pelas instituições organizadoras, notícias em veículos midiáticos dos governos (federal, municipais e estaduais) e artigos acadêmicos sobre esses sistemas de avaliação.

Houve uma grande dificuldade em coletar dados do sistema próprio de avaliação municipal do Rio de Janeiro nos documentos oficiais. Por isso, foram consideradas informações fornecidas pelo responsável técnico da Coordenação de Supervisão e Avaliação da Secretaria Municipal de Educação. A consulta foi feita por meio de correio eletrônico.

Em um segundo momento, seguiu-se para a coleta de dados com os diferentes participantes da pesquisa. Para isso, os instrumentos de coleta de dados, o guião de entrevista e o questionário foram construídos com base na matriz de investigação. A fundamentação teórica para a construção das questões do guião da entrevista e do questionário se deu a partir de uma ampla revisão bibliográfica de publicações nacionais e internacionais, e análise documental (Alavarse, 2013; Bauer et al., 2015; Fernandes, 2007, 2010a, 2013; Freitas, 2007; Kellaghan & Madaus, 2003; Madaus et al., 2009; Madaus & Stufflebeam, 2000; Mons, 2009; Nevo, 1997; Patton, 2005; Stecher, 2002;).

Na elaboração dos instrumentos, foram considerados critérios como clareza, objetividade, relevância, precisão e credibilidade.

Para atender ao propósito da pesquisa, realizou-se entrevista semiestruturada, seguindo um guião previamente estabelecido (Anexo 3), buscando aprofundar os dados relacionados à percepção dos participantes. De acordo com Duarte (2004), as entrevistas

são fundamentais quando se precisa/deseja mapear práticas, crenças, valores e sistemas classificatórios de universos sociais específicos, mais ou menos bem delimitados, em que os conflitos e contradições não estejam claramente explicitados. Nesse caso, se forem bem realizadas, elas permitirão ao pesquisador fazer uma espécie de mergulho em profundidade, coletando indícios dos modos como cada um daqueles sujeitos percebe e significa sua realidade e levantando informações consistentes que lhe permitam descrever e compreender a lógica que

preside as relações que se estabelecem no interior daquele grupo, o que, em geral, é mais difícil obter com outros instrumentos de coleta de dados. (p. 215)

A entrevista semiestruturada possibilitou seguir um roteiro pré-estabelecido, no qual as perguntas fossem semelhantes em todas as entrevistas, para assim realizar uma compreensão ampla dos pontos estabelecidos na matriz de investigação e também possibilitar uma posterior comparação tanto dentro do mesmo grupo de participantes quanto entre os grupos. No entanto, outras perguntas e comentários surgiram ao longo das entrevistas, devido ao caráter de interação que as permeia, deixando-as mais ricas.

Decidiu-se pela utilização do questionário, pois é um instrumento reconhecidamente útil para recolher informações referentes a opiniões, percepções ou reflexões dos participantes e pela sua facilidade em trabalhar com um elevado número de inquiridos em um reduzido período de tempo. O questionário incluiu itens de resposta fechada porque assim propicia maior uniformidade na análise das respostas e permite uma melhor contextualização da questão (Anexo 4). Foi utilizada uma escala do tipo Likert, com quatro escolhas possíveis: concordo totalmente, concordo, discordo totalmente e discordo. Decidiu-se por não colocar uma proposição neutra, entendendo que seria difícil defini-la claramente (Komorita, 1963) e, além do mais, sendo os participantes conhecedores das matérias em causa, é dispensável o chamado ponto neutro (Coelho & Esteves, 2007; Johns, 2005).

A qualidade de um instrumento, segundo Pilatti, Pedrosa e Gutierrez (2010), passa pela sua capacidade para produzir resultados consistentes e válidos. Com o intuito de garantir a qualidade dos instrumentos, submeteram-se a um painel de quatro especialistas, professores universitários com pesquisas e publicações no domínio da avaliação. Foram escolhidos três pesquisadores portugueses e um brasileiro, que apreciaram a validade de conteúdo, a abrangência dos instrumentos, a formulação de cada item e a relação das questões com o objeto de pesquisa. Seguidamente, foi feita uma administração piloto do questionário junto de 30 participantes, para identificação de possíveis problemas na aplicação. A triangulação foi utilizada para garantir o rigor, a validade interna, a confiabilidade e a consistência dos resultados da pesquisa (Cohen, Manion & Morrison, 2000).

Participantes da Pesquisa

Ao realizar uma escolha, deve-se ficar claro que se abandonam outras opções disponíveis. Com isso, é relevante destacar que, ao escolher os participantes para esta investigação, foram deixados de lado outros participantes que também estão envolvidos de alguma forma no processo de avaliação externa, tais como os alunos, os responsáveis pelos alunos e os gestores escolares. A escolha pelos três grupos de participantes, contudo, não foi tomada de forma aleatória; deu-se em razão de que, além de serem grupos com forte poder de influência ou decisão sobre o objeto analisado, também se alinham no sentido de exercerem possível influência mútua de ações, isto é, os especialistas geram conhecimento que pode vir a embasar as ações dos decisores políticos que, por sua vez, geram ações que influenciam a prática dos professores da educação básica. Entende-se que esse movimento não se dá apenas nesse sentido, mas também pode ocorrer de forma inversa. No entanto, o ponto fundamental é o efeito de alinhamento entre os grupos escolhidos.

Nesta pesquisa, entende-se como especialistas aqueles que possuem uma visão teórica importante para a construção do objeto de pesquisa, podendo ser eles professores universitários ou pesquisadores da área de educação. Para além disso, muitas vezes, sua produção serve como base para sustentar medidas políticas, tendo em vista que os administradores públicos, com o intuito de assegurar a qualidade e a credibilidade de suas decisões, acabam por procurar legitimar suas ações a partir de estudos, relatórios, estatísticas e falas de autoridades, o que Alves e Canário (2002) chamam de “magistratura de influência”. Assim, nos últimos tempos, vem se observando, nos materiais divulgados pelo poder público, a presença de discursos acadêmicos, a fala de pesquisadores e/ou as parcerias com universidades renomadas. Como afirma Gatti (2008, p.61),

Alguns desses administradores [públicos] já vinham tomando medidas para tentar garantir certa qualidade a esses programas na seara pública, com estabelecimento de critérios, em editais e resoluções executivas, para as instituições que se responsabilizariam pelos trabalhos, investindo nas mais credenciadas, . . . Assim, encontramos universidades bem qualificadas e instituições com tradição de qualidade envolvidas nesses projetos em associação com o MEC e com Secretarias de Educação estaduais ou municipais, como a Universidade Federal de Juiz de Fora (UFJF), a Universidade de Ijuí (UNIJUÍ), Universidade Federal do Rio Grande do Norte (UFRN), a Universidade Federal de Mato Grosso

(UFMT), a Universidade Federal de Pernambuco (UFPE), a USP, a UNICAMP, a UNESP, a Fundação Getúlio Vargas, a Fundação João Pinheiro, entre tantas outras.

Parece conveniente dizer que, mesmo colocando os participantes dentro de um mesmo grupo, entende-se que eles podem ter opiniões distintas e até mesmo opostas, pois o tema estudado está longe de ser consensual dentro da comunidade acadêmica e educacional. Isso ocorre porque há aqueles que reconhecem a utilidade das avaliações externas, porém há uma parcela significativa que desconsidera as contribuições das medidas educacionais em larga escala (Bauer et al., 2015).

A escolha por contemplar os decisores políticos nesta pesquisa se deu ao entender que a política é um lugar de “construção, por um dado conjunto social, de um modo de compreender o mundo e de agir sobre ele” (Carvalho, 2006, p. 37). Há de se considerar que a realidade educativa se compreende também a partir de políticas educacionais e que as avaliações externas estão diretamente ligadas a essas políticas.

Como decisores políticos, entende-se aqueles que possuem autoridade em matéria de políticas públicas educacionais, situados em órgãos nacionais ou subnacionais, organizações locais, que tomam decisões e produzem materiais de caráter legislativo ou normativo para o setor educacional (Carvalho, 2006). É oportuno destacar que são poucos os trabalhos no campo da educação que consideram os decisores políticos como participantes. Na verdade, a maioria das pesquisas consultadas realiza apenas uma análise documental a partir de materiais produzidos por esse grupo. No que se refere a avaliação externa, pouco se sabe sobre o uso de resultados pelos formuladores políticos (Stecher, 2002).

Por fim, o grupo de professores da educação básica. De acordo com Tardif (2002), a pesquisa que tem por objetivo compreender a natureza do ensino deve ser baseada em um diálogo fecundo com os professores, entendendo-os não apenas como quem aplica conhecimentos produzidos por outros, mas sim como atores no

sentido forte do termo, isto é, sujeito que assume sua prática a partir dos significados que ele mesmo lhe dá, um sujeito que possui conhecimentos e um saber-fazer provenientes de sua própria atividade e a partir dos quais ele a estrutura e a orienta (p.230).

A partir dessa perspectiva, o autor considera que toda pesquisa que tenha como campo de estudo o ensino deve registrar o ponto de vista dos professores, considerando a experiência relativa ao trabalho dos professores, com suas tensões, seus dilemas e suas rotinas.

Dessa forma, considerar o ponto de vista do professor significa levar em consideração seu conhecimento e saber-fazer, ou seja, entender que os professores pensam, dão sentido e significado aos seus atos. Isso indica que realizar uma análise da avaliação a partir da perspectiva desses atores representa compreender que o docente não é aquele que apenas cumpre ou executa ordens ou conhecimento já prontos, provindos de universidades ou esferas políticas, mas também aqueles que constroem conhecimento e uma cultura própria, que vêm pautar suas ações cotidianas. Por conseguinte, entende-se que o ensino

aproxima-se bastante, assim, daqueles ofícios e daquelas profissões cujo espaço cotidiano de trabalho é marcado por uma grande autonomia e em que as atividades são desenvolvidas de acordo com representações, muitas vezes, renovadas, móveis, imprevisíveis na sua concretização e onde, por fim, a personalidade do trabalhador torna-se parte integrante do processo de trabalho (Tardif, 2009, p.45).

Com o intuito de observar o objeto de pesquisa de diferentes ângulos e de realizar uma compreensão profunda sobre o tema, decidiu por fazer também uma recolha de dados quantitativos com os professores da educação básica, para além das entrevistas. Para tal, a seleção se deu a partir de alguns critérios.

Como a pesquisa parte do internacional ao local, foi necessário escolher um estado e um município para a análise da avaliação local. O estado do Rio de Janeiro foi o escolhido, pois desde os anos 2000 até 2016, apresentou avaliações em larga escala de forma intermitente. A partir do estado, foi escolhido o município do Rio de Janeiro, já que possui a maior rede de educação municipal da América Latina. Em dados fornecidos pela Secretaria Municipal de Educação (SME), no ano de 2018, havia 1.537 unidades escolares no município do Rio de Janeiro, com 641.397 alunos matriculados e um total de 38.894 professores, sendo 2.987 professores do ensino fundamental II das disciplinas de Português e Matemática (Rio de Janeiro, 2018).

Optei por realizar inquéritos por entrevista e por questionário aos professores da rede pública de ensino, tendo em vista que essas escolas são oficiais, no sentido de serem dirigidas pelo governo.

Diante dessa realidade, é oportuno destacar que o processo de seleção dos participantes é uma etapa de fundamental importância para o delineamento da pesquisa. Para além da seleção da unidade da amostra, devem-se conjugar também as informações de dimensão e de procedimento, levando em consideração que tais informações são fatores básicos para a validade do estudo. Para se definir a dimensão do número dos participantes, algumas variáveis devem ser analisadas, tais como recursos financeiros, disponibilidade de tempo e a análise estatística. Porém, mesmo entendendo que amostragem probabilística apresenta um maior grau de confiabilidade de resultado e o poder de generalização em um estudo quantitativo, devido à grande dificuldade de acessibilidade aos elementos da população e à disponibilidade de tempo, optei por realizar uma amostra não probabilística intencional, que permite a delimitação da população. De acordo com Churchill (2006), a característica desse tipo de amostra é “que os elementos da população são selecionados intencionalmente. Esta seleção é feita considerando que a amostra poderá oferecer as contribuições solicitadas.” (p. 301).

Dessa forma, entendeu-se que apenas professores de Português e Matemática participariam da pesquisa, por serem disciplinas que são contempladas em todas as avaliações externas apresentadas. Como o sistema de ensino brasileiro é administrado por diferentes instâncias, a depender do ciclo de escolaridade⁶, definiu-se para a seleção professores apenas do ensino fundamental II (6º ao 9º ano), que estão sob responsabilidade do município. Esse recorte dos participantes seguiu o critério de selecionar atores que estão mais diretamente envolvidos com o fenômeno em questão.

Tal como se ilustra na Figura 6, a Secretaria Municipal de Educação do Rio de Janeiro está organizada em coordenadorias regionais de educação (CREs). Assim, o critério utilizado para a escolha das CREs foi o de proximidade geográfica das escolas, possibilitando uma melhor recolha de dados.

⁶ Segundo a Lei de Diretrizes e Bases da Educação (Lei nº 9.394/94), a educação básica é dividida em três etapas: educação infantil (creche e pré-escola), ensino fundamental I (1º ao 5º ano), ensino fundamental II (6º ao 9º ano) e ensino médio (1ª a 3ª série). A responsabilidade de cada etapa é dividida prioritariamente da seguinte forma: cabe ao município responder pela oferta da educação infantil e do ensino fundamental, e o estado pela oferta do ensino médio.

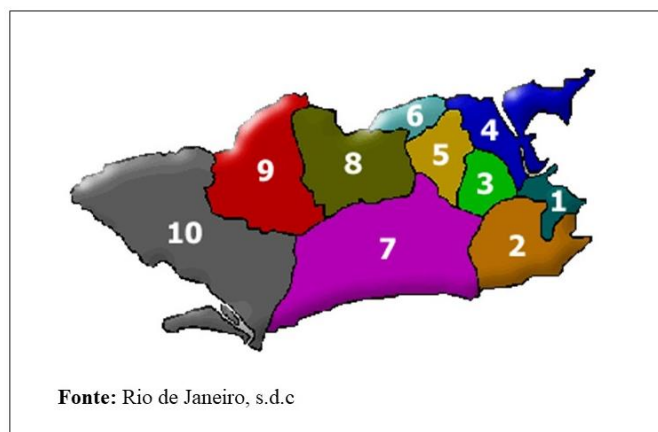


Figura 6

Divisão das CREs no município do Rio de Janeiro

Por isso, pediu-se autorização para as 2ª e 3ª CREs que, por sua vez, selecionaram um total de 21 escolas sob suas responsabilidades, respeitando os critérios definidos pela pesquisa.

Coleta e Registro de Dados

O primeiro passo desta pesquisa foi realizar um estudo documental, tanto em documentos oficiais quanto em relatórios técnicos, para a construção de uma trajetória histórica com o intuito de realizar um desenho das avaliações em larga escala empregadas nos diferentes âmbitos governamentais, desde o internacional até o local, para poder direcionar o estudo. Dessa forma, buscou-se descrever, a partir dos princípios norteadores desses documentos, as dimensões do fenômeno estudado. Guba e Lincoln (1981, citados por Lüdke & André, 1986, p. 39) reconhecem que os documentos são “uma fonte estável e rica, . . . surgem num determinado contexto e fornecem informações sobre esse mesmo contexto, . . . são fontes não reativas, permitindo a obtenção de dados quando o acesso ao sujeito pode alterar seu comportamento, ou seus pontos de vista”. Os autores chegam a afirmar que, por serem uma fonte tão repleta de informações sobre a natureza do contexto, nunca devem ser ignorados. No entanto, deve-se também destacar que os documentos são amostras não representativas dos fenômenos pesquisados, haja vista que nem sempre o que é apresentado em um documento é aplicado na prática. Por isso, realizou-se também uma análise da trajetória histórica com base em pesquisa bibliográfica. Essa etapa da pesquisa também auxiliou na elaboração dos instrumentos de coleta de dados e na posterior análise dos dados coletados pelo questionário e pela entrevista.

Nas fases seguintes, mesmo havendo uma sequência de apresentação da coleta e do registro de dados nessa sessão, cabe ressaltar que a estratégia utilizada foi a concomitante, na qual os dados qualitativos e quantitativos foram recolhidos simultaneamente.

No que tange aos professores da educação básica, em abril de 2018, deu-se início ao pedido de autorização para a realização das entrevistas e aplicação do questionário na Secretaria Municipal de Educação do Rio de Janeiro. Após dois meses de análise, o pedido foi aceito; no entanto, as escolas estavam em período de recesso de julho, havendo, por conseguinte, necessidade de se esperar até o mês de agosto para dar início à coleta de dados.

Foi utilizado o Google Forms como ferramenta para a elaboração e preenchimento dos questionários. O recurso empregado apresenta algumas vantagens, tais como: a possibilidade de se ter acesso em qualquer lugar e horário, a facilidade de envio para um grande número de participantes, a gratuidade e a facilidade de uso. O *link* foi disponibilizado para os professores a partir do contato e aceitação da escola em participar da pesquisa, ficando disponível para resposta por um total de, aproximadamente, três meses.

No questionário aplicado aos professores, para além das perguntas referentes ao objeto de pesquisa, havia também questões para obtenção de dados pessoais, para que fosse possível caracterizar devidamente os inquiridos.

No total, foram 124 questionários respondidos, 59,7% de professores de Português e 40,3% de professores de Matemática. De acordo com as respostas do perfil dos professores, observou-se que a maior parte dos respondentes possui especialização (40,3%), seguida por mestrado (31,5%), graduação (25%), doutorado (1,6%) e pós-doutorado (1,6%). No tocante à atuação docente, na distribuição por ano de escolaridade (6º ao 9º ano), observa-se que 61,6% atua no 6º ano, 56,5% no 7º ano, 55,6% no 8º ano e 50,8% no 9º ano. É importante destacar que o total ultrapassa 100% porque um professor pode lecionar para mais de um ano de escolaridade. Apenas 19,4% afirmaram já ter feito parte da equipe de gestão escolar em sua trajetória profissional.

A Tabela 3 registra o tempo de atuação como professor e o tempo de atuação na rede municipal do Rio de Janeiro.

Tabela 3

Tempo de atuação dos docentes participantes da pesquisa

Anos	Tempo de atuação como professor	Tempo de atuação como professor na rede municipal do RJ
Menos de 6	12,1%	25,8%
6-15	45,2%	45,2%
16-25	20,2%	16,9%
Mais de 25	22,6%	12,1%

Os dados obtidos através do inquérito por questionário permitiram determinar as principais tendências de resposta quer através do cálculo das frequências relativas, quer através do cálculo das respectivas médias. Desse modo, foi possível complementar os dados de natureza qualitativa, obtidos através do inquérito por entrevista, com os dados de natureza quantitativa.

As entrevistas foram acordadas com os professores na escola de forma individual e no horário de melhor conveniência ao profissional. Todas as entrevistas foram gravadas, com a permissão dos participantes, a partir do Termo de Consentimento (Anexo 5). Para a operacionalização das entrevistas, houve a transcrição na íntegra dos áudios e a transformação da linguagem oral em escrita. A fidedignidade das informações foi garantida pela conferência das transcrições após todo o processo.

Alguns professores não se sentiram confortáveis em gravar entrevista, com medo da identificação. Para atender a esse grupo, as perguntas foram enviadas pelo Google Forms, junto com um termo de consentimento sem identificação do participante. Entende-se que se perde muito ao enviar as perguntas de forma eletrônica, tendo em vista que não há interação entre o entrevistador e o participante. Porém, como o tema envolve questões de prática do profissional, bonificação salarial e opiniões sobre gestores e políticos, julgou-se importante considerar esses participantes de maneira anônima, pois estariam mais confortáveis em responder às questões de forma mais próxima da realidade existente na escola.

No total, nove entrevistas foram realizadas, cinco com professores de Português e quatro com professores de Matemática.

A entrevista com os outros dois grupos decorreu de forma semelhante, porém devido à dimensão geográfica do Brasil, a dificuldade financeira e de tempo para

deslocamento para outros estados, as entrevistas fora do estado do Rio de Janeiro ocorreram por Skype ou pelo telefone, atendendo à disponibilidade dos entrevistados.

Para o grupo de especialistas, foram selecionados 13 participantes e enviado um *e-mail* convite para participar da pesquisa. Ao todo, tive 11 respostas favoráveis que acabaram por permitir a concretização de nove entrevistas. No entanto, utilizei apenas sete entrevistas, pois não obtive dados suficientes em duas delas, optando assim por desconsiderá-las.

Os especialistas foram escolhidos a partir do critério de relevância das suas publicações no domínio da avaliação em educação. Pode-se considerar que os participantes que compuseram o grupo de especialistas são professores e pesquisadores de grande prestígio no Brasil, cujo trabalho é reconhecidamente incontornável para quem pretende desenvolver qualquer pesquisa nesse domínio.

Posso destacar que são especialistas conceituados no campo da avaliação no Brasil, tendo em vista que possuem publicações em prestigiados periódicos; produções de livros ou capítulos de livros; são coordenadores de grupos de pesquisa das melhores universidades brasileiras e conferencistas em diversos eventos nacionais e internacionais. A maioria dos especialistas tem mais de 20 artigos publicados na área de avaliação. Há alguns com mais de 40 publicações, sendo citados mais de 3.000 vezes em outros artigos e diversos trabalhos acadêmicos⁷. Dentre os especialistas, há aqueles que também contribuem com secretarias municipais e estaduais de educação e também com setores do MEC, como consultores, colaboradores ou pesquisadores.

Com relação aos decisores políticos, em um primeiro momento, pensei em realizar as entrevistas com as diferentes instâncias analisadas nesta pesquisa: nacional, com o Ministério de Educação (MEC); estadual, com a Secretaria de Educação do Estado do Rio de Janeiro (SEEDUC); e municipal, com a Secretaria de Educação do Município do Rio de Janeiro (SME-RJ). No entanto, apenas a SME-RJ deu parecer favorável em tempo hábil para a realização das entrevistas. A Secretaria de Educação Básica do MEC afirmou que não responde sobre o tema de investigação, informando que deveriam ser procurados o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) para determinado fim. A Capes não se encontra no escopo desta pesquisa, por ser direcionada apenas para o nível superior, por isso não foi procurada. Houve uma grande dificuldade em realizar

⁷ Dados coletados a partir do Currículo Lattes dos especialistas e de informações do Google Scholar.

contato com a Diretoria de Avaliação da Educação Básica do Inep, mas, quando consegui, a resposta foi positiva para a participação. No entanto, depois de várias tentativas, não foi possível marcar um dia para a realização da entrevista, por falta de espaço na agenda da pessoa responsável.

No que se refere ao âmbito estadual, um processo para o pedido de autorização para a realização da entrevista foi aberto na SEEDUC, no entanto, depois de nove meses, não obtive nenhum posicionamento. O processo tramitou em diferentes instâncias, entrei em contato algumas vezes para saber um posicionamento, no entanto, a única resposta era que estava sendo analisado. Com a demora em dar uma resposta e a inviabilidade com relação ao tempo de elaboração da investigação, decidiu-se por não considerar mais os decisores políticos da SEEDUC como participantes.

Por fim, no âmbito municipal, o pedido foi feito à SME-RJ e aceito em tempo hábil para a realização da pesquisa. Foram selecionados três decisores políticos da secretaria e recebi parecer favorável para a realização, no entanto, apenas dois participaram da entrevista, tendo em vista que um não conseguiu horário na agenda.

Dessa forma, apenas foram realizadas duas entrevistas das sete planejadas inicialmente. Diante da dificuldade encontrada, pensei em um novo desenho para contemplar a percepção de um número maior de pessoas nesse grupo. O que resultou no convite de pessoas que mesmo hoje não ocupando cargos políticos, já o fizeram. E também representantes das diferentes agências de avaliação, entendendo que também podem influenciar políticas públicas. Assim, a partir desse novo modelo, foram convidadas mais cinco pessoas, o que possibilitou mais quatro entrevistas, resultando em um total seis participantes no grupo de decisores políticos, como pode ser observado na Tabela 4.

Tabela 4
Perfil dos decisores políticos

Cargo	Atuação
Ex-Secretário-Executivo do MEC e Ex-presidente do Inep	Governo Federal
Ex-Secretário de Educação	Secretaria Municipal
Subsecretaria de Ensino	Secretaria Municipal
Coordenação de Supervisão de Avaliação	Secretaria Municipal
Coordenação de Centro de Avaliação	Agência de Avaliação
Coordenação da Unidade de Avaliação	Agência de Avaliação

No total, foram feitas 24 entrevistas, porém cabe salientar que nem todas proporcionaram informações consistentes em todas as perguntas realizadas. Assim, foram consideradas um total de 22 entrevistas.

Com o sentido de garantir o anonimato de todos os participantes e uniformizar a forma de tratamento dos entrevistados, optou-se por referir-se a todos no gênero masculino, quando necessário. No entanto, cabe ressaltar que houve presença feminina em todos os grupos.

Apresentação e Discussão dos Principais Resultados

A Matriz de Investigação, apresentada na Figura 4, mostra claramente que o estudo principal desta pesquisa é a avaliação externa da aprendizagem dos alunos. A partir disso, foram destacados dois objetos de análise: a avaliação externa da aprendizagem e a qualidade da educação. Para cada objeto, foi criado um conjunto de dimensões, que se propõe a ajudar na caracterização de cada objeto. A construção das dimensões da Matriz foi pensada com base na teoria e na literatura analisadas, nas questões de investigação e no problema de pesquisa.

Com base na Matriz de Investigação, foi realizada a transformação dos dados, com um exercício de síntese e de integração de informação. Optei por olhar para esses dados a partir das recomendações de Wolcott (1994). O autor apresenta três momentos fundamentais para a realização da transformação dos dados, a saber: descrição, análise e interpretação.

A descrição é a fase inicial, na qual o pesquisador realiza a escrita de textos resultantes dos dados originais registrados pelo investigador. Nesse momento, ocorre o

exercício de redução das informações, ou seja, cabe ao pesquisador perguntar-se se cada detalhe considerado para a inclusão é relevante para a análise. Trata-se de um significativo esforço de síntese e decisão do que entra para compor a descrição e em qual nível de detalhe será contemplado; e o que será excluído, tendo em consideração o propósito da pesquisa. Ou seja, a essência do todo é revelada e não os detalhes.

Para atender a essa fase de descrição, decidiu-se por construir uma narrativa para cada dimensão da Matriz de Investigação, sendo composta pela narrativa de cada entrevistado. Dessa forma, cada entrevistado recebeu um código, ou seja, os especialistas receberam a letra E mais uma letra do alfabeto (E-A, E-B, E-C...). Os outros grupos seguiram a mesma lógica, decisores políticos (DP-A, DP-B, DP-C...) e professores da educação básica (P-A, P-B, P-C...).

A fim de compor as narrativas, foi necessário certificar-se de que a descrição continha as informações necessárias para análise e interpretação subsequentes, tendo como preocupação não levantar dados que suportassem apenas uma estrutura preconcebida, isto é, realizar descrição com suposições antecipadas, sem base objetiva e real nas entrevistas.

A análise é o segundo momento, no qual se objetiva a organização dos dados com vistas a salientar os aspectos essenciais, a identificação dos fatores-chave e a descrição sistemática de inter-relações entre eles. O intuito é levantar questões a partir dos dados descritos, buscando coerência e uma direção a ser seguida. Nesse momento, ocorre uma seletividade maior, tendo em vista que alguns dados recebem mais atenção.

Na parte de análise, busquei organizar as informações trazidas nas narrativas de modo a revelar as propriedades, estruturas subjacentes e relações existentes entre elas. Assim, construiu-se a opinião do grupo sobre a dimensão trabalhada. Nesse instante, algumas falas das entrevistas foram evidenciadas, com o intuito de exemplificar o ponto em análise. Optei por apresentar um novo material dos dados, para trabalhar de forma refinada os detalhes. Novas falas foram acrescentadas, material considerado rico para a exemplificação e compreensão. Isto posto, cabe salientar que essa fase busca não apenas transmitir informações, mas também enfatizar aspectos específicos do estudo.

Por fim, a parte de interpretação diz respeito ao processo de obtenção de significados e contextos, ir além das fronteiras dos casos particulares levantados, buscando, por conseguinte, um significado mais amplo dos dados. Empenha-se, neste ponto, em responder às questões de investigação, realizando, de forma conjunta com a teoria levantada, uma reflexão sobre o objeto de estudo.

É importante destacar, assim como afirma Wolcott (1994), que a descrição, a análise e a interpretação não são etapas mutuamente exclusivas, uma vez que há, normalmente, uma convivência entre as três, possuindo limites, muitas vezes, tênues.

Ao longo dessas últimas considerações, foi possível considerar que a transformação de dados desta pesquisa foi realizada de forma artesanal, ou seja, não se baseou em algoritmo pré-definido ou em processamento de codificação excessivamente analítico. Houve um esforço na direção de um estudo mais aprofundado nos dados, com organização da informação de forma intuitiva, assim como proposto por Wolcott (1994).

A organização do capítulo que compreende a análise e a discussão dos dados foi pensada de forma que pudesse evidenciar a percepção dos diferentes intervenientes do processo educativo sobre a avaliação externa da aprendizagem dos alunos. Dessa forma, busquei descrever, analisar e interpretar os dados da pesquisa dentro do contexto global do discurso de cada grupo. Explicando de outra forma, os dados foram trabalhados em quatro partes, uma para cada grupo de análise, totalizando três grupos; na última parte, foi realizado uma síntese das evidências mais importantes encontradas em cada grupo participante.

Com essa divisão, busca-se um olhar particular da narrativa de cada participante da pesquisa; uma visão por grupo, com base na análise das respostas de cada participante; e por fim, a construção de uma narrativa a partir do cruzamento dos dados dos três grupos. Esse desenho possibilita uma gradação de olhar sobre o objeto, desde a particularidade de um especialista, de um decisor político ou de um professor até um quadro geral da avaliação externa, proporcionando uma visão ampla sobre o assunto.

Capítulo IV

Do Internacional ao Local: Uma Discussão Acerca das Avaliações Externas Internacionais, Nacionais e do Estado do Rio de Janeiro

Em 1951, a UNESCO dá início a uma publicação para a elaboração de um estudo comparativo sobre a educação, tendo como base os sistemas educacionais do mundo. Para tal, 57 países participaram preenchendo um questionário básico que resultou na publicação *World Survey of Education*, em cinco volumes, de 1955 a 1971, com o objetivo de oferecer um instrumento que viesse a cooperar entre as autoridades e os educadores. Na década seguinte, em 1960, a International Association for the Evaluation of Educational Achievement (IEA) inicia o primeiro estudo que realizou uma avaliação do desempenho escolar, um projeto-piloto, que avaliou o desempenho nas áreas de leitura, matemática, ciências, geografia e habilidade não verbal, simultaneamente, em diversos países. O estudo disponibilizou dados comparáveis e abrangentes, de caráter estatístico e descritivo, sobre a organização e administração da educação, produzindo categorias universalmente aceitas para a geração de informações educacionais possíveis de serem comparadas e usadas para a melhoria da qualidade da educação (Almeida & Wolyneec, 2007).

Estudos desta natureza e os sistemas educativos nacionais possibilitaram a utilização de uma metodologia comparada, tornando-se um instrumento analítico dos sistemas em questão. Essa ação viabilizou realizar análises com o intuito de identificar semelhanças e diferenças, estudar a realidade educacional de um determinado país frente a outros. A comparação internacional propicia que os países reconheçam tanto seus pontos fracos quanto os fortes, que, em geral, são ignorados ou não percebidos dentro de seus próprios sistemas. Na literatura, há diversas publicações sobre a história da Educação Comparada (cf. Carvalho, 2014; Kazamias, 2012; Nóvoa, 2009), sendo sua linha teórica dividida por períodos e perspectivas. Não cabe aqui discorrer sobre todos os períodos históricos, mas, em linhas gerais, situar como a Educação Comparada, os sistemas educativos e as organizações internacionais se harmonizam no contexto de avaliação externa.

Por volta do final da década de 1950 e início de 60, surge uma perspectiva que visa à formulação de leis gerais para a evolução dos sistemas educativos, com tendência

a isolar alguns de seus componentes, que se tornam objeto de uma análise específica (por exemplo, as questões do financiamento, da estrutura dos prédios). Essa tendência foi chamada de historicista por Nóvoa (2009). Em ruptura com essa perspectiva, surge uma nova concepção de Educação Comparada, que busca “uma abordagem que seria mais sistemática do que subjetiva, mais empírica e analítica do que puramente especulativa e descritiva; mais ‘microcósmica’ do que ‘macrocósmica’ e olímpica, e mais ‘científica’ do que ‘histórica’ ou ‘filosófica’” (Kazamias, 2012, p. 179). Com isso, inicia-se a vontade de fornecer soluções científicas aos decisores, por meio de relações de causa e efeito. Como exemplo de tal abordagem da Educação Comparada, pode-se destacar a criação da IEA, que surgiu em 1958 e teve um projeto-piloto aplicado em 1960. Outro exemplo é a expansão das organizações internacionais (UNESCO, OCDE, Banco Mundial etc.), que vieram a desenvolver-se em espaços essenciais para a formulação das políticas educativas (Nóvoa, 2009). É nesse contexto que se estabelecem sistemas de avaliação com o objetivo de produzirem informações mais sistemáticas e empíricas sobre os sistemas educativos.

O principal propósito deste capítulo é descrever e analisar alguns sistemas internacionais de avaliação das aprendizagens como forma de proceder igualmente a uma descrição e análise dos sistemas nacional e local que são utilizados no Brasil. Na primeira parte, são discutidas avaliações externas internacionais, regionais e nacionais numa perspectiva comparada. Num segundo momento, apresenta-se o sistema de avaliação do estado e do município do Rio de Janeiro, finalizando com uma comparação entre as avaliações externas locais e nacionais.

Avaliações Externas Internacionais

Algumas instituições internacionais se destacam com reconhecidos sistemas de avaliação internacionais, com base em indicadores educacionais com rigor técnico e monitoramento com possíveis comparações entre diferentes sistemas de ensino. Dentre as organizações mais conhecidas, encontram-se a Associação Internacional para Avaliação do Desempenho em Educação – IEA, sendo reconhecida pelo TIMSS (Trends in International Mathematics and Science Study) e pelo PIRLS (Progress in International Reading Literacy Study); a Organização para a Cooperação e o Desenvolvimento Econômico – OCDE, que, a partir do Programa Mundial de Indicadores, desenvolve o Programa Internacional de Avaliação de Estudantes – PISA (Programme for International Student Assessment); por fim, no âmbito da América Latina, a OREALC/UNESCO desenvolveu o Laboratório Latinoamericano de Evaluación de la Calidad de la Educación

(LLECE), que realizou, até o momento, três avaliações na região. De acordo com Díaz-Barriga (2018), essas organizações internacionais tornaram-se protagonistas a partir de meados de 1980, com um discurso educacional que conduz várias políticas educacionais estabelecidas por diversos governos.

O Brasil teve sua primeira participação em um estudo com base em indicadores quando foi convidado pela UNESCO para fazer parte do Projeto WEI – World Education Indicator. O projeto teve início em 1997, originalmente com 12 países convidados que receberam, desde o início, apoio do Banco Mundial. O principal objetivo foi desenvolver indicadores na área da educação, para medir o estado do ensino de forma comparável internacionalmente. Os resultados do programa incluem um conjunto de dados comparativos dos indicadores de educação e o mapeamento detalhado dos sistemas educacionais nacionais. O Projeto WEI não conta com provas padronizadas para avaliação de alunos, mas sim com levantamentos estatísticos.

Os dados gerados antes dos anos de 1990, como financiamento e gasto com educação, eram levantados por agências internacionais, principalmente UNESCO e OCDE. Contudo, identificaram-se dificuldades na comparação dos dados, tendo em vista que havia problemas metodológicos e técnicos nos países que levavam à produção de dados não confiáveis. Assim, esses dados eram de pouca utilidade para a tomada de decisão de políticas públicas educacionais. Em decorrência desse fato, a OCDE viu a necessidade de realizar esforços com o intuito de melhorar a qualidade da informação e, conseqüentemente, a sua comparabilidade. A partir disso, consolidou-se o projeto Indicators of Education Systems (INES) com o objetivo de coletar dados de educação para a produção de indicadores educacionais comparáveis entre os países-membros e auxiliar os participantes do projeto no desenvolvimento de dados confiáveis. O resultado dessa ação levou à primeira publicação do *Education at a glance*, em 1991. A participação do Brasil nesses projetos internacionais possibilitou conhecer e produzir indicadores educacionais seguindo técnicas dos estudos internacionais (Almeida, 2008).

Também no ano de 1997, um estudo no âmbito da América Latina começa a se desenvolver. O Laboratório Latino-americano de Avaliação da Qualidade da Educação (LLECE) realiza o Primeiro Estudo Regional Comparativo e Explicativo (Perce), o qual contou com a participação de 13 países da América Latina, incluindo o Brasil. Já no ano 2000, a OCDE iniciou a aplicação do Programa Internacional de Avaliação de Estudantes (PISA). Trata-se de uma avaliação comparada, que é aplicada de forma amostral a estudantes de 15 anos de idade que frequentam a escola. Atualmente, o PISA é o único

programa de avaliação externa das aprendizagens e competências dos alunos de que o Brasil participa.

De acordo com Fernandes (2010b), esse espaço de importância que os estudos internacionais da aprendizagem do aluno vêm ganhando é uma resposta à necessidade de os países melhorarem sua formação, atender às aspirações e motivações dos jovens e aos desafios da competitividade da mobilidade mundial. No entanto, o autor destaca que

os estudos internacionais podem ser importantes elementos de pressão para que se alterem práticas e procedimentos, sem, contudo, nos indicarem em que direcção. . . . Podem ajudar os governos a justificar ou a fundamentar mudanças nos currículos, nos sistemas de formação de professores, nos sistemas de alocação de recursos às escolas ou nos sistemas de gestão e administração escolar. Neste sentido e só neste sentido, pelo menos assim me parece, podem pressionar os próprios governos para que alterem as suas políticas educativas. Mas sempre numa base algo abstracta e pouco sólida (Fernandes, 2010b, p. 9).

Com base no desenvolvimento e importância das avaliações externas internacionais nas agendas políticas dos países participantes, busca-se, na sequência, desenvolver a descrição e análise das avaliações internacionais, tanto de âmbito mundial quanto regional, e a participação do Brasil nelas.

Programa Internacional de Avaliação de Estudantes (PISA).

A OCDE apresenta como escopo dados econômicos, porém também cobre áreas diversas, como educação, saúde, comércio, emprego, mercado financeiro e energia, com a possibilidade de realizar um estudo comparativo e trocas de experiências entre os países participantes.

No que se refere à educação, a OCDE busca um trabalho de melhoria dos indicadores internacionais de desempenho educacional, o que leva a um investimento na área de educação através da organização de pesquisas internacionais comparáveis (Inep/PISA, 2001). O principal programa da organização nesse sentido é o PISA, um estudo internacional comparado que apresenta como objetivo a produção de indicadores que possam contribuir com a discussão da qualidade da educação nos países participantes. Sendo assim, serve como dado para a formulação de políticas educacionais, permitindo que administradores e gestores políticos possam comparar o funcionamento de seus sistemas educativos com outros países e tomar decisões no âmbito nacional. Além disso,

pode ajudar a impulsionar e orientar reformas educacionais, para que venham a melhorar a qualidade da educação de seus países. Tal fato, só é considerado possível pela organização, por não estar vinculado ao currículo, facilitando, assim, uma comparação mais assertiva entre os países.

Na construção de seus objetivos, o PISA apresenta uma maior ênfase nas informações do desempenho dos alunos, sendo o ponto norteador para três frentes: avaliar conhecimentos e habilidades; relacionar o desempenho a temas de políticas públicas; e realizar um monitoramento regular dos padrões de desempenho. Nos objetivos propostos, encontram-se os verbos avaliar, relacionar e realizar, o que configura um caráter mais informacional em suas contribuições, ou seja, há mais informações e diretrizes a serem cumpridas do que possibilidade de discussão sobre os dados gerados. Os três verbos fecham uma tríade proposta pela avaliação, ou seja, avaliar o desempenho dos alunos para relacionar com outros países e realizar um monitoramento dos padrões de desempenho. A OCDE não especifica como as matrizes de referências do PISA são elaboradas, afirma apenas que são pensadas com especialistas em educação.

A avaliação é aplicada em estudantes na faixa etária dos 15 anos, a partir do 7º ano do ensino fundamental, que se pressupõe estarem no final da educação básica obrigatória na maioria dos países. A primeira avaliação ocorreu em 2000 e envolveu 265.000 alunos de 32 países. As avaliações acontecem a cada três anos, tendo até o momento sete edições realizadas. Elas abarcam três áreas do conhecimento, a saber: leitura, matemática e ciências. Em cada ciclo de administração das provas, dá-se ênfase a uma dessas três literacias testadas. No primeiro ano de aplicação, a ênfase ocorreu sobre o domínio de leitura; dito de outra forma, havia mais questões que avaliavam o desempenho em leitura. Para além dos testes, são também aplicados questionários de contexto, tendo como objetivo relacionar o desempenho dos alunos a variáveis demográficas, socioeconômicas e educacionais. Os questionários são aplicados aos alunos, aos professores e às escolas.

Com relação ao conteúdo cobrado na avaliação, o PISA se diferencia de outros testes, uma vez que avalia como os jovens são capazes de utilizar os conhecimentos, capacidades e atitudes aprendidos na educação básica para resolver problemas, ou seja, busca-se compreender como os sistemas educacionais estão desenvolvendo as competências consideradas básicas, ao longo dos 9 anos de escolaridade, em três domínios. Assim, não avaliam diretamente o domínio de conhecimentos escolares curriculares por parte dos alunos, como ocorre no TIMSS da IEA, que são baseados no

currículo em vigor nos países participantes, considerando o currículo proposto, o ensinado e o aprendido.

Para tal, no TIMSS, são analisados outros fatores que podem interferir no desenvolvimento do currículo, como condições das escolas e das salas de aula, processos metodológicos, as características da comunidade à qual a escola pertence, entre outros. Há a concepção de que se pode conseguir uma compreensão adequada dos elementos que influenciam a aprendizagem mediante uma análise cuidadosa do rendimento dos estudantes, suas características, o currículo que seguem, as metodologias didáticas dos professores e os recursos disponíveis na sala de aula e na escola. Entende-se que tais características são reflexos da comunidade e do sistema educativo, que, por sua vez, também refletem a sociedade do país em que vivem os alunos. Nesse sentido, o PISA e o TIMSS são avaliações muito diferentes (Díaz-Barriga, 2005; Fernandes, 2010b; Hutchison & Schagen, 2007).

A IEA tem o objetivo de contribuir com a melhora na qualidade da educação, proporcionando informação para os decisores políticos e a comunidade acadêmica acerca da qualidade da educação ofertada. Além disso, propõe-se a apoiar os países participantes na compreensão das diferenças observadas entre si e as que ocorrem em seus próprios sistemas educacionais. Não há um posicionamento direto sobre a orientação de políticas educativas, embora tal propósito possa estar implícito em seus objetivos (Fernandes, 2010b).

Essas diferenças são construídas, em grande parte, pela estrutura das organizações. A OCDE é uma organização econômica, com fins governamentais, que visa à condução de políticas públicas com o intuito de realizar um desenvolvimento econômico e bem-estar social, havendo uma estabilidade financeira – principalmente entre os países-membros da organização – e o fortalecimento da economia global. Por sua vez, a IEA é uma organização com mais fins pedagógicos, não apresentando ligação governamental.

Em estudo realizado por Andrews, Ryve, Hemmi e Sayers (2014), evidencia-se como a diferença das duas avaliações – PISA e TIMSS – pode ser vista para além dos resultados e da estruturação dos testes. Os autores realizaram uma pesquisa com o desempenho dos alunos finlandeses em ambos os testes. Em contraste ao sucesso mundialmente conhecido apresentado pela Finlândia nos exames do PISA, o mesmo não ocorre no TIMSS, tendo em vista que apresentam resultados modestos no domínio de matemática. Para lançar luz sobre essa diferença, o estudo realiza uma análise a partir das

práticas em sala de aula de matemática com base em entrevistas com professores e gravação de vídeos de sequências de lições ensinadas.

Os autores concluem que as evidências de sala de aula observadas e as entrevistas com os professores parecem ser incompatíveis com os resultados do PISA, verificando que o ensino dos professores estaria mais voltado para o desempenho modesto no TIMSS. Dito de outra forma, as suposições feitas por políticos públicos e promovidas pela imprensa mundial, que associam o sucesso do PISA à qualidade pedagógica finlandesa, parece ser um equívoco. Algumas indicações para o sucesso dessa avaliação são levantadas pelos autores, como: ensino direcionado para o teste, participação dos pais no ensino de matemática e alta proficiência em leitura dos alunos.

Segundo Wu (2010), há uma grande relação entre o desempenho em leitura e as habilidades de resolução de problemas matemáticos da vida cotidiana na avaliação do PISA. Dessa forma, o elevado desempenho em matemática estaria mais relacionado à habilidade de compreensão, pois em países onde o desempenho em leitura é relativamente maior, há uma exposição a ambientes para além do espaço escolar que facilita as habilidades de resolução desses problemas matemáticos. É importante evidenciar que a Finlândia possui uma taxa de alfabetização de 100%⁸, um dado relevante para esse tipo de análise. Com relação à participação da família, Carnoy, Khavenson, Loyalka, Schmidt e Zakharov (2016) versam sobre a influência dos recursos acadêmicos da família nos ganhos dos testes do PISA. Esses dois fatores contribuem para a percepção de que a formulação de políticas públicas estritamente voltadas para a formação de professores em prol da melhoria da qualidade da educação é falha, tendo em vista que há outros fatores relevantes associados ao desempenho.

A pesquisa de Andrews et al. (2014) evidencia que há uma grande necessidade de realizar mais estudos para compreender essa diferença, mas conclui que ela está mais relacionada com o currículo e a cultura do país do que com relação à qualidade do ensino em sala de aula.

Em contrapartida, há estudos que destacam que, se há diferenças nas classificações entre os resultados do PISA e TIMSS, isso ocorre porque alguns países não prepararam os alunos para a aplicação da matemática escolar na vida cotidiana. Por outro lado, existem países que não prepararam os alunos também em áreas especializadas da matemática, como álgebra e geometria, pois os prepararam para resolver matemática para

⁸ Disponível em: <https://finland.fi/pt/?fact=100-taxa-de-alfabetizacao-na-finlandia>. Acesso em: 14 out. 2019.

problemas na vida cotidiana (Hutchison & Schagen, 2007; Wu, 2010). Isso leva a uma análise apenas de conteúdo para diferenciar as duas avaliações – uma análise mais simplista, sem pensar em outros fatores que influenciam os resultados das avaliações.

O estudo de Andrews et al. (2014) corrobora com algumas críticas levantadas pela literatura ao exame do PISA. Para alguns, a avaliação internacional faz parte de uma imposição cultural, na qual se toma como modelo de cidadania os países desenvolvidos (Díaz-Barriga, 2018) e geram-se comparações entre situações não comparáveis, tendo em vista a diversidade de países, com diferentes economias, culturas e língua (Díaz-Barriga, 2018; Fernandes, 2010b; Harlen, 2001). Além disso, esse exame não seria um instrumento que permite melhorar as aprendizagens, tendo em vista a sua amplitude, não impactando, assim, o trabalho desenvolvido em sala de aula (Díaz-Barriga, 2018). Não se produzem também orientações claras que possam orientar as políticas educativas (Fernandes, 2010b), de modo que o exame é visto como uma competição, com vencedores e perdedores, na qual se destacam em *ranking* os melhores e piores (Lange, 2006). Por fim, o conteúdo desenvolvido é um ponto de preocupação que recebe pouca atenção (Lange, 2006).

É de extrema importância um olhar crítico sobre os testes internacionais, pois hoje eles ganharam uma força simbólica no sentido de seu desempenho estar estritamente relacionado com o crescimento econômico e prestígio nacional. Faz-se necessário realizar um estudo sobre o que realmente eles medem e qual é a relação entre suas medidas e as necessidades econômicas e sociais do país. Porém, o que se vê atualmente é um pacote pronto de sucesso a ser alcançado, sem muitas considerações sobre como foi construído, em qual contexto foi produzido e o que anuncia.

Não é uma boa ideia usar resultados de testes internacionais transversais, como os resultados do PISA, para fazer generalizações abrangentes sobre o que funciona na educação. . . . O estudo também mostra que, se os formuladores de políticas devem investir esforço e dinheiro em reformas, devem ter estimativas muito mais precisas e menos tendenciosas do que os resultados internacionais e nacionais transversais podem fornecer (Carnoy et al., 2016, p. 30, tradução nossa).

Em contrapartida, pontos favoráveis ao exame são considerados pela literatura, tais como: possui uma ótima coleção de itens que compõem os instrumentos (Lange, 2006); possui uma inovadora, ambiciosa e confiável estrutura de avaliação (Lange, 2006);

contribui com novos modelos de governança global em educação (Niemann, Martens & Teltemann, 2017; Sellar & Lingard, 2014).

Em pesquisa realizada por Michel (2017), verifica-se como o PISA, com base na análise de suas avaliações, apresenta influência crescente sobre as políticas educacionais no âmbito nacional e subnacional em diferentes países. O autor destaca três razões que fizeram o programa ter tamanha influência: a avaliação de competências básicas em alfabetização, matemática e ciências de estudantes de 15 anos de idade, ou seja, por volta do final do ensino obrigatório em muitos países; a avaliação ser baseada numa confiabilidade metodológica e os testes serem embasados em pesquisas qualitativas e estudos; e os resultados levarem a recomendações e divulgação pela mídia na maioria dos países. Destaca-se, assim, a confiabilidade, as recomendações sugeridas pela organização e a pressão midiática fundamentada nos resultados, como influências para a mudança nas políticas públicas educacionais dos países envolvidos, levando, por exemplo, à reformulação de currículos, maior autonomia das escolas e busca por inovações, como observado na França.

Outro estudo recente sobre o impacto do PISA na política educacional de um país foi realizado por Niemann et al. (2017). Os autores afirmam que, por meio da avaliação, a OCDE conseguiu influenciar de forma significativa os sistemas nacionais de educação de diferentes países, tendo em vista que os Estados foram confrontados com aconselhamentos externos baseados em dados empíricos. O estudo toma como caso a Alemanha, entendendo que o PISA pode ser considerado um “divisor de águas” na formulação de políticas educacionais do país. Com um sistema educacional tradicional, sendo desenhado com base em legados históricos profundamente arraigados, após a divulgação dos resultados do PISA no final de 2001, houve uma surpreendente reforma educacional no país, ao perceberem que, comparados com outros países europeus, o sistema educacional alemão estava ficando para trás. A Alemanha, então, passou a pensar o currículo com foco nas competências adquiridas; os métodos de governança passaram a ser orientados para resultados, refletindo uma abordagem política de responsabilização; a governança de cima para baixo, monitorada centralmente, foi reduzida, reforçando a autonomia da escola; e vários projetos foram lançados para combater discrepâncias na educação. Assim, mudanças significativas, com base em uma governança mais global, foram instituídas no país, para alcançar melhoria nos resultados da educação.

Dada a relevância dessa avaliação e os diferentes posicionamentos, é pertinente afirmar que ela possui o seu lugar – o de espelhar o que ocorre num determinado

momento, a partir do que se propõe avaliar. É uma referência, dentre diversas outras, que os governos nacionais podem utilizar para olhar para o seu sistema educacional e verificar, com base no cruzamento de diferentes dados, quais mudanças devem ser feitas.

O PISA no Brasil.

O Brasil participou de todas as sete edições aplicadas até o momento do PISA, o que possibilita um estudo comparativo ao longo dos anos. Outros países da América Latina também foram avaliados ao longo das edições, no entanto, não participaram de forma constante de todas as edições como o Brasil. Tal fato pode ocorrer devido ao alto custo financeiro exigido para a participação na avaliação ou por não considerarem o diagnóstico oferecido algo necessário para a melhoria de seu sistema educacional. Cabe destacar que, segundo Araújo e Tenório (2017), a entrada no Brasil no PISA ocorreu num momento em que havia um interesse do país em ser membro da OCDE. Com isso, participar do PISA seria um meio necessário para alcançar *status* de membro da organização.

A realização do PISA no Brasil buscou alcançar os seguintes objetivos:

1. obter informações para situar o desempenho dos alunos brasileiros no contexto da realidade educacional, nacional e internacional;
2. fomentar a discussão sobre indicadores de resultados educacionais adequados à realidade brasileira;
3. participar das discussões sobre as áreas de conhecimento avaliadas pelo PISA em fóruns internacionais de especialistas;
4. promover a apropriação de conhecimentos e metodologias na área de avaliação educacional;
5. disseminar as informações geradas pelo PISA, tanto em termos de resultados quanto em termos de conceitos e metodologias, entre diversos atores do sistema educacional, governamentais e não governamentais. (Inep/PISA, 2001)

Em sua última edição, o PISA contou com a participação de mais de 60 países de diferentes economias, tornando-se uma importante referência em avaliação em larga escala. O comparativo entre as médias das avaliações e os fatores associados pode ser realizado entre quaisquer países, mas tende a ocorrer mais entre os países com o Índice de Desenvolvimento Humano (IDH) semelhantes. É divulgado um *ranking* do

desempenho de todos os países participantes a cada edição. O Brasil encontra-se nas últimas colocações, juntamente com outros países da América Latina. Dois ciclos de ênfase em cada disciplina já ocorreram, o que possibilita uma comparação do desempenho entre duas avaliações. Em todas as três disciplinas, o Brasil conseguiu aumentar as médias. No entanto, de acordo com Araújo e Tenório (2017, p. 362),

É superficial alardear que somos o país que apresentou maior crescimento nos resultados do PISA, principalmente no que se refere à Matemática, sem divulgar a real situação da educação brasileira mesmo após esse crescimento – ainda no nível 1 da escala de proficiência –, e sem discutir os ajustes feitos no processo de execução do PISA.

A avaliação oferece resultados de desempenho da federação, das regiões (sudeste, nordeste, norte etc.) e das unidades da federação, o que possibilita uma maior análise dos resultados, tendo em vista a extensão e diversidade do território brasileiro. No relatório nacional, em que se realiza uma leitura dos dados por região e unidades da federação, a comparação é feita entre as unidades da federação e/ou entre outros países. Como exemplo, na análise do desempenho do nível de dificuldade dos 181 itens da avaliação do PISA 2015, os jovens brasileiros, tiveram um alto nível de dificuldade nos itens.

Verifica-se que, em média, o nível de dificuldade dos 181 itens do PISA 2015 para os jovens brasileiros (15,03) foi maior que para os de outros países, exceto os da República Dominicana e do Peru. Em geral, os alunos do Uruguai e do Chile foram os que obtiveram os maiores percentuais de acerto em comparação com os outros da América Latina, com índices Delta de 14,36 e 14,13, respectivamente. Dentre as unidades da Federação brasileiras, os jovens do Espírito Santo e do Distrito Federal apresentaram as médias de respostas corretas mais próximas desses dois países latino-americanos (Inep/PISA, 2015, p. 52).

O PISA é uma avaliação externa internacional que apresenta um elevado impacto na sociedade, tanto no que diz respeito à divulgação governamental como à midiática. Contudo, no Brasil, percebe-se uma maior divulgação da mídia em detrimento da divulgação governamental. Isso pode ocorrer devido às baixas médias do país em todas as edições dessa avaliação, o que favorece um posicionamento pessimista da mídia sobre a educação brasileira e a não utilização por parte do governo em campanhas. No que diz respeito à mídia, Araújo e Tenório (2017) classificam como um uso indevido, tendo em

vista que há uma preocupação em alardear a posição do Brasil no *ranking* da OCDE, porém, em geral, não se apresentam pesquisas ou dados mais aprofundados para oferecer ao público subsídios necessários à realização da avaliação dos resultados.

Com relação ao governo, de acordo com Díaz-Barriga (2018), deve haver uma clareza por parte dos países que realizam a aplicação do PISA, no sentido de não somente realizar a divulgação de resultados, mas sim de uma responsabilização pelo compromisso de impulsionar modificações no âmbito pedagógico-didático no seu sistema educacional. Caso contrário, o ciclo de resultados insatisfatórios se manterá.

De acordo com Bieber e Martens (2011), a OCDE deriva uma série de recomendações de políticas a partir dos resultados do PISA, no entanto, o impacto dessas recomendações nos países participantes ainda não foi analisado com profundidade suficiente. Os autores afirmam que a recepção dessas recomendações difere entre os países. Ao realizarem uma comparação entre a Suíça e os EUA, os pesquisadores concluíram que, para aquele país, o PISA possibilitou um aprendizado no nível dos especialistas governamentais, levando, assim, a mudanças em suas políticas educacionais, como também observado anteriormente nos casos da França e da Alemanha (Michel, 2017; Niemann et al., 2017). Já nos EUA, o exame foi considerado apenas mais um estudo avaliativo do desempenho dos sistemas educacionais, não causando nenhum impacto em suas políticas educacionais.

Lingard (2016), em pesquisa realizada sobre o PISA, para além de uma análise do programa, descreve como as nações fundamentam a participação na avaliação e analisa os usos que são feitos do PISA em três países: EUA, Inglaterra e Austrália. O autor conclui que os resultados são utilizados mais no sentido de legitimar reformas políticas que são estruturadas por imperativos políticos internos e ideologias do que como um aprendizado político provindo de estudo crítico dos resultados.

Ao realizar uma pesquisa sobre os usos do PISA no Brasil, Araújo e Tenório (2017), com base em análise documental e entrevistas semiestruturadas com gestores da educação básica e envolvidos com o PISA no Brasil, concluíram que um dos usos mais evidentes consiste no estabelecimento da meta para o Ideb, e, em segundo lugar, o ganho em termos de *know-how* em avaliação, ambos classificados como usos indiretos da avaliação. Com relação à formulação de políticas públicas, o estudo evidencia que não houve qualquer análise específica sobre os resultados da avaliação para o direcionamento de políticas educacionais brasileiras. Os resultados da avaliação podem até ter sido

considerados no momento da elaboração de propostas governamentais, contudo, de forma genérica, sem uma verdadeira análise dos resultados e seus significados.

Outros estudos realizam uma análise da estrutura da prova e do desempenho dos estudantes brasileiros: análise dos itens e desempenho em matemática (cf. Ortigão, 2018; Tenório & Araújo, 2016); análise dos itens e desempenho em ciências (Antunes & Galvão, 2015; Waiselfisz, 2009); letramento científico no exame (cf. Vieira, 2017; Duboc, 2007); evolução do desempenho dos alunos ao longo dos anos (cf. Soares & Nascimento, 2012);

Contudo, apesar de haver estudos sobre as três áreas contempladas pela avaliação, Dickel (2010), em análise sobre o impacto do PISA na produção acadêmica brasileira, afirma que há um número reduzido de produções encontradas que problematizam com maior profundidade e rigor o efeito dessa avaliação no currículo escolar.

Outros estudos confrontam o PISA com o Saeb, realizando comparações entre um sistema de avaliação internacional e outro nacional. Na área de linguagem, encontra-se o estudo de Bonamino, Coscarelli e Franco (2002), que busca contribuir para a compreensão dos critérios de letramento que estão subjacentes às provas de língua portuguesa do PISA e do Saeb. No campo da matemática, Nunes et al. (2014) realizam um estudo comparativo entre as avaliações com o objetivo de destacar semelhanças e diferenças na avaliação da competência matemática, analisando, por conseguinte, o conceito de letramento matemático presente nas duas avaliações. Por fim, em ciências da natureza, Godinho e Farias (2013) desenvolvem um estudo com a finalidade de realizar uma análise dos conteúdos de ciências nessas avaliações e seus reflexos nas políticas públicas voltadas para a formação continuada dos docentes no Brasil.

No que tange aos questionários contextuais, Santos (2015) realiza uma análise dos questionários de ambas as provas para compreender em que medida são capazes de fornecer informações e características sobre elementos que constituem o ensino, tais como: o conhecimento, a experiência, o envolvimento do professor, o clima da classe, a gestão da classe e a gestão da matéria.

Os resultados das avaliações também são utilizados como base para estudos. Carnoy, Khavenson, Fonseca, Costa e Marotta (2015) realizam um estudo comparativo com resultados das duas avaliações, com o objetivo de verificar se a educação brasileira está melhorando. Soares e Candin (2007) utilizam os resultados do PISA e do Saeb para medir o efeito das escolas brasileiras de educação básica em dois sentidos: sua capacidade

de elevar o desempenho dos alunos e a de reduzir a desigualdade de desempenho entre eles.

Com isso, destaca-se a importância do PISA no contexto brasileiro, porém muito mais com relação aos seus resultados, sem muita reflexão, do que como uma proposta de uso da avaliação como um fator de indução da política educacional, tal qual assevera o programa.

Avaliação Externa regional – América Latina.

O Laboratório Latino-americano de Avaliação da Qualidade da Educação (LLECE) surgiu com base em um esforço regional de ampliar a cobertura dos sistemas educativos, reduzir o analfabetismo e introduzir reformas para a melhoria da qualidade da educação. Trata-se de “um recurso técnico que a UNESCO coloca à disposição dos países latino-americanos, constituindo-se um âmbito técnico-político de discussão da problemática do aprendizado e de suas variáveis relacionadas.” (Castro, Pestana & Marte, 1995, p. 51), contribuindo com o sistema de avaliação externa na região. O laboratório apresenta como principais objetivos: a identificação e a avaliação dos padrões do aprendizado escolar para a região; o fomento às mudanças educativas que permitam alcançar tais padrões; a formação de recursos humanos que possibilitem as mudanças (Castro et al., 1995, p. 52). Para tal, o LLECE desenvolve cinco componentes principais:

1. redesenhar, elaborar e colocar em prática um sistema de avaliação da qualidade dos aprendizados em língua e em matemática, nos níveis de educação básica dos países da região;
2. estabelecer um sistema de acompanhamento, monitoração e disseminação dos resultados da avaliação, de modo a motivar e a capacitar os dirigentes educacionais na utilização desses resultados, na tomada de decisões sobre reformas e políticas educativas;
3. desenvolver um programa de investigações sobre os fatores associados à qualidade da educação básica;
4. fortalecer a capacidade técnica dos Ministérios da Educação, na área de avaliação da qualidade;
5. fortalecer o Laboratório Latino-americano de Avaliação da Qualidade da Educação como instância de execução do projeto, de coordenação dos países participantes e de prestação de assistência técnica a estes. (Castro et al., 1995, p. 52)

O laboratório entendeu que um estudo comparativo entre os países latino-americanos era necessário, sendo um instrumento importante para compreender o estado da educação nos diferentes países dessa região. Tal esforço contribuiria para a integração cultural desses países, levando em consideração que a maior parte dos estudos internacionais realizados não inclui os países latino-americanos ou não considera suas matrizes curriculares (UNESCO-Santiago, 1998).

O Primeiro Estudo Regional Comparativo e Explicativo (Perce) ocorreu em 1997, com 13 países da América Latina. Foram aplicadas provas de língua (português ou espanhol) e matemática para alunos dos últimos anos do ensino fundamental I, 4º e 5º anos, tendo como base os elementos do currículo escolar comuns aos países da região. Também foram aplicados questionários aos alunos, professores, diretores e tutores dos alunos. O estudo objetivava oferecer dados para diferentes grupos: aqueles que elaboram e gerem as políticas educacionais; os sujeitos educativos, no que tange à família e à escola; por fim, a comunidade acadêmica e os pesquisadores.

O Segundo Estudo Regional Comparativo e Explicativo (Serce) ocorreu nove anos após o primeiro, em 2006. Nesse estudo, como novidade, para além da avaliação de língua e matemática, também foi contemplada a disciplina de ciências da natureza. Os anos de escolaridade também mudaram, a amostra correspondeu aos alunos do 4º e 7º anos. Houve a participação de 16 países mais o estado de Nuevo León, no México (UNESCO-Santiago, 2008).

O Terceiro Estudo Regional Comparativo e Explicativo (Terce) ocorreu apenas em 2013, com a participação de 15 países mais o estado de Nueva León, no México. O Terce continuou com as mesmas disciplinas e anos da segunda avaliação.

Apesar de haver uma constância de objetivos, normativas e metas com relação à gestão, fatores externos e aprendizagens, não há possibilidade de realizar comparação direta entre as três avaliações do LLECE, tendo em vista o conjunto de inovações presentes no Serce.

O laboratório constrói os objetivos de suas avaliações com destaque para o estudo como um todo, não apenas para o desempenho do aluno nas avaliações, mas levando em consideração também os fatores associados. Sua essência evidencia uma construção de caráter mais colaborativo, na medida em que pressupõe produzir informações para gerar conhecimento e contribuir com novas ideias para as unidades de avaliação. Essa colaboração pode ser vista na construção de suas matrizes de referência. No documento de análise curricular para o Terce, há um relato de como ocorreu a atualização das

matrizes de referências para o terceiro ciclo de avaliação, informando não apenas as alterações, mas as opiniões e posicionamentos de cada país participante (UNESCO-LLECE, 2014).

No que diz respeito às políticas públicas, o LLECE realiza uma análise do desempenho dos países e orienta políticas públicas. De acordo com Comar (2017), há orientações políticas de alfabetização, reestruturação curricular, formação e certificação de professores, aquisição de equipamentos tecnológicos, entre outras sugestões que compõem o cenário de reformulações no campo educacional.

O LLECE no Brasil.

O Brasil apresentou como maior motivação para participar das avaliações do LLECE a oportunidade de complementar o Saeb através da incorporação de modernas técnicas e metodologias que o Laboratório coloca à disposição dos países-membros, além do enriquecimento do banco de itens nacional, modelos de formulários e questionários, maior debate sobre padrões de qualidade e criação de marcos comparativos.

Nas avaliações aplicadas pelo LLECE, cerca de 10 países participaram da última edição. Com IDH e realidade próximas, a comparação entre as nações participantes torna-se mais significativa. Não é possível realizar uma análise do desempenho dos países nas três avaliações aplicadas ao longo dos anos, pois a escala de proficiência não foi a mesma nas três edições, apresentando, por conseguinte, médias e desvio padrão diferentes. O que se pode realizar é uma comparação com relação à média geral de cada avaliação. O Brasil esteve, na maioria das vezes, no grupo de países que obteve pontuação um pouco acima da média geral.

Na avaliação desenvolvida pelo LLECE, não ocorre uma análise por região ou unidades da federação, as comparações são realizadas apenas entre os países, com os resultados da federação. Nas três avaliações aplicadas pelo LLECE, apenas na segunda edição o Inep desenvolveu um relatório nacional. Com isso, observa-se que os dados fornecidos pelo PISA são mais detalhados e utilizados pelo governo brasileiro.

Em análise sobre a concepção de avaliação adotada pelo LLECE, Comar (2017) realiza também um estudo de como o Brasil se insere nessa concepção de avaliação. Como resultado, o autor entendeu que, no Brasil, a avaliação em larga escala converge para as orientações do laboratório nas últimas décadas, a consolidação de um perfil meritocrata e gerencial para as práticas avaliativas.

Não foram encontrados estudos que realizem comparação entre o PISA ou o Saeb e as avaliações realizadas pelo LLECE. Acredita-se que se deva ao fato de não ter uma periodicidade definida e/ou não apresentar um padrão estatístico nas diferentes edições, não tendo dados consistentes para realizar comparação. Até mesmo com relação ao PISA, uma avaliação amplamente divulgada, com mesmo padrão em todas as edições e periodicidade definida, considera-se que há poucos estudos que realizam comparações com o Saeb.

Os países latino-americanos, ao participarem das avaliações internacionais, acabam por “sofrer” duplamente com os resultados da avaliação, tendo em vista que veem comprovados os seus problemas de qualidade educativa, que na maioria das vezes já são evidenciados no sistema nacional, bem como expostos a um juízo duplo (interno e externo) e a uma comparação com seus pares na região. Assim, os governantes devem não apenas explicar e justificar o baixo rendimento de seus alunos, mas também por que têm menor rendimento que outros países da América Latina. Para além disso, há o custo elevado na realização das provas e os possíveis efeitos que tais avaliações impõem aos países menos desenvolvidos. Por outro lado, a participação nas avaliações internacionais tem proporcionado maior *know-how* e o fortalecimento da capacidade técnica do sistema de avaliação nacional, realizando um grande desenvolvimento tecnológico e metodológico, o que contribui de forma positiva (Carrasco & Torrecilla, 2016; Perez & Passone, 2007).

Por isso, a decisão por continuar a participar do sistema de avaliação internacional não é fácil, pois “se continuam, correm sempre o risco de ser ‘mal avaliados’, enquanto a autoexclusão implica ficar à margem de um importante espaço de reflexão, de troca e de aprendizagem em matéria de avaliação da qualidade educativa.” (Carrasco & Torrecilla, 2016, p. 36).

Cabe, por conseguinte, ressaltar a necessidade de um maior estudo e reflexão sobre os dados produzidos pelas avaliações internacionais, tendo em vista que o Brasil participa de duas grandes avaliações, uma no âmbito mundial e outra regional. O objetivo é que, com isso, venha estabelecer metas e prioridades estratégicas, identificar problemas e potencialidades de seu sistema. Infelizmente, o que se vê hoje é a baixa contribuição dos dados para a formulação de políticas públicas e mudanças efetivas nas escolas. Talvez um pouco mais no que tange ao aporte técnico das avaliações, no entanto, é importante destacar que ter ou participar de sistema de avaliação com alto nível técnico não garante a melhora da qualidade da educação no país.

Por outro lado, convém ressaltar, que os dados internacionais são mais abstratos, por isso, como afirma Fernandes (2010b), deve-se olhar para esses estudos internacionais com interesse, mas também com os devidos cuidados. Não se pode exigir o que eles não são capazes de dar e que dificilmente darão, isto é, orientações concretas para o desenvolvimento de políticas educativas e práticas escolares. Deve-se, por conseguinte, realizar estudos e análises contextualizadas, a partir de cruzamento de dados, para que, assim, possam contribuir com o sistema nacional de educação.

Avaliações Externas Nacionais – Brasil

Até a década de 80, não houve movimentos de avaliação do sistema de ensino no Brasil, tendo em vista que a avaliação estava centrada no aluno. Uma das primeiras iniciativas a serem pensadas como avaliação em larga escala foi o Programa de Expansão e Melhoria do Ensino no Meio Rural do Nordeste Brasileiro – EDURURAL (Gatti, Vienna & Davis, 1991; Horta Neto, 2007). O estudo foi de caráter transversal e amostral e ocorreu nos anos de 1981, 1983 e 1985, em 60 municípios dos estados de Piauí, Pernambuco e Ceará. A amostra integrou cerca de 600 escolas e 6.000 alunos. O primeiro ano serviu como base para a verificação da aprendizagem escolar, gerando investimento na região nos anos subsequentes. Como base para tal verificação, levou-se em consideração o desempenho dos alunos nas provas de rendimento acadêmico, que foram aplicadas nas disciplinas de português e matemática em alunos das 2ª e 4ª séries do ensino fundamental. Outro ponto que a avaliação objetivava era apreender a forma de atuação da escola em seu contexto; para tal fim, foram realizados seis estudos de caso.

Em 1984, iniciou-se o Projeto Nordeste pela Secretaria Nacional de Educação Básica, que estabeleceu parceria entre o Banco Mundial, os nove governos estaduais nordestinos e o governo federal. O projeto tinha duração prevista de cinco anos, e desenvolveria nove subprojetos, um para cada estado, e um nacional. No projeto nacional, havia uma componente de avaliação sistemática dos resultados escolares (Horta Neto, 2007).

Em 1987, deu-se início ao Sistema de Avaliação do Ensino Público de Primeiro Grau (Saep), com a finalidade de “estabelecer um sistema de controle dos conteúdos mínimos e dos padrões de qualidade do ensino básico” (MEC/Inep, 1994), podendo, assim, dar subsídio às Secretarias de Estado da Educação com informações sobre os programas de ensino. Foram avaliadas, no total de três fases, 238 escolas em 69 cidades e 27.455 alunos. Não era objetivo do programa abarcar a totalidade de alunos

matriculados na rede de ensino, sendo as escolas selecionadas a partir da indicação das Secretarias de Educação, com critérios pré-estabelecidos.

Não havendo um currículo único nacional adotado por todas as escolas, os programas para a realização das avaliações foram pensados juntamente com uma amostra de docentes do 1º grau e, subsequentemente, ratificados pela Secretaria de Educação. A avaliação foi composta por provas de português e matemática para os alunos do 1º grau. Devido a problemas orçamentários, a avaliação nacional ocorreu apenas em 1990 (Gatti et al., 1991). Com a finalidade de seguir a nova nomenclatura adotada na Constituição de 1988, o Saep, a partir de 1991, passa a designar-se Sistema de Avaliação da Educação Básica (Saeb), tendo em conta que, posteriormente, poderia incluir alunos do 2º grau. Nesses termos, a avaliação externa acaba por ganhar força no Brasil na década de 90, seguindo uma tendência internacional relacionada com o monitoramento da qualidade da educação básica pública.

Sistema de Avaliação da Educação Básica (Saeb).

Assim como afirma Pestana (1998, p. 15), o desenvolvimento do Saeb ao longo dos anos veio a sofrer muitas alterações e, até hoje, como qualquer outro processo, “não está pronto, tão cedo não estará; talvez nunca esteja pronto e acabado”. O programa, até 2018, foi composto por três avaliações externas em larga escala, a saber: Avaliação Nacional da Educação Básica (Aneb); Avaliação Nacional do Rendimento Escolar (Anresc/Prova Brasil); e Avaliação Nacional da Alfabetização (ANA).

Atualmente, o Saeb apresenta como principal objetivo

avaliar a Educação Básica brasileira e contribuir para a melhoria de sua qualidade e para a universalização do acesso à escola, oferecendo subsídios concretos para a formulação, reformulação e o monitoramento das políticas públicas voltadas para a Educação Básica. Além disso, procura também oferecer dados e indicadores que possibilitem maior compreensão dos fatores que influenciam o desempenho dos alunos nas áreas e anos avaliados (Inep, n.d.).

Outras avaliações em larga escala ocorrem no Brasil, como: Provinha Brasil, Exame Nacional para Certificação de Competências de Jovens e Adultos (Encceja), Exame Nacional do Ensino Médio (Enem), Prova Docente, entre outras. Essas avaliações não serão objeto de estudo desta pesquisa, tendo em vista que o objetivo é abordar o Saeb,

no que se refere ao Aneb/Saeb e a Prova Brasil, e as avaliações do estado e do município do Rio de Janeiro, que são aplicadas no ensino fundamental II.

Desde a primeira avaliação aplicada pelo Saeb, já se passaram quase 30 anos e 15 ciclos de avaliação. Ao longo desse período, em busca de um aperfeiçoamento do sistema de avaliação, reformulações foram feitas. Com isso, para uma melhor compreensão da configuração atual, acredita-se ser pertinente uma análise, de forma sucinta, de alguns dados relevantes ao longo de sua história.

A primeira avaliação consistiu numa bateria de testes semiobjetivos, de português, matemática, ciências da natureza e redação. Para além disso, foram colhidas informações no âmbito da gestão escolar e da situação profissional do docente. O primeiro ciclo considerou a participação de professores e especialistas para a composição dos currículos mínimos dos conteúdos avaliados para a elaboração dos instrumentos de medida e para a correção das provas.

Nas diretrizes apresentadas nos dois primeiros ciclos de aplicação da avaliação, observa-se um movimento mais descentralizado no que se refere à elaboração e aplicação do Saeb. Isso ocorre porque há efetiva participação dos professores regentes, por estarem ligados mais estritamente ao processo em sala de aula; de especialistas das secretarias estaduais, por apresentarem maior ligação com o sistema educativo; e de especialistas acadêmicos de diferentes universidades, vindo a corroborar com análises e pesquisas educacionais (Bonamino & Franco, 1999).

A elaboração do currículo mínimo foi feita de forma conjunta, buscando a participação dos docentes e a análise das diferentes propostas curriculares dos estados, considerada como diretriz indutiva (Bonamino & Franco, 1999). A implantação do questionário sobre oportunidades de aprendizagem estabeleceu um canal de comunicação entre a secretaria estadual e os professores, visto que teve por objetivo o aperfeiçoamento das propostas curriculares através da análise dos conteúdos abordados em sala de aula e o que efetivamente encontrava-se no currículo proposto. A análise interligada do rendimento do aluno, da gestão escolar e da situação e competência do professor buscou uma sistemática de avaliação mais ampla.

A orientação para uma ação do sistema avaliativo como um recurso para avaliação dos resultados de políticas e estratégias educacionais foi um movimento que começou a se delinear no segundo ciclo. Em análise dos objetivos do primeiro ciclo, pode-se notar a busca pela implantação de uma cultura de avaliação. Não apareceu, num primeiro

momento, a utilização do sistema de avaliação como recurso no que diz respeito às políticas educacionais.

O terceiro ciclo, aplicado em 1995, foi um momento de mudanças para o sistema de avaliação, tendo em vista que foi o primeiro ciclo em que todas as unidades da federação participaram voluntariamente e houve aplicação na rede particular de ensino, fatos que levaram a um crescimento da cultura de avaliação no país. Nova metodologia de elaboração de testes, técnicas estatísticas de análise e interpretação de resultados se estabeleceram. Passou-se a utilizar a TRI, um modelo estatístico que emprega, no programa analisado, três parâmetros de análise: o grau de dificuldade de cada item; a capacidade que o item tem de discriminar diferentes grupos de estudantes que o acertaram ou não; e a possibilidade de acerto ao acaso. Uma grande vantagem dessa teoria é a possibilidade de se comparar resultados entre diferentes avaliações, o que possibilita um estudo longitudinal.

Ao longo dos outros ciclos, a concepção e a estrutura se mantiveram, em linhas gerais, constantes. A partir do ano de 2001, o Saeb passou a avaliar apenas português e matemática em todos os anos. No mesmo ano, houve também uma atualização das matrizes de referência. Em 2005, o Saeb foi reestruturado e passou a ser composto por duas avaliações: Avaliação Nacional da Educação Básica (Aneb) e Avaliação Nacional do Rendimento Escolar (Anresc). Nas provas e nas publicações, as avaliações são conhecidas como Saeb e Prova Brasil, respectivamente. Em 2013, foi incorporada ao Saeb a Avaliação Nacional de Alfabetização (ANA), aplicada anualmente nas turmas de 3º ano do ensino fundamental, de caráter censitário. E assim seguiu sem grandes alterações até a aplicação de 2018.

No entanto, em junho de 2018, o MEC divulgou que o sistema de avaliação nacional sofreria mudanças para o ano de 2019. Dentre algumas mudanças anunciadas, encontram-se: a avaliação da educação infantil; a inclusão das disciplinas de ciências da natureza e ciências humanas para os alunos do 9º ano; os diferentes nomes das avaliações que compõem o sistema deixarão de existir e todas as avaliações passam a ser identificadas pelo nome de Saeb (MEC, 2018).

Ao longo dos sete primeiros ciclos – de 1990 a 2003 –, o Saeb corroborou com dados substanciais no que tange à realidade brasileira, servindo como base para a formulação e reformulação de políticas públicas. Porém, ao longo desses mesmos treze primeiros anos, o programa também veio a receber críticas da comunidade acadêmica. Cabe destaque às considerações de Vianna (2002) de que, mesmo após o primeiro decênio

de levantamento de dados, o Saeb não conseguiu instituir uma educação de qualidade; devido ao seu caráter amostral, seus resultados reverberam apenas no sistema.

Com base em considerações diversas e necessidades apresentadas ao longo dos anos, o Inep decidiu pela expansão do Saeb, apresentando sua divisão em dois processos de avaliação: a Avaliação Nacional da Educação Básica (Aneb) e a Avaliação Nacional do Rendimento Escolar (Anresc/Prova Brasil). No que diz respeito à Aneb, ficou estabelecido que se manteriam os objetivos e procedimentos já adotados pelo Saeb. No que corresponde à Anresc, o seu direcionamento se dá para as unidades escolares. Com isso, o que outrora vinha como dados que mais favoreciam ao sistema, ou seja, o Saeb como fonte de informação para gestores públicos, com ênfase nas políticas públicas, a Anresc/Prova Brasil passa a fornecer dados por unidade escolar. Não há divulgação de resultado por aluno, mas sim, pela unidade escolar. Tal fato contempla duas dimensões, na medida em que auxilia os gestores públicos com instrumento para tomada de decisões e também a escola, pois a nova forma de apresentação dos resultados possibilita uma análise individual e comparativa.

As matrizes de referência e as escalas de proficiências das avaliações do Saeb – Aneb e Anresc – são iguais. Nos anos em que ocorrem as avaliações, as secretarias estaduais e municipais de educação, como também as escolas públicas de educação básica que possuem os anos de escolaridade avaliados pelo sistema, recebem um material que contém as matrizes de referência, temas, tópicos e descritores. Para além desses conteúdos, o material apresenta informações para professores e gestores sobre alguns sistemas de avaliação externa, como PISA, Enem, Encceja, Provinha Brasil e Saeb; os pressupostos teóricos que embasam a avaliação; e uma série de exemplos de itens das disciplinas avaliadas.

Assim, a Prova Brasil inicia com o objetivo de afunilar o olhar do sistema para a escola. Tal característica, segundo Souza (2013), foi um forte poder indutivo para a implementação de sistemas próprios de avaliação por parte dos estados e municípios, pois colocou em evidência os resultados obtidos por redes de ensino e escolas, levando as secretarias locais, por conseguinte, a uma busca por melhores resultados.

No ano de 2007, referente às avaliações externas nacionais, foi criado o Índice de Desenvolvimento da Educação Básica (Ideb), um indicador nacional que permite o acompanhamento da qualidade da educação. Ele é calculado a partir dos dados sobre aprovação escolar, obtidos no Censo Escolar, e das médias de desempenho nas avaliações

do Inep, do Saeb para as unidades da federação e para o país, e da Prova Brasil para os municípios. O índice serve para traçar metas de qualidade educacional dos sistemas.

A utilização dos dados das avaliações externas da aprendizagem nos estudos comparativos pode contribuir para o alcance de diferentes objetivos que visam à melhoria da qualidade da educação. Nesse caminho, o Brasil desenvolve, de forma constante, avaliações externas nacionais há aproximadamente 30 anos e participa das avaliações regionais e internacionais há aproximadamente 20 anos, um tempo médio considerável para estabelecer um comparativo de desempenho entre as médias do país ao longo do tempo e também com relação às médias de outros países.

O Saeb é anterior ao PISA e as avaliações aplicadas pelo LLECE. Desde 1990, ocorrem avaliações nacionais no Brasil. A primeira avaliação do LLECE, o Perce, surgiu 7 anos depois do Saeb e o primeiro exame do PISA, apenas dez anos mais tarde. Um dos objetivos de se participar das avaliações internacionais é o de fortalecimento das avaliações nacionais, tanto em dados metodológicos quanto em discussões sobre a área de conhecimento em avaliação. O LLECE apresenta um objetivo específico no que diz respeito ao fortalecimento das capacidades avaliativas dos países participantes. O PISA não apresenta esse objetivo, mas o Brasil o definiu em sua justificativa para participar do exame. Um objetivo comum tanto nas avaliações internacionais como no sistema nacional é a avaliação como forma de fornecimento de dados para as políticas públicas.

As avaliações internacionais e nacionais avaliam as mesmas áreas de conhecimento: linguagem, matemática e ciências da natureza. As avaliações do LLECE e o Saeb apresentam conteúdos que fazem parte do currículo escolar. O LLECE, através do diálogo e aprovação dos conteúdos com os diferentes países, leva em consideração só os elementos comuns das propostas curriculares dos países participantes. O Saeb realiza as avaliações com base numa matriz de referência própria, com conteúdos mínimos propostos para cada série avaliada. Diferente do PISA, que não está relacionado diretamente com o currículo escolar.

Como já referido, a avaliação nacional realizou uma indução de cultura de avaliação, objetivo proposto no primeiro ciclo do Saeb, realizando um aumento das avaliações nos estados e municípios. Tal fato acentuou-se, principalmente, depois da criação da Prova Brasil, que passou a apresentar o resultado por escola, aumentando, assim, a visibilidade do desempenho das unidades educacionais. Nas linhas que se seguem, realiza-se uma descrição e análise de como o Rio de Janeiro construiu seu

sistema próprio de avaliação, tanto no âmbito estadual quanto no municipal, e como esse sistema veio a se desenvolver ao longo dos anos.

Avaliações Externas Estaduais – Estado do Rio de Janeiro

A política descentralizada no primeiro ciclo do Saeb propiciou um direcionamento para a criação de sistemas próprios de avaliação dos estados brasileiros. Com isso, no que diz respeito à avaliação em âmbito estadual, dois estados deram início ao sistema próprio de avaliações no ano de 1992: Minas Gerais e Ceará. E, ao longo dos anos, outros estados seguiram na mesma direção.

O estado do Rio de Janeiro iniciou seu sistema próprio de avaliação em larga escala em janeiro de 2000, com o Programa Nova Escola (PNE), que buscou melhorar a gestão das escolas estaduais e responsabilizar seus diretores e funcionários pelos resultados dos alunos, sendo considerado um modelo pioneiro de responsabilização. No período de vigência do programa, de 2000 até 2008, houve cinco aplicações, nas quais ocorreram diversas modificações, apresentando dificuldades em se estabelecer objetivos e métodos ao longo do programa.

No ano 2000, sob a responsabilidade da Fundação Cesgranrio, foi avaliado o desempenho dos alunos da 3ª e da 6ª séries do ensino fundamental e do 1º ano do ensino médio, com o objetivo de iniciar um estudo longitudinal que pudesse separar os efeitos do nível socioeconômico e da aprendizagem anterior dos alunos e os efeitos da escola propriamente dita. Devido a esse objetivo, em 2001, os testes foram aplicados aos alunos da 4ª e da 7ª séries e do 2º ano do ensino médio.

A aplicação seguinte foi realizada em 2003, tendo novos instrumentos, pois houve a necessidade de implementar os modelos de itens e da escala do Saeb, para que assim pudesse haver dados comparativos do estado do Rio de Janeiro com as médias nacionais. Em consequência disso, o caráter de estudo longitudinal do programa foi deixado de lado e foram avaliados os alunos da 4ª e da 8ª séries e do 3º ano do ensino médio. Nesse ano, também foi elaborado o índice de Gestão Escolar, de Eficiência Escolar e de Aprendizagem, com o intuito de se evitar a comparação entre escolas com nível socioeconômico muito diferentes. Para isso, estabeleceram-se grupos de referência com base na média de notas padronizadas de cinco índices criados, por exemplo, a renda dos alunos.

Em 2004, a instituição que ficou responsável pelo programa foi o CAEd, e a avaliação das escolas passou a se concentrar em três análises: desempenho, fluxo escolar

e gestão. Não houve mais a utilização de grupos de referência, sendo considerados para análise cinco níveis de qualidade com base nas notas. Cada nível corresponderia a uma gratificação de valor diferente no salário, podendo chegar a R\$ 500 (Brooke, 2006).

Por utilizar o desempenho cognitivo do aluno numa avaliação externa como base para formulação de critérios para pagamento individual ou coletivo a professores ou funcionários da escola, o programa Nova Escola configura-se como uma política de alto impacto, pois impõe bônus às escolas e/ou aos profissionais de forma direta, podendo, em consequência, interferir mais na dinâmica intraescolar (Haney & Madaus, 1986; Heubert & Hauser, 1999; McDonnell, 1994).

O programa recebeu muitas críticas por parte dos professores, representados pelo Sindicato Estadual dos Profissionais da Educação (Sepe). Em agosto de 2005, o sindicato moveu três ações contra o governo do estado do Rio de Janeiro com relação ao programa, todas devido a alguma discordância em relação à gratificação dos professores. De acordo com Brooke (2006), com relação à técnica empregada, houve crítica no que diz respeito à classificação de escolas, pois não houve referência ou padrões claros que indicassem que as escolas com melhores classificações alcançaram efetivamente esse nível de qualidade. Outra crítica foi direcionada à utilização de dados do ano anterior para pautar a gratificação, o que seria um erro devido à alta rotatividade de professores de um ano para o outro. Assim, “por gerarem tanta animosidade, essas mudanças acabam por ofuscar o propósito da avaliação, que é oferecer informação aos membros da equipe escolar e permitir um diagnóstico do funcionamento da escola” (Brooke, 2006, p. 393).

Em 2008 surge o Sistema de Avaliação da Educação do Estado do Rio de Janeiro, também conhecido por Saerj. O programa teve como objetivo a monitoração do padrão da qualidade do ensino e buscou colaborar com sua melhoria. O sistema de avaliação foi dividido em dois estágios: o Programa de Avaliação Externa (Saerj)⁹ e o Programa de Avaliação Diagnóstica do Desempenho Escolar (Saerjinho). “Embora com perspectivas diferentes, os resultados dessas avaliações são complementares e, para que possam fazer a diferença na qualidade da educação oferecida, devem ser integrados ao cotidiano do trabalho escolar.” (Seeduc, n.d.b).

⁹ Como o Sistema de Avaliação de Educação do Estado do Rio de Janeiro e o Programa de Avaliação Externa são conhecidos pela mesma sigla Saerj, neste trabalho, optou-se por aquele ser sempre chamado pelo nome completo e este utilizar a sigla, para que não haja confusão.

Programa de Avaliação Externa – Saerj.

Em novembro de 2008, em consonância com a política nacional de avaliação externa, foi aplicada a primeira avaliação do Saerj, instituído pela Secretaria de Estado de Educação do Rio de Janeiro (Seeduc/RJ). O programa, ao longo de todo o período de aplicação, compreendeu uma avaliação externa em larga escala com caráter anual, a ser aplicada durante o último bimestre do ano letivo, nas disciplinas de língua portuguesa e matemática, tomando como referência os níveis das escalas de desempenho do Saeb. A amostra era composta pelos anos finais de cada ciclo de ensino: 5º e 9º anos do ensino fundamental, 3ª ano do ensino médio e suas fases equivalentes na educação de jovens e adultos (EJA). O objetivo apresentado pelo Saerj na primeira edição foi “possibilitar um acompanhamento mais preciso do desempenho dos nossos alunos ao longo dos anos e, assim, nortear o monitoramento das ações em curso e a programação de novas ações que visem à melhoria constante da qualidade da educação ofertada à população fluminense.” (Rio de Janeiro, 2008, p. 7). Assim, observa-se que a avaliação está mais voltada para o monitoramento, a avaliação e a formulação de políticas públicas (Madaus & Russell, 2010).

Desde a primeira edição, através de parceria com o Seeduc/RJ, o Centro de Políticas Públicas e Avaliação da Educação (CAEd) ficou responsável pela elaboração, aplicação, correção e devolutiva dos resultados. Tal parceria foi instituída através de licitação pública, gerando uma terceirização operacional do programa.

Em linhas gerais, serão analisados alguns pontos importantes ao longo dos oito anos de aplicação das avaliações do Saerj.

A avaliação apresentou, em sua primeira edição, um quantitativo representativo de participantes – foram previstos 295.604 alunos, sendo que efetivamente fizeram a prova 191.966 alunos, isto é, 64,9% da rede pública de ensino. A partir dos dados colhidos, a Seeduc/RJ, em conjunto com o CAEd, elaborou quatro materiais para apresentar às escolas os resultados obtidos: revista do professor, revista do gestor, roteiro básico para a discussão e apropriação dos resultados e boletim contextual.

Com base na análise dos documentos oferecidos pelo CAEd, na primeira aplicação, o Saerj já se apresenta como uma avaliação com propósito de utilização não apenas pelo sistema, mas também pela escola. Tal perspectiva se confirma com a elaboração de materiais voltados para o professor que não apenas elucidam e explicam os resultados, mas também apresentam como podem ser utilizados diretamente em sala de

aula, com sugestões de atividades e a indução de aplicação na sala de aula das atividades sugeridas, como um desafio a ser cumprido. Para o professor, é dito na revista:

Os resultados que você [professor] descobriu na estação anterior são as coordenadas necessárias para esse trecho da caminhada. Os números passarão, então, a ter um significado cada vez mais próximo de você, até ser possível identificar aqueles alunos que necessitam de uma atenção especial. Será uma transformação e tanto (CAEd, 2008, p. 15).

Diante dessas considerações, fica evidente a tentativa de trazer o professor para junto dos resultados, auxiliando-o a conhecer melhor a natureza da avaliação e sua “linguagem”. O material apresentou exemplos de como realizar a ligação entre o que foi visto na avaliação externa e os recursos pedagógicos utilizados pelo professor em sala de aula, com o objetivo de para minimizar o fato de os professores apresentarem dificuldades para entender as matrizes de referência dos testes e associá-las aos conteúdos ministrados e trabalhados em sala de aula (Brooke & Cunha, 2011; Nevo, 1997).

No ano de 2010, ocorreu a terceira edição do programa, com modificações significativas. Houve uma extensão da amostra de alunos avaliados, pois a avaliação cobriu os alunos do 4º ao 9º ano do ensino fundamental e todas as séries do ensino médio. As disciplinas continuaram sendo português e matemática. Como objetivo, o programa pretendia “produzir informações que devem subsidiar a implementação de políticas públicas voltadas à melhoria da educação e à igualdade de oportunidades bem como o processo de intervenção pedagógica nas escolas.” (Rio de Janeiro, 2010, p. 15). Nesse ciclo de avaliação, já aparece a utilização no processo de intervenção pedagógica. Os resultados não foram divulgados apenas por escola, mas também passaram a ser no âmbito de cada aluno, tendo em vista que o Saerj passou a produzir “um diagnóstico do desempenho de cada aluno avaliado, bem como da evolução desse desempenho ao longo de sua trajetória escolar” (Rio de Janeiro, 2010, p. 5), para que assim pudesse cumprir o objetivo de ser um “instrumento útil no apoio às discussões pedagógicas, na elaboração de projetos educativos, no planejamento das aulas, na análise do material didático e na reflexão sobre a prática educativa.” (Rio de Janeiro, 2010, p. 5). Começa, com isso, a ganhar um caráter diagnóstico em um nível menor, passando da escola para o aluno.

Na abertura da *Revista do Sistema de Avaliação*, há um texto da Secretaria de Educação. O texto, em linhas gerais, ressalta a importância dos resultados do Saerj como instrumento para apoiar a implantação e o desenvolvimento de políticas educacionais;

além de possibilitar discussões pedagógicas, no âmbito de projetos educacionais, planejamento de aulas, análise de material didático e reflexão sobre a prática educativa. Com isso, em 2010, aparece como objetivo, e de forma clara no discurso, a possibilidade de intervenção pedagógica da avaliação, o que nos outros exames só era possível perceber de forma implícita na condução do material apresentado para o professor.

No ano de 2012, nos documentos de análise da avaliação, aparece, pela primeira vez, a política de bonificação por desempenho, colocando essa política no objetivo principal. Na fala de abertura na *Revista do Sistema de Avaliação*, o secretário de educação do estado afirma que o Saerj além de

colaborar para a implantação de políticas de reforço escolar, para a redefinição de trajetórias e para a melhoria nas práticas escolares, a avaliação também serve como base para premiar os profissionais que vêm trabalhando em conjunto nas escolas, melhorar o desempenho dos alunos e atingir suas metas. (Rio de Janeiro, 2012, p. 4)

Para além da bonificação ao docente, o secretário comenta sobre a premiação de alunos, com computadores portáteis e viagens à cidade do Rio de Janeiro para aqueles que atingirem os melhores resultados. Essa política não é nova no estado, tendo em vista que a avaliação estadual Programa Nova Escola já a utilizava.

Na sexta edição do Saerj, no ano de 2013, o secretário de educação evidencia que a avaliação está a serviço do professor:

permite aos sistemas de ensino prestar contas à sociedade que, afinal, é quem paga a educação que recebe e realimentar o processo educativo que a escola desenvolve para redirecionar práticas e reformular as estratégias que devem levar aos objetivos visados. Nesse caso, a avaliação externa é uma ferramenta pedagógica a serviço do professor e de sua prática docente (Rio de Janeiro, 2012, p. 5).

Apresenta, por conseguinte, duas ideias para a avaliação externa: ser utilizada como instrumento para a prestação de contas e para o redirecionamento da prática docente. O secretário também destaca que o Sistema de Avaliação da Educação do Estado do Rio de Janeiro, compreendido pelo Saerj e o Saerjinho, vem mostrando eficiência no norteamento das políticas públicas implantadas no ano de 2010, o que contribui para destacar o Rio de Janeiro nas avaliações em larga escala de nível nacional: Enem, Saeb e Prova Brasil. Configura-se, com essa ponderação, a primeira vez que aparece nos

documentos oficiais a ideia do sistema próprio do Rio de Janeiro comparado com os sistemas de avaliações de âmbito nacional.

Ao longo dos sete anos de aplicação, o Saerj teve um aumento progressivo de participação dos alunos das redes de ensino, chegando a atingir 83,2% de estudantes avaliados. Pode-se notar, pelo discurso apresentado pela Secretaria de Educação no material de divulgação, a permanência do Saerj, desde a primeira edição, como um instrumento para o acompanhamento do desempenho dos alunos com vistas à monitoração das ações em curso e à promoção de novas ações, sendo, por conseguinte, um instrumento para formulação e reformulação de políticas públicas. No entanto, na condução do material, já é possível verificar a avaliação como forma de intervenção da prática docente. Em seguida, de forma oficial, surge, como objetivo, a utilização da avaliação para ajudar na melhoria do trabalho pedagógico da escola, com a finalidade de gerar reflexões pedagógicas. Com o passar do tempo, observam-se outras ações vinculadas aos resultados do programa, tais como: gerar dados para comparabilidade; bonificação por desempenho de professores e alunos; forma de prestação de contas à sociedade; e instrumento para gestão pedagógica.

A edição de 2016 não chegou a acontecer, pois o governo acabou com o programa de avaliação no início do ano. Após muitas críticas do sindicato dos professores e dos próprios professores, o governo anunciou, em 2015, que o Saerj sofreria mudanças. Porém, em 2016, os estudantes pediram o fim do programa, como uma das exigências para acabar com a ocupação de diversas escolas públicas. O governo aceitou algumas reivindicações para terminar com a manifestação estudantil, sendo uma delas o fim do Sistema de Avaliação da Educação do Estado do Rio de Janeiro e, conseqüentemente, do Saerj. Assim, a rede de ensino estadual passaria a ser avaliada apenas pelo Saeb.

Programa de Avaliação Diagnóstica do Desempenho Escolar.

O Programa de Avaliação Diagnóstica do Desempenho Escolar (Saerjinho), que fez parte do Sistema de Avaliação da Educação do Estado do Rio de Janeiro, foi implementado em abril de 2011 e ocorreu até o primeiro bimestre de 2016. Configurou-se no âmbito do Sistema Estadual de Ensino, sendo uma avaliação censitária, diagnóstica-formativa, que era aplicada três vezes ao longo de um ano letivo, uma prova a cada

bimestre¹⁰. As informações geradas pela avaliação correspondiam à consolidação de habilidades e competências esperadas para o ano escolar avaliado. Eram avaliados os alunos do 5º e do 9º anos do ensino fundamental, nas disciplinas de português, matemática e ciências da natureza, e todos os anos do ensino médio, ensino médio integrado e curso normal, em português, matemática, ciências da natureza e ciências humanas. As avaliações tinham como base a matriz de referência do Saerjinho, que apresentava as competências e habilidades por disciplina e anos avaliados e os pré-requisitos de cada bimestre¹¹, que foram elaborados levando em consideração o currículo mínimo do estado, a matriz de referência do Saerj e as diretrizes curriculares dos anos em questão. Diferentemente do Saeb e do Saerj, que utilizam a TRI como modelo estatístico, a medida do Saerjinho considerava o percentual médio de acerto por descritor, a partir da Teoria Clássica dos Testes (TCT). Assim como o Saerj, apresentava parceria com o CAEd, que era responsável pela elaboração, logística, leitura, processamento de dados e elaboração de relatórios e boletins das avaliações.

O programa tinha como objetivo a identificação de deficiências que podiam ser corrigidas ao longo do ano letivo. Com isso, apresentava uma maior atuação na escola, pois considerava que os resultados, por serem gerados a cada bimestre, poderiam dar informações mais precisas, de forma ágil e fiel, aos docentes e aos alunos, servindo, por conseguinte, como uma ferramenta pedagógica. No entanto, também apresentava como característica traçar políticas públicas para a melhoria da qualidade da educação ofertada, além de ser uma ação que pretendia melhorar a colocação do Rio de Janeiro no Ideb até 2014.

Não há, no *site* oficial do programa, material disponível com a proposta e apresentação de resultados aberta ao público, tendo em vista que o acesso a ele é restrito. Por isso, as informações coletadas encontram-se, em sua grande maioria, nos informes do *site* da Seeduc/RJ.

Uma matéria com o título “Saerjinho ajuda a treinar para o Enem, para vestibulares e para entrevistas de emprego” foi publicada no *site* da Seeduc/RJ em novembro de 2011. A matéria apresenta o objetivo do programa; o seu caráter de simulado

¹⁰ No Brasil, o ano letivo, na grande maioria das escolas, é composto por quatro bimestres. O Saerjinho era aplicado nos três primeiros bimestres e o Saerj, no quarto bimestre.

¹¹ Na matriz, há um espaço denominado “detalhamento”, que apresenta, dependendo da habilidade, um pré-requisito a ser cumprido. Exemplo: para o 5º ano, na disciplina de Língua Portuguesa, deve-se trabalhar com relato de viagem / literatura de informação na habilidade “Identificar a finalidade de textos de diferentes gêneros” no 1º bimestre.

com a finalidade de treinamento para avaliações externas nacionais (Enem, Prova Brasil), para vestibulares e entrevistas de emprego; o bônus por desempenho através dos resultados; e exemplos de ações de correções que podem decorrer, como: reforço escolar, material pedagógico, formação continuada para os docentes, acompanhamento de cada unidade escolar, estabelecimento do currículo mínimo e projeto de correção de fluxo.

A matéria também apresenta os critérios para que a unidade escolar receba a bonificação:

cumprir 100% do currículo mínimo; participar de todas as avaliações internas e externas; efetuar o lançamento das notas do alunado no formato e prazo estabelecidos; alcançar, no mínimo, 95% de resultado em cada meta de Iderj do ensino regular da Unidade Escolar; alcançar, no mínimo, 80% de resultado de cada meta de ID da Educação de Jovens e Adultos presencial da Unidade Escolar; e ter, pelo menos, 70% de frequência presencial no período de avaliação, que corresponde ao ano letivo (Seeduc, n.d.b).

Numa outra notícia da Secretaria, há a afirmação de que a prova é de caráter fundamental para que os educadores venham a elaborar estratégias pedagógicas para as metas das escolas poderem ser alcançadas (Seeduc, n.d.c).

Havia a divulgação de relatórios, que apresentavam informações sobre a evolução da aprendizagem dos alunos, a produtividade das atividades curriculares e a qualidade do trabalho escolar (Rio de Janeiro, n.d.a). O resultado da avaliação era divulgado por escola, por turma e por aluno no *site* da Secretaria de Educação, tendo acesso o diretor, o professor e o secretário da escola. O professor era obrigado a utilizar a nota do Saerjinho para a composição da nota bimestral do aluno, conforme norma estabelecida pela Portaria Seeduc/Sugen nº 419, de setembro de 2013, que apresentava no capítulo 2, artigo 4º, inciso V. No entanto, não fica claro como ocorreu tal controle e se realmente os professores seguiam essa normativa.

A Avaliação Diagnóstica Bimestral do Processo Ensino-Aprendizagem - Saerjinho, aplicada nos níveis de ensino, anos/séries, disciplinas e bimestres definidos pela Seeduc, é um dos instrumentos avaliativos obrigatórios para composição da nota bimestral do discente, com valor/nota definido (a) pelo Professor, e deverá ser registrada no Diário de Classe ou outro instrumento

indicado pela Seeduc, bem como no Sistema Eletrônico de Registro Escolar (Rio de Janeiro, 2013).

Como parte do projeto do Saerjinho, a Subsecretaria de Gestão de Ensino elaborou um banco de itens que ficava à disposição para os professores da rede de ensino. O objetivo do banco era contribuir com os docentes no processo de aprendizagem e avaliação. Havia sugestão para o professor trabalhar com os itens segundo as necessidades de seus alunos, após a análise dos resultados da avaliação. O professor, ao acessar o banco, podia selecionar os itens ou as habilidades que desejava trabalhar e o sistema montava a prova.

No final de 2011, os professores de língua portuguesa e matemática participaram de uma oficina de elaboração de itens, promovida pelo CAEd. A oficina apresentou como objetivo capacitá-los para elaborarem itens para o Saerjinho e também se tornarem multiplicadores do conhecimento. Contudo, tal iniciativa não se restringiu apenas a elaborar itens para a avaliação externa, mas também a “servir de subsídio para a preparação dos exames nas próprias escolas.” (Seeduc, n.d.d). Para participar, o professor ou o diretor de sua unidade deveria manifestar interesse realizando o cadastro no *site* da Seeduc/RJ.

Outro projeto vinculado ao Saerjinho foi o Jovens Turistas, que consistia em premiar os alunos e os professores do ensino médio da Rede Estadual de Ensino que mostraram bom desempenho nas avaliações. O critério empregado foi a evolução da nota da instituição em comparação com o último exame. Assim, seriam contempladas as escolas que apresentassem maior evolução. “O Jovens Turistas está em consonância com a atual política de meritocracia adotada pela Secretaria de Estado de Educação, ou seja, valoriza o mérito de estudantes, professores e funcionários da rede que se destacam pela qualidade pedagógica.” (Rio de Janeiro, n.d.b).

O programa de avaliação conquistou o Prêmio Anual 2015 de Gestão por Resultados da Comunidade de Profissionais e Especialistas em Gestão Pública, uma ação do Banco Interamericano de Desenvolvimento (BID). Foram oito categorias premiadas, sendo que o Saerjinho venceu em duas: Avaliação e Monitoramento e Gestão para Resultados em Setores (Seeduc, n.d.e).

Porém, desde o primeiro ano de implantação, há movimentos contrários ao programa. Tal fato fica claro em uma notícia divulgada em junho de 2011, no portal da Secretaria, com o título “SEEDUC repudia atitudes que prejudicam alunos e professores”.

A secretaria considera “lamentáveis manifestações como as tentativas de boicote ao Saerjinho”, afirmando que isso é ser contra alunos e professores. Identifica-se o Sindicato Estadual dos Profissionais de Educação do Rio de Janeiro (Sepe) como incentivador do boicote. Salienta, também, que repudia as inverdades contadas a respeito do Saerjinho (Seeduc, n.d.f).

Em 2016, juntamente com o Saerj, o Saerjinho foi encerrado após uma manifestação estudantil.

Avaliações Externas Municipais – Município do Rio de Janeiro

No ano 2000, houve uma grande mudança na política de organização do tempo, do currículo e da avaliação do ensino fundamental no município do Rio de Janeiro, pois começou então a vigorar o ciclo de formação, ou seja, períodos de escolarização que ultrapassam a seriação anual, pois são organizados em blocos com durações variáveis, podendo até mesmo ultrapassar a totalidade de anos previstos para um determinado período de ensino. Geralmente, a estruturação de tempo no sistema de ciclos se faz em unidades maiores e flexíveis, com isso, há uma tentativa de se superar a fragmentação do currículo durante o processo e a avaliação perde o seu caráter de retenção, pois durante o ciclo não há reprovação do aluno. Essa política foi utilizada em vários estados e municípios brasileiros na década de 1990, para que pudesse ocorrer uma regularização do fluxo no ensino fundamental.

Num primeiro momento, o ciclo de formação abrangia o 1º, o 2º e o 3º anos do ensino fundamental. A intenção era que, de forma gradativa, o sistema de ciclo fosse adotado por todo o ensino fundamental, ou seja, até o 9º ano, no entanto, apenas em 2007 esse objetivo foi alcançado. Nesse momento, há uma normatização da avaliação da aprendizagem dos alunos através da Resolução nº 946, de 25 de abril de 2007, que suspende a reprovação entre os anos do ensino fundamental, uma vez que a avaliação seria formativa e continuada. Uma repercussão negativa instalou-se na cidade do Rio de Janeiro, com ampla divulgação da mídia, que considerava o sistema de ciclo como sinônimo de aprovação automática dos alunos.

Em 2009, atendendo aos pedidos, em grande parte, dos professores, o então prefeito recém-empossado, suspende o sistema de ciclos, a partir do Decreto nº 30.340, considerando que tal sistema era inadequado para alcançar a melhoria do ensino público municipal e considerando ser “imprescindível o estabelecimento de parâmetros objetivos de aferição do rendimento escolar dos alunos.”.

A partir dessa nova política, a avaliação, no município do Rio de Janeiro, passa a estar vinculada ao conceito de qualidade, o que Oliveira e Araújo (2005) denominaram de terceiro momento da qualidade na educação no Brasil. De acordo com a Resolução da SME nº 1.010/2009, é possível observar que se estabelece o início da avaliação em larga escala em nível local no município do Rio de Janeiro, seguindo uma tendência nacional e internacional, na medida em que o segundo artigo apregoa que “o Nível Central enviará às escolas, bimestralmente, provas a serem aplicadas a todos os alunos, visando o acompanhamento do seu processo de aprendizagem.”. Para além do estabelecimento de um novo sistema de avaliação externo à escola, a resolução também determinava que o peso da avaliação externa fosse o dobro do peso da avaliação aplicada pelo professor. A diferenciação de peso da avaliação não foi aceita por diferentes segmentos da comunidade escolar, com isso, a SME/Rio revogou essa ação e estabeleceu o mesmo peso para os instrumentos de avaliação externos e internos. Neste momento, fica evidente o posicionamento governamental de maior valor à avaliação externa.

As Provas Bimestrais vêm sendo aplicadas, desde o ano de 2009, aos alunos do 2º até o 9º ano de escolarização, e têm por objetivo contribuir para uma avaliação do processo de desenvolvimento e aprendizagem. Os alunos fazem questões de múltipla escolha, com preenchimento de cartão resposta, das disciplinas de português, matemática e ciências, que são elaboradas com base nas orientações curriculares do município, por uma equipe pedagógica da SME/Rio, composta por professores. De acordo com Cláudia Costin (2009), secretária de Educação da época, as Provas Bimestrais permitiam com que “cada escola soubesse se aquela criança estava evoluindo positivamente ou negativamente, e que a rede soubesse, inclusive, como cada escola estava evoluindo comparativamente com as outras, com a média da rede.”.

É de responsabilidade da escola corrigir e lançar a nota no Sistema de Desempenho Escolar (DESESC) e da SME/Rio a elaboração de um relatório, com dados por Coordenadoria Regional de Ensino (CRE), realizando comparação entre elas – por ano de escolaridade, por questão da prova, por descritores, entre outros. As avaliações bimestrais não são consideradas, pela SME, como avaliações externas da aprendizagem dos alunos, pois

não possuem as mesmas características deste tipo de avaliação. As Provas Bimestrais têm um número reduzido de questões, não calibradas, e são compostas por cadernos de provas comuns a todos os alunos do mesmo ano de escolaridade.

Todos respondem às mesmas questões que não estão dispostas numa escala de proficiência. São questões que atendem a descritores correspondentes às habilidades previstas para o bimestre avaliado. O principal objetivo desta prova é permitir que, na medida do possível, as habilidades para as quais for aferido baixo desempenho por parte dos alunos sejam observadas e subsidiem um replanejamento das ações, no nível escolar, por parte dos professores e da equipe gestora. As Coordenadorias Regionais de Educação e o Nível Central da SME devem atuar no sentido de promover ações de formação e produção de material de apoio pedagógico com base nos resultados dessas Provas (Mendonça, 2018, p.1).

Assim, optou-se, para este trabalho, não considerar a avaliação bimestral da SME como uma avaliação externa.

A avaliação externa da rede municipal do Rio de Janeiro é a Prova Rio, também iniciada na gestão de 2009, tendo como objetivo “apontar a qualidade do ensino na Rede Municipal. Sua importância reside em recolher indicadores comparativos de desempenho que servirão de base para futuras tomadas de decisões no âmbito da escola e nas diferentes esferas do sistema educacional.” (Secretária Municipal de Educação – RJ, 2010). Uma avaliação de caráter censitário e anual, que avalia habilidades das disciplinas de português e matemática. A matriz de referência tem por base as Orientações Curriculares da Rede e a matriz de referência da Prova Brasil. Os resultados são analisados segundo a escala de proficiência do Saeb. Além das provas, há a aplicação de questionários que coletam informações dos alunos, dos professores e dos diretores sobre temas relacionados às questões pedagógicas.

A sua primeira edição foi aplicada aos alunos do 3º e do 7º anos do ensino fundamental, em outubro de 2009. Os resultados foram divulgados no mês de janeiro do ano seguinte. Nessa edição, os 20 melhores alunos foram premiados com uma bicicleta e seus professores com um *netbook*. A partir dos resultados da primeira edição, foi criado, no dia 22 de junho, o Índice de Desenvolvimento da Educação no Município do Rio de Janeiro (IDE-Rio), pelo Decreto nº 32.416/10.

O índice foi criado com o intuito de servir como base para o estabelecimento de metas e premiações dos servidores instituídos pelo Prêmio Anual de Desempenho, ou seja, a premiação destina-se não apenas para professores, mas também para todos os funcionários da escola. Antes do índice próprio do município, a premiação tinha como

base o Ideb. O Índice de Desenvolvimento da Educação do Rio de Janeiro (IDE-Rio) é calculado levando em consideração o desempenho dos alunos na Prova Rio e o fluxo escolar, forma semelhante ao Ideb.

Na edição de 2010, as provas passaram a ser aplicadas aos alunos do 3º, do 4º, do 7º e do 8º anos, avaliando as mesmas disciplinas. Neste ano, uma nova avaliação externa foi incorporada ao calendário do município, a Alfabetiza Rio, aplicada aos alunos do 1º ano, com o objetivo de identificar quantos alunos foram alfabetizados no final do 1º ano do ensino fundamental. Houve também premiação para os alunos que obtiveram as maiores notas. As empresas que foram responsáveis pela elaboração, aplicação e emissão de relatórios analíticos dessas avaliações foram o Cespe/Cebraspe e a Consulplan (Prova Rio) e o CAEd (Alfabetiza Rio).

Nesse contexto de reorganização das políticas educacionais do município do Rio de Janeiro, para além das avaliações externas, fundamentam-se também outras frentes de ação, como a elaboração do currículo básico de todas as disciplinas do ensino fundamental; a disponibilização de cadernos pedagógicos, com orientações para o desenvolvimento dos conteúdos escolares; e as avaliações externas, que vieram com o propósito de lançar luz, com precisão, sobre os pontos de atenção e sobre a evolução dos estudantes, assegurando que a melhoria do ensino fosse garantida, além de indicar quais alunos necessitavam participar do Programa Reforço Escolar, também desenvolvido pela nova gestão.

De acordo com Fernandes (2015), apesar de os documentos oficiais da SME/Rio trazerem uma concepção de avaliação de carácter processual, formativa, dialógica, investigativa e transformadora, não há uma materialização dessa concepção nos instrumentos e registros de avaliação elaborados pela Secretaria, muito em decorrência do sistema de premiação e responsabilização das escolas. Para Mendes e Frangella (2014), é um sistema de avaliação com carácter centralizador, no qual se fundamenta uma proposta oficial em que se determina o que se deve aprender, ensinar e avaliar, considerada por Freitas (2007) como pressão verticalizada.

Análise das Diferentes Avaliações Externas Presentes no Rio de Janeiro

Na primeira edição do Saeb, em 1990, dos três objetivos propostos, dois deles estavam direcionados para a implantação de sistemas próprios de avaliação pelos estados e municípios. Com isso, desde o primeiro ciclo de aplicação do Saeb, já se planejava o desenvolvimento das capacidades avaliativas dos estados e municípios, tornando-se a

avaliação um instrumento de planejamento ou monitoração e um provocador de mudanças qualitativas, como pode ser observado no primeiro objetivo apresentado, no que se refere às capacidades avaliativas.

1. desenvolver e/ou aprofundar as capacidades avaliativas das unidades gestoras do sistema educacional (o próprio Ministério, as Secretarias Estaduais e Órgãos Municipais de Educação), através da constituição progressiva de quadros especializados quanto aos diferentes aspectos da avaliação enquanto instrumento seja de planejamento ou monitorização, seja provocador de mudanças qualitativas;
2. regionalizar a operacionalização do processo avaliativo, de forma a possibilitar sua gestão diretamente pelas instâncias locais (equipes técnicas das Secretarias, Instituições de Pesquisa, Universidades etc.), criando nexos e estímulos para o desenvolvimento local de uma infra-estrutura de pesquisa e avaliação educacional;
3. propor uma estratégia metodológica em condições de articular e relacionar os resultados dos diversos trabalhos de pesquisa e avaliação já realizados ou em vias de implementação (Inep/Pnud, 1992, p. 7).

Em resposta a essa política federal, muitos estados e municípios começaram a desenvolver seus sistemas próprios de avaliação. Em 2015, havia 18 estados federados com sistemas próprios de avaliação, tendo mais 2 estados com planos de implantação em 2016, totalizando, por conseguinte, 20 estados das 27 unidades federativas (CAEd, 2015). De acordo com Bonamino (2013), as avaliações estaduais tendem a apresentar modelo censitário, avaliando todos os alunos das séries finais de ciclo (5º e 9º ano do ensino fundamental e 3º ano do ensino médio); utilização da escala Saeb a partir do ano 2000; e utilização das avaliações para compor um índice que referencia a definição de metas por escola, com iniciativas de responsabilização de professores e gestores.

Observa-se, na linha do tempo (Anexo 1), que, depois do início da Prova Brasil em 2015, houve uma maior continuidade dos sistemas de avaliação estaduais. Por ser uma avaliação censitária e não amostral como o Saeb, a Prova Brasil acabou por ter maior impacto na estruturação das avaliações estaduais. Há uma única matriz de referência e escala de proficiência para as provas do Saeb (Prova Brasil e Aneb/Saeb); com isso, verifica-se que quase a totalidade das avaliações estaduais utiliza a escala de proficiência do Saeb. A utilização da escala de proficiência do Saeb nos estados passa a ocorrer de

forma acentuada também depois da implantação da Prova Brasil. O fato de os estados utilizarem uma escala de proficiência igual à do Saeb possibilita não só uma comparação dos estados com o sistema de avaliação nacional, mas também uma possível comparação entre os estados.

Após dez anos do início do sistema nacional de avaliação em larga escala, o estado do Rio de Janeiro dá início ao seu sistema próprio de avaliação, tendo, em dezenove anos, três avaliações: Programa Nova Escola (2000-2008), Saerj (2008-2016) e Saerjinho (2011-2016). O município do Rio de Janeiro, por sua vez, de 2009 até o momento, tem a Prova Rio e Alfabetiza Rio.

No tocante às avaliações estaduais, Souza (2013), com base em pesquisas, considera a presença de duas fases. A primeira fase, que ocorreu até meados de 2005, apresentou frágeis vínculos entre resultados das avaliações e políticas e práticas, tendo em vista que havia um anseio de que a avaliação viesse a iluminar a formulação e/ou reformulação das políticas públicas educacionais, com foco em questões relativas à gestão, currículo, infraestrutura, desenvolvimento do currículo e formação de professores. No entanto, o uso do resultado das avaliações externas com tal objetivo era tênue, tendo em vista que “as gestões das redes apoiavam-se nas estruturas burocráticas não se orientando pelos resultados de desempenho escolares” (Souza, 2013, p. 70). Havia, até mesmo, dificuldades em compreender os resultados, desde os atores escolares até os técnicos das Secretarias de Educação. Também ocorria a sensação de que se repetia, nos estados, o que já ocorria em nível nacional com a Prova Brasil. Tais fatos geraram uma frágil articulação da avaliação com as políticas e práticas.

Em resposta a esse fato, surge uma segunda fase, na qual há uma associação entre resultados das avaliações e gestão educacional, tendo em vista que tiveram movimentos mais consistentes na direção de avaliar e orientar a política educacional; de informar as escolas sobre a aprendizagem dos alunos e definir as estratégias de formação continuadas; de informar ao público; de alocar recursos; de criar políticas de incentivos salariais; de desenvolver a avaliação docente; e de certificar alunos e escolas (Brooke & Cunha, 2011). Esse momento encontra-se mais alinhado com a fase de gestão gerencial na política pública, com base nos princípios de gestão estratégica e de controle de qualidade, havendo uma maior descentralização, o que possibilita o envolvimento dos atores educacionais na consecução dos objetivos propostos (Castro, 2008).

No entanto, observa-se que há, ao longo dos anos, um desenvolvimento maior com foco no resultado por aluno, começando na Prova Brasil (Figura 7), que dá início à

divulgação do resultado por escola, e se multiplica no estado e município do Rio de Janeiro, que passa a divulgar o resultado por aluno, para que se possa, se necessário, mudar a condução do que é ensinado e como é ensinado, do que é aprendido e como é aprendido, passando a uma função de regulação do processo de ensino e aprendizagem, com maior efeito dentro de sala de aula (Madaus & Russell, 2010).

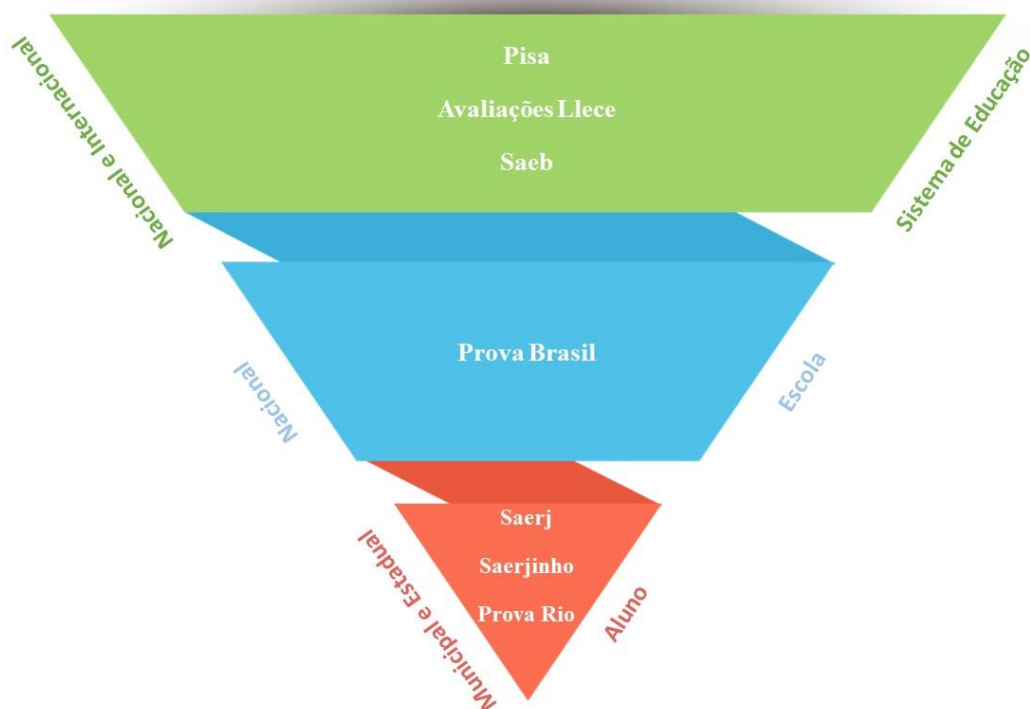


Figura 7

Desenvolvimento dos sistemas de avaliação presentes no Rio de Janeiro

Observa-se que as avaliações externas, desde a Prova Brasil até a avaliação municipal do Rio de Janeiro, possuem a configuração com dupla função, pois apresentam caráter político e pedagógico, principalmente no que tange à prestação de contas e à regulação da aprendizagem dos alunos, muito mais voltado para uma pressão verticalizada (Freitas, 2007) do que com um caráter de subsidiar a regulação e a melhoria dos processos de ensino e aprendizagem.

A avaliação externa administrada por instâncias governamentais passa a ter função formativa, mais precisamente no que se refere à instância local, assumindo algumas características distintas das que são pensadas numa avaliação formativa realizada pelo

professor, porém não necessariamente perdendo seu efeito. De acordo com Wiliam (2006), as avaliações externas também podem apresentar característica de avaliação formativa, desde que utilizadas de forma a identificar o que precisa ser aprendido e ajustar estratégias de ensino. Assim, o que torna uma avaliação formativa não seria o tempo de retorno do *feedback*, onde é realizada ou por quem é realizada, mas sim, que os dados sejam coletados, interpretados segundo as necessidades de aprendizagem e usados para realizar ajustes no que tange à melhor atender a essas necessidades.

ANO	AVALIAÇÕES		
	Nacional	Estadual	Municipal
1º EF			Alfabetiza Rio
2º EF	Provinha Brasil		
3º EF	ANA		Prova Rio
4º EF			Prova Rio
5º EF	Saeb / Prova Brasil	PNE/ Saerj / Saerjinho	
6º EF		Saerjinho	
7º EF			Prova Rio
8º EF			Prova Rio
9º EF	Saeb / Prova Brasil	PNE/ Saerj / Saerjinho	
1º EM		Saerjinho	
2º EM		Saerjinho	
3º EM	Saeb	PNE/ Saerj /Saerjinho	

■ Último ano do ciclo

Figura 8

Avaliações externas ao longo da história no Rio de Janeiro

A Figura 8 evidencia todas as avaliações externas que ocorrem ou já ocorreram no Rio de Janeiro, da nacional até a local, separadas por ano de escolaridade e instância governamental que gerencia a avaliação. Tal quadro possibilita ver com maior clareza como cada governo vem, ao longo do tempo, distribuindo suas avaliações.

Observa-se que os sistemas nacional e estadual concentram suas principais avaliações externas nos anos finais de cada ciclo, ou seja, no final do ensino fundamental

I (5º ano), do ensino fundamental II (9º ano) e do ensino médio (3º ano). Isso se caracteriza com maior ênfase na avaliação com função somativa, com caráter político, tendo em vista que busca verificar como os alunos chegam ao final de uma determinada etapa, para que, assim, possam avaliar, formular e reformular políticas educacionais.

Contudo, para além dessa função, há evidências de que tanto a Prova Brasil quanto o Saerj também apresentam como objetivo produzir informações para auxiliar na intervenção pedagógica em sala de aula. Por isso, a Prova Brasil apresenta resultado por escola e o Saerj apresentava resultado por aluno. Iniciam-se, com a Prova Brasil, as avaliações externas com dupla função, isto é, política e pedagógica.

O estado vê a necessidade, então, de acrescentar mais uma avaliação e passa a aplicar o Saerjinho, com o objetivo de identificar e corrigir, dentro de um mesmo ano letivo, as deficiências dos alunos. Dessa forma, essa avaliação apresentava maior função pedagógica. No entanto, o programa também considerava a função de traçar políticas públicas de melhoria da qualidade da educação ofertada.

Torna-se evidente que o estado possuía um sistema de avaliação cobrindo quase todo o ensino fundamental II e todo o ensino médio, com dupla função. Chegando até mesmo a ter séries com avaliação externa nos quatro bimestres do ano letivo.

O município foi o último a instituir um sistema próprio de avaliação, repetindo em grande parte o modelo do estado. Apesar de não termos considerado as Provas Bimestrais neste trabalho, elas também são aplicadas em três bimestres no ano letivo, culminando no último bimestre com a Prova Rio. Essa avaliação ocorre nas séries em que não há avaliação nacional ou estadual. Da mesma forma, apresenta tanto função política quanto pedagógica.

Com isso, observa-se que, do período de 2011 até 2015, todas as séries escolares foram avaliadas no Rio de Janeiro, por diferentes instâncias governamentais, chegando, até mesmo, a mais de quatro avaliações externas em um mesmo ano letivo. Com exceção do Saeb, todas as outras avaliações apresentam mais de uma função.

Ficou evidente que as avaliações em larga escala no Brasil seguem diferentes desenhos, dependendo do nível de governo que a administra e de seus objetivos, no entanto, praticamente todas apresentam função política e pedagógica. Um sistema de avaliação pode apresentar diferentes finalidades, que não são necessariamente excludentes, podendo, assim, o mesmo sistema combinar várias delas (Ravela et al., 2008). Porém, deve haver um esforço no sentido de compreender que cada opção apresenta um determinado desenho, exigências técnicas e custos diferentes. Contudo, de

acordo com Brooke e Cunha (2011, p. 20), no Brasil, “fica evidente que os mesmos instrumentos estão sendo usados com objetivos diversos de modo que a finalidade precípua ou original do instrumento pouco esclarece sobre seus desdobramentos posteriores”. Para além disso, constantes mudanças no desenho da avaliação ao longo do tempo e a descontinuidade de algumas avaliações – o que ocorre com maior frequência nos sistemas de avaliação estadual e municipal – geram maior desconfiança e resistência (Gatti, 2013).

No âmbito nacional, há duas avaliações com desenhos diferentes, a Aneb/Saeb de caráter amostral, que apresenta resultado apenas para o sistema educacional, e a Prova Brasil de caráter censitário, com resultados por escola. Porém, apesar de possuírem nomenclaturas e objetivos distintos, elas são aplicadas juntas. No estado e no município do Rio de Janeiro, encontram-se avaliações de caráter censitário e com resultado divulgado por aluno, com maior impacto na escola. Também, como visto, com políticas de responsabilização forte, na medida em que, ao longo de sua história, apresentou bonificação e premiação para escola, funcionários, professores e/ou alunos, assumindo, segundo Brooke (2006), um modelo pioneiro de política de responsabilização no país. Possivelmente, esses foram alguns dos motivos de grande resistência dos atores escolares no Rio de Janeiro à avaliação externa, desde o sindicato aos alunos, com ações judiciais, a manifestações com ocupação de escolas. Assim, as diferentes instâncias gestoras, os diferentes desenhos encontrados e a quantidade de avaliações externas presentes no cenário brasileiro, mais especificamente no Rio de Janeiro, constroem um quadro em que há uma diversidade de avaliações externas, com propósitos nem sempre claros e aceitos pelos sujeitos escolares.

Deve-se considerar que os bons resultados, a partir da prática de políticas de avaliações externas, estão estritamente ligados à compreensão e à aceitação dos profissionais da educação, devendo-se aprimorar mais a relação entre as diferentes esferas hierárquicas que fazem parte do sistema de ensino brasileiro. É necessário desenvolver um movimento de ação-reflexão-ação, no qual se deve inserir atores que tradicionalmente são colocados ao largo do processo de avaliação – na medida em que são considerados apenas como objetos e não como sujeitos da avaliação externa –, o que poderia vir a contribuir para a diminuição de tanta resistência (Alavarse et al., 2012; Cerdeira & Almeida, 2012).

Esse movimento de maior interação entre os atores escolares e a avaliação externa pode ser visto no início do Saeb, quando havia maior participação dos docentes no

desenho da avaliação, desde a participação colaborativa na construção das matrizes de referência até o estabelecimento de um canal de comunicação entre a secretaria estadual e os professores. No entanto, essa interação foi perdendo a força ao longo dos anos, prevalecendo uma política de maior indução do que de colaboração. Esse desenvolvimento do desenho da avaliação nacional acabou por influenciar os outros sistemas de avaliação, não apenas no que tange às matrizes e ao processo, mas também à forma de gerenciamento da avaliação.

As avaliações estaduais seguem uma linha muito próxima das avaliações nacionais, com uma política maior de indução, a partir de materiais de divulgação e oficinas de formação de professores. Políticas essas mais voltadas para a informação e regulação docente do que para um diálogo e colaboração.

Diferentemente do que ocorre no âmbito municipal, a partir de estudos realizados por Alavarse et al. (2012), constatou-se uma maior participação de diferentes profissionais no processo avaliativo, tais como gestores escolares e professores. Tal fato é de suma importância tanto no que diz respeito ao desenvolvimento da política local, quanto para a apropriação e validação da avaliação externa. Contudo, devido a políticas de bonificação e/ou premiação e responsabilização das escolas, o sistema municipal passa também a ser visto como centralizador (Fernandes, 2015; Mendes & Frangella, 2014).

Com isso, evidencia-se que há uma maior necessidade, no sistema de avaliação externa brasileiro, de passar de um discurso de descrição e julgamento, com políticas de responsabilização verticalizada, para um discurso de maior diálogo. Tendo em vista que, dessa forma, a avaliação deixa de ter um caráter coercitivo, vindo a ter maiores possibilidades de uso do seu desenho, processo e resultado, o que influencia na qualidade da educação (Freitas, 2007; Nevo, 1997).

Capítulo V

A Avaliação Externa sob a Apreciação dos Diferentes Intervenientes: Descrição, Análise e Interpretação dos Dados

Neste capítulo, faz-se a descrição, análise e interpretação dos dados obtidos através das entrevistas com os três grupos de participantes na investigação – especialistas, decisores políticos, docentes –, do questionário aplicado a um grupo de professores da educação básica e da análise do sistema de avaliação no Brasil e no Rio de Janeiro.

Como se referiu no capítulo anterior, no Brasil há uma diversidade de avaliações externas da responsabilidade de diferentes instâncias – internacional, federal, estadual e municipal (Bauer et al., 2017; Sousa & Oliveira, 2010). Nesses termos, é possível que a percepção dos diferentes intervenientes do processo educacional relativa à avaliação externa possa variar de acordo com a instância a que se refere a questão formulada. Porém, o intuito deste trabalho foi estudar o fenómeno da avaliação externa globalmente considerado, e não analisar especificamente cada uma das diferentes avaliações externas que se desenvolvem no Brasil ou no Rio de Janeiro. Isso não significa, contudo, que se desconsidere a relevância de estudar a especificidade de cada uma das avaliações externas. Significa apenas que o propósito desta pesquisa foi descrever e analisar as percepções dos diferentes intervenientes acerca das avaliações externas, para poder interpretar os resultados tendo em conta o marco teórico que enquadrou conceitualmente este trabalho. Sendo assim, não se buscou caracterizar as diferentes avaliações externas na fala dos participantes da pesquisa.

Optou-se por dividir este capítulo a partir dos três grupos de participantes na investigação, respectivamente: especialistas, decisores políticos e professores da educação básica. Em cada parte, serão abordados os objetos da Matriz de Investigação (Figura 4). Ao final, realiza-se uma síntese das evidências mais importantes encontradas em cada grupo participante e um cruzamento de informações entre grupos, para que assim se possa delinear, de forma interpretativa, cada dimensão proposta e possam ser respondidas as questões levantadas nesta tese.

Narrativa dos Especialistas

Nesta seção, a avaliação externa da aprendizagem dos alunos será vista a partir das narrativas dos especialistas participantes da pesquisa. Como especialistas, são considerados professores universitários que apresentam como linha de pesquisa a avaliação educacional e pesquisadores de diferentes instituições, privadas ou públicas.

Avaliação externa da aprendizagem.

Diante do quadro de grande diversidade de avaliações externas no Brasil, entender a natureza e os propósitos dessa avaliação é o primeiro passo para uma análise mais aprofundada. É importante, por conseguinte, considerar a sua essência e a finalidade, com base na apreciação dos diferentes participantes envolvidos no processo educacional, para que assim possamos identificar o que mais se destaca.

Compreende-se por natureza aquilo que compõe a substância do ser, sua essência. Já como propósito, o que se busca alcançar quando se faz alguma coisa; objetivo, finalidade, intuito. Cabe destacar que a importância dessa dimensão consiste em entender como os diferentes intervenientes concebem a avaliação externa, pois, como afirmam Brook e Cunha (2011, p. 19), é a “natureza da tarefa que vai determinar a forma da avaliação e o uso posterior dos resultados”. Sendo assim, o uso dos resultados já é dado pela natureza e pelo propósito do instrumento.

Toma-se como base para analisar essa dimensão da avaliação externa, como também para as outras, os elementos apresentados por Alavarse (2013). O primeiro é o sujeito da avaliação, compreendendo quem a conduz, quem decide sobre os objetivos e o que será avaliado. Outro elemento é o objeto da avaliação, o que se refere ao que será avaliado. Há também a finalidade da avaliação, que representa, possivelmente, o ponto de contato do processo de avaliação propriamente dito com o projeto pedagógico da escola, no qual surge a questão da avaliação formativa e somativa. O usuário dos resultados da avaliação é também um elemento que deve ser considerado. Para além dos resultados, levou-se em conta também o possível uso de processos e o desenho da avaliação externa. Por fim, a metodologia, que, neste trabalho, chamaremos de mecanismos de processamento, que diz respeito aos procedimentos, técnicas, instrumentos e meios que são utilizados para realizar a avaliação.

Há de se acrescentar o papel político da avaliação externa, visto que não apenas a dimensão técnica conjuga o quadro de análise da natureza e dos propósitos da avaliação,

mas também suas dimensões políticas e ideológicas, na medida em que condicionam a interpretação que é dada pelos participantes, a utilidade creditada à avaliação e o uso que se dá aos seus resultados (Afonso, 2005; Bauer et al., 2015; Worthen, Sanders & Fitzpatrick, 2004).

Questionados sobre as características que melhor definem as avaliações externas, o grupo de especialistas apresentou respostas diversas com relação à natureza e aos propósitos da avaliação externa. No entanto, é possível verificar que, para os especialistas participantes desta pesquisa, as avaliações externas são de natureza essencialmente somativa em detrimento da formativa.

De acordo com Harlen e Gardner (2010), a avaliação não é em si mesma formativa ou somativa, mas é a partir de seus propósitos que há a distinção entre ambas. De acordo com os autores (citado por Santos, 2016), a avaliação, com a mesma informação e recolhida no mesmo modo, será nomeada de formativa, caso seja usada como apoio à aprendizagem e ao ensino, ou somativa, se for apenas utilizada para registrar e reportar.

Uma avaliação somativa é considerada como um “balanço final”, uma análise sobre o todo, tendo como possibilidade a verificação dos resultados da aprendizagem e a possibilidade de classificação (Harlen & James, 1997; Ribeiro, 1991). De acordo com Santos (2016), a informação produzida pela avaliação somativa dirige-se, sobretudo, a entidades externas aos alunos, isto é, à comunidade escolar (professores e gestores escolares), aos gestores políticos e à sociedade. Isso faz com que essa avaliação seja caracterizada por uma dimensão social.

Os propósitos considerados pelos especialistas reforçam sua natureza somativa, pois a consideram pontual, na medida em que realizam um balanço, num determinado momento do processo, para identificar o que os alunos sabem e são capazes de fazer, além de uma medição do desempenho dos professores. Com isso, seria possível mensurar o desempenho dos alunos, monitorar objetivos de ensino e metas dispostas por políticas públicas; avaliar estas políticas e a qualidade da educação. A natureza de ordem mais somativa da avaliação externa apresentada pelos especialistas participantes coaduna com a ideia defendida por Nevo (2010), na qual os avaliadores externos executam melhor uma avaliação somativa.

Fica evidente, nas narrativas dos especialistas, que a avaliação externa está sendo empregada apenas como um instrumento de medição. Cabe salientar que a medição de uma determinada realidade não corresponde à sua avaliação, nem mesmo a somativa. Para se conhecer dada realidade, deve-se realizar um conjunto de medidas que trazem

informações sobre a realidade que se deseja conhecer, porém essas medidas sozinhas têm pouco significado. Para que haja significado, é necessário que ocorra uma comparação a uma determinada situação anterior, que já tenha sido objeto de medição, ou que as duas sejam comparadas com uma situação ideal. Com isso, pode-se fazer um julgamento sobre a situação atual. A realidade só será avaliada se for emitido um juízo de valor. Por fim, cabe sublinhar que realizar uma avaliação de um sistema educacional que tenha um rendimento abaixo de um índice considerado ideal, por si só não melhora o desempenho, é necessário que se identifiquem as causas e se realizem ações para superar as dificuldades encontradas (Horta Neto, 2010).

A medição, ou verificação como apresentada por Luckesi (1998), é um processo que se encerra no momento em que se obtêm os dados ou as informações que se busca, sendo assim um processo estático e definitivo. Ao se colocar como o sujeito que realiza apenas a medição, o governo transfere para as instituições escolares a responsabilidade por fazer alguma ação no sentido de utilização dos resultados gerados e melhora dos índices. No entanto, como afirma Gatti (2012, p. 32), ao olhar o modelo das avaliações externas utilizadas universalmente “. . . há pouca informação que possa alimentar e orientar processos de ensino. O dado fica como provocador, supondo-se, parece, que cada escola, cada rede se ‘vire’ para atingir metas teóricas propostas”.

Na construção da natureza da avaliação externa, foram levantados também alguns mecanismos de processamento, possibilitando conhecer mais sobre a natureza da avaliação externa, a saber: sujeito responsável por avaliar externo à escola; provas padronizadas; comparabilidade entre as informações coletadas; mecanismos de processamento recentes; algumas etapas do exame são públicas (conteúdos, critérios de correção e resultados); processo inteligível para o público; validade e fidedignidade. Considerar esses mecanismos de processamento evidencia a necessidade de haver um instrumento que tenha métodos que sejam confiáveis, com validade e garantia de qualidade, sendo essas características necessárias a uma avaliação somativa (Harlen & James, 1997), principalmente no que tange à avaliação somativa externa. Há, por conseguinte, uma preocupação com o aspecto técnico da avaliação.

Ao caracterizar as avaliações externas, os especialistas destacaram seus efeitos, porém o fizeram evidenciando os efeitos negativos, como se pode destacar na fala de E-C: “Ela hoje tem este impacto ruim.”. Tal postura coaduna com a pesquisa realizada por Stecher (2002), um estudo da experiência com testes de alto impacto nos EUA e um estado da arte sobre o efeito desses testes. O autor considera que os testes podem

apresentar efeitos positivos, negativos e nulos / ambíguos. O estudo conclui que a maior parte das pesquisas americanas é realizada em nível de sala de aula, ou seja, em relação ao currículo, controle do professor, prática de ensino; e pouco se conhece com relação aos efeitos dos testes sobre os decisores políticos. Outra conclusão do estudo diz respeito à incerteza do efeito líquido dos testes, tendo em vista que

. . . os pesquisadores não têm documentado as consequências desejáveis dos testes – de mais instruções, trabalhando mais e de forma mais eficaz –, tão claramente quanto os indesejáveis – como alocação negativa do tempo de aula para enfatizar tópicos abordados pelo teste, treinamento excessivo e trapaça (Stecher, 2002, p. 99, tradução nossa).

Observa-se que os especialistas destacaram apenas efeitos negativos provenientes da avaliação externa, também com ênfase no efeito no contexto de sala de aula, como: o estreitamento do currículo, a responsabilização dos professores e a naturalização entre os diferentes intervenientes de que o teste de desempenho é uma boa medida para se avaliar a qualidade da educação.

De acordo com Bauer et al. (2015), há uma parcela significativa da comunidade acadêmica que desconsidera as eventuais contribuições que podem se originar das medidas educacionais de larga escala, havendo uma maior incidência da difusão de posições críticas. Essa postura dificulta a verificação dos reais impactos positivos e negativos da avaliação externa, tanto no contexto de sala de aula quanto em decisões de políticos públicos. Assim como afirma Stecher (2002), mais pesquisas devem ser elaboradas para determinar algo sobre o assunto.

Há uma grande discussão, no campo de políticas públicas, sobre o que ocorre nos documentos governamentais, na elaboração dessas políticas, e o que acontece na prática. Com isso, qualquer análise de objeto relacionado a políticas públicas devem compreender que há uma distinção entre o que o governo pretende fazer e o que efetivamente faz, gerando, assim, uma divisão das etapas das políticas públicas. No campo da educação, essa relação entre o planejado e o que efetivamente ocorre não é diferente, tanto no contexto da escola quanto na relação entre políticas públicas e o dia a dia escolar.

A partir da fala do especialista E-G, pode-se perceber que há uma mensuração, uma análise dos efeitos produzidos pelas políticas públicas, sendo, portanto, uma avaliação *a posteriori* (Saraiva, 2006), tendo em vista que trabalha com a diferença entre o que se propõe e o que realmente acontece ou o que se espera de uma avaliação. Refere-

se às características presentes no plano ideal, marcando, dessa forma, que há avaliações que não cumprem com as características esperadas e/ou apresentam características que não deveriam ocorrer, o que concretiza a ideia de marcar uma diferença entre o ideal e o real. “Então todo discurso é um discurso, mas a prática é outra. (E-B)”. “A única coisa que eu queria esclarecer é que aí a gente está falando de um plano ideal . . . essa é minha opinião sobre as características que deveriam definir uma avaliação externa (E-G)”.

Outro ponto que se considera oportuno destacar, com base na fala dos especialistas, é o fato de todos os níveis de governança no Brasil possuírem avaliações externas. Isso faz com que se possa apresentar várias avaliações com diferentes naturezas e propósitos, como se visto na fala de E-G: “. . . eu acho que as avaliações externas do governo federal têm algumas características, as avaliações externas estaduais têm outras características, avaliações externas municipais têm outras.”.

Essa fala marca o posicionamento de que, a depender de quem está gerindo as avaliações, ou seja, o nível governamental que conduz a avaliação, as características da avaliação externa podem mudar, especificamente com relação às consideradas pelo especialista, isto é, validade, fidedignidade, consistência para com o fim proposto. O que pode levar a uma maior credibilidade, a uma determinada avaliação em detrimento de outra.

Os especialistas afirmaram que, por ter uma divulgação pública e uma série histórica do processo e dos resultados, as avaliações federais tendem a ter uma qualidade maior do que as avaliações estaduais e municipais. Tal fato pôde-se verificar ao longo da recolha dos dados deste trabalho, pois houve uma grande dificuldade em localizar informações e dados de avaliações do estado do Rio de Janeiro, principalmente as do município. No entanto, Alvarse et al. (2012), em pesquisa de diferentes avaliações dos governos federal, estaduais e municipais no Brasil, afirmam que há maior participação de vários segmentos profissionais nos processos de avaliação municipais, como gestores e professores, o que proporciona maior aceitação, validação e potencialidade para o desenvolvimento de política local. A qualidade da avaliação, assim, pode ser vista não apenas pelo seu aspecto técnico, mas também pelo grau de envolvimento dos participantes.

No entanto, outro especialista participante referiu que, independentemente de qual nível governamental as esteja conduzindo, todas as avaliações externas seguem um determinado modelo com idênticos propósitos e não têm feito mais do que medir o desempenho dos alunos.

Na maior parte das vezes, aquilo que está acontecendo aqui no Brasil, de uma forma geral, seja nos estados, seja no governo federal, seja nos municípios, a avaliação está se resumindo a medir desempenho de aluno. Ela não tem passado disso. . . . de uma forma geral, todas as avaliações repetem o mesmo modelo. Então é aquilo que a gente comenta, a gente está fazendo mais do mesmo, aumenta-se o número de avaliações, testam-se sempre com base na Prova Brasil, na matriz de referência da Prova Brasil, do Saeb (E-B).

Com isso, entende-se que as avaliações se diferenciam a depender da instância gestora, ou seja, se é a federação, o estado ou o município. Todavia, há uma percepção de repetição de modelos das avaliações para um único fim – medir o desempenho do aluno, não indo além de uma medição, seguindo apenas uma lógica de gerar resultado.

. . . de modo geral, o que eu percebo é uma grande neurose, uma grande preocupação com os resultados, e para isso vale tudo, desde você abduzir aluno até responder pelo aluno. . . . Por isso, se eu fosse ministro da Educação, pelo menos por uns cinco anos, eu pararia com isso aí, vamos trabalhar em cima dos dados que a gente já tem, tem tanto dado aí . . . (E-E).

Os especialistas salientam, por conseguinte, em conformidade com Sousa e Oliveira (2010) e Bottani (1998), que há um acúmulo de informações, sem precedentes na história da educação, como resultado do processo de avaliação externa. O desempenho dos estudantes acaba por ser o foco dos sistemas, vindo a diluir a importância das informações sobre variáveis de contexto escolar que são coletadas e outros fatores escolares.

O sistema de avaliação utiliza tanto um teste de rendimento do desempenho dos alunos quanto um questionário, em que buscam realizar o levantamento dos fatores que influenciam o desempenho dos discentes. Os fatores levantados pelos questionários não estabelecem uma relação de causalidade entre os dados levantados e o desempenho do aluno, contudo, podem indicar quais as áreas que necessitam de maior atenção governamental e quais as iniciativas políticas que deveriam ser tomadas (Horta Neto, 2010). A não utilização desses fatores acaba por empobrecer o resultado obtido pelo sistema de avaliação, visto que se perde o contexto da geração do nível de rendimento acadêmico.

As perspectivas defendidas pelos especialistas participantes na pesquisa estão sintetizadas na Figura 9. A análise da Figura permite destacar que para esse grupo de participantes, as avaliações externas são de natureza essencialmente somativa. Os propósitos apresentados confirmam essa natureza, já que para os especialistas a avaliação externa é vista como pontual, na medida em que realizam um balanço, num determinado momento do processo, para identificar o que os alunos sabem e são capazes de fazer, além de uma medição do desempenho dos professores. Com isso, seria possível mensurar o desempenho dos alunos, monitorar objetivos de ensino e metas dispostas por políticas públicas; avaliar estas políticas e a qualidade da educação.

Natureza e propósitos	
E-A	A avaliação externa possibilita informação comparável sobre um conjunto grande de sujeitos, utiliza mecanismos de processamento mais recentes e algo simples que consegue transformar o processo algo inteligível para o público.
E-B	A avaliação externa está se resumindo a medir o desempenho dos alunos; não tem passado disso. Então, entender o processo de avaliação externa da aprendizagem significa reconhecer que hoje a gente faz medida de desempenho dos alunos dentro de um teste. Ela também tem servido para estreitar o currículo, na medida em que se corre atrás de aumentar o indicador, tomando como referência a matriz, deixa de praticar o currículo e corre atrás daquilo que vai ser cobrado no teste. É naturalizado entre os professores, entre os secretários de educação, entre os gestores de uma forma geral e na sociedade que o teste de desempenho, o teste cognitivo, é uma boa medida para se avaliar a qualidade da educação.
E-C	A avaliação externa tem uma forte preocupação com a medida, pelo fato de ser externa à escola. Ela não avalia aspectos particulares das escolas, pois tem uma perspectiva ampliada, panorâmica da avaliação. O que tem acontecido atualmente é que ela está atrelada à ideia de responsabilização do professor. Hoje ela tem esse impacto ruim, toma um caminho ruim aqui no Brasil.
E-D	O sujeito que é responsável por essa avaliação da aprendizagem está fora desse espaço.
E-E	Acho que o que caracteriza a avaliação externa é essa questão da padronização.
E-F	A avaliação externa é um instrumento de medição de desempenho de alunos, professores e/ou escolas, planejada por instâncias localizadas fora das escolas, e destinada a monitorar objetivos de ensino e/ou metas dispostas por políticas públicas de educação. A avaliação externa avalia política pública e, nesse sentido, não necessita ser censitária, bastando avaliações nacionais ou estaduais amostrais.
E-G	Quanto as características, ela tem que ter validade, fidedignidade, ser consistente e ser consequente para o fim que ela foi proposta.

Figura 9

Dimensão “Natureza e os propósitos”: especialistas

Com base na análise da Figura 9, cabe evidenciar que a avaliação externa da aprendizagem vem ocupando, com maior ênfase, um lugar de medição do desempenho dos alunos, segundo os especialistas, com uma multiplicidade de avaliação de diferentes níveis governamentais, vindo, em grande medida, a gerar dados sem reflexão e ação, o que leva a um processo estático, sem desdobramentos.

De acordo com Leviton e Hughes (1981), há um entendimento de que a utilização da avaliação não é um conceito unitário, mas sim há diferentes tipos, alicerçados nos propósitos a que a avaliação serve. Haja vista a diversidade de avaliações externas no cenário brasileiro, tendo a maioria dupla função, ou seja, objetivando atender tanto à demanda política quanto à pedagógica, é imprescindível discutir a relevância da avaliação externa para o processo de ensino e aprendizagem, bem como salientar o conjunto das características necessárias para isso. Isso se deve ao fato de que, ao longo do tempo, a avaliação externa teve um ganho progressivo de espaço nas agendas e discursos políticos, mas não necessariamente nas escolas.

De modo geral, todos os especialistas participantes consideram que as avaliações externas são relevantes para que haja outro olhar sobre o processo de ensino e aprendizagem. No entanto, há algumas relações que necessitam ser pensadas para que isso ocorra de forma efetiva.

A primeira relação a que se deve atentar para que os resultados sejam relevantes para o processo de ensino e aprendizagem diz respeito aos mecanismos de processamento da avaliação externa e à política educacional na qual ela está inserida. Assim, a responsabilidade estaria com quem conduz a avaliação, ou seja, “seria necessário gerar *feedback* compreensível pela escola” (E-A); “a avaliação ser mais do que simplesmente medida de resultado” (E-B); “na medida em que ela favoreça, facilite e de fato seja incorporada às discussões pelos professores.” (E-D); “na medida em que ela traz elementos, outros dados além daquele que o professor já tem” (E-E); “depende da política educacional na qual ela está inserida e de sua própria tecnologia de preparação.” (E-F).

Dessa forma, segundo os especialistas participantes, o sujeito que conduz a avaliação é determinante no sentido da relevância dela para o processo de ensino e aprendizagem. Com isso, é importante destacar que a medição por si só não melhora o desempenho dos alunos. Não é apenas de responsabilidade dos gestores políticos a medição, mas sim uma atuação após o processo, realizando análises, possibilitando incorporações e elaborando ações com os dados coletados. Assim como salienta E-B, que destaca a necessidade de haver, depois da medida, um julgamento de valor e a partir disso realizar um trabalho para modificar os resultados, com base em discussões e estratégias para desenvolver o processo de ensino. Ou seja, deve ser muito mais do que simplesmente a medida de resultados.

Segundo Souza e Oliveira (2010), para possíveis contribuições da avaliação externa na melhoria do ensino, havendo, por conseguinte, relevância e importância no

processo de ensino e aprendizagem, têm sido discutidos dois tipos de argumentos: a) a lógica interna dos processos de avaliação, realizando uma análise de seus princípios e procedimentos; e b) a utilização dos resultados na condução das políticas educacionais, examinando a validade, a relevância e a oportunidade das iniciativas adotadas como consequência do conhecimento obtido mediante o processo avaliativo. Ou seja, relaciona-se com as características levantadas pelos especialistas participantes como requisitos para que a avaliação externa seja relevante para o processo de ensino e aprendizagem, como para os mecanismos de processamento da avaliação externa e a política educacional na qual a avaliação está inserida.

A relação entre a avaliação externa e a realizada pelo professor, isto é, a interna, também é considerada pelos especialistas participantes como outro ponto que constrói a relevância da avaliação externa para o processo de ensino e aprendizagem. Trata-se de um diálogo a partir de duas perspectivas diferentes, em que a avaliação externa proporcionaria uma visão mais ampla e a interna, uma visão mais pontual, como afirma E-C. No sentido da avaliação externa, se relaciona com a interna para que, dessa forma, possa ter sentido para os atores escolares (E-F). É uma combinação entre as duas avaliações (Nevo, 1997), para que, assim, possa haver uma utilidade da avaliação externa para a escola.

A Figura 10 apresenta as perspectivas defendidas pelos especialistas, com base na Figura pode-se evidenciar que todos os especialistas participantes consideram que as avaliações externas são relevantes para o processo de ensino e aprendizagem, contudo deve haver uma maior análise de algumas relações presentes, como os mecanismos de processamento da avaliação externa, a política educacional na qual ela está inserida e a relação entre a avaliação externa e a interna.

Relevância e importância	
E-A	A relevância está no fato de gerar <i>feedback</i> compreensível pelas escolas. Se esse processo não for feito, a avaliação externa provavelmente não vai ter impacto relevante no aprendizado e nas práticas pedagógicas nas escolas.
E-B	É relevante para refletir sobre o processo, algo que vai além da medida. Depois da medida, é feito um julgamento de valores, sobre os quais, você trabalha para modificar o resultado. Você discute os resultados depois do seu julgamento de valor e estrutura estratégias para desenvolver o processo educacional. A avaliação externa é extremamente relevante.
E-C	Olhar para essa avaliação me faz pensar esses processos de aprendizagem, por isso é relevante. Quando eu penso na avaliação da minha sala de aula, dos meus alunos, eu estou pensando a particularidade da minha aula e a particularidade de cada aluno. A avaliação externa não me possibilita isso, é como se eu estivesse olhando uma fotografia panorâmica, então é uma outra perspectiva de olhar. Na combinação dessas duas perspectivas, o olhar que vem da sala de aula, da discussão mais particular, mais pontual, e a discussão mais ampliada, eu acho que tem sido bastante interessante.
E-D	A avaliação externa é bem relevante na medida em que ela favoreça, facilite, e de fato seja incorporada nas discussões pelos professores. Quando falo dessa incorporação, isso pode se dar, por exemplo, com relação às matrizes, aos resultados, aos instrumentos. Um outro aspecto é quando os resultados dessas avaliações podem ser de alguma maneira comparados, articulados com os resultados que os professores também produzem sobre aprendizagem dos seus alunos, um diálogo entre as avaliações externas e as avaliações internas da aprendizagem. Não necessariamente para uma cópia, não necessariamente para um condicionamento absoluto, mas um diálogo no sentido de olhar e, enfim, fazer essa crítica, essa análise mais aprofundada. Havendo isso, é possível começar a se dizer que há uma relevância da avaliação externa.
E-E	Eu acho que, na medida em que ela traz elementos, outros dados além daqueles que o professor já tem, sobre esse processo de ensino e aprendizagem. Isso tanto do ponto de vista interno quanto dos gestores da instituição da rede. Então, acho que ela pode ser um elemento para ajudar a pensar esse processo de ensino e aprendizagem.
E-F	A relevância da avaliação externa depende da política educacional na qual ela está inserida e de sua própria tecnologia de preparação. A sua eficácia dependerá da existência de uma cultura de avaliação própria nas escolas, construída internamente, à qual poderão se somar os resultados das avaliações externas, que podem vir a ganhar sentido para os atores escolares. Sem esta mediação, as informações provenientes de fora da escola são perdidas ou usadas com propósito de mero controle.
E-G	As avaliações externas podem tanto estar mais destinadas a um gestor de sistema quanto aos coordenadores e professores nas unidades escolares, contribuindo para repensar a forma de organizar o trabalho na escola, fazer adequações no currículo, perceber questões que têm a ver com a formação de professores. Em tese, as avaliações são instrumentos de diagnóstico importantes para dar subsídio para tomada de decisão e para readequação dos processos de ensino e aprendizagem que ocorrem nas escolas.

Figura 10

Dimensão “Relevância e importância”: especialistas

O motivo pelo qual a avaliação externa, ao longo dos quase 30 anos no Brasil, veio progressivamente ganhando espaço nas agendas políticas, nas discussões midiáticas, nas pesquisas, mas não necessariamente confirmada uma relevância no âmbito escolar, mais especificamente no processo de ensino e aprendizagem, é ainda uma questão que merece estudo. Deve-se considerar que uma das condições fundamentais para que as avaliações externas possam merecer o devido reconhecimento por parte dos diferentes intervenientes nos sistemas escolares e dos grupos sociais em geral é a sua credibilidade.

A validade técnica é um fator de suma importância nessa discussão, vindo a ser largamente abordada na literatura. No entanto, no sistema de avaliação brasileiro, ela vem sendo, na maioria dos casos, apenas inferida, o que acaba por limitar o instrumento como um indicador de qualidade e minar a confiança entre os diferentes intervenientes

(Cronbach & Meehl, 1955; Messick, 1995; Miller & Linn, 2000; Vianna, 1983;). Para além do caráter técnico, faz-se necessário ressaltar que outro critério levado em consideração pelos autores escolares é o grau de conhecimento do contexto escolar pelo avaliador externo.

Para os especialistas participantes, a credibilidade das avaliações externas para o seu grupo está associada a três fatores principais: a) aos mecanismos de processamento; b) às políticas públicas que embasam a avaliação externa; e c) às perspectivas teóricas seguidas pelo pesquisador.

Com relação aos mecanismos de processamento, podem-se observar diferentes perspectivas. A primeira é que há pesquisadores do campo da educação que não possuem conhecimento sobre o processo das avaliações externas, por isso a baixa credibilidade está ligada a uma resistência ao desconhecido. Outra visão está relacionada às críticas realizadas aos mecanismos de processamento, isto é, aos aspectos de natureza técnica (e.g., questionamento da Teoria de Resposta ao Item, da metodologia de cálculo, dos instrumentos utilizados). O último ponto estaria direcionado a uma credibilidade ligada a quem produz a avaliação. Por exemplo, as avaliações federais têm um histórico de estarem disponíveis ao público, ou seja, é possível verificar sua credibilidade. Diferentemente das avaliações estaduais e, principalmente, das municipais, que não têm um histórico tão consistente e transparente, pois, na maioria dos casos, não estão acessíveis ao público, o que já foi discutido anteriormente.

Para os especialistas participantes, as políticas públicas que embasam as avaliações externas também estariam relacionadas à percepção de credibilidade, pois muitos criticam o uso que se faz dos resultados e apresentam desconfiança com relação aos efeitos que as avaliações externas podem provocar. Deve-se destacar que há uma influência da divulgação dos resultados sobre os seus usos, uma relação entre os objetivos das avaliações e as informações por elas produzidas, assim como uma discussão política e ideológica em torno do assunto, vindo, assim, a sublinhar seu caráter político (Bauer et al., 2015). Há uma discussão de falta de credibilidade vinculada a esse assunto quando ocorrem, geralmente, políticas com fins de maquiar resultados, mascarar realidades, ganhar votos, realizar pressão e competição nas escolas etc.

Por fim, a credibilidade estaria relacionada, para além da avaliação em si, com a perspectiva teórica dos pesquisadores da área de educação. Como afirma E-G, haveria três grupos com posicionamentos diferentes. Um grupo é fortemente contrário à avaliação externa, independentemente de seu mecanismo de processamento e das políticas que a

embasam, sempre num movimento de crítica, evidenciando apenas os seus efeitos negativos sobre o processo de educação. Outro grupo acredita cegamente nas avaliações, sem discussões fundamentadas na teoria e em pesquisas empíricas, evidenciando a sua contribuição com a qualidade da educação. O terceiro grupo questiona alguns usos dados aos resultados, porém acredita que elas são pertinentes ao processo, consideram suas contribuições, contudo não deixam de referenciar seus efeitos negativos. A percepção dessa diferenciação entre os especialistas coaduna com a literatura sobre distintos posicionamentos sobre a avaliação externa (cf. Alavarse, 2013; Bauer et al., 2015; Oliveira & Araújo, 2005).

Com relação ao público em geral, os especialistas afirmam que há uma maior credibilidade, pelo fato de não haver conhecimento necessário do assunto para realizar questionamentos ou críticas. Isso evidencia uma falta de comunicação sobre a avaliação externa para esse grupo.

A qualidade do instrumento da avaliação externa é reconhecida pelos especialistas, estando relacionada com os mecanismos de processamento adotados nas avaliações, como: bom pré-teste, banco de questões pré-testado, clareza do processamento, matriz de referência, TRI. Porém, cabe ressaltar que, assim como afirma E-B, para poder verificar a qualidade, é necessário que os instrumentos e o processo sejam públicos, que venham possibilitar a consulta ao longo dos anos de aplicação das provas e a divulgação dos resultados, podendo assim fazer uma análise da série histórica. De acordo com Alavarse (2013, p. 148), deve ocorrer a meta-avaliação, pois “dado o problema de sua legitimidade política, as avaliações externas, inscritas na delimitação de seu sujeito e de suas finalidades, devem ser controladas e acompanhadas, no mínimo, em suas etapas discriminadas pelos aspectos técnicos.”. Nesse sentido, as avaliações não podem ficar sob o absoluto controle do governo, sem um controle social. Muitas acabam por não ser credíveis por falta de acesso ao processo, faltando-lhes clareza.

Outro ponto importante sobre a qualidade do instrumento é a falta de discussões no cenário brasileiro sobre suas limitações (Almeida, Dalben & Freitas, 2013; Bauer, 2012; Gatti, 2012, 2014; Horta Neto, 2012). Falta compreensão do antes, durante e depois da aplicação, marcando seu processo, suas potencialidades e, principalmente, suas dificuldades e limitações. Entender que as avaliações externas são um recurso, não um instrumento perfeito acima de qualquer problema e que irá revelar a verdade sobre a educação.

Agora não é simplesmente a questão de ter um instrumento adequado, no caso dessas grandes empresas ou mesmo do governo federal, não é simplesmente um instrumento adequado, bem feito, uma análise bem-feita dos resultados . . . mas também como é que esses instrumentos são aplicados, quais são os problemas de aplicação que aconteceram, se esses problemas podem ou não interferir no resultado final. Isso é muito pouco discutido. Outra questão que é muito pouco discutida é o erro padrão. Toda medida tem um erro e pouco se trabalha em relação aos erros-padrão, qual é o tamanho desses erros. Se esse erro pode ou não impactar mais ou menos um determinado resultado. Quando você fala de abrangência nacional, o impacto é pequeno, mas quando você começa a falar de abrangência local, escola por escola, aí o negócio muda de figura (E-B).

A Figura 11 apresenta as percepções mais relevantes que os especialistas participantes evidenciaram no contexto desta pesquisa sobre a credibilidade. A análise da Figura permite verificar que, para esses especialistas, a credibilidade das avaliações externas para o seu grupo está associada a três fatores principais: a) aos mecanismos de processamento; b) às políticas públicas que embasam a avaliação externa; e c) às perspectivas teóricas seguidas pelo pesquisador.

Credibilidade	
E-A	No meio acadêmico e para os professores da educação básica, a credibilidade é baixa, possivelmente por falta de conhecimento do processo, há muita desconfiança. O público em geral tende a considerar como válida. A qualidade da avaliação externa depende do processo de construção da prova, por exemplo, se você tem um bom pré-teste, um banco de questões já pré-testado, uma clareza do processamento.
E-B	Há uma resistência, por parte dos especialistas, que vem diminuindo com relação ao tempo. Hoje a crítica está maior no uso que se faz dos resultados. É muito difícil você ir contra o resultado da medida, você pode ter críticas na forma como foi feita a medida, nos instrumentos que foram utilizados para fazer essa medida, você pode fazer críticas com relação à metodologia de cálculos desses resultados, questionar se a TRI é tudo aquilo que se fala sobre ela, se os resultados são realmente comparados... Tem uma série de questões que podem ser discutidas e existem estudos que reforçam isso, mas acho que não existe uma descrença por parte dos pesquisadores. O público em geral acredita. Sobre a qualidade da avaliação externa, as avaliações federais têm um histórico, a maior parte daquilo que é utilizado, os instrumentos que são utilizados, as formas de se chegar a esses instrumentos, como são realizados os cálculos, tudo isso é público, então é possível você checar e verificar se existe qualidade ou se não existe. Porém, como é que esses instrumentos são aplicados, quais são os problemas de aplicação que aconteceram, se esses problemas de aplicação podem ou não interferir no resultado final... Isso é muito pouco discutido.
E-C	Entre os especialistas, depende da perspectiva do pesq, principalmente da perspectiva teórica, o envolvimento com esse tipo de estudo, esse tipo de abordagem, eu sinto que ainda tem certo preconceito em ter um envolvimento com a avaliação em larga escala ou com as questões das estatísticas. O que é divulgado para o público em geral é o resultado, o ranqueamento, então a credibilidade é diferente. Eu acho que a avaliação externa é construída com bastante credibilidade, mas eu não acho que ela é capaz de informar o que o aluno sabe e é capaz de fazer, porque o que a avaliação externa propõe é um recorte pequeno em relação a isso. Por isso que eu acho que é muito perverso tomar resultado da avaliação para definir o perfil do aluno. Eu acho que os processos teóricos metodológicos são muito bem cuidados.
E-D	No nível técnico, ela tem credibilidade. O problema, muitas vezes, não está na avaliação em si, está na política que lança mão dela. Entre os acadêmicos, tem uma clivagem importante, você tem um segmento que é fortemente contrário às avaliações, contra qualquer perspectiva psicométrica, aliás dizem que só os professores podem avaliar. Ao mesmo tempo, você tem também acadêmicos que acreditam cegamente nessas avaliações e utilizam seus resultados para fazerem uma série de inferências e suposições. Você tem esses dois grandes gráficos, no mundo acadêmico, que consideram essas avaliações como algo pertinente no processo educacional e outros que cegamente acreditam que não. Mas também há um grupo que tem uma série de críticas a essas avaliações, sem negá-las como pertinentes ao processo, ou seja, nós acreditamos que sim, elas podem contribuir, elas têm potencial, embora identifiquemos suas restrições. O público em geral não tem essa capacidade de questionamento, eles muitas vezes acabam admitindo aquilo que é divulgado pelos meios de comunicação. Então desse ponto de vista, talvez possa até assumir que acreditam mais do que os professores e os acadêmicos. Sim, são construídas com qualidade, a partir daquilo que elas delimitam como matriz, mas naquilo que elas anunciam que vão avaliar. Via de regra, elas têm um processo, a maioria delas tem uma boa consistência, ou seja, elas têm validade, avaliam aquilo que elas dizem que vão avaliar.
E-E	Com relação aos especialistas, eu vejo que há um questionamento sobre a validade dessas avaliações externas da maneira como elas são realizadas. Porque uma coisa é a gente pensar no sentido em si da avaliação externa, outra como elas são aplicadas, como elas são realizadas. Eu acho que não há muita credibilidade, sinto mais uma desconfiança. Não tanto do instrumento, mas em relação ao efeito que ela provoca. Eu sinto que há, por parte da população, por parte dos pais sobretudo, uma aceitação. Eles têm uma confiança maior nesse tipo de avaliação. Talvez os públicos que mais acreditam são os que aplicam e os pais. Também alguns gestores do sistema, que têm nela uma credibilidade maior, mas varia. Eu acho que, de modo geral, há um cuidado na preparação da avaliação externa, acho que existem teorias fundamentando, como a TRI. A minha preocupação é quando começar a ter muitas avaliações, por exemplo, as redes municipais começam a fazer as suas. Ai já não tem tanto conhecimento, já não tem tanta certeza. Mas pensando um pouco nas grandes avaliações externas, elas têm qualidade dentro daquilo a que elas se propõem. Eu acho que de modo geral elas têm qualidade.
E-F	Em geral, para os especialistas, entendo que as avaliações externas gozam de pouca credibilidade, considerando o que pretendem fazer, ou seja, ser um espelho do desenvolvimento do estudante. Os pesquisadores são mais céticos pois estudam cientificamente a questão e, portanto, têm mais informações do que os pais e o público em geral. O público em geral segue o senso comum e a opinião da grande imprensa que usualmente acredita nelas e faz das avaliações matéria-prima de suas manchetes. O grande público associa avaliação à meritocracia e eficiência. Quando as avaliações usam TRI, inspiram mais confiança. No entanto, mesmo nesses casos, a medição de habilidades e competências de nível alto não encontra uma formulação adequada. Portanto, o melhor avaliador do desenvolvimento dos estudantes é o próprio professor. As avaliações externas devem ser destinadas a medir políticas públicas, ou seja, a ação dos próprios governos.
E-G	A credibilidade vai depender da avaliação, se ela é uma avaliação estadual, municipal... Quem faz, a depender de quem gera essa avaliação, a confiança dos dados produzidos é maior ou menor. Tem muita gente às vezes que faz críticas às avaliações sem compreender também as suas especificidades técnicas, os seus próprios objetivos. Tem muita gente que contesta, parece não confiar ou não concordar com o processo avaliativo, mas não necessariamente tem conhecimento de causa do que está acontecendo realmente no contexto escolar. Então eu acho que as perspectivas são diversas. Com relação ao professor, eu acho que essa credibilidade varia, você vai ter profissionais que vão sempre ter um olhar muito desconfiado. Eu nem sei se é credibilidade. Então eu diria que isso vai mudar muito de agente para agente. Sobre a qualidade da avaliação externa, eu acho que você tem casos e casos. O que eu posso te dizer é que há uma diversidade de ações avaliativas acontecendo no momento, algumas eu afirmaria com maior qualidade técnica do que outras.

Figura 11

Dimensão “Credibilidade”: especialistas

Como já referido, há um grande número de avaliações externas, geridas por diferentes instâncias do sistema educacional no Brasil. Tal fato permite o desenvolvimento de diferentes propósitos, ou seja, a avaliação externa gerida pelo governo federal apresenta pontos semelhantes, porém também propósitos e usos diferentes de uma avaliação administrada pelo governo estadual ou municipal, por exemplo. Outra questão é a utilização de uma mesma avaliação para diferentes propósitos ao mesmo tempo. Há, por conseguinte, diferentes usos, desde gerar informações para a análise da rede, a utilização do desempenho do aluno na avaliação externa para compor a sua nota escolar, até a regulação do trabalho docente.

No entanto, para além dos usos que estão sendo dados, se faz relevante entender qual é a utilidade da avaliação para os diferentes participantes, no sentido de aprofundar o estudo e para verificar se a avaliação está sendo útil. A utilidade de uma avaliação é vista a partir da necessidade que um determinado grupo de interessados tem em relação aos seus resultados, devendo a avaliação ser julgada com base em sua utilidade real para os diferentes intervenientes do processo educacional (Patton, 2005).

A utilidade para os especialistas participantes pode ser vista tanto no âmbito do sistema educacional quanto no da escola. No sistema, a utilidade estaria voltada para a tomada de decisões, desenvolvimento de políticas públicas, monitoramento para gerar indicadores, estando, assim, voltada para melhoria das redes de ensino. Para a escola, a utilidade da avaliação externa seria a de um instrumento pedagógico, para reflexão e apoio ao processo de ensino e aprendizagem. Desse modo, prevê que elaborem medidas que visem à melhoria de ações pedagógicas ou à verificação do rendimento dos alunos ao longo de um determinado período.

No entanto, observa-se que há uma maior tendência dos especialistas em considerarem apenas a utilidade para o sistema. Isso pode ocorrer pelo fato de não acreditarem numa avaliação que tenha utilidade no âmbito escolar, no sentido de se obter resultado por escola ou por aluno, o que confirma a natureza somativa da avaliação para esse grupo.

Você falar em avaliação educacional para apresentar resultados por escola, para mim não faz nenhum sentido. Quando você está falando de sistema, faz algum sentido, na medida em que você tem uma rede relativamente grande, dispersa, com uma série de peculiaridades, com muita diferença e você quer saber o que está acontecendo nessa rede. Então a avaliação faz algum sentido. . . . trabalhar

em termos do que vai acontecer em cada uma das escolas, não faz sentido nenhum (E-B).

Outro ponto seria não haver um trabalho com os atores educacionais, o que seria fundamental para que a avaliação externa fosse verdadeiramente útil para a escola. Sem isso, a avaliação externa torna-se um recurso sem utilidade para os usuários no âmbito escolar.

Eu acho que você não consegue realizar essa função se você não propõe junto todo um aparato de apoio ao professor, de material, de formação. Porque senão parece que os professores não melhoram a situação porque não querem, às vezes eles simplesmente não sabem como naquele contexto (E-G).

Para Nevo (1997), deve haver um treinamento básico com os profissionais da escola, com o objetivo de fornecer conhecimento sobre a avaliação externa e a sua linguagem, para que, assim, os sujeitos escolares possam compreender as suas vantagens potenciais, para as necessidades específicas de suas escolas. O envolvimento dos professores e gestores escolares com as avaliações externas depende, em grande parte, da percepção e do conhecimento que possuem sobre estas.

No entanto, há um posicionamento de que a avaliação externa estaria direcionada para uma política a serviço do controle e desmoralização de professores e escolas públicas, sendo utilizada para esses fins pelo governo. A divulgação pela mídia dos constantes baixos resultados obtidos por essas escolas, acarretaria a uma descrença da população, gerando, assim, evidências da necessidade de acabar com o sistema público de ensino em prol de uma educação de qualidade privada.

Nas políticas de responsabilização (*accountability*), elas servem para controlar os professores e as escolas. Também estão inseridas em exigências que se destinam a desmoralizar as escolas públicas e seus professores, criando as bases (na grande mídia e junto aos pais) para que haja a privatização do sistema educacional (E-F).

A análise da Figura 12 evidencia que há uma maior tendência dos especialistas participantes em considerarem apenas a utilidade para o sistema, com baixa utilidade para a escola.

	Utilidade
E-A	A avaliação externa serve para orientar a tomada de decisão em nível de política pública, tanto formação de serviço quanto os estudos curriculares. Ela dá muitas informações que são muito úteis para políticas públicas. E dependendo do tipo de divulgação que você consegue fazer na escola, ela pode ser usada como instrumento pedagógico.
E-B	A avaliação externa só tem sentido quando você trabalha avaliação educacional no nível do sistema educacional. Você falar em avaliação educacional para apresentar resultados por escola, para mim não faz nenhum sentido. Quando você está falando de sistema, faz algum sentido, na medida em que você tem uma rede relativamente grande, dispersa, com uma série de peculiaridades, com muita diferença e você quer saber o que está acontecendo nessa rede. Fazer uma avaliação nacional, como a gente faz, com instrumento a partir do governo federal, e chegar o resultado até a escola, não faz nenhum sentido. Porque o governo federal não é o gestor da rede, portanto ele tem pouca governança sobre aquilo que vai acontecer na rede. Não é ele que constrói escola, capacita professor, contrata professor, define as políticas locais. Simplesmente mostrar o resultado é muito pouco para tentar mudar alguma realidade. Então, do nível federal para o local, a influência para mim é zero. Nos municípios, com avaliação dos sistemas educacionais, eu tenho como intervir rapidamente no processo e ir verificando as condições locais, porque eu tenho outras informações além daquelas que os testes me dão. Então é possível o governo municipal intervir de alguma forma. Nos estados, fica também complicado imaginar uma ação do governo estadual para suas escolas, mesmo que sejam só as escolas da rede. Tem uma distância muito grande, então pode pegar esses resultados e, por exemplo, trabalhar com formação de professores, trabalhar com algumas políticas de incentivo, faz sentido.
E-C	Tem uma função de monitoramento, para apontar indicadores para a política pública no campo de educação, mas indicadores mais gerais, e não pressionar a mudança de currículo ou os professores como tem acontecido. Eu acredito que a avaliação tem uma função importante para possibilitar esse diagnóstico mais amplo sobre o processo educacional.
E-D	A grande utilidade da avaliação externa é permitir comparações. E as comparações não são necessariamente ruins, pode chamar de ranqueamento, ordenamento, isso não é um problema em si, o problema é o que eu quero fazer com isso. Há preocupação central em encontrar a melhor e a pior escola, ou ter ideia de onde aquilo que está sendo avaliado está com maior ou menor grau de proficiência. Desse ponto de vista, pensa-se como um gestor que vê um conjunto. Uma outra utilidade, que em geral pouco se vê, é que elas poderiam ensinar, permitir, facultar ações para orientar políticas públicas. Por exemplo, detectar dificuldades, necessidades que algumas escolas têm mais do que outras. Então você pode ter uma política como se fosse uma bonificação invertida, por exemplo, você vai dar mais para quem tem menos. Há um terceira utilidade que é desenvolver políticas de formação dos professores em avaliação educacional.
E-E	Eu acho que, em relação àquilo a que ela se propõe, é justamente trazer novos dados para reflexão do processo de ensino e aprendizagem, o que possibilita a tomada de decisão, a melhoria da qualidade. Acho que seria o horizonte desejável, seria mais um elemento além dos processos internos de avaliação que a escola já tem. Mas historicamente eu acho que não é para isso que ela vem servindo, mas sim para engessar o sistema. Se ela fosse feita por amostragem, eu acho que ela poderia perfeitamente cumprir essa função. Mas, na medida em que ela tem essa pretensão de universalizar, de avaliar cada aluno, ela acaba funcionando como grande instrumento de engessamento das escolas.
E-F	Nas políticas de responsabilização (<i>accountability</i>), a avaliação externa serve para controlar os professores e as escolas. Também está inserida em exigências que se destinam a desmoralizar as escolas públicas e seus professores, criando as bases (na grande mídia e junto aos pais) para que haja a privatização do sistema educacional. De fato, as avaliações obrigam professores e escolas a se limitarem ao currículo padronizado previamente prescrito. Tudo que não cabe na avaliação externa fica prejudicado.
E-G	A motivação inicial ou essencial da proposta da avaliação externa era propiciar uma análise, fazer um diagnóstico sobre o rendimento escolar dos alunos em geral, fazer um acompanhamento, propiciar um retrato daquele sistema para que, principalmente, o gestor pudesse pensar as intervenções mais adequadas para sua rede de ensino, a partir do diagnóstico. Para esse tipo de ação, a gente tinha os modelos amostrais, que em tese dariam conta. Há uma certa tendência – sobre a qual eu realmente tenho dúvidas – de entender a avaliação como função de acompanhamento de políticas, o que significa propiciar intervenção direta na escola. Eu acho que você não consegue realizar essa função se você não propõe junto todo um apoio ao professor, de material, de formação. Porque senão parece que os professores não melhoram a situação porque não querem. Às vezes eles simplesmente não sabem como fazer isso naquele contexto, então existe essa função, mas eu não sei se ela se concretiza.

Figura 12

Dimensão “Utilidade”: especialistas

De acordo com Patton (2003), a “avaliação centrada na utilização está preocupada com a forma como as pessoas reais no mundo real aplicam os resultados da avaliação e experimentam o processo de avaliação. Portanto, o seu foco está no uso pretendido pelos usuários.” (Patton, 2003, p. 223, tradução nossa). Assim, é relevante verificar como se dá a utilização da avaliação pelos diferentes intervenientes do processo.

A utilização da avaliação externa ocorre de duas formas: a partir dos resultados, proveniente das descobertas da avaliação; e atrelada ao processo ou desenho da avaliação (Alkin & Taut, 2003; Patton, 1997, 2003; Silva et al., 2013). Neste trabalho, optou-se por dar maior foco à utilização da avaliação externa com base nos resultados, nos diferentes níveis, desde a escola, com o professor e o gestor escolar, até o governamental, com os decisores políticos. A escolha decorreu pelo fato de haver pouca divulgação dos processos ou desenho da avaliação, o que levaria a uma utilização menor que a de resultados. A utilização dos resultados da avaliação apresenta maior enfoque hoje com relação à avaliação externa, tanto no que concerne à pesquisa acadêmica quanto ao acesso dos diferentes intervenientes. De acordo com os especialistas, há uma baixa utilização dos resultados da avaliação pelos diferentes intervenientes do processo, como observado.

Por parte dos professores, a utilização é baixa ou até mesmo nenhuma, sendo levantados três fatores que justificariam tal postura. O primeiro decorre de problemas na comunicação dos resultados, tendo em vista que estes não chegam à escola ou não chegam em tempo hábil para a utilização ou, ainda, o professor não consegue compreendê-los. Isso tem a ver com a forma de comunicação do resultado, que provém de um problema de quem conduz a avaliação e da política pública que a embasa, na medida em que os resultados demoram a chegar ou não ocorre uma formação adequada do docente para entendê-los e utilizá-los.

. . . os resultados das avaliações externas, usualmente processados por terceirizadas que vendem serviços, chegam muito atrasados para os professores e não podem ser levados em conta no seu trabalho (E-F).

Em algumas escolas, eles [professores] até tentam, então eles pegam aqueles resultados e tentam decodificar, mas aí tem uma série de problemas, a forma como é divulgado o resultado da avaliação externa a partir da TRI, a partir dos níveis que são colocados, dos níveis que são escritos, a própria descrição é muito abrangente, ela é muito generalista. Então não dá para você ter certeza do que está acontecendo (E-B).

O segundo fator consiste na não utilização dos resultados na dinâmica escolar, por entender que não há nenhum ganho em utilizá-los, não havendo, assim, uma apropriação do resultado como um apoio pedagógico. Dessa forma, a avaliação externa seria considerada apenas pela sua natureza somativa, não ocorrendo sua incorporação na

dinâmica escolar, o que direciona para o fato de ser um instrumento de utilização de gestores de sistema e políticos, e não de professores.

. . . na academia ficam brigando, um acusando o outro [...], que as avaliações são decisivas para o bem do Brasil ou que elas estão acabando com o Brasil. Você vai para as escolas, a imensa maioria dos professores, claro que sabem que elas existem, até olham para elas, mas se tirar as avaliações externas ninguém vai sentir falta e o mundo vai continuar (E-D).

. . . tenho até feito assim uma série de discussões com professores ou oficinas de como entender o resultado, porque a avaliação externa trabalha com uma outra lógica que é diferente da que o professor está habituado a usar (E-C).

Outro posicionamento que surge em relação à utilização da avaliação externa é a preocupação de que ela venha a substituir a avaliação interna, ou seja, ser considerada mais eficaz, numa concepção de que o professor não teria a competência de avaliar seu próprio aluno. Representa, nesse caso, a não utilização por resistência, por causa da política pública que embasa a avaliação, em geral uma política de *accountability*.

Há uma falsa ideia de que os professores não conhecem os seus estudantes e que precisam, portanto, de avaliações externas imparciais para que possam reprogramar sua forma de lidar com os estudantes que não estejam aprendendo. Nenhuma avaliação externa pode substituir ou ser mais competente que o próprio professor para saber o que seu estudante precisa que ele faça. Ele convive diariamente com o aluno (E-F).

Cabe ressaltar que, não havendo um diálogo, evidencia-se a falta de motivação dos professores por usar a avaliação externa, que lhes é apresentada através de um discurso coercivo de uma só via. É importante que ocorra uma compreensão por parte dos professores dos meandros da avaliação e também das alternativas de seus usos para que possa ser utilizada como recurso pedagógico (Alvarose et al., 2017; Nevo, 1997).

Quando ocorre a utilização dos resultados pelos professores, o intuito é realizar um treinamento para a prova, não havendo, assim, um movimento de melhoria da aprendizagem dos alunos, mas sim uma busca por melhoria dos resultados (Darling-Hammond, 1994), para atender às exigências do gestor escolar, que, por sua vez, busca responder a pressões superiores.

. . . eu já fiz pesquisa que identificou que os professores pouco estavam preocupados com aquela avaliação, tirando os momentos em que estava na época de preparar alunos para avaliação. . . (E-G).

Com relação aos gestores escolares, os especialistas consideram que também há um baixo nível de utilização, mas quando utilizam é para controle e cobrança, tendo em vista que há uma pressão verticalizada (Freitas, 2007; Moe, 2003), sendo, desse modo, uma imposição por meio de poder hierárquico.

. . . o pouco que eu tenho lido, que eu tenho conversado é sempre na perspectiva de pressionar os professores, para melhorar os resultados, então é um uso, no meu entendimento equivocado (E-C).

O objetivo das avaliações externas não é avaliar o professor ou a escola, mas sim avaliar as políticas públicas dos governos. No entanto, elas têm seus objetivos invertidos, como se pudessem avaliar o trabalho dos professores e os estudantes. Portanto, para os professores, elas são inúteis (E-F).

No que tange aos decisores políticos, os especialistas consideram que eles utilizam as avaliações para medir ou pressionar por resultados e para a formulação de políticas públicas. Neste caso, considera-se que a utilização ocorre com o intuito de obter melhoria da qualidade da educação ou por uma política de *accountability*.

Boa parte delas utiliza a avaliação como forma de controle e cobrança no interior das escolas, principalmente quando estão associadas a consequências fortes como pagamento de bônus por desempenho. Mas, mesmo quando não estão, há uma pressão via escala hierárquica – Secretaria da Educação – Coordenadores regionais – Supervisores – Coordenadores – Diretores da escola (E-F).

A partir de duas pesquisas voltadas para o levantamento da utilização da avaliação (Brooke & Cunha, 2011; Silva et al., 2013), pode-se diferenciá-la segundo os usuários – por instância de gestão educacional governamental, como também por instância local, ou seja, a escola – e segundo o tipo – de acordo com o processo ou desenho da avaliação e dos resultados. Para esta tese, optou-se por realizar um levantamento da utilização da instância de gestão educacional governamental, para a verificação do uso com o propósito de controlar o currículo e regular e reorientar o trabalho docente. A escolha decorre do

fato de que, ao se introduzirem políticas de responsabilização baseadas em consequências simbólicas e materiais, ou seja, avaliações de segunda e terceira geração no Brasil, o propósito é fazer com que professores trabalhem mais em prol da aprendizagem dos alunos. Contudo, evidências nacionais e internacionais apontam que o uso de resultados das avaliações de terceira geração (*high-stakes testing*) para informar iniciativas de responsabilização forte pode envolver riscos para o currículo escolar e realizar práticas de ensino voltadas para o teste (Bonamino & Souza, 2012; Nichols, Glass & Berliner, 2012).

Os especialistas consideram que a utilização da avaliação externa para controle de currículo não ocorre, o que não passaria de discurso:

Então existe até um discurso de que a avaliação de alguma forma conseguiria controlar os processos educacionais, e algumas pessoas vêm com discurso moralista, garantido que todos os alunos, em qualquer escola, em qualquer situação, sejam apresentados aos mesmos conteúdos, e seja feito todo um trabalho para que os alunos apreendam esses conteúdos e tal, então isso faz parte do discurso (E-B).

No entanto, pode-se perceber que o que os especialistas estariam dizendo sobre o fato de não acontecer o controle do currículo refere-se à ideia de as avaliações externas não estarem impulsionando mudanças em currículos inadequados (Au, 2007; Madaus, Russell & Higgins, 2009).

. . . mas aí eu vejo no sentido mesmo de vigiar e punir, de controle mesmo e não de incentivo de subsídio, de melhoramento. . . (E-E).

Estudos sobre a articulação entre avaliação externa e currículo escolar no Brasil são muito limitados e recentes. No entanto, há evidências de que há uma relação de causa e efeito, isto é, há efeitos perversos no currículo em decorrência da avaliação externa, como o fato de professores realizarem os seus trabalhos com ênfase nas habilidades, competências e conteúdos das matrizes de referência da avaliação, com vistas à preparação dos alunos para os testes (Bonamino & Souza, 2012; Cerdeira & Almeida, 2012; Esteban & Fetzener, 2015; Sousa, 2003; Sousa & Arcas, 2010).

Com relação à regulação e à reorientação do trabalho docente, os especialistas foram unânimes em considerar que a avaliação externa tem potencial para contribuir com isso, o que ocorre porque é possível realizar uma comparação da escola com as demais

da rede, para verificação de pontos a serem trabalhados. Assim, é possível induzir determinadas ações pedagógicas, discussões com professores com base nos resultados, para que assim possam pensar sobre as potencialidades e as fragilidades do trabalho pedagógico. Trata-se de uma ajuda a partir de um olhar externo em diálogo com os dados produzidos de forma interna.

. . . fazer uma discussão [sobre] como é que o professor compreende aquilo, que sentidos ele dá, como ele avalia os alunos dele em relação aquilo que está sendo colocado, como os alunos da sua escola estão posicionados em relação à avaliação. Eu acho que essa discussão pode favorecer. Eu já trabalhei em várias oficinas em que nós fazíamos isso, uma discussão sobre os resultados pedagógicos e como, a partir dessa leitura, eu posso pensar as práticas pedagógicas dentro de uma escola (E-C).

Além das possibilidades diretas, como citadas, há a possibilidade de fazer a regulação e a reorientação do trabalho docente com base na formação dos professores, o que é preferido pelo especialista E-B.

. . . é possível você agregar alguma coisa ao trabalho do professor, mas, para mim, o que faz sentido em termos de políticas públicas da avaliação educacional é a formação de professores, é o desenvolvimento de políticas para apoiar o desenvolvimento das escolas. Se a gente focar, quiser forçar o trabalho do professor em sala de aula, vai dar com os burros n'água, porque não tem como chegar até lá. O máximo que a gente consegue, com base nas informações, é formação de professores, identificar problemas que estão acontecendo com algum grupo de professores e, a partir daí, trabalhar com formação continuada com eles. (E-B)

Na Figura 13 pode-se verificar os pontos de vista defendidos pelos especialistas participantes na pesquisa. De acordo com esse grupo, há uma baixa utilização dos resultados da avaliação pelos diferentes intervenientes do processo, por diferentes razões, como: problemas na comunicação dos resultados; não entender que haja ganho com a utilização dos resultados; preocupação de que a avaliação externa venha a substituir a interna. Quando ocorre a utilização, apresenta-se mais no sentido de controle, para medir ou pressionar por resultados.

Utilização	
E-A	A avaliação externa não tem nenhuma utilidade para os docentes, porque há um problema na divulgação. Já os gestores das unidades escolares têm um nível básico de utilização, mas quando utilizam, tentam identificar os problemas que foram detectados na avaliação na escola ou na rede e tentam agir sobre eles. A utilização também é muito baixa para os decisores políticos, mas estes buscam desenvolver estímulos para resultados melhores. A avaliação externa não é usada para controlar o cumprimento do currículo e a formulação de políticas públicas.
E-B	Sobre utilidade, para o professor é muito difícil. Não é questão de dizer que os professores são mal-intencionados ou que estão querendo burlar o sistema, não é nada disso. É porque a pressão em cima deles é tão grande, que eles têm que apresentar algum resultado. Eles não conseguem entender o resultado da avaliação exatamente por causa da linguagem, que é muito hermética, não foi feita para o professor. Ela foi feita para o gestor da rede, e o gestor tenta se livrar disso. Os gestores das unidades escolares também tentam pegar aqueles resultados e decodificar, mas aí tem uma série de problemas, por exemplo, a forma como é divulgado o resultado da avaliação externa a partir da TRI, a partir dos níveis que são colocados, dos níveis que são escritos, a própria descrição é muito abrangente e generalista. Para os decisores políticos, a utilização é medir resultados. Sobre o cumprimento do currículo, é apenas discurso, não ocorre. As políticas públicas são mais no que se refere à política de indução, por exemplo a política de bonificação de professores.
E-C	A utilidade que os docentes dão para as avaliações é o treinamento para as provas, para melhorar o resultado. Os gestores pressionam os professores, para melhorar os resultados. Os decisores políticos utilizam para realizar pressão por resultados. Eu acho que, até o momento, o controle do currículo não vem acontecendo, a pressão é pelo resultado e não pelo controle do que está no currículo em si. Sobre a utilização da avaliação externa para formação de políticas públicas, são mais no sentido de política de controle de professor.
E-D	A avaliação externa não tem nenhuma utilidade para os docentes. As equipes gestoras, na verdade, acabam reforçando isso, pois nem elas próprias sabem ou utilizam adequadamente e, é claro, não repassam isso, não incentivam, não articulam no interior das escolas. Os decisores políticos também não fazem nada. A equipe vem e trabalha seriamente, entrega o resultado e não tem ninguém na rede para constituir de fato um núcleo de avaliação para ir se apropriando disso. Eles quase que fazem um culto, dão mitificação dessas avaliações e seus resultados, mas pouco investem para que estes sejam apropriados de forma política ou prática às escolas. E menos ainda se preocupam em criar, no interior das redes, um grupo de professores que vai se apropriando dos conhecimentos gerados pelas avaliações. Não há controle do currículo, o que nós temos é, via os questionários contextuais que muitas vezes são aplicados juntos com as provas, informações sobre o que os professores ensinam e de como estão desenvolvendo o seu currículo. E, até onde eu sei, pouco se tem feito a partir desses levantamentos. Com relação às políticas públicas, muito pouco nesse sentido.
E-E	Eu não vejo muita utilização por parte dos professores. Os gestores, eu acho que eles utilizam mais como cobrança em relação aos professores, as escolas que utilizam, porque eu acho que também tem isso, eu acho que utilizam pouco. Eu acho que existem os casos de orientação realmente, mas o que eu sinto é mais no sentido de cobrança. Com relação aos decisores políticos, eu diria que há um pouco essa questão de desencadear algumas políticas de formação em termos da rede e algumas formas de intervenção de orientação das unidades escolares, tentando essa melhoria no resultado. Essa questão do resultado é muito forte. Eu acho que existe mesmo essa tentativa do controle do currículo, mas eu vejo isso no sentido mesmo de vigiar e punir, de controle e não de incentivo, de subsídio, de melhoramento. Mas é aí que você pensa sobre o grau de interferência no cotidiano, eu vejo a rede municipal como um caso forte, o número de provas que os alunos estão fazendo. Era prova, a prova para a rede municipal, a prova para preparar para a prova da rede municipal, então não dava mais tempo de dar aula. Isso você está mexendo no currículo. Políticas públicas? Eu não sei, difícil isso, eu acho que formalmente sim, mas sinceramente não.
E-F	Para os professores, as avaliações externas visando <i>accountability</i> são inúteis. Boa parte dos gestores escolares utiliza a avaliação como forma de controle e cobrança no interior das escolas, principalmente quando estão associadas a consequências fortes como pagamento de “bônus” por desempenho. Mas mesmo quando não estão, há uma pressão via escala hierárquica – Secretaria da Educação – Coordenadores regionais – Supervisores – Coordenadores – Diretores da escola. Já os decisores políticos utilizam nas políticas de <i>accountability</i> verticalizada, que são as mais frequentes, são usadas como forma de pressionar as escolas e seus atores para a obtenção de metas previstas nas políticas públicas. Visam também criar legislação que privatize a escola pública. O controle do currículo é o principal objetivo com que são vistas pelos governos. Não há uma avaliação voltada para as políticas públicas e sim para a ponta, para as escolas. Em países periféricos como o Brasil, a política pública é copiada via OCDE, Banco Mundial e BID e seus interlocutores internos, como o movimento de empresários do Todos pela Educação.
E-G	Há duas tendências, eu já fiz pesquisa em que se constatou que os professores pouco estavam preocupados com aquela avaliação, tirando os momentos em que você estava na época de preparar alunos para avaliação. Eu também já vi casos em que a escola toda para, fazendo reflexão para tentar fazer adequações no trabalho, discutir currículo. Ela movimenta tanto os professores como o projeto pedagógico, eu diria que a escola se movimenta a partir das avaliações. O gestor da escola pode até pensar redução de aulas para determinado professor, considerar o perfil do professor em relação ao tipo de trabalho, tendo como referência a avaliação, então coloca os professores nos anos que vão ser avaliados, esse tipo de aspecto a gente também já observou. Atualmente, os decisores políticos têm um movimento de maior utilização dessas avaliações para proposição de ações, como materiais de apoio do professor, o incentivo de dar certo direcionamento sobre o que vai ser ensinado, a adoção de sistemas de gestão das informações. Pelo menos nos grandes centros você tem uma preocupação maior, inclusive no sentido de propor ações de política pública a partir das avaliações, uma verificação do processo. Sobre o controle do currículo, a gente tem algumas pesquisas que vêm mostrando que isso realmente ocorre, mas eu também não generalizaria os resultados. Sobre as políticas públicas, alguns gestores municipais afirmam positivamente que utilizam, mas nem todos dão exemplos desse tipo de utilização.

Figura 13

Dimensão “Utilização”: especialistas

Para Shulha e Cousins (1997), um dos fatores que pode influenciar a utilização da avaliação externa é o envolvimento dos usuários. Ao identificar uma baixa utilização dos resultados da avaliação pelos diferentes intervenientes do processo, como mostrado na Figura 13, cabe destacar que deve haver um baixo nível de satisfação, posto que, para que haja uma efetiva utilização, deve haver uma satisfação com o sistema de avaliação externa por parte dos usuários, o que não ocorre.

Contudo, é importante salientar que há outras formas de utilização, mesmo quando não há satisfação por parte dos usuários, como quando ocorre a responsabilização ou pressão verticalizada (Freitas, 2007; Moe, 2003). No entanto, tal política acaba por causar efeitos negativos, como deixar de trabalhar para a melhoria da qualidade e equidade da educação, para trabalhar no sentido de melhorar os resultados da avaliação, realizando decisões como gerenciamento de tempos e conteúdos a serem ensinados, influenciando negativamente professores e alunos ou o investimento nos melhores alunos, podendo induzir ao aumento da desigualdade (Darling-Hammond, 1994; Iaies, 2003; Madaus, Russell & Higgins, 2009; Oliveira et al., 2013). Assim, a busca por melhoria do resultado não é necessariamente uma relação direta com a melhoria da educação.

Se considerar que o processo avaliativo tem por objetivo a obtenção de informações que sejam úteis aos seus usuários, fornecendo informações importantes no sentido de sustentar tomadas de decisões, seja no âmbito da sala de aula ou de uma rede de ensino, os usuários devem reconhecer nos resultados uma relação com a realidade em que vivem. Caso isso não ocorra, dificilmente a avaliação externa conseguirá conduzir às transformações almejadas. Por isso, a satisfação dos usuários é um fator importante no que concerne à avaliação externa, pois está diretamente relacionada com a apropriação e utilização dos resultados, do processo e do desenho da avaliação.

Para os especialistas participantes, os resultados aparecem como um fator de insatisfação. Para os professores, devido à pressão proveniente da divulgação dos baixos resultados e pela cobrança. E para todos os intervenientes, a quantidade de resultados sem nenhum uso efetivo.

Acho que eles [professores] não têm uma satisfação com o processo, eu acho que as formas de divulgação são perversas. As divulgações de ranqueamento e as pressões que os professores estão sentindo, acho que eles se sentem pressionados com essa forma, que isso não deve gerar nenhuma satisfação (E-C).

Eu sou da tese de que deveria parar por uns cinco anos, acho que a gente já tem dados demais, vamos levar isso a sério e parar de fazer avaliação externa, e acho que se parasse ninguém iria à rua ou à internet para reclamar (E-E).

Na fala dos especialistas, o que mais apareceu foi a satisfação dos gestores políticos. Para os outros grupos, acredita-se que não exista um consenso sobre a satisfação com relação à avaliação externa, como pode ser visto na Figura 14.

Satisfação dos diferentes intervenientes	
E-A	Difícil dizer sobre a satisfação dos diferentes intervenientes. Tem rede que está satisfeita, tem rede que não. E na academia não é consolidado.
E-B	Sim, eu acho que existe um consenso entre os gestores de que a avaliação externa é o principal instrumento para melhoria da qualidade da educação. Entendo que o que eles entendem por melhoria de qualidade da educação é um resultado alto nos testes, então isso é consenso. Nos outros grupos acho que não há um consenso.
E-C	A satisfação com relação aos gestores talvez exista, mas com os professores não.
E-D	Com os governos, parecem que sim, os que adotam essas avaliações, que vão e realizam. Agora em geral, nas escolas, há mais um espírito de descrença, de descontentamento com essas avaliações ou então indiferença.
E-E	Não, eu acho que quem está satisfeito são os que fazem as avaliações externas e os pais. Os outros talvez, alguns gestores municipais, alguns gestores escolares, mas eu não vejo assim como uma coisa desejada.
E-F	Penso que não. Uns e outros, por diferentes motivos, estão insatisfeitos. Os gestores dos países que usam <i>accountability</i> verticalizada, já há algum tempo, não obtiveram os resultados que esperavam (veja Chile e USA). Os professores estão desmoralizados com estas políticas. Os estudantes pressionados por uma enxurrada de testes. Os pais preocupados com a sobrecarga de testes sobre seus filhos gerando doenças. Os únicos que ganham com esta política são os empresários, que, ao privatizarem as escolas, as transformaram em comércio que gera mais lucro.
E-G	No caso dos gestores, se eles propõem, é porque eles têm uma certa expectativa de utilizar aqueles resultados. Falar de satisfação de professor é mais difícil, pois tem professor que acha adequado, tem professor que não acha adequado, não tem como quantificar o fenômeno.

Figura 14

Dimensão “Satisfação dos diferentes intervenientes”: especialistas

Esta primeira parte tentou caracterizar a avaliação externa sob a perspectiva dos especialistas. Pode-se notar que as dimensões apresentam interligam entre si, o que propicia um quadro coeso sobre o objeto de pesquisa analisado. As linhas que se seguem versam sobre a qualidade da educação, buscando entender como se dá a sua relação com a avaliação externa.

Qualidade da educação.

Pensar a avaliação externa como indicador para nortear políticas e programas educacionais ou verificar a qualidade da educação não é um consenso na literatura. Há três grupos que se destacam em distintos posicionamentos acerca desse tópico. O primeiro considera que a avaliação externa tem a capacidade de revelar a verdade, tendo, por isso, um caráter de garantia absoluta. O segundo grupo desconsidera a avaliação externa como um indicador com potencial de fornecer contribuições para o sistema educacional. Por fim, um terceiro grupo reconhece possíveis contribuições advindas da avaliação externa, mas se contrapõe à forma como vêm sendo difundidas, principalmente no que tange a ser considerado como indicador único e principal de qualidade do ensino (Alavarse, 2013; Bauer et al., 2015; Oliveira e Araújo, 2005).

Há ainda um posicionamento de que, a partir da análise dos últimos anos de avaliação em larga escala, verificou-se que não há potencial para essas avaliações melhorarem a qualidade do ensino, visto que pouco se alterou no quadro educacional nos locais que a utilizam como instrumento de gestão (Sousa & Oliveira, 2010; Tedesco, 2003). Segundo Tedesco (2003, p. 13), “a informação dos resultados não se transformou num insumo para a melhoria das políticas, os dados não se desenvolveram nas escolas e o impacto sobre as estratégias de ensino foi muito débil.”

De acordo com Nevo (1997), mudar a educação é mudar a escola. Assim, se a avaliação externa tem como objetivo o aprimoramento de um sistema educacional, deve ocorrer de forma a transformar a escola. Entretanto, o autor afirma que as avaliações externas se mostram mais direcionadas para problemas conceituais relacionados à avaliação de programas e projetos, ou seja, acabam por produzir informações que são mais úteis para os gestores de redes ou para políticas do que na ajuda aos professores ou no fornecimento a pais e a alunos de alguma informação. O trabalho da avaliação externa deveria, dessa forma, ser pensado no sentido de promover uma articulação com o trabalho pedagógico escolar para aprimorá-lo. Com isso, há um entendimento de que essa avaliação pode contribuir no processo de ensino e aprendizagem, porém, para tal, deve-se pensar num aprofundamento do conhecimento pelos professores, com relação aos fundamentos, objetivos e resultados das avaliações externas.

Entendendo que a avaliação externa produz dados que podem apontar problemas e potencialidades tanto na intervenção da Secretaria de Educação quanto dos gestores escolares e também da ação do professor na sala de aula, é importante analisar como os

diferentes intervenientes pensam a relação entre a avaliação externa e o processo de ensino e aprendizagem. Como afirma Vianna (2005, p. 16):

A avaliação não é um valor em si e não deve ficar restrita a um simples rito da burocracia educacional; necessita integrar-se ao processo de transformação do ensino/aprendizagem e contribuir, desse modo, ativamente, para o processo de transformação dos educandos.

Os especialistas participantes pontuaram limitações tanto na relação entre as avaliações externas e o processo de ensino e aprendizagem quanto na ideia de que a avaliação externa permite saber o que os alunos sabem e são capazes de fazer. Essas limitações podem ocorrer por causa de alguns fatores, segundo esses especialistas, como: por ser um recorte do momento; por causa das especificações do instrumento; e por conta da relação entre a avaliação externa e os atores escolares.

É importante evidenciar que as avaliações externas podem apresentar limitações que são inerentes ao processo e aquelas que decorrem por decisões políticas ou por escolhas de determinados mecanismos de processamento. Conhecer essas limitações é muito importante, dado que “os usuários potenciais precisam estar cientes das prováveis limitações da avaliação e ter a capacidade de criar demanda efetiva que possa ser realisticamente satisfeita.” (Ala-Harja & Helgason, 2000).

Como exemplo de limitação inerente ao processo da avaliação externa, pode-se destacar o fato de o currículo da escola ser inevitavelmente mais amplo do que a matriz de qualquer avaliação, pois deve abranger diferentes tipos de aprendizagens que nenhuma avaliação externa conseguiria medir. O fato de boa parte das avaliações se propor a medir a proficiência dos alunos em apenas duas ou três disciplinas e considerar essa medida como representação de todo o trabalho realizado nas escolas é um exemplo de uma limitação por decisões políticas ou por escolhas de determinados mecanismos de processamento (Dalben & Almeida, 2015; Gatti, 2014).

De acordo com as respostas dos especialistas, pode-se observar que os dois tipos de limitações foram considerados, no entanto, com maior ênfase para limitações que decorrem de decisões políticas e escolhas de mecanismos de processamento.

. . . a gente sempre vai ter uma fotografia, e sempre nessa fotografia a gente vai escolher alguns ângulos. A gente não vai conseguir pegar todos os ângulos . . . tem-se a fotografia daquele momento. Por melhor que seja a medida estatística,

continua sendo uma fotografia de um determinado momento. Por mais que você faça um monte de controle, sempre vai continuar sendo uma fotografia, então esse é o primeiro limite. O segundo é o próprio instrumento, ele tem algumas dificuldades (E-A).

Então é possível você fazer diversos tipos de medidas, o problema é que a gente entra num modelo que é muito pasteurizado, que você usa o mesmo tipo de item para fazer a mesma coisa. . . Não que a qualidade seja ruim, mas o potencial da avaliação é muito maior do que aquilo que a gente consegue chegar. Exatamente porque a gente está repetindo, fazendo mais do mesmo (E-B).

. . . imagina um professor que está nesta escola este ano, e no ano que vem está em outra escola, então aquele que recebe a avaliação já é outro professor, que não é aquele que estava na escola quando a avaliação foi aplicada. Então aquele dado, pra ele, parece que não é dele (E-E).

A relação entre a avaliação externa e o processo de ensino e aprendizagem pode passar por duas dimensões: a de colaboração com o processo de ensino e aprendizagem e a de avaliação do processo de ensino. Na primeira, a avaliação dá subsídios para a qualidade do processo, no sentido de estruturar ou reestruturar, podendo ser usada tanto por professores e gestores nas unidades escolares quanto por gestores de redes. A segunda dimensão está relacionada à avaliação do desempenho do professor, na medida em que a avaliação externa seria usada para a verificação do desempenho dos alunos, para que, assim, pudesse verificar a qualidade do ensino. Essa segunda dimensão foi levantada de forma crítica, ou seja, os resultados da avaliação externa não podem ser considerados com base num determinismo nas relações entre eles e o trabalho dos professores, como se esses fossem os únicos responsáveis pelos resultados escolares (Alavarse, 2013; Oliveira & Araújo, 2005).

. . . sobre o ensino é mais complexo, é claro que há circunstâncias quando a gente começa analisar alguns itens, é possível supor que existam problemas de ensino, em relação aos alunos. Quando a gente começa a perceber certos tipos de erros que os alunos cometem, a gente pode até inferir que existam problemas de ensino, ou porque se ensina mal ou porque não se ensina. . . não posso pegar o resultado e fazer comparação de trabalho de professores, como fazem em redes que adotam

esses resultados para bonificar o trabalho de professores, é preciso de um monte de refinamentos que essas avaliações por si só não fornecem (E-D).

A qualidade técnica foi destacada por vários especialistas, mostrando, assim, que a clareza do processo da avaliação é um ponto importante, pois através de uma análise do processo é possível verificar a qualidade da avaliação com relação ao nível técnico. O que possibilita verificar se houve qualidade ou se a falta dessa característica pode vir a ser uma limitação do uso do desenho da avaliação e dos seus resultados. Desse modo, pode-se considerar que os especialistas participantes concordam que a avaliação externa possibilita conhecer o que o aluno sabe e é capaz de fazer, no entanto, depende de sua qualidade técnica.

Eu acho que uma prova bem feita te dá informações boas disso. Claro, você tem que ver referente à aplicação, essa coisa toda. Mas, digamos, se o processo segue um manualzinho, vai bem (E-A).

. . . o problema da avaliação é como se prepara. Estão supondo que a preparação tem qualidade, tem fundamentação teórica, mas depois como é que você aplica? Aqui já começa o problema, já é uma coisa externa, uma coisa diferente, muitas vezes isso já é um ruído forte. Depois como é que você corrige? Depois como você devolve? Demora para chegar. E, sobretudo, o que você faz com o resultado? Eu acho que, em princípio, tem um potencial de ajudar, se ela for aplicada com cuidado, por que não? É mais um elemento, mas nada que seja assim, salvadora da pátria, nada disso (E-E).

Vai depender do grau da qualidade técnica da avaliação (E-F).

A análise da Figura 15 permite destacar que para esse grupo de participantes, há limitações tanto na relação entre as avaliações externas e o processo de ensino e aprendizagem quanto na ideia de que a avaliação externa permite saber o que os alunos sabem e são capazes de fazer.

Avaliação externa e qualidade do processo de ensino e aprendizagem	
E-A	Acredito que, com <i>feedback</i> do processo, a partir de uma linguagem matemática, a avaliação externa estaria relacionada com o ensino dos professores e com a aprendizagem dos alunos, ou seja, o que você mede é feito com razoável fidedignidade. Sobre o que os alunos sabem e são capazes de fazer, acredito que seja razoável, ela consegue dar boas informações.
E-B	Não acredito que a avaliação externa esteja relacionada com o ensino dos professores e com a aprendizagem dos alunos e que consiga medir o que os alunos sabem e são capazes de fazer, pois é apenas uma fotografia de um momento.
E-C	Eu acho que tem uma relação, mas é uma relação frágil. Acho que é frágil porque a avaliação externa não é uma medida da aprendizagem, ela é uma medida de um momento, de um resultado, de uma proficiência num determinado momento. Outra coisa, a avaliação trabalha com amostra, então ela não deveria ser o resultado da escola, mas sim daqueles alunos da escola. Mas há um uso equivocado de que aquilo seja um resultado da escola. O aluno é avaliado por um determinado processo, mas ele pode ter mudado da escola, traz uma defasagem de aprendizagem de anos anteriores, traz reprovação que não ocorreu naquela escola, mas em outra escola, como é que você toma esse resultado do aluno como uma medida do professor? Por isso eu digo que é frágil. Há muito mais uma pressão para que esses usos impactem esses processos do que de fato o desenho da avaliação permitiria fazer inferências dessas naturezas. A avaliação externa me permite só um aspecto limitado de resposta, de como o aluno está fazendo, não sei se isso é uma medida do que ele sabe, mas uma medida de como ele respondeu.
E-D	Sobre a aprendizagem dos alunos, a relação ocorre a partir do processamento das respostas deles. Dá para fazer boas indicações das aprendizagens dos alunos, ou seja, aqueles resultados eles indicam coisas importantes. Agora, sobre o ensino, é mais complexo. É claro que há circunstâncias, quando a gente começa analisar alguns itens, que nos levam a supor que existam problemas de ensino. Quando a gente começa a perceber certos tipos de erros que os alunos cometem, a gente pode até inferir que existam problemas de ensino, ou porque se ensina mal ou porque não se ensina. Mas não posso pegar o resultado e fazer comparação de trabalho de professores, como fazem em redes que adotam esses resultados para bonificar o trabalho de professores. É preciso de um monte de refinamentos que essas avaliações por si só não fornecem. Outra coisa é como um professor entende isso, e como ele pode fazer uma leitura desses dados interpretados e conectá-los com os processos de ensino. Esse salto os professores em geral não conseguem fazer. Há um hiato, uma fossa abissal, entre como os professores trabalham, como ensinam e como se avalia. As avaliações externas permitem conhecer o que os alunos sabem e são capazes de fazer apenas na medida da escala do Saeb.
E-E	Sobre o processo de ensino e aprendizagem, eu vejo pouca relação. Acho isso por toda essa questão do cotidiano do professor e da sobrecarga de trabalho, e a coordenação pedagógica nem sempre adequada. A desconfiança dele em relação à avaliação externa, a rotatividade dos professores na escola, tudo isso vai tornando esse possível efeito positivo da avaliação externa em algo tênue. Eu acho que, na medida que a avaliação externa é mais um elemento do que os alunos sabem e são capazes de fazer, traz mais dados que a gente espera que sejam significativos em relação aos alunos. Acho que ela possibilita entender melhor o que está se passando com os alunos. Assim, em princípio, tem um potencial de ajudar, se for aplicada com cuidado, por que não? É mais um elemento, mas nada que seja assim salvadora da pátria, nada disso.
E-F	Para que haja esta relação, são feitas as bases nacionais comuns curriculares que padronizam o ensino e a aprendizagem e permitem criar um “mercado” em que as escolas competem entre si por determinados benefícios que estão associados às metas atingidas. Pela natureza dos instrumentos e a quantidade de aprendizagens acumuladas, as avaliações em geral se valem da forma de múltipla escolha nos testes, o que as limita a aprendizagens mais simples. As avaliações externas medem aprendizagens mais simples e não as mais complexas – devido ao tempo que elas levam para ser respondidas e à grande quantidade de informações acumuladas durante algumas séries escolares, objeto da medida. São, portanto, limitadas em espaço e tempo. Além disso, são avaliações que fazem uma única medição, de tipo pontual, e não medem o desenvolvimento longitudinal dos estudantes, ao longo das séries. São fotografias. Sendo únicas e pontuais, não têm condições de abarcar o desenvolvimento do estudante.
E-G	Com o ensino de professores eu acho que vai depender dos sistemas. Você vai ter professores mais atentos a esses resultados e professores que realizam o seu trabalho independentemente dos resultados. Então eu acho que o fenômeno é variável, mas eu diria que tem muita pesquisa que mostra que tem professores que consideram os resultados das avaliações na hora de pensar o seu planejamento, a sua organização curricular. Em relação à aprendizagem dos alunos, essa pergunta é mais difícil de responder, porque poucas são as pesquisas que tentam avaliar o impacto da avaliação em larga escala sobre a aprendizagem dos alunos. Eu acho que há avaliações que não permitem conhecer o que os alunos sabem e são capazes de fazer, porque não têm qualidade técnica que permita afirmar. Eu estou falando de algumas avaliações, principalmente de contextos municipais, que são feitas, às vezes, de uma forma muito caseira. Mas há avaliações que têm uma ótima qualidade técnica, certamente o constructo medido é adequado.

Figura 15

Dimensão “Avaliação externa e qualidade do processo de ensino e aprendizagem”: especialistas

Ao realizar uma análise dos objetivos das diferentes avaliações externas, apresentados nos documentos oficiais, há um movimento no sentido de considerar a avaliação externa como uma forma de subsidiar diferentes níveis do sistema no que tange à tomada de decisão com vistas à melhoria da qualidade do ensino. Isto é, apresentam função política e função pedagógica. Com isso, espera-se que os resultados das avaliações sejam apropriados pelos gestores governamentais, nas secretarias educacionais, e também pela comunidade escolar, ou seja, gestores escolares, professores, alunos e pais. Dessa forma, de acordo com Souza e Oliveira (2010), “há referências de que a avaliação deve iluminar e trazer consequências para a formulação e/ou reformulação das políticas educacionais . . . assim como às propostas e práticas das escolas...” (pp. 803-804). Para além da utilização do resultado da avaliação externa com o objetivo de tomada de decisões, deve-se considerar também o conhecimento proveniente do processo e do desenho da avaliação. Ao longo de 30 anos, o conhecimento sobre recursos conceituais e metodológicos foi acumulado pelos avaliadores de programa, no entanto, pouco se utilizou na escola (Nevo, 1997).

Atualmente, vem crescendo um movimento no sentido de utilizar esse conhecimento para o benefício dos alunos, dos professores e dos gestores escolares, principalmente nas políticas das avaliações estaduais e municipais, como pode ser visto nos materiais de divulgação das avaliações, como manuais, livretos e boletins. Com isso, é necessário haver uma abordagem baseada na escola, passando de um discurso de descrição e julgamento para um diálogo.

Assim, Nevo (1997) apresenta alguns princípios de avaliação que teriam uma abordagem baseada na escola.

- 1) Os alunos e suas realizações não devem ser os únicos objetos da avaliação escolar.
- 2) Os resultados ou impactos não devem ser o único fator a ser examinado durante a avaliação de um programa, projeto, ou de qualquer outro objeto de avaliação dentro da escola.
- 3) A avaliação escolar tem de atender tanto à função formativa da avaliação quanto à somativa, proporcionando informações para o planejamento e para o aprimoramento, mas também para a seleção, certificação e responsabilização.
- 4) As necessidades internas de avaliação de uma escola podem ser mais bem atendidas através de uma equipe de professores e de outros educadores, para quem a avaliação é apenas parte da definição de suas ocupações, apoiada por treinamento adequado e assistência técnica externa.
- 5) A avaliação baseada na escola deve ser fundamentada em uma combinação de avaliações internas e externas, mas a avaliação interna é uma pré-condição para a avaliação externa que tenha utilidade (Nevo, 1997, pp. 90-91).

É importante destacar que ainda há uma dificuldade na utilização da avaliação externa, pelo uso dos resultados, do processo e desenho da avaliação, para a orientação da prática de ensino. Segundo Fontanive (2013, p. 12), “constata-se que os resultados das avaliações parecem não ter sido adequadamente incorporados à prática docente e capazes de promover a melhoria da aprendizagem dos alunos da Educação Básica, no Brasil.”. Por isso, é importante verificar como as relações entre a avaliação externa, a prática de ensino e a avaliação da aprendizagem se estabelecem.

Os especialistas participantes consideraram que a avaliação externa influencia pouco as práticas dos professores, acontecendo apenas quando há uma cobrança do sistema, resultando em movimentos que busquem educar para o teste, causando afunilamento do currículo, o que estaria relacionado a um efeito negativo. Assim, pode-se considerar que, para esses especialistas, os professores não têm envolvimento com o sistema de avaliação, por isso este não estaria causando efeitos positivos em suas práticas de ensino, como haver uma reflexão a partir dos resultados e da realização de ações provenientes dessa reflexão.

Eu gostaria muito que impactasse nas práticas dele [o professor] pensar esses resultados, [no sentido] de ele entender a escala pedagógica, relacionar com aquilo que ele sabe relacionar. Que impactasse a própria perspectiva de ele repensar as construções de questões para uma prova . . . (E-C).

No entanto, deve-se considerar a diferenciação levantada pelo especialista E-E, que considera a influência da avaliação externa não nas práticas de ensino dos professores, mas sim nas ações em sala de aula. Ou seja, a didática do professor seria pouco influenciada pela avaliação externa, a forma como ele ensina, sua relação com o aluno, mas o tipo de atividade e a ênfase nos conteúdos trabalhados na avaliação externa sofreriam maior influência (Firestone et al., 1998). Isso estaria mais relacionado a uma preparação para o teste do que a uma busca por melhorar a qualidade do ensino (Darling-Hammond, 1994).

. . . se a gente for falar das práticas em sala de aula, eu acho que elas influenciam, mas não nas práticas de ensino se a gente pensar na questão de o professor alterar sua didática, alterar sua metodologia, eu acho que tem pouco efeito . . . eu vejo uma influência muito maior na questão de dar ênfase a esses componentes curriculares, de fazer provinhas preparatórias, muito mais nesse sentido do que propriamente ali na alma da relação professor-aluno no processo de ensino e aprendizagem mesmo (E-E).

Nesta cultura de avaliação, com ênfase em classificação e responsabilização, a preocupação dos professores estaria voltada para o rendimento dos alunos; com a credibilidade da prova; com a culpabilização e consequências provenientes de resultados inadequados; com a regulação do trabalho docente a partir das avaliações externas. São preocupações, em grande parte, com relação a ações externas, providas dos órgãos governamentais, direcionadas aos professores com base nos resultados da avaliação externa. Essas preocupações podem vir a gerar consequências como preparação para o teste, fraude e segregação dos alunos.

Demonstrar que as metas foram atingidas para não serem massacrados pela mídia e/ou ganhar bônus por desempenho. Os casos de indução à fraude não são poucos. Neste processo, são obrigados a abandonar os estudantes que estão muito abaixo da média (e os que estão muito acima), colocando sua atenção naqueles que estão

mais próximos da pontuação média, na tentativa de melhorar a avaliação da turma. Com isso, a segregação aumenta (E-F).

Com relação à consistência das avaliações externa e interna, os especialistas apresentaram diferentes posicionamentos, podendo ser divididos em três grupos. O primeiro grupo considerou que não há estudos nesse campo, tornando-se, então, difícil dizer sobre a consistência das avaliações. Salienta, por conseguinte, a falta de pesquisas que tenham como objeto de estudo a consistência dos diferentes tipos de avaliação, vindo a evidenciar a necessidade de discussão nessa área.

O segundo e o terceiro grupos, de certa forma, evidenciam os mecanismos de processamento com relação à consistência das avaliações. Aquele afirmou que as avaliações externas e internas são diferentes, sendo, assim, difícil realizar uma análise. Esta construção acaba por demonstrar que os mecanismos de processamento são considerados no momento em que se fala sobre consistência da avaliação, devendo ser analisados no que tange às especificidades e objetivos de cada avaliação. Já o terceiro grupo considerou a avaliação externa mais consistente que a avaliação interna, em decorrência de mecanismos de processamentos mais claros e possíveis de serem analisados.

. . . quando ele [professor] vai dar a nota de Geografia, ele começa a colocar comportamento, caderno, participação etc. E aí a nota de Geografia quer dizer o quê? Um monte de coisas, mas eu quero saber de Geografia e às vezes tem pouco de Geografia . . . então o aluno que é mais comportado, que participa da aula, ele acaba tendo chances maiores de ter nota do que um aluno que, mesmo sabendo mais, não participa, masca chiclete, é cabeludo, essas coisas. Dar ponto por participação, por caderno, isso deveria ser proibido. Eu não estou aqui discutido a importância do comportamento, do caderno . . . a Prova Brasil eu posso discutir claramente, eu posso chegar pra Prova Brasil e perguntar: minha senhora, o que a senhora está avaliando? Olha, aqui está aqui minha matriz, estão aqui os meus itens. Pergunte a um professor o que ele avalia em História, Geografia, Matemática, eles gaguejam pra dizer, então isso faz com que as avaliações dos professores tenham uma consistência muito menor (E-C).

Essa consideração revela que os professores em suas avaliações muitas vezes levam em conta elementos que não correspondem à habilidade avaliada, julgando até

mesmo comportamento e esforço, o que diminui a consistência da avaliação interna (Alavarse et al., 2017; Bercier-Larivière & Forgette-Giroux, 1999).

A falta ou a baixa consistência da avaliação interna foi considerada também pelo terceiro grupo por conta da fraca formação de professores na área de avaliação, indicando que as universidades e os cursos de formação continuada não estariam trabalhando, de forma efetiva, a avaliação (Bercier-Larivière & Forgette-Giroux, 1999).

Eu acho que avaliação tem sido muito desprestigiada nos cursos de formação de professores, formação inicial. Ela tem pouco espaço de discussão nos cursos de pedagogia, nos cursos de licenciaturas, então isso aí tem sido ruim. . . . não tem uma discussão sobre as questões teóricas sobre avaliação, questões de consistência, dos instrumentos de avaliação, como eles são construídos, como eles são aplicados, como eles são corrigidos, como eles são analisados, como eles são devolvidos (E-C).

A relação entre a avaliação externa e as ações escolares é vista como limitada pelos especialistas, tendo em vista que consideram que não há uma utilização efetiva pelos sujeitos escolares para a melhoria do processo de ensino e aprendizagem, mas sim, na maioria dos casos, como algo imposto, podendo resultar em um efeito negativo.

Como exemplo, podemos destacar a falta de relação entre a avaliação externa e o estudo dos alunos. Todos os especialistas consideraram que essa relação não ocorre, no máximo o que poderia acontecer seria o treino com base em simulados. O fato de a avaliação externa não ser aliada do aluno em sua aprendizagem, coloca em xeque um dos critérios mais importantes da avaliação formativa, que seria o *feedback* (Sadler, 1989). O *feedback* consiste na distribuição de informação sobre como algo foi ou está sendo feito, apresenta uma dupla retroação, uma vez que se dirige aos alunos, para apoiá-los nas aprendizagens, e aos professores, como forma de auxiliar o desenvolvimento de seu programa pedagógico (Noizet & Caverni, 1985; Sadler, 1989).

Para os especialistas participantes, também não há relação da avaliação externa com a prática de avaliação interna do professor, tendo em vista que as duas avaliações apresentam mecanismos de processamento diferentes. Por isso, as avaliações externas não condicionariam à prática de avaliação interna do professor. No entanto, alguns levantaram a possibilidade de maior utilização de itens fechados pelos professores e realização de simulados, como um efeito negativo.

Em resumo, os especialistas participantes consideraram que a avaliação externa apresenta baixa ou nenhuma influência nas práticas de ensino e nas práticas avaliativas dos professores; e nos estudos dos alunos, como fica evidente na Figura 16.

Avaliação externa, prática de ensino e avaliação da aprendizagem	
E-A	Em geral, a avaliação externa pouco influencia nas práticas dos professores, mas se o sistema cobra muito, pode haver algumas distorções, como preparação para o teste e afunilamento de currículo. As maiores preocupações dos docentes estariam relacionadas com o rendimento de seus alunos e com a desconfiança em relação à prova. Sobre a consistência das avaliações, não sei dizer, pois no Brasil não há estudos sobre isso. Não acredito que as práticas de avaliação dos professores sejam influenciadas pela avaliação externa, pois são processos separados que conversam pouco entre si. A avaliação externa pouco condiciona o estudo dos alunos, só se houver muita cobrança, aí terá um movimento de educar para o teste.
E-B	A influência que a avaliação externa tem na prática dos professores resume-se à preparação para o teste. Os professores estariam preocupados, em um contexto de avaliação externa, com obter o melhor resultado, que eles não sejam culpabilizados, que eles não sejam responsabilizados por um resultado inadequado, um resultado ruim. No geral, é difícil dizer sobre a consistência das avaliações. Tem avaliações boas e ruins nas duas avaliações. Um bom trabalho deveria fazer com que os resultados das avaliações externas iluminassem também o resultado das avaliações internas, havendo possibilidade de um diálogo entre elas. A avaliação externa influencia diretamente nas práticas avaliativas dos professores, na medida em que repetem o modelo das avaliações externas. Para o estudo dos alunos, a avaliação externa não faz diferença.
E-C	A avaliação externa influencia na prática de ensino dos professores porque tem impactado na perspectiva de que professor tem que ensinar aquilo correndo, tem que treinar o aluno. Sobre a consistência das avaliações, são naturezas completamente diferentes. A avaliação externa tem uma abordagem metodológica completamente diferente, não analisa a prova como um todo, mas o item. Do ponto de vista estatístico, metodológico e do processo de construção da prova, eu acho que ela tem consistência interna, pois é aplicada, analisada, testada, então tem toda uma configuração de medições que são feitas, que vão indicar se são bons itens ou maus itens. Sobre a avaliação interna que o professor faz, eu acho que esse tema tem sido muito desprestigiado nos cursos de formação de professores, então ela tem pouco espaço de discussão. Não acredito que as práticas de avaliação dos professores sejam influenciadas pela avaliação externa. Para o estudo dos alunos, só se for em um movimento para treinar para os testes.
E-D	A influência da avaliação externa nas práticas dos professores está muito mais para efeito nulo do que para efeito positivo ou negativo. Mas talvez mais professores tenham passado a utilizar os itens de múltipla escolha. Mas aí vem o efeito colateral, os itens que os professores produzem são muito ruins, carecem de uma série de propriedades. Os professores se sentem desautorizados, eles sentem certa pressão. Claro que tem rede que é mais e tem rede que é menos, mas eles sentem uma certa pressão com relação aos resultados, pelo menos na época da divulgação, depois com o tempo isso passa. Acredito que isso seja uma preocupação para eles. Na consistência das avaliações, entra o tal do diálogo. Não é para desautorizar o professor, mas sim fazer ele repensar o que ele está fazendo. A consistência das avaliações internas é muito menor do que a das avaliações externas. Então isso tem permitido a gente fazer um trabalho comparando as notas que os professores deram para cada aluno e o que foi apurado na avaliação externa. Por incrível que pareça, em geral, há concordância. A avaliação externa não influencia as práticas de avaliação dos professores, quando muito eles começam a usar um pouquinho mais de itens de múltipla escolha. Então o condicionamento do estudo dos alunos também é muito baixo, porque essas avaliações externas têm um problema, que algumas superam, mas a maioria não. Entre o momento que o aluno faz a prova e quando chega o resultado na escola, o aluno nem está mais na escola.
E-E	Eu acho que influencia pouco na prática de ensino, mas se a gente for falar das práticas em sala de aula, eu acho que elas influenciam. Mas não nas práticas de ensino se a gente pensar na questão de o professor alterar sua didática, alterar sua metodologia, aí eu acho que tem pouco efeito. Eu vejo a influência muito maior na questão de fazer provinhas preparatórias, muito mais nesse sentido do que propriamente ali na alma da relação professor/aluno, no processo de ensino e aprendizagem mesmo. Acho que a grande preocupação dos professores é com relação ao resultado. Eu acho que as avaliações externas, de um modo geral, são mais consistentes, porque elas são testadas, existem critérios e tudo mais. A avaliação interna é menos consistente, mas ela é mais significativa para o professor, porque foi ele quem fez. Eu acho que a avaliação externa condiciona as práticas avaliativas dos professores na medida em que é cada vez mais comum o professor fazer simulados como forma de preparação para avaliação externa. De forma alguma a avaliação externa condiciona o estudo dos alunos. A percepção que eu tenho é que influencia em quase nada, a não ser em função de preparação para a prova, mas isso não é estudo.
E-F	A avaliação externa retira a autonomia dos professores para adequar o que ensinam às características específicas dos seus estudantes, pois a padronização imposta pelas bases nacionais curriculares associadas às avaliações externas definem o que ensinar. Neste sentido, impedem o desenvolvimento do pensamento crítico e criativo de alunos e professores. A preocupação dos professores é demonstrar que as metas foram atingidas para não serem massacrados pela mídia e/ou ganhar bônus por desempenho. Se as avaliações internas forem feitas como forma de antecipar as externas, então elas tendem a ser consistentes, especialmente se preparadas na forma de simulados que treinam para as avaliações externas. Caso sejam avaliações conduzidas autonomamente pelos professores, elas têm objetivos diferentes, pois são de caráter formativo. Neste caso, podem não ser consistentes com as externas, sendo, no entanto, mais consistentes com os objetivos de aprimoramento do ensino por parte do professor. Há uma influência da avaliação externa nas práticas de avaliação dos professores. Quando as avaliações internas são feitas como preparação para as externas, elas têm formatos e objetivos idênticos. Os alunos estudam em função do histórico dos exames anteriormente realizados. "Aprendem" como se comportar frente ao formato e tipo de avaliações.
E-G	Eu acho que haverá professores que serão influenciados pelas avaliações externas e também outros que não vão modificar sua prática. Eu acho que os professores se preocupam em relação ao bônus, essa é uma preocupação. A depender da idade da tradição da avaliação, eles se preocupam com possíveis consequências dos resultados, com a sua própria carreira. Eu já fiz pesquisas em que o professor falou que começou a utilizar mais item fechado, porque só usava questão aberta e via que os alunos tinham dificuldades, começaram a diversificar mais o instrumento de avaliação. Mas isso acontece com todo mundo? Eu diria que não. Algumas escolas são mais inspiradas pelas avaliações externas, outras menos. Sobre condicionar o estudo dos alunos, o que eu posso dizer é que tem escola que faz simulado, ensina aluno a estudar, mas isso vai depender do sistema. Às vezes tem até uma orientação do sistema de ensino para que as escolas e os professores tenham alguma ação no sentido de estimular os alunos a se prepararem para a avaliação.

Figura 16

Dimensão “Avaliação externa, prática de ensino e avaliação da aprendizagem”: especialistas

Com isso, fica claro que os resultados das avaliações, de alguma forma, não estão sendo apropriados pela comunidade escolar, ou seja, gestores escolares, professores, alunos e pais, tendo em vista que não apresentam impacto nas propostas e práticas das escolas. No entanto, apesar de o resultado de uma avaliação externa, no desenho existente no Brasil, ser geralmente o foco do debate, muito se fala de resultados, em detrimento da fundamentação da avaliação, do seu desenho, da sua aplicação e de suas limitações. E mesmo tendo toda visibilidade, na maioria das vezes, ao implementar um programa de avaliação,

. . . há uma preocupação maior em organizar diferentes equipes para fins diversos: administrar, elaborar manuais, construir instrumentos . . . contudo, quase sempre se omite a equipe responsável pela disseminação dos resultados . . . criando-se, assim, um vácuo nas comunicações, talvez o responsável maior pela ausência de um efetivo transformador (Vianna, 2005, p. 19).

No entanto, é sabido, que em qualquer tipo de avaliação, a redução de seu propósito à sistematização de informação quantitativa, como apenas a apresentação de notas, dados, gráficos etc., ou a simples apresentação de informação não representa a totalidade de uma avaliação. Existe a necessidade de haver não apenas um conhecimento dos resultados, mas uma análise dos dados obtidos, uma reflexão sobre eles, um diálogo com outras fontes de avaliação, uma produção de juízo de valor sobre esses dados. Ter os dados não garante uma apropriação deles, não garante a utilização no sentido de direcionar ações, sejam políticas, sejam pedagógicas. Dessa forma, falar sobre resultados vai muito além de apresentar parâmetros, índices e notas, é discorrer sobre divulgação, apropriação e análise também.

De uma forma geral, para os especialistas, a apropriação da avaliação externa pelos diferentes intervenientes e interessados no processo é baixa. Foi levantado que a apropriação está diretamente relacionada à política do estado e ações da escola com relação à avaliação. Ou seja, se há uma política ou ação que envolva os sujeitos escolares no processo, a chance de haver uma apropriação dos usuários é maior.

Com relação aos alunos, ponderaram que não há apropriação dos resultados, visto que não faz diferença para eles os resultados da avaliação. Para a família, entende-se que há um poder limitado dessa apropriação, porque a pressão da família por uma escola pública de qualidade é algo ainda muito no campo do ideal, tendo em vista que estamos falando de um público de baixo poder aquisitivo, com pouco impacto na reivindicação.

. . . por mais que você diga que os resultados dos testes, na hora que eles são divulgados, ajudariam na responsabilização dos professores e numa pressão das famílias para melhoria da escola, existem limites. A gente sabe que a escola pública brasileira é frequentada pelos pobres, existe limite para os pobres poderem tirar os seus filhos de uma escola que tem um desempenho ruim e colocá-los em uma outra que tem um desempenho bom. Por causa de distância, por causa de tempo... uma série de fatores. Então, essa mobilidade que se imaginaria que fosse possível, que é pregada... então se imaginaria que a família tira o aluno dessa escola, aí a escola começa a perder aluno, aí ela vai melhorar, para recuperar os alunos, certo? Bobagem, não existe (E-B).

Sobre o corpo escolar, mais especificamente os professores e os gestores escolares, considera-se que a apropriação venha por meio da pressão por parte dos resultados. Foi considerado um grupo com uma apropriação mais por meio de pressão do que por análise crítica, por envolvimento com o processo de avaliação externa.

Especificamente por parte dos professores, foi levantada a questão de troca de escola com base nos resultados, como um determinante para a busca de uma escola melhor, devido à dificuldade em ensinar e à culpabilização do professor.

Isso de alguma forma tem uma pressão sobre aquele corpo escolar, as pessoas que frequentam aquela escola, forte. Apesar de eu não ter lido nenhum estudo, mas imagina-se que existe um efeito grande (E-B).

São apropriados segundo a lógica da política educacional que cerca a avaliação. Em geral são políticas de *accountability* verticalizada. Neste caso, interessa que as metas sejam atingidas (usualmente metas numéricas) (E-F).

Então existe uma busca dos professores para sair dessa escola problema, porque é difícil ensinar nessa escola. “Os testes sempre mostram que o desempenho da escola está ruim, então vou buscar uma outra escola porque essa aqui eu não aguento mais. Dá muito trabalho ensinar para esse povo. Não consigo e todo

mundo me culpa. Eu sei que a culpa não é só minha.” Então ele corre atrás de uma outra escola (E-B).

Por parte dos gestores do sistema, foram levantados três movimentos: uma apropriação com base na realização de uma pressão verticalizada; com base numa gestão educacional com transferência de responsabilidade (Freitas, 2007; Moe 2003), os resultados sendo analisados pelos gestores para a realização de pressão nas escolas por melhorias da qualidade da educação; uma apropriação ingênua da avaliação e seus resultados, visto que a realização da avaliação estaria fundamentada na crença de que esta poderia melhorar a qualidade da avaliação por si só; e uma apropriação com buscas na realização de políticas públicas.

Os gestores, via de regra, aumentam a pressão em cima das escolas que têm mau resultado, e louvam aquelas escolas de bom resultado (E-B).

. . . os gestores, eu acho que têm um aspecto que tem que ser diferenciado, acho que tem gestores que levam a sério, que se dedicam a isso, que tentam qualificar, tentam desencadear políticas de formação, aquelas ações na escola. E a questão dos políticos, eu acho que é essa crença meio ingênua, que qualquer coisa vai melhorar pela avaliação . . . então vamos fazer mais uma avaliação (E-E).

Como sintetizado na Figura 17, fica claro que os especialistas consideram que os diferentes intervenientes e interessados apresentam uma baixa apropriação da avaliação externa da aprendizagem dos alunos.

Resultado da avaliação externa	
E-A	A apropriação dos resultados da avaliação externa por todos é pequena. No atacado, os professores não conhecem os resultados das avaliações externas. Acredito que os docentes não analisam os resultados. Os gestores escolares realizam a análise dos resultados somente se o sistema realizar alguma pressão.
E-B	Com relação à apropriação, para os alunos, os resultados não fazem muita diferença; família não tem poder com esta informação, efeito limitado; professores se apropriariam no sentido de procurar melhores escolas para lecionar; gestores aumentam a pressão ou louvam suas escolas devido aos resultados. Os professores sabem pouca coisa sobre os resultados da avaliação externa, sabem se o Ideb aumentou ou baixou. Eles não analisam os resultados, sabem apenas o resultado final. Os gestores escolares realizam a análise de resultados a depender da pressão, analisam no sentido de cobrança por resultados.
E-C	A apropriação dos resultados no caso de decisores da política, no nível da gestão, tem tido impacto dessas bonificações, premiações. Com relação ao aluno, eu acho que é tão pouco. Os professores conhecem só o que é divulgado, ou seja, as notas. Eu acho que os docentes não analisam os resultados das avaliações externas.
E-D	A apropriação, em geral, é muito baixa, alunos nem contam praticamente. Os pais dependem de alguma escola que faça alguma reunião, que divulgue isso. A apropriação não quer dizer aceitação, não é isso, apropriação tem a ver com a própria crítica que se faz, mas nem isso às vezes eles fazem. Eu acho que os professores conhecem muito pouco dos resultados das avaliações externas. Análise do resultado pelos professores, nem pensar. Eles sabem que o Ideb é 4, isso, aquilo e tal... mas decompor, analisar, estritamente falando, não. Com relação à análise de resultados pelos gestores das escolas, não há de forma generalizada. Poderia ser interessante ver a série histórica, tentar recompor o que teria acontecido, para explicar por que subiu ou por que desceu.
E-E	Por parte dos alunos, eu acho que a apropriação é quase nada. Os professores apresentam uma relação como esse elemento externo de controle, que tem que apresentar resultado. Os gestores têm um aspecto que deve ser diferenciado. Acho que tem gestores que levam a sério, que se dedicam a isso, que tentam qualificar, tentam desencadear políticas de formação, aquelas ações na escola. E a questão dos políticos é essa crença meio ingênua de que qualquer coisa vai melhorar pela avaliação. Eu acho que de um modo geral os professores não conhecem os resultados das avaliações externas, mas também varia muito de acordo com os gestores. Eu acho que os gestores das escolas fazem uma análise, em primeiro lugar, por uma questão de <i>marketing</i> , mas eu aí vejo muito no sentido de que já analisam os resultados visando resultado.
E-F	São apropriados segundo a lógica da política educacional que cerca a avaliação. Em geral, são políticas de <i>accountability</i> verticalizada. Neste caso, interessa que as metas sejam atingidas (usualmente metas numéricas). Os professores são instados a conhecer os resultados quantitativos pelos gestores, como forma de alcançar as novas metas. Não penso que os professores analisem os resultados. Em geral os gestores recebem os relatórios prontos das Secretarias de Educação, fornecidos por terceirizadas que desenvolvem o processo de avaliação das redes. O entendimento dos gestores foca na "responsabilização" do professor e não nas condições de trabalho ou nas condições de vida dos estudantes.
E-G	Eu não sei se os alunos chegam a se apropriar desses resultados, porque vai depender do sistema. Tem sistema que vai boletim para casa, mas eu nunca fiz pesquisa sobre isso. Também acho que conhecer o resultado da avaliação externa depende de quem está na gestão da escola. Eu acho que tem escolas que vão fazer um trabalho de divulgação desses resultados, tem sistemas de ensino que induzem que a escola faça esse trabalho de divulgação dos resultados, vira tema de hora de trabalho pedagógico, de discussão. Mas também vai ter escola que não faz isso. Se a escola faz a análise de resultados, me parece mais possível que o docente faça análise também, eu não diria que é um fenômeno generalizado do docente por si ir atrás desses resultados. Eu acho que os gestores escolares em geral fazem essa análise de resultados. Nas minhas pesquisas, a gente percebe uma análise dos gestores. Eu acho que vai depender um pouco do perfil do gestor, da qualidade desse sistema, de que sistema você está falando.

Figura 17

Dimensão "Resultado da avaliação externa": especialistas

Essas foram as percepções apresentadas pelos especialistas entrevistados sobre a avaliação da aprendizagem e sua relação com a qualidade do ensino, estando dividida nas dimensões desses objetos de pesquisa. Na parte a seguir, será discutida a percepção dos decisores políticos.

Narrativa dos Decisores Políticos

Nesta seção, a avaliação externa da aprendizagem dos alunos será vista a partir das narrativas dos decisores políticos participantes da pesquisa. Como decisores políticos, são considerados aqueles que estão em cargos ou já estiveram, em órgãos nacionais e subnacionais, que tomam decisões e produzem materiais de cunho legislativo para o setor educacional. Além desses, também as pessoas ligadas à coordenação de agências de avaliação, que são contratadas, na maioria dos casos, para gerenciar os programas de avaliações subnacionais.

Avaliação externa da aprendizagem.

Com base nas respostas dos decisores políticos participantes desta pesquisa, relativamente à natureza e aos propósitos da avaliação externa, o que se constata é a presença da natureza somativa dessa avaliação, tendo em vista que evidenciam como propósitos verificar a qualidade da educação e obter informações sobre o sistema educacional, para monitorar a qualidade da educação. Há também uma natureza formativa, já que entendem que a avaliação externa apresenta como propósitos subsidiar políticas públicas e contribuir com a qualidade da educação, com a promoção da equidade e com a intervenção pedagógica.

De acordo com Black e Wiliam (2001), na avaliação formativa “a evidência é na realidade usada para adaptar o trabalho de ensino para atender às necessidades dos alunos.” (p. 2, tradução nossa), e sua prática produz ganhos significativos na aprendizagem, destacando as seguintes características: *feedback*, envolvimento dos alunos, resultados da avaliação sendo usados para ajustar o ensino e autoavaliação.

As avaliações externas também podem apresentar característica de avaliação formativa, segundo Wiliam (2006), desde que utilizadas de forma a identificar o que precisa ser aprendido e ajustar estratégias de ensino. Assim, o que torna uma avaliação formativa não seria o tempo de retorno do *feedback*, onde é realizada ou por quem é realizada, mas sim, que os dados são coletados, interpretados segundo as necessidades de aprendizagem e usados para realizar ajustes no que tange a melhor atender a essas necessidades.

A natureza formativa da avaliação externa pode ser vista na fala de DP-A, que considera que essa avaliação pode contribuir com o processo de intervenção pedagógica,

com o propósito de ajudar o professor. Há a ideia de que o resultado da avaliação por aluno é mais rico, vindo a possibilitar um auxílio no processo de intervenção pedagógica para reverter dificuldades encontradas pelos alunos.

Eu acho que inclusive a avaliação externa tem ajudado o professor a organizar o seu pensamento, porque ele já é dono dessa realidade, até mais do que os dados da avaliação externa, ele já tem o diagnóstico da sala de aula, mas a avaliação vem a contribuir de uma forma mais pontual, mostrando o conjunto de habilidades que o menino precisa vencer e não está conseguindo e os que já venceram, os que estão em condições avançadas . . . Esse resultado que vem por regional, que vem por escola, que vem por turma, por turno e muitos programas estaduais chegam ao aluno, uma visão mais rica, uma visão mais da escola, de um processo de intervenção pedagógica para reverter o quadro de precárias condições (DP-A, grifos nossos).

No entanto, há um equívoco no sentido de considerar que a avaliação formativa deve possibilitar um *feedback* imediato ao professor, para que assim ele possa intervir, se necessário, no processo de ensino e aprendizagem. Dessa forma, a avaliação externa seria apenas de natureza somativa, pois não conseguiria atender a esse requisito. Como se pode perceber na fala de DP-F, “. . . a característica é somativa, ela não dá um *feedback* imediato para o professor sobre cada aluno, mas permite entender uma lógica sistêmica”.

No entanto, como afirma Wiliam (2006), a avaliação externa pode ser de natureza formativa, desde que utilizada de forma a identificar o que precisa ser aprendido e ajustar estratégias de ensino, ou seja, se a avaliação externa tiver como propósito contribuir para a melhoria da qualidade da educação, subsidiar políticas públicas etc., ela pode apresentar natureza formativa.

O propósito de monitorar a qualidade da educação oferecida por um sistema educacional vem ao encontro do terceiro momento que a educação brasileira está vivenciando, que é a utilização dos índices das avaliações externas como forma de medir a qualidade da educação. Considerando esse momento, a fala de DP-A corrobora com a ideia de algumas correntes que consideram que a avaliação externa é o único meio de atestar a qualidade da educação, que ela teria a capacidade de revelar a “verdade” (Alavarse, 2013), quando afirma:

Hoje nós estamos num cenário que se discute muito qualidade de educação, mas o que é qualidade da educação? Eu só vou ter uma entrada e só vou atestar a

qualidade da educação ofertada, de um modo muito especial da educação pública, por meio de uma avaliação externa (grifos nossos).

Outro ponto levantado pelos decisores políticos é o de a avaliação externa ser um instrumento para subsidiar políticas públicas. Tomando como base a dimensão da utilização, que será apresentada mais adiante, pode-se afirmar que todos eles consideram que as avaliações externas são utilizadas para formular políticas públicas de educação. Como exemplos, são referidas as políticas públicas voltadas para a promoção da equidade, para a melhoria dos programas de formação de professores e para a melhoria dos currículos escolares.

Os decisores políticos participantes citaram alguns mecanismos de processamento das avaliações externas, o que possibilita conhecer, de forma mais específica, a natureza da avaliação externa, a saber: sua adequação à população em termos do seu grau de dificuldade (complexidade compatível com a idade e séries avaliadas), a sua congruência com o currículo e os seus domínios, a possibilidade de comparabilidade dos resultados, a fiabilidade ou consistência dos resultados, a análise dos resultados a partir de uma escala de proficiência e a padronização dos instrumentos. Há assim uma preocupação com os aspectos técnicos da avaliação.

Apesar de haver pouca discussão da literatura brasileira sobre os aspectos técnicos empregados nas avaliações externas, cabe ressaltar que vem se pensando sobre eles ao longo dos anos, com melhora significativa do desenho técnico dos instrumentos no Brasil, principalmente no que diz respeito à teoria da medida, o que acaba por impactar na confiabilidade dos resultados (Brooke & Cunha, 2011; Sousa & Oliveira, 2010). Tal fato pode ser considerado como um conhecimento provindo da participação do Brasil em avaliações internacionais, visto que um dos objetivos de participar dessas avaliações é a construção de maior *know-how* e o fortalecimento da capacidade técnica do sistema de avaliação nacional.

As perspectivas defendidas pelos decisores políticos participantes na pesquisa estão sintetizadas na Figura 18. A análise da Figura permite destacar que para esse grupo de participantes, as avaliações externas são de natureza somativa e formativa. Os propósitos apresentados confirmam a presença das duas naturezas, visto que para os decisores políticos a avaliação externa objetiva verificar a qualidade da educação e obter informações sobre o sistema educacional, para monitorar a qualidade da educação; e

subsidiar políticas públicas e contribuir com a qualidade da educação, com a promoção da equidade e com a intervenção pedagógica.

Natureza e propósitos	
DP-A	A avaliação externa tem como objetivos subsidiar políticas públicas, contribuindo para a melhoria da qualidade da educação e com a promoção da equidade; monitorar a qualidade da educação pública do país; e contribuir no processo de intervenção pedagógica.
DP-B	A principal característica seria a função diagnóstica, no sentido de verificar a qualidade da educação que está sendo oferecida. Mas há também a função formativa, tendo em vista que seus resultados podem ajudar o professor a criar processos educacionais para melhorar a aprendizagem das crianças. Assim, as avaliações devem ter uma complexidade compatível com a idade e devem ser suficientemente abrangentes para cobrir todo tipo de currículo praticado nas escolas.
DP-C	A principal característica da avaliação externa e, talvez, a mais importante, é avaliar a rede ou o sistema de ensino para além da sala de aula. Com base nos resultados obtidos, é possível planejar ou replanejar políticas públicas mais efetivas para as realidades encontradas, sendo necessária uma metodologia e instrumentos específicos, tais como: comparabilidade e confiabilidade dos resultados; padronização de construção dos instrumentos; escala de proficiência.
DP-D	A avaliação externa é um instrumento para obtenção de informações sobre aprendizagens específicas dos alunos.
DP-E	Oferecer subsídios para refletir sobre e reformular as políticas educacionais de um projeto de ensino e aprendizagem, sendo, por conseguinte, seu principal objetivo subsidiar políticas de melhoria da aprendizagem. Não é para fazer <i>ranking</i> , não é para avaliar aluno individualmente, porque a avaliação externa, em geral, não serve para isso.
DP-F	O que mais bem define a avaliação externa é a sua característica de avaliação somativa, que permite ver o que se desenvolveu de competências ou o que se obteve de aprendizagem ao longo de um certo período, o que permite avaliar se aquele sistema tem um desempenho bom. Ela não dá um <i>feedback</i> imediato para o professor sobre cada aluno, mas permite entender uma lógica sistêmica.

Figura 18

Dimensão “Natureza e propósitos”: decisores políticos

Há um entendimento de que a avaliação externa da aprendizagem apresenta tanto uma natureza somativa quanto formativa, porém, como destacado na fala dos decisores políticos, é importante entender o papel dessa avaliação na função formativa, pois como visto na Figura 18, há alguns equívocos. Para que possa haver relevância em sua função pedagógica, a função formativa da avaliação externa deve ser melhor esclarecida.

Os decisores políticos consideram que a avaliação externa deve apresentar um diagnóstico claro e comunicação do resultado de forma eficaz (DP-A e DP-B), contudo destacam que os mecanismos de processamento da avaliação interferem na relevância da avaliação externa para o processo de ensino e aprendizagem. Dessa forma, também consideram que a responsabilidade seria das entidades responsáveis pelo processo de avaliação.

De acordo com Fontanive (2013), em reflexão sobre a divulgação dos resultados das avaliações externas, mesmo havendo tentativas de simplificar informações

complexas, tanto em matéria de divulgação (manuais, boletins, folhetos etc.) quanto na realização de eventos (capacitação de professores, seminários, videoconferências etc.), parece que tais ações não têm sido bem-sucedidas, em virtude de não provocarem impactos esperados na qualidade e na equidade da educação.

Tem que ter várias estratégias de comunicação de resultados. A gente vê que talvez esse seja um ponto que as avaliações externas, pelo menos as brasileiras, tiveram pouca competência em fazer. A gente não vê o impacto disso muito claramente nas escolas, como se os professores ao verem aqueles resultados, não se sentissem representados (DP-B).

No entanto, mesmo entendendo que há uma necessidade de diagnóstico claro e comunicação do resultado de forma eficaz, há uma compreensão de que a relação entre aprendizagem e medida de aprendizagem não é algo fácil. O que deixa clara a existência de dois processos, um de medida da aprendizagem e outro de utilização dessa medida para o processo de ensino e aprendizagem. Assim, a avaliação externa já saberia fazer a medição, porém ainda não conseguiu transformá-la num instrumento de auxílio pedagógico.

. . . se eles não sabem fração no 5º ano, na próxima avaliação os alunos deveriam dominar fração, Foi dito que eles não sabem fração, mas a gente não vê essa relação. Até porque, às vezes, é difícil mesmo, a relação entre a aprendizagem e a medida da aprendizagem não é tão linear (DP-B).

A perspectiva dos decisores políticos com relação ao processo de ensino e aprendizagem está voltada para o sistema, isto é, a avaliação externa contribui para os gestores pensarem nesse processo e realizar ações que possam influenciá-lo. Algumas ações citadas foram: formação de professores, investimento em boas práticas pedagógicas, melhora de material didático e pedagógico. Dessa forma, corrobora com a perspectiva de Nevo (1997, p. 90), que afirma que

Os avaliadores em educação vêm-se mostrando mais interessados no trabalho com os administradores e no atendimento das necessidades de informação dos responsáveis pelas políticas, do que na ajuda aos professores a aprimorar os métodos de avaliação que vêm sendo usados há décadas, ou no fornecimento a pais e a alunos do tipo de informação de que eles precisam, como “clientes da avaliação”.

Considerando tal afirmativa, pode-se concluir que, ao destinarem seus esforços para gerar informações que propiciem planejamento, avaliação e ações do sistema educacional para os gestores políticos, os avaliadores não consigam trabalhar os resultados para os professores. Não que sejam movimentos excludentes, mas apresentam processos de comunicação de resultados distintos.

No que tange ao investimento em boas práticas pedagógicas, o participante DP-F considera que a avaliação propicia identificar boas práticas, podendo, por conseguinte, haver investimentos mais efetivos, evitando gastos financeiros com “abordagens que não impactam na aprendizagem” (DP-F). A identificação de boas práticas seria realizada a partir de um contato com a escola em que os alunos tiveram um bom desempenho na avaliação externa, para que, assim, a equipe gestora governamental pudesse identificar as práticas realizadas e multiplicá-las na rede educacional.

Com base na Figura 19, pode-se evidenciar que os decisores políticos consideram que pode haver relevância da avaliação externa para o processo de ensino e aprendizagem, no entanto os mecanismos de processamento da avaliação interferem no impacto dessa relevância. Para este grupo, a maior ênfase nessa relação estaria no sentido de a avaliação proporcionar aos gestores uma reflexão desse processo e realizar ações que possam influenciá-lo.

Relevância e importância	
DP-A	A avaliação externa pode ser relevante para o processo de ensino e aprendizagem se houver um diagnóstico rico e apurado, e as escolas conseguirem aprender com este diagnóstico, pensando algumas medidas como reelaboração do Projeto Político-Pedagógico da escola.
DP-B	Uma avaliação externa tem que saber comunicar os resultados. Porém, no Brasil, talvez não se tenha conseguido realizá-la de maneira competente até então. Não se vê o impacto disso muito claramente nas escolas, como se os professores, ao verem aqueles resultados, não se sentissem representados. Os resultados poderiam contribuir com a formação de professores, para que o professor repense sua prática. Agora o que a gente percebe é que ela não tem tido poder de informar para mudar a prática pedagógica. Essa avaliação externa tem que ser comunicada de uma maneira que o professor veja a sua prática, o que ele ensina naqueles resultados.
DP-C	Os resultados podem orientar decisões voltadas à melhoria no sistema de ensino e nas escolas. Permitem a percepção do que é esperado e o que ainda precisa ser desenvolvido em relação à aprendizagem dos alunos destas unidades.
DP-D	Os resultados são relevantes na medida em que oferecem informações para os gestores planejarem ações de formação ou de investimento.
DP-E	São relevantes quando se utilizam os resultados das avaliações para melhorar a formação de professores, melhorar o currículo, melhorar os materiais didáticos e pedagógicos, e identificar pontos que realmente a escola não está conseguindo melhorar.
DP-F	A avaliação externa é relevante na medida em que evidencia boas práticas, mostra abordagens que não impactam na aprendizagem, o que é particularmente importante, quando a gente desperdiça recursos educacionais e emprega recursos na adoção de práticas que não trazem efetividade.

Figura 19

Dimensão “Relevância e importância”: decisores políticos

Assim, observa-se que há uma maior relevância para ações desenvolvidas pelos decisores políticos, ações que venham a interferir no processo de ensino e aprendizagem, utilizando os resultados da avaliação externa para realizar formação de professores, investir em boas práticas pedagógicas, melhorar de material didático e pedagógico. Com isso, a maior relevância estaria para esse grupo, até mesmo porque é o grupo que considera que a credibilidade da avaliação externa é alta.

Eu acho que hoje [a credibilidade] é total, até porque as avaliações externas hoje, com o Ideb, é dinheiro na escola. Então, é essa relação do financeiro, do impacto social dos resultados, do impacto da minha política que rende votos, que rende projeção . . . Aquilo que está lá é a única medida comparativa que eu tenho, não pra me situar, mas pra eu estar entre os melhores, pra eu ter um ideal de melhoria . . . (DP-B).

Diante de tal situação, pode-se concluir que a credibilidade da avaliação externa para os decisores políticos está na aceitação da sociedade, pois é a partir do impacto social causado pelos resultados que se consegue voto e implementação de políticas públicas. Torna-se evidente que, ao desconsiderar os mecanismos de processamento para falar da

credibilidade da avaliação externa, se está diante de uma crença absoluta nesse tipo de avaliação, entendendo que não há erros ou esses são mínimos. Não há, por conseguinte, uma discussão sobre suas limitações (Almeida et al., 2013; Bauer, 2012; Gatti, 2012, 2014).

Com relação ao público em geral, os decisores participantes acreditam que tende a acreditar mais na avaliação externa por não ter informações suficientes ou clareza do processo. Já os professores apresentam baixa confiança na avaliação externa, muito em razão da política pública que a embasa, sendo a principal a bonificação por resultado. Como salienta DP-A

Tem algo que eu acho que compromete muito a credibilidade da avaliação, é quando o Estado implementa um sistema de avaliação junto com uma proposta de meritocracia, de bônus. A avaliação chega mal, porque já vem com um objetivo de responsabilização muito forte . . . Acho que essa questão da responsabilização precisa ser repensada, só pode ser usado o sistema de responsabilização depois que se tenha credibilidade, que as pessoas usem os dados, que acreditem na avaliação, que se envolvam com aquela avaliação, que ela já faça parte do cotidiano da escola, que eles já compreendam os dados, já usem os dados na sua interpretação pedagógica, ou seja, [depois que] eles acreditem na eficácia dos resultados daquela avaliação. Então, quando isso acontece, é quando você tem o uso positivo da avaliação. . . . ela pode acontecer sim, a meritocracia pode acontecer, mas posterior a essa fase. Eu acho que quando acontece tudo junto, é um risco muito grande (DP-A).

Isso indica que a credibilidade também está ligada às políticas públicas que embasam a avaliação externa, tendo em vista que, como afirma DP-A, haveria uma baixa credibilidade entre os professores por causa do emprego de responsabilização forte (*high-stakes testing*) com base em resultados da avaliação externa. É pertinente dizer que, além de o uso da responsabilização forte poder originar efeitos não desejáveis, há estudos que revelam que os incentivos ou ameaças pouco ou nenhum impacto têm no desempenho dos alunos ao longo dos anos. Por isso, tais testes são fortemente criticados por diversos setores sociais e da educação, como também por prestigiadas associações científicas (Afonso, 2009; Nichols et al., 2012;).

No entanto, a fala de DP-A está em consonância com pesquisa apresentada por Ahn e Vigdor (2010). Os autores concluíram que os estados que proporcionavam

recompensas monetárias aos professores para o crescimento acadêmico escolar faziam com que os professores se esforçassem mais. O maior esforço do professor era traduzido em melhora do desempenho dos alunos. No entanto, os autores afirmam que, devido à necessidade de confiabilidade estatística das pontuações dos testes, as autoridades educacionais podem ter que esperar anos para poder recompensar professores.

Na *Revista do sistema de avaliação* da rede estadual do Rio de Janeiro – Saerj 2012, o secretário de educação Wilson Risolia evidencia a utilização do resultado da avaliação externa como forma de premiação de professores e alunos.

. . . a avaliação também serve como base para premiar os profissionais que vêm trabalhando em conjunto nas escolas, melhorar o desempenho dos alunos e atingir suas metas. Ela também continua premiando com computadores portáteis os alunos que atingem melhores resultados nas provas, assim como com viagens à cidade do Rio de Janeiro. Vagas em projetos como o Pronatec também são obtidas com a participação no Saerj. Essas iniciativas são o merecido reconhecimento da dedicação de alunos e professores.¹²

Para além de prêmios atrelados aos resultados do Saerj, outra ação promovida pelo governo foi a criação de bonificação por resultados, um programa que tinha como objetivo remunerar os servidores que atingissem ou superassem as metas estabelecidas. Tal ação visava ao aumento do grau de participação dos servidores com o processo de ensino e aprendizagem.

Entretanto, o sistema de bonificação não era algo novo no Rio de Janeiro, já que o estado já havia se utilizado desse recurso no Programa Nova Escola (PNE), configurando-se uma política de alto impacto desde o início de seu sistema de avaliação, tendo sua primeira aplicação em 2004. Desde então, veio recebendo muitas críticas por parte dos sindicatos e resistência dos professores.

Com isso, fica claro que não houve um tempo para que a avaliação externa se consolidasse no estado do Rio de Janeiro, para que assim pudesse construir um terreno sólido, buscando incorporar um sistema confiável não apenas tecnicamente, mas também na relação entre os atores escolares e os avaliadores externos. Isso pode ter influenciado de forma significativa a aceitação da relevância do instrumento no espaço escolar.

¹² Disponível em: http://www.avaliacaoexternasaerj.caedufjf.net/wp-content/uploads/2013/07/SAERJ_SISTEMA_2012.pdf. Acesso em: 28 jul. 2019.

No que diz respeito à qualidade do instrumento, os decisores políticos, em sua maioria, ou acreditam que as avaliações têm qualidade ou afirmam que elas melhoraram muito ao longo do tempo. A qualidade para esse grupo está vinculada aos mecanismos de processamento adotados, tais como: matrizes de referências, TRI, pré-teste, índice de dificuldade e de discriminação.

Apenas um entrevistado (DP-B) considerou uma limitação das avaliações, que seria o fato de basicamente todas serem de múltipla escolha, tendo em vista que o aluno não produz a resposta, mas reconhece a correta. Isso corrobora com a ideia de baixa discussão das limitações da avaliação externa entre os decisores políticos.

Cabe salientar que a falta de reflexão sobre as limitações da avaliação externa reduz sua utilização, como afirma Bauer (2012, pp. 74-75),

... considera-se que a falta de debate e reflexão sobre a validade desses resultados, por si só, seria uma limitação para o seu uso que deveria ser considerada no momento da elaboração de ações a serem realizadas a partir deles. Tais questões não parecem estar sendo consideradas pelos gestores educacionais e por aqueles que elaboram as avaliações, já que a intenção de usar os resultados para o planejamento de ações de formação docente, ou mesmo para a oferta de informações educacionais para que o professor gerencie seu trabalho, tem sido apontada como objetivo em vários sistemas de avaliação do Brasil.

Os decisores políticos participantes consideram que a credibilidade da avaliação externa para o seu grupo é alta, como se pode observar na Figura 20, principalmente porque os resultados geram um impacto social, no sentido de render aceitação do público com relação às políticas que estão sendo implementadas. Porém, entendem que essa os professores não compartilham desta opinião, principalmente por causa da política de bonificação.

Credibilidade	
DP-A	Há um quadro diversificado com relação à credibilidade da avaliação externa. Com relação aos decisores políticos, há aqueles que acreditam e utilizam os resultados, porém também existem aqueles que não têm uma política de intervenção com base na avaliação. A credibilidade da avaliação para os professores tende a ficar comprometida, devido à implementação de propostas de meritocracia a partir dos resultados da avaliação. Acredito que as avaliações são construídas com qualidade, utilizam-se a TRI, as matrizes de referências etc. No entanto, é muito complicado quando a matriz passa a ser o currículo da escola.
DP-B	Há uma credibilidade total por parte dos decisores políticos, tendo em vista que há um impacto social dos resultados. Com relação aos outros intervenientes, eu acredito que a maioria também possui credibilidade nas avaliações externas. O que deve ser destacado, é que ocorre uma tendência política de o professor negar os resultados, mas não corresponde ao critério de credibilidade. Acredito que as avaliações são construídas com qualidade, pois fazem pré-testes, apresentam índice de dificuldade e de discriminação. Porém, há também limitações, pois são basicamente provas de múltipla escolha, em que o aluno reconhece a resposta certa ao invés de produzi-la.
DP-C	Os decisores políticos veem com bons olhos a avaliação externa, talvez seja o público que melhor se utiliza desta e nela mais acredita. Já os outros intervenientes variam. O público em geral não tem clareza com relação ao papel das avaliações externas. Os professores estão começando a encará-las de forma mais positiva, devido ao processo de formação de professores. E os pesquisadores as veem com bons olhos, mesmo identificando alguns riscos. Com relação às provas, a sua qualidade está associada a uma boa matriz de referência. Porém, não acredito que ela seja capaz de identificar o nível de desenvolvimento do aluno, apenas a proficiência da rede de ensino e, no máximo, da unidade escolar.
DP-D	Há uma parcela significativa de decisores políticos que atribuem um alto valor às avaliações externas, devido ao impacto dos resultados. Porém, muitas vezes, as discussões e posições ficam polarizadas entre os que defendem e os que criticam, tanto entre os gestores quanto entre os acadêmicos e professores. Já o público em geral não recebe uma informação qualificada para uma tomada de posição clara. Já sobre a credibilidade do instrumento, identifiquei problemas ou inadequações na padronização de material para coletas em nível nacional ou internacional.
DP-E	Os prefeitos e os governadores se preocupam muito, porque há impacto nas mídias em geral e a população identifica no Ideb, por isso dão um grande valor. A sociedade em geral atribui um grande valor às avaliações externas no momento, enquanto os professores em geral têm uma percepção mais crítica. Por fim, já com relação aos acadêmicos, vai depender muito de sua vinculação teórica. Quanto ao processo de desenvolvimento de itens de provas, o Brasil melhorou muito, principalmente nos casos do Saeb, do Ideb e da Prova Brasil.
DP-F	Os decisores políticos apresentam uma atenção muito forte para o Ideb. Ao criar um índice que permite identificar aprendizagem sem exclusões, criou-se um certo mecanismo de controle social para saber se as crianças estão aprendendo. Os professores têm uma relação ambígua com as avaliações. Quando essa avaliação é associada a bônus, por exemplo, gera-se, nas representações dos professores, principalmente nos sindicatos, uma visão de que estão culpando os professores – mesmo que seja um dinheiro a mais, fora do salário – por um desempenho educacional que não depende completamente deles. Quando se consegue superar as metas do Ideb, então a celebração é genuína. Acredito que, no caso brasileiro, nós ainda estamos aprendendo a construir avaliações. Se olharmos no início, as primeiras provas eram mais fracas como instrumentos de avaliação do que as atuais.

Figura 20

Dimensão “Credibilidade”: decisores políticos

Diante de tal situação, com a alta credibilidade da avaliação externa para os decisores políticos, fica claro que “se ainda há dificuldade no uso dos resultados para orientar a prática dos professores, ela não tem impedido sua aplicação em uma nova geração de políticas de gestão” (Brooke & Cunha, 2011). À vista disso, os decisores políticos participantes consideram que o maior foco estaria na utilidade para o sistema, no sentido de avaliar a qualidade da educação e subsidiar políticas públicas, como pode ser visto na Figura 21.

	Utilidade
DP-A	Avaliar o conjunto de competências e habilidades, básicas e essenciais, para entender determinadas etapas de escolaridade essenciais para que o aluno prossiga o seu estudo com sucesso. O foco principal da avaliação externa é a escola, analisar a escola e a turma, como eu posso trabalhar com aquela turma de modo que eu consiga ajudá-los a ter um processo de aprendizado mais eficaz.
DP-B	A avaliação externa vê de uma maneira mais geral a qualidade da educação oferecida no país, numa região, com um instrumento comparável de resultados, com uma boa avaliação. Para isso eu faria a avaliação externa, não para avaliar o aluno em particular, mas para que todo mundo saiba qual é a qualidade da educação oferecida no país, um instrumento democrático.
DP-C	Diagnosticar, por meio dos resultados, as lacunas na aprendizagem de uma rede. Considero a principal utilidade das avaliações externas o suporte para definições de políticas públicas cada vez mais eficientes e eficazes.
DP-D	A avaliação externa poderia ser um instrumento interessante para uma análise cruzada entre vários instrumentos ou informações de outras fontes.
DP-E	Ela serve para subsidiar as políticas, no que se refere à melhoria da qualidade, à formação de professores, à melhoria dos materiais didáticos e de apoio ao ensino.
DP-F	A utilidade da avaliação externa seria de avaliar se um determinado sistema educacional está tendo um desempenho bom no que se refere à aprendizagem. Ela não consegue dar um retorno imediato do aluno para o professor.

Figura 21

Dimensão “Utilidade”: decisores políticos

Com isso, os decisores políticos participantes da pesquisa entendem que os professores não utilizam os resultados das avaliações externas ou há uma baixa utilização. Tal fato decorre de uma baixa valorização no sistema de avaliação externa, por não saber usar esse tipo de dados, isto é, uma questão ligada à forma de divulgação do resultado, o que provém de um problema de quem conduz a avaliação e da política pública que a embasa, na medida em que não ocorre uma formação adequada do docente para entender e utilizar o resultado. Outra questão decorre do fato de as informações não terem utilização para os professores, não ocorrendo, por isso, uma apropriação do resultado como um apoio pedagógico.

Mas o grande desafio tem sido a pergunta que você me fez, como o docente utiliza isso. A gente tem se esforçado muito com material de divulgação, com itinerários informativos, com curso a distância, com curso presencial, oficinas. . . . mas eu não posso afirmar que os professores usam esses resultados para buscar

estratégias de aprendizagem, para verificar as dificuldades dos alunos. Eu tenho comigo que ainda tem algo muito complicado nas escolas. Aquelas escolas que saem bem, o boletim informativo chega até o professor, na rede, na comunicação. Mas aquelas que não saem bem, eu tenho a impressão que o boletim é arrebatado, que não há uma divulgação, que não há uma retomada (DP-A).

Com relação aos gestores escolares, os decisores políticos consideram que há uma utilização, no sentido de realizar comparações entre escolas, análise de resultados juntamente com os professores e repensar as práticas pedagógicas. A relação da equipe gestora como responsável dentro da escola para dialogar com os professores sobre os resultados das avaliações externas, aparece na fala dos decisores políticos.

O professor . . . é o fator escolar mais importante para o sucesso da aprendizagem. Ele fechou a porta, [a sala] é só dele. Ou a escola se abre para contar com ele, numa dinâmica de respeito e de troca, ou nada vai dar certo. Então, a equipe pedagógica, ao pegar os resultados da avaliação, tem que conversar com os professores. Tem às vezes de sensibilizar, às vezes os resultados são ruins e a tendência é culpabilizar o professor . . . (DP-B).

Dessa forma, quem mais teria utilização dos resultados das avaliações seriam os próprios decisores políticos, no sentido de planejamento ou replanejamento de políticas públicas, capacitação de professores, análise do desempenho das redes de ensino. No entanto, há uma percepção de que depende do estado, pois há os que utilizam, mas há outros em que os decisores políticos nem tomam conhecimento dos resultados, não sendo, assim, algo unânime em todas as gestões dos estados brasileiros. Um dos entrevistados, ex-secretária de educação da cidade do Rio de Janeiro, exemplificou como utilizava os resultados.

Eu vou contar um pouco como eu fiz como secretária de educação do Rio. . . . como era uma rede municipal muito grande, com 1.083 escolas, não dava para fazer uma reunião com todos, então a gente reunia os coordenadores das CREs, analisava quais eram as melhores escolas e quais eram as práticas que elas adotavam. A partir daí, por coordenadoria, a gente sentava com as dez melhores escolas e as escolas que mais evoluíram, que para mim é um dado muito mais importante. . . E a gente discutia um pouco as práticas, na percepção dos diretores,

a que eles atribuíam a melhora especificamente. O que eles fizeram que levou a isso. . . . Depois a gente identificava as escolas com maiores desafios em termos de desempenho, e a gente pareou essas escolas com escolas na mesma área e com o mesmo tipo de desafio. A gente chamou de escolas madrinhas, que trabalhavam juntas em um plano de recuperação da aprendizagem (DP-F).

No entanto, em pesquisa dos sistemas estaduais de avaliação, Sousa e Oliveira (2010) concluem que os resultados das avaliações externas são apropriados por serem considerados mais um indicador pelos decisores políticos, não identificando políticas específicas. O gerenciamento do sistema é realizado com base nas estruturas burocráticas, não encontrando evidências que pudessem afirmar que são utilizados os resultados de desempenho escolar. Evidenciam, assim, que o esforço é feito no sentido de produção de informações e não em ações que estimulem e apoiem o uso dos resultados. A pesquisa verificou que as iniciativas realizadas indicam a escola como grande usuária, devido às políticas de responsabilização de resultados.

Todos os decisores políticos consideraram que as avaliações são utilizadas para regular e reorientar o trabalho docente e que são uma função importante ou até mesmo a principal função da avaliação externa. Exemplos dessa regulação e reorientação seriam a formação de professores e gestores, a criação de projetos especiais, a elaboração de material didático, entre outros.

Sobre o controle do currículo, os decisores não consideram que essa utilização ocorra. Há respostas no sentido de que não seria o perfil de uma avaliação diagnóstica controlar o currículo e sim de uma avaliação formativa (DP-A). Mas também foi pontuado que não ocorre um controle porque os gestores não saberiam fazer, não conseguiriam traduzir as matrizes em conteúdos por ano escolar (DP-B).

Com relação à utilização para a formulação de políticas públicas, os decisores políticos foram unânimes em afirmar que as avaliações apresentam essa utilização, com o objetivo de melhorar a qualidade da educação. Porém, poucos apresentaram exemplos concretos. Uma política pública citada foi o trabalho de capacitação dos docentes, o que aparece constantemente na fala dos decisores políticos como ação para melhoria da qualidade da educação.

No nosso município, por exemplo, temos grupos de trabalho dos professores de cada componente curricular, para que, por meio de discussões sobre currículo,

resultados das avaliações, entre outros assuntos, possamos pensar nas estratégias de atuação para a melhoria na qualidade do ensino (DP-C).

A falta de mais exemplos concretos, para além da formação de professores, como forma de utilização da avaliação externa para a formulação de políticas educacionais, corrobora com pesquisas nacionais e internacionais que evidenciam que tais avaliações são mais utilizadas no sentido de realizar um estudo avaliativo do desempenho dos sistemas educacionais e legitimar reformas políticas, do que efetivamente para formular políticas educacionais (Araújo & Tenório, 2017; Bieber & Martens, 2011; Lingard, 2016).

Pode-se ver na Figura 22, as perspectivas defendidas pelos decisores políticos participantes na pesquisa. A análise da Figura permite destacar que para esse grupo de participantes, a avaliação externa se configura mais como um instrumento de utilização de gestores de sistema, e não de professores, como vem sendo delineado nas outras dimensões.

	Utilização
DP-A	Eu acho ainda que nem o professor nem o gestor da escola usam os resultados para buscar estratégias de aprendizagem, para verificar as dificuldades dos alunos. Já os decisores políticos, uns usam e outros não têm nem conhecimento dos resultados. Acredito que a principal função da avaliação externa seja a de regular e reorientar o trabalho pedagógico do professor. Mas quando fala em monitorar o currículo, não é muito o perfil, eu posso falar que estou monitorando o conjunto de habilidades que o aluno precisa aprender. O controle do currículo estaria mais voltado para a avaliação formativa. A avaliação externa é utilizada para formulação de políticas públicas, na criação de programas para a melhoria da qualidade da educação.
DP-B	Os docentes valorizam pouco a avaliação externa, por isso há uma baixa utilização. Eles não se sentem representados nessas avaliações e as criticam por não avaliar toda a aprendizagem. Não sei como as equipes de gestores das escolas as utilizam, mas acredito que deveriam discutir os resultados com os professores. Quanto aos decisores políticos, vejo que estão dando ênfase à capacitação dos professores a partir do resultado das avaliações. Acho que a avaliação externa tem sim potencial para regular e reorientar o trabalho docente, porém há limites, pois tem vários aspectos educacionais que ela não avalia e isso deve ser levado em consideração. Sobre o controle do currículo, eu nunca vi acontecer. Acredito que os gestores não saibam fazer. Isso não é fácil, tem que ter uma intermediação pedagógica. Com certeza a avaliação externa é utilizada na formulação de políticas públicas. Na realidade, acho que quase todos os programas de melhoria da qualidade da educação passam pela utilização de seus resultados.
DP-C	Não podem pensar que esta avaliação dá ao professor insumos mais importantes que suas próprias avaliações. Até porque a avaliação externa tem um caráter de informar mais aos gestores de uma rede de ensino. Por isso, acredito que há baixa utilização. A orientação é para que os gestores escolares usem os dados das avaliações internas e os analisem em conjunto com os resultados da avaliação externa. Com essas informações reunidas, as principais ações que podem realizar são a formação de seus professores e o replanejamento de suas ações pedagógicas. Os decisores políticos utilizam na tomada de decisões acerca do replanejamento de políticas públicas. Penso que regular e reorientar o trabalho docente seja um objetivo importante da avaliação externa, sendo feito a partir de criação de programas e na formação de professores. Sobre o currículo, acredito que apenas em âmbito nacional, pois a matriz de referência da avaliação externa é baseada no que há em comum nos currículos dos estados brasileiros. Se não há formulação e replanejamento de ações de políticas públicas a fim de melhorar o desempenho dos estudantes, com foco na sua aprendizagem, a avaliação externa perde seu sentido.
DP-D	Na educação básica, ainda há pouca discussão e aprofundamento de informação sobre as avaliações externas, por isso é difícil falar sobre efetivamente como os professores as utilizam. Os gestores as utilizam para comparar escolas ou grupos de escolas e, em alguns casos, apenas questionam sobre a veracidade das informações. Já os decisores políticos as utilizam como instrumento de análise do desempenho de sua própria rede e destacam os aspectos que consideram mais graves para um direcionamento da sua política. A avaliação externa poderia ser um instrumento interessante para regular e reorientar o trabalho docente, porém deveria ser feita levando em consideração análise cruzada com outros instrumentos ou informações de outras fontes. Não percebo que são usadas para controlar o currículo de cada escola. Quando as avaliações externas destacam os aspectos que consideram mais graves para um direcionamento da sua política, estão sendo utilizadas com direcionamento para as políticas públicas, o que eu acredito que ocorra.
DP-E	Os professores dão muita importância às avaliações externas, utilizando para a melhoria da aprendizagem de seus alunos, juntamente com os gestores das unidades escolares, com análise dos resultados. Até porque em muitos estados brasileiros há a atribuição de bônus para aqueles que conseguiram alcançar a meta. Os decisores políticos se preocupam no sentido de mostrar se o estado ou município estão fazendo um bom trabalho, estão melhorando o Ideb. Eu acredito que sim, a avaliação externa pode ser utilizada para reorientar o trabalho pedagógico dos professores e das escolas, de alguma maneira são os resultados das avaliações e a dificuldade de melhoria das aprendizagens que levam à definição de muitas políticas educacionais, como por exemplo na formação de professores. Sobre o monitoramento do currículo, acredito que as AE podem provocar a melhoria do currículo, mas lembrando que as avaliações externas são apenas um recorte do currículo, elas não dão conta de todo o currículo. A avaliação externa é utilizada para subsidiar as políticas públicas, principalmente as voltadas para a melhoria da qualidade, da formação de professores e dos materiais didáticos.
DP-F	A minha sensação é que hoje, como uma média no Brasil, o professor não usa as avaliações, nem as externas e nem as formativas. Eu acho que ainda está dissociado o ensino da aprendizagem, e o professor brasileiro não sabe usar dados para o seu processo de ensino. Acho que crescentemente as equipes gestoras usam para repensar as suas práticas. No caso dos decisores políticos, normalmente eles olham inicialmente com as suas equipes os dados e depois se reúnem com os diretores das escolas para repassar as informações. Estou falando especificamente o caso de avaliações censitárias, que é o caso da Prova Brasil, mas também é o caso de algumas avaliações estaduais. Podem e devem ser usadas para orientar o trabalho pedagógico. Regular, eu acho que não faz mais sentido com a Base Nacional Curricular. Porque a Base vai estipular as aprendizagens, estão reguladas as aprendizagens mínimas. Eu acho que a avaliação externa dá uma série de informações, inclusive até que ponto eu estou usando o currículo e como estou trabalhando. A gente tem mais na avaliação como um saber, primeiro se o currículo está sendo cumprido e segundo com qual qualidade ele está sendo cumprido. Sobre as políticas públicas, eu acho que não são suficientes. Eu acho que existem boas práticas nacionais. Se a gente fala do Brasil como um todo, o que aconteceu com a Prova Brasil de 2015 até 2017, nos anos iniciais, a cada ano houve progresso. Podem não ser progressos acelerados, mas são progressos importantes.

Figura 22

Dimensão “Utilização”: decisores políticos

Com base na Figura 22, há uma maior utilização dos resultados das avaliações seriam os próprios decisores políticos, no sentido de planejamento ou replanejamento de políticas públicas, capacitação de professores, análise do desempenho das redes de ensino, o que fica claro para o grupo de decisores políticos. Contudo, esses não conseguem responder sobre o grau de satisfação das avaliações externas. Não realizam uma relação entre a utilização e o grau de satisfação dos usuários. Isso pode ocorrer pelo motivo de haver pouca discussão em relação a outros fatores da avaliação, tendo em vista que a maior parte das discussões está voltada para resultados e efeitos, como afirma DP-D.

Eles entendem que há um grupo satisfeito, porém também há uma insatisfação por parte de outros. Não há um consenso no que tange à satisfação dos usuários, nem mesmo dentro de uma mesma categoria, ou seja, não se pode afirmar que todos ou a maior parte dos gestores políticos estão satisfeitos com a avaliação.

Convém destacar alguns pontos levantados pelos decisores políticos participantes. O primeiro diz respeito ao não questionamento sobre a realização da avaliação externa, como afirma DP-C “. . . todos estão de acordo com a necessidade de uma avaliação externa.”, mas sim que se deve pensar em um novo caminho para esse tipo de avaliação “. . . pois estes testes já não estão satisfazendo as expectativas dos gestores.”, como destaca DP-A.

Então o que eu acho é que esses estados que têm um sistema de avaliação consolidado e usam o resultado da avaliação já estão usando novas formas de avaliação. Tanto no que diz respeito à avaliação formativa; à avaliação digital, que ficam vários testes disponíveis para o professor, para o próprio professor avaliar o desempenho de seus alunos ao longo do tempo; à avaliação de competências sócioemocionais (DP-A).

Dois grupos foram considerados pelos decisores políticos participantes como críticos às avaliações externas, os especialistas das universidades públicas e os sindicatos dos professores.

De acordo com Bauer et al. (2015), a aceitação das avaliações externas entre os acadêmicos está longe de ser consensual, tendo em vista que ainda se encontra uma parcela significativa que desconsidera as eventuais contribuições oriundas das medidas educacionais em larga escala. No entanto, há também aqueles pesquisadores que não negam a sua contribuição, porém realizam uma análise crítica ao modo como as

avaliações vêm se difundindo e ao uso que se vem realizando de seus resultados. Este último posicionamento é percebido por DP-C, pois afirma que

Acredito que os pesquisadores veem a avaliação externa com “bons olhos”. A avaliação é um indicador para os gestores, e os educadores em geral. Mas uma característica forte dos pesquisadores é ir fundo nas questões investigadas. Por isso podem enxergar também alguns riscos na avaliação externa. A supervalorização da avaliação em detrimento do processo de ensino e aprendizagem, a distorção entre o conceito de matriz curricular e matriz de referência, a comparabilidade e o “ranqueamento” de escolas de uma rede, podem ser fatores encarados como fatores de risco por eles.

Com relação aos sindicatos, há um movimento contrário às avaliações externas, pois acreditam que elas reduzem a autonomia docente e escolar, além de responsabilizar injustamente os profissionais pelos resultados dos alunos (Cerdeira, Almeida & Costa, 2014). Com isso, o discurso sindical vem no sentido de abandonar a avaliação externa em prol da valorização do trabalho docente (Rosistolato & Viana, 2014). Como exemplo, a nota emitida no site do Sepe-Rio, que afirma não conhecer os objetivos e os critérios da avaliação, mas considera que qualquer avaliação diagnóstica deve ser realizada pela comunidade escolar, não por agentes externos, que não conhecem a realidade da escola, realizando um movimento de boicote à avaliação. Fica evidente que o sindicato se opõe a qualquer tipo de avaliação externa diagnóstica.

O Sindicato Estadual dos Profissionais de Educação, na Assembleia da Rede Estadual do dia 6 de abril, aprovou o boicote à “Avaliação Conhecer” da SEEDUC. A decisão dos profissionais de educação foi a de não participar das avaliações externas organizadas pela Secretaria de Educação. Apesar da falta de informações sobre os objetivos e os critérios da avaliação, entendemos que qualquer tipo de avaliação diagnóstica tem que ser organizada pela comunidade escolar, respeitando suas especificidades. As avaliações externas acabam produzindo dados que não refletem as particularidades das escolas. Além disto, são pensadas em uma lógica meritocrática, tanto para os discentes quanto docentes. Tal lógica é ineficaz para o processo de ensino e aprendizagem (Sepe-Rio, 2019).

Assim, como visto na fala dos decisores políticos, evidenciado na Figura 23, e na literatura, a avaliação externa recebe diferentes críticas, “por diferentes segmentos e entidades representativas, como sindicatos dos professores e importantes lideranças do meio acadêmico...” (Castro, 2009, p. 274).

Satisfação dos diferentes intervenientes	
DP-A	Então nós estamos tendo uma nova gama, uma nova era, um novo tipo de avaliação, porque estes testes já não estão satisfazendo as expectativas dos gestores. Sobre os outros intervenientes, é mais difícil saber.
DP-B	Eu não sei responder a esta pergunta, diferentes intervenientes não sei. Eu acho que muitos gestores anseiam pelos resultados para tomarem decisões, mas muitos também criticam que avaliação é muito parcial, não avalia todo o processo pedagógico. Mas com relação aos outros intervenientes, realmente para mim é difícil dizer.
DP-C	Eu diria que todos estão de acordo com a necessidade de uma avaliação externa. Mas acredito que não há nada que não possa ser melhorado.
DP-D	Não encontramos muito uma discussão específica sobre o processo de avaliação externa, mas sim dos seus resultados e os efeitos nas redes de ensino. Por isso, sobre a satisfação ficaria difícil responder.
DP-E	É difícil responder. A maioria dos secretários de educação acha importante a avaliação externa e em geral não pedem mudanças. Porém, do ponto de vista das faculdades e das universidades públicas, há um maior questionamento, é difícil responder.
DP-F	Há uma forte desconfiança de sindicatos, por exemplo, sobre a avaliação externa. E o sindicato acaba, de certa maneira, pautando a narrativa dos professores. Mas esses mesmos professores, se a escola recebe algum tipo de reconhecimento ou visibilidade, pelo bom Ideb, pelo bom resultado no Saesp ou Saerj, eles comemoram os avanços. Há, portanto, uma desconfiança.

Figura 23

Dimensão “Satisfação dos diferentes intervenientes”: decisores políticos

Neste primeiro momento, buscou-se realizar a caracterização da avaliação externa com base nas narrativas dos decisores políticos. Com base nessa caracterização, o segundo momento releva-se como os decisores políticos compreendem a relação da avaliação externa com a qualidade da educação, um conceito muito presente nos discursos políticos.

Qualidade da educação.

Com relação ao processo de ensino e aprendizagem, os decisores políticos consideram que há uma relação. Contudo, reconhecem também que há uma limitação nessa relação, por considerarem ser parcial, na medida em que é um recorte do que se ensina na escola, dando apenas um indicativo de onde pesquisar mais, como aponta DP-F.

A relação direta da avaliação externa com o processo de ensino e aprendizagem foi levantada com base no currículo, ou seja, a avaliação como algo que espelha o currículo escolar, estando, assim, diretamente relacionada com o que se pretende ensinar, o que é ensinado, o que é aprendido e o que é avaliado. Dessa forma, observa-se que há diferentes dimensões do currículo: o proposto, normalmente designado por oficial; o ensinado, o que é efetivamente posto em prática pelos professores e pelas escolas; e o aprendido, o que é efetivamente assimilado pelos alunos (Díaz-Barriga, 2005; Fernandes, 2006, 2010b; Hutchison & Schagen, 2007).

Eu me lembro bem da Prova Brasil de 2005, quando saíram os resultados, eu acompanhei muito de perto, pois passou a ser censitária . . . aí eu fui olhar o resultado e vi que 65% dos alunos de 9º ano não sabiam porcentagem. . . . Isso deveria acender uma luz vermelha fortíssima, e fazer com que as redes comesçassem a prestar atenção. Porque essa é uma competência que vai ser decisiva, não só para etapas posteriores de aprendizagem, mas pra comprar crédito, fazer coisas básicas que um cidadão precisa, para além de ser alfabetizado (DP-F).

Para tanto, de acordo com os decisores políticos participantes, a avaliação externa consegue conhecer o que está sendo ensinado e aprendido nas escolas, podendo, com isso, servir como apoio ao professor. Porém, para isso, deve ter menos caráter punitivo e ser vista mais como apoio pedagógico, dado que não há motivação para a utilização pedagógica caso a avaliação seja apresentada através de um discurso coercivo de uma só via, da descrição e/ou julgamento, ou de seus resultados serem incorporados a partir de um determinismo entre eles e o trabalho docente (Nevo, 1997; Oliveira & Araújo, 2005).

A questão da qualidade técnica da prova e a metodologia estatística foram consideradas como pontos favoráveis da avaliação externa, no sentido de conseguir conhecer o que os alunos sabem e são capazes de fazer. Isso demonstra uma forte confiança na qualidade técnica dos instrumentos.

A análise da Figura 24 permite destacar que para os decisores políticos, há limitações tanto na relação entre as avaliações externas e o processo de ensino e aprendizagem, principalmente porque essa avaliação é parcial. Outro ponto levantado é o caráter punitivo da avaliação, o que acaba por ser um forte obstáculo na utilização dela como apoio pedagógico.

Avaliação externa e qualidade do processo de ensino e aprendizagem	
DP-A	É muito difícil falar qual é a medida dessa correlação, mas que existe uma correlação forte, existe. Se o professor ensinar e o aluno aprender, a proporção de acerto na avaliação externa tende a ser grande. Acredito que a avaliação permite conhecer o que os alunos sabem e são capazes de fazer, inclusive a própria metodologia estatística, pois faz com que o resultado tenha eficácia.
DP-B	Elas costumam espelhar o currículo que é dado nas escolas, os itens são feitos para avaliar aquele aspecto do currículo, nesse aspecto eu acho que elas estão diretamente relacionadas com o que é ensinado. Porém, não avaliam todo o processo pedagógico, então são uma parte do que é ensinado na escola. Nesse ponto, elas são falhas. As avaliações são medidas cognitivas da aprendizagem, avaliam um tipo de habilidade, mas não todas as habilidades cognitivas. Por isso, elas não podem ser o único instrumento que a escola tem para avaliar a qualidade do que está sendo ensinado.
DP-C	O que se verifica com a aplicação de uma avaliação externa deve corresponder a um “currículo mínimo” utilizado pelos professores, ou seja, a relação deveria ser total, verificar se o processo de ensino e aprendizagem foi concluído de forma adequada. É complicado, pois dizer que é possível chegar no nível de proficiência do aluno é extremamente arriscado, pois o erro na medida é muito grande. Para se obterem informações tão detalhadas, os instrumentos aplicados pelos professores são mais adequados. Com as avaliações externas, podemos ter diagnósticos gerais, que nos ajudam a entender as fragilidades comuns entre os alunos.
DP-D	A relação existe quando um instrumento de avaliação busca informações sobre a aprendizagem dos alunos, indica que há relações sobre o trabalho de ensinar do professor. A avaliação externa permite conhecer em parte o que os alunos sabem e são capazes de fazer quando constroem instrumentos objetivos na busca dessas informações.
DP-E	Há pouca relação, eu gostaria que os resultados das avaliações influenciassem mais os projetos de ensino para os alunos. Eu acho que está começando, mas ainda é pouco utilizado no dia a dia das escolas. Acredito também que as faculdades de educação poderiam utilizar os resultados das avaliações externas para melhorar o processo de formação dos professores. Há várias evidências mostrando que as avaliações permitem conhecer o que os alunos sabem e são capazes de fazer. Por exemplo, mostram que, ao final do 9º ano, os alunos não aprenderam aqueles conteúdos básicos previstos, em língua portuguesa e matemática. Ou seja, elas permitem, agora falta melhorar o processo de ensino e os processos de formação de professores.
DP-F	A aplicação e a obtenção de resultados em si são um exercício de formação continuada de professor se bem utilizadas. Mas o resultado dela pode indicar onde os professores precisam de ajuda adicional. E esse caráter menos punitivo e mais de apoio ao professor, acho que ainda não está sendo utilizado o suficiente. Uma avaliação bem feita permite conhecer o que os alunos sabem e são capazes de fazer. Evidentemente, ela não consegue isso com tudo. Ela me dá indicativo de onde eu devo pesquisar mais.

Figura 24

Dimensão “Avaliação externa e qualidade do processo de ensino e aprendizagem”: decisores políticos

Para os decisores políticos, a avaliação externa não influencia as práticas de ensino ou influencia pouco, pois os professores não conseguem utilizá-la como um recurso para a melhoria de suas práticas. Assim, ela teria um potencial, mas é necessário maior investimento na formação inicial dos professores, em cursos de capacitação e em treinamentos. A questão, então, é posta na falta de habilidade dos professores em utilizar

a avaliação externa como apoio à sua prática, havendo, por conseguinte, um problema na formação destes.

Partindo desse ponto, cabe evidenciar que se desconsideram quaisquer questões da avaliação externa que possam vir a impedir ou minimizar sua influência na prática docente, como a natureza e propósito da avaliação, os mecanismos de processamento, a política pública que embasa a avaliação. Tendo em vista que o problema está na falta de habilidade do professor em utilizar a avaliação a seu favor.

Há uma distância do que é avaliado e o que eu faço no dia a dia. O que eu faço no dia a dia é carregado de coisas, de emoção de crianças problemáticas, de família... E a prova chega naquele momentinho uma única vez, então eu acho que ainda tem uma dificuldade de o professor transpor as habilidades que a avaliação externa mostra que os alunos adquiriram ou não com aquilo que ele tem que fazer pro aluno melhorar. Acho que essa transposição teria que ser muito mediada pelas faculdades, pelos cursos de capacitação (DP-B).

O grupo de decisores políticos presume que a avaliação externa cumpre o seu papel de coletar, produzir e difundir informações que possibilitem uma análise da realidade educacional, das potencialidades e dificuldades existentes, porém o professor não possui recursos para transpor tais dados para a sua prática, sendo necessário trabalhar essa relação no nível de formação de professores.

As preocupações dos professores em contexto de avaliação externa estariam relacionadas à política de bônus atrelada ao resultado; à culpabilização e às consequências provenientes de resultados inadequados; à falta de relação entre o que é verificado na avaliação externa e aquilo que a escola ensina. Essas preocupações vêm em resposta a um contexto de responsabilização, no qual órgãos governamentais pressionam gestores escolares e professores com base nos resultados da avaliação externa. Movimento que acaba por gerar consequências como a rejeição da avaliação por parte dos professores, desacreditando-a.

No entanto, para o decisor político DP-F, a avaliação externa é um instrumento que possibilita dar visibilidade para aqueles que realizam um bom trabalho. Aqueles que não apresentam um desempenho satisfatório estariam reclamando de uma pressão proveniente da avaliação externa, afirma DP-F.

. . . em todas as profissões, os bons profissionais ganham visibilidade. E uma ação contra dar visibilidade a bons resultados prejudica o reconhecimento da

profissão de professor. É como se professor não pudesse ser pressionado para um bom desempenho ou não pudesse se diferenciar dos colegas, todo mundo na mesma mediocridade. Eu estou exagerando para me fazer mais clara. Os professores que tiveram reconhecimento, a partir do momento que o Brasil passou a fazer avaliação externa, adoram o fato de isso ter acontecido, e vários ganharam visibilidade . . . Reconhecimento é fundamental em qualquer profissão, mas, pra ter reconhecimento, não basta ter a narrativa de pais felizes ou de uma professora que eventualmente é amiga da secretária ou amiga de não sei quem . . . Tem que ser centrada em dados e a avaliação externa permite o reconhecimento de professores (DP-F).

Dessa forma, a avaliação externa é vista como um instrumento de avaliação do trabalho docente. Fernandes (2008), com base em Medley (1982), categoriza esse tipo de avaliação como referente apenas à eficácia dos professores, pois toma como base a apreciação dos resultados dos seus alunos. A eficácia do professor é um conjunto que agrega a sua competência e o seu desempenho, somados ao que os alunos forem capazes de responder em determinadas situações.

Como afirma Sousa (2008), a eficácia é um novo elemento que começa a se delinear no Brasil, para além de fatores que comumente são avaliados, como assiduidade, pontualidade, incremento na formação, produção intelectual. Em defesa dessa nova política de avaliação docente, surge o argumento de que, assim, podem-se evidenciar boas práticas que dificilmente seriam vistas e, a partir daí, motivar outros profissionais a melhorarem. Com isso, a partir da avaliação da eficácia dos professores, seria possível chegar a uma avaliação da qualidade do ensino, visto que analisa o desempenho do professor, ou seja, aquilo que o professor efetivamente faz quando está em sala de aula (Fernandes, 2008).

Em pesquisa realizada em cinco secretarias estaduais e dez secretarias municipais de educação no Brasil, André (2015) apresenta, com base no eixo de políticas de valorização do trabalho docente, três principais tipos de iniciativas dos órgãos gestores das secretarias de educação: socialização de práticas exitosas, prêmios ou bônus por bom desempenho, e licenças ou bolsas para qualificação profissional dos docentes. Geralmente, com relação à avaliação externa, há uma maior utilização de premiação e bônus por bom desempenho, havendo pouco no sentido de socialização de práticas exitosas, o que necessitaria de mais estudos para verificar a eficácia de tal política.

Verifica-se uma maior ocorrência para a descoberta de tais práticas por meio de inscrição em concursos ou eventos, pelo próprio professor ou pela escola. A autora conclui que

há uma tendência a compensarem o professor individualmente, pois tanto a socialização de práticas exitosas quanto a concessão de licenças e afastamentos dirigem-se ao indivíduo isoladamente. Para que essas medidas atinjam resultados mais amplos, seria conveniente desenvolver propostas de incentivo ao trabalho coletivo, às comunidades de aprendizagem e ao desenvolvimento profissional do coletivo escolar. . . . No caso das experiências e práticas exitosas, pode-se indagar se não ficam restritas a momentos pontuais do exercício docente, pois o estímulo à socialização de práticas em congressos ou em publicações não parece ser suficientemente forte para sua consolidação e continuidade (André, 1015, pp. 227-228).

No que se refere à consistência das avaliações externa e interna, os decisores políticos consideraram que as duas avaliações são consistentes, desde que elas forneçam informações corretas ou que haja um cruzamento das informações entre as duas avaliações. No entanto, a formação inicial dos professores foi novamente abordada pelos decisores políticos, considerando que não há uma formação adequada na área de avaliação, o que vem a prejudicar a consistência da avaliação interna.

. . . nós não sabemos quais são as avaliações internas que as escolas adotam, eu acho que falta para o Brasil, em geral, uma prioridade nas avaliações formativas internas. Os nossos professores não sabem fazer avaliação formativa e avaliações processuais. As faculdades de educação, nos programas de formação, não orientam os professores nesse sentido. Isso é um problemão. (DP-E).

Os decisores políticos acreditam que há um condicionamento da avaliação externa sobre a prática de avaliação dos professores, ocorrendo de forma positiva, no sentido de os professores pensarem melhor no instrumento que estão elaborando e verificarem se a questão que eles estão fazendo permite realmente alcançar aquilo que eles desejam aferir. Isso seria um efeito do desenho da avaliação, a partir de uma reflexão sobre a consistência da avaliação. Por outro lado, há o levantamento de situação negativa, quando a avaliação externa condiciona os professores a utilizarem itens das avaliações externas para treinar os alunos.

Dois decisores políticos informaram que há oficinas que são destinadas a professores para que possam aprender a elaborar itens. Outro movimento é a criação de bancos de itens pelo governo, como recurso para os professores terem acesso a itens e poderem utilizá-los em suas aulas.

Quando ele participa de oficina de elaboração de itens, há de influenciar na prática sim. Em outro sentido, é o professor que busca na internet questões de prova estilo CAEd, Prova Brasil e aplica para os alunos. Eu não acredito que isso estimule a aprendizagem (DP-A).

Como eu tenho um diálogo constante com professores, eles não vão dizer isso, mas eles passam a pegar itens, quando tem alguns itens que estão disponíveis no banco de dados. E nós mesmo fizemos banco de itens para as avaliações unificadas . . . e muito professor passou a usar item. Algumas redes deram aulas de como fazer itens. Então, nós fizemos duas formações de professores que quiseram aprender a fazer itens e eles passaram a usar itens motivados pela avaliação externa (DP-F).

Como exemplo, podem-se destacar as políticas desenvolvidas pelo Sistema Mineiro de Avaliação da Educação Básica (SIMAVE). O programa de Minas Gerais conta com um banco de itens *on-line* de geração de provas. O banco, de acordo com o programa, é um suporte para as avaliações do professor em sala de aula, possibilita agilidade na aplicação de provas e obtenção de relatórios estatísticos do desempenho dos alunos, sendo, assim, um recurso para um diagnóstico pedagógico. Outra iniciativa do programa mineiro é a oficina de análise e elaboração de itens, com o objetivo de capacitar os professores da rede estadual para desenvolver habilidades de elaboração de itens para as avaliações escolares.

Entende-se que como é difícil para os atores escolares desenvolverem seus próprios testes. Por isso, os municípios e os estados realizam a criação de bancos de itens e oficinas como forma de assistência técnica às escolas. Assim, os professores são formados para a elaboração de itens e também têm um recurso *on-line* para a geração de provas, desenvolvendo seus próprios testes com parte dos mecanismos de processamento de uma avaliação externa.

Essas iniciativas objetivam disponibilizar os recursos da avaliação externa para a prática docente, provocando uma maior reflexão do professor com relação à avaliação por ele elaborada, o que poderia resultar numa maior consistência da avaliação interna,

vindo a servir “como mais um mecanismo para avançar na qualidade da educação pública oferecida.” (Alavarse et al., 2017, p. 1.369).

No entanto, essas práticas podem provocar efeitos negativos, na medida em que podem induzir à prática de simulados para a preparação de alunos para as provas e à utilização de mais itens de múltipla escolha na avaliação interna. Não foram encontradas pesquisas que objetivam um estudo sobre o impacto de banco de itens na prática docente, o que torna oportunas as pesquisas dessa natureza.

No que tange ao condicionamento do estudo dos alunos, os decisores políticos não consideram que tal fato ocorra; acreditam que não é algo que chegue com força no aluno a ponto de alterar sua forma de estudar.

Em síntese, os decisores políticos participantes consideram que a avaliação externa apresenta baixa ou nenhuma influencia nas práticas de ensino, sendo justificado, em grande parte, pelo fato de os professores não conseguirem utilizá-la como um recurso para a melhoria de suas práticas. Com relação à influência sobre as práticas avaliativas dos professores, esse grupo considerou que ocorrendo de forma positiva, pois os professores passam a pensar melhor na avaliação interna e verificam se a questão que eles estão fazendo permite realmente alcançar aquilo que eles desejam aferir. Por fim, no que tange ao estudo dos alunos, acreditam que tal fato não ocorra, como fica evidente na Figura 25.

Avaliação externa, prática de ensino e avaliação da aprendizagem	
DP-A	<p>Acredito que a avaliação externa não influencia nas práticas dos professores. O que podemos destacar sobre este assunto é o fato de alguns professores estarem treinando seus alunos para a realização das avaliações, o que, a meu ver, não gera eficácia. Os professores tendem a se preocupar com políticas de bônus atrelado ao resultado. Esta política pode causar rejeição por parte dos professores, mas também, em alguns lugares, mostrou-se efetiva no envolvimento para melhorar o desempenho dos alunos. A avaliação externa tem o foco de análise na escola; já [para] a avaliação interna, o objetivo é avaliar o processo de aprendizagem. Elas se completam. Acredito que a avaliação externa condiciona as práticas de avaliação interna no sentido de melhoria dos instrumentos utilizados pelos professores. Eles começam a melhorar o tipo de prova que estão elaborando. Com relação ao estudo dos alunos, a influência depende do movimento do professor. Não acredito que o aluno fará algo por si mesmo.</p>
DP-B	<p>Penso que a avaliação externa tem o potencial de influenciar as práticas dos professores, mas eu acho que talvez isso só se dê com a mediação pedagógica da escola e dos professores, com curso de capacitação. Eu não acho que hoje isso seja uma prática. A maior preocupação do professor hoje é ele ser culpado pelos resultados ruins. Pois há uma tendência em culpar o professor, então, às vezes, o professor rejeita a avaliação externa porque ela não é justa com ele. As avaliações externas são consistentes, eu acho que tem uma tecnologia de elaboração, ela tem estatísticas produzidas com confiabilidade, ninguém nega isso. As internas eu não sei, depende do professor, o problema é que na sala de aula o que o professor faz é uma caixa-preta. Não acho que a avaliação externa tem o poder de condicionar as práticas de avaliação dos professores, de eles fazerem avaliações só com o tipo de itens da avaliação externa. Eu acho que de nenhuma maneira condicionam os estudos dos alunos, se eles não influenciam os professores na sala de aula, muito menos os alunos.</p>
DP-C	<p>A prática dos professores pode ser repensada e o planejamento de suas ações revisitado, a partir dos resultados das avaliações tanto externas, quanto próprias. Para isso, a formação destes professores é um fator importantíssimo para que possam entender a lógica das avaliações externas e como tirar proveito dos seus resultados. Acredito que a maior preocupação dos professores seja a possível responsabilização pelos resultados, um ranqueamento entre as escolas, dando origem, assim, a uma impressão ruim desse tipo de avaliação. Acredito que ambas as avaliações sejam consistentes, a partir do momento em que consigam fornecer as informações corretas sobre o que precisa ser feito com base nos resultados obtidos. Não acredito que as avaliações externas condicionem as práticas de avaliação, pois são operações diferentes, com finalidades diferentes. Eventualmente, podem ocorrer simulados apenas para preparar para esta operação diferenciada. Não acredito que elas condicionem o estudo dos alunos.</p>
DP-D	<p>A avaliação externa é um instrumento para influenciar as políticas e não as práticas dos professores, por isso acho que não influencia a prática do professor. Acredito que o professor corre o risco de ser o único a ser responsabilizado pela não aprendizagem dos alunos, essa é a sua maior preocupação. A consistência dos instrumentos (internos/externos) é possível de ser identificada desde a sua elaboração até a sua análise, desde que haja um cruzamento dos resultados e uma análise pelos envolvidos. A avaliação externa pode influenciar na construção e diversificação de instrumentos mais objetivos das avaliações internas, mas predominam os instrumentos mais personalizados e punitivos. Assim como não influenciam as práticas dos professores, também não condicionam o estudo dos alunos, pois são instrumentos para influenciar as políticas.</p>
DP-E	<p>As práticas dos professores ainda são tradicionais e não utilizam nem os resultados das avaliações externas e nem há uma integração entre aquilo que se ensina e a prática didática. Eu acho que precisa de um grande trabalho nos programas de formação para que essas práticas de ensino dos professores melhorem, esse é um grande desafio da educação brasileira. A maior preocupação dos professores é que a avaliação externa não avalia aquilo que a escola ensina. Nós não sabemos quais são as avaliações internas que as escolas adotam, falta para o Brasil uma prioridade nas avaliações formativas internas. Os nossos professores não sabem fazer avaliação formativa, avaliações processuais. As faculdades de educação, nos programas de formação de professores, não orientam nesse sentido. As avaliações externas, muitas vezes, condicionam as avaliações internas, e condicionam de modo errado. Por exemplo, utilizam itens da avaliação externa para treinar alunos. Isso não resolve nada, professores precisam é melhorar as avaliações processuais e formativas, para ver o que os alunos estão aprendendo. Não acredito que a avaliação externa condicione o estudo dos alunos.</p>
DP-F	<p>Todo mundo fala muito que há um risco na avaliação externa de o professor ensinar para o teste. Se for viável ensinar para o teste, esse teste é mal feito. Os professores temem que avaliação externa coloque uma pressão exagerada sobre eles, eles temem que isso seja usado de uma maneira injusta com eles. As avaliações internas são inconsistentes, pois o professor, em muitos casos, enfatiza mais um elemento que não é visto como tão importante dentro da base ou do currículo municipal ou estadual. A outra coisa é quando o professor usa a reprovação como um instrumento de poder. Infelizmente, a avaliação externa influencia as práticas de avaliação dos professores, é impressionante como influenciam. Como eu tenho um diálogo constante com professores, eles não vão dizer isso, mas eles passam a pegar itens, quando disponíveis no banco de dados, para utilizar. Os alunos sabem que a avaliação externa não vai ser usada para nota, por isso condicionar o estudo não aparece com força.</p>

Figura 25

Dimensão “Avaliação externa, prática de ensino e avaliação da aprendizagem”: decisores políticos

Para os decisores políticos participantes, a apropriação de diferentes intervenientes depende de alguns fatores de caráter externo, isto é, política de incentivo ou direcionamento para a apropriação, seja ela uma política de estado, da rede, seja da escola. Desse modo, deve haver uma ação voltada para isso, por pressão ou como recurso para gestão ou planejamento. Visto isso, para que haja apropriação de resultados, a avaliação deve ter utilidade.

. . . eu acredito que hoje, no país, 90% dos secretários fazem sua gestão com base nos resultados (DP-A).

Aí depende muito da escola e da rede física. Algumas redes dão um grande valor ao resultado da avaliação, discutem esses resultados com os docentes, outras não. Eu acredito que seria importante que as redes analisassem os resultados das avaliações externas, discutissem isso com os professores, identificassem os problemas, fizessem as margens dos erros e pudessem, a partir daí, melhorar a formação dos professores e o processo de ensino e aprendizagem (DP-E).

Eu acho que conhecer bem, eles conhecem, nas redes que têm algum tipo de bônus. Depende muito, na nossa rede [Rio de Janeiro], eles sabiam bem, mas não é uma realidade de Brasil. Nas redes que celebram os resultados do Ideb, eles conhecem (DP-F).

Entendeu-se também que a apropriação se dá de forma diferente para cada interveniente, os professores e os gestores teriam maior interesse; já para os alunos e as famílias, não haveria muita apropriação.

Com relação ao conhecimento e à análise dos resultados pelos professores, dois pontos merecem destaque. O primeiro versa no sentido de relacionar as ações dos diferentes níveis governamentais (Federal, Estadual e Municipal) e a apropriação dos professores.

. . . depende um pouco da estratégia, as avaliações externas do governo federal têm um alcance, as avaliações externas das redes estaduais e municipais têm outro. A gente tem visto que as redes municipais e estaduais tendem a fazer relatórios pedagógicos com os resultados apresentados na escala de proficiência e passam para as escolas, para os professores. Aí eu vejo que os professores analisam. Nas avaliações nacionais, você entra no site do Inep, está lá o que é o nível 250 . . . está longe do professor. Então, eu vejo assim, quando as avaliações externas são das redes estaduais, eu vejo muitos seminários para o professor interpretar o relatório pedagógico . . ., o desempenho dos alunos, a explicação do desempenho, os exemplos de itens que eles sabem resolver. Isso eu acho que causa impacto nos professores, as avaliações mais próximas deles. No nível estadual, as secretarias de educação convocam os diretores e os coordenadores pedagógicos para discutir. Eles têm o resultado de cada escola, de cada turma . . .

É diferente do resultado nacional, são resultados abrangentes . . . Então, os níveis municipal e estadual geralmente fazem essa análise, com o relatório técnico-pedagógico . . . é o que eu vejo. E a nível federal, há um só resultado . . . que não está tão perto do professor como você vê nos níveis estaduais e municipais . . . (DP-B).

É importante ter presente que, devido ao fato de que o Saeb apresenta baixa interferência na vida das escolas e no currículo escolar, o programa acabou por receber críticas, dado que não conseguiu instituir uma educação de qualidade e seus resultados reverberaram apenas no sistema (Bonamino & Sousa, 2012; Vianna, 2002). Por isso, muitos estados e municípios sentiram a necessidade de criar seus sistemas próprios de avaliação, além de o governo federal implementar a Prova Brasil, que passa a fornecer dados por unidade escolar.

No âmbito subnacional, observa-se uma maior participação de gestores escolares e professores, principalmente no que se refere às avaliações municipais (Alavarse, Bravo & Machado, 2012). Há uma incidência maior de fornecimento de *feedback* a esses atores escolares, muito em decorrência de sua abrangência local, se propondo a ser um instrumento de apoio pedagógico. Assim, as secretarias disponibilizam material informativo que visa traçar um diagnóstico do nível de aprendizagem dos alunos, geralmente através de boletins pedagógicos, que são direcionados para os professores, e disponibilizam os resultados específicos de cada escola e os relatórios, destinados aos gestores, que resumem de forma mais técnica os resultados da rede (Brooke & Cunha, 2011). Além disso, promovem seminários, oficinas e encontros regionais para a apresentação e discussão de resultados.

Outro ponto apontado pelos decisores participantes é a maior apropriação pelos gestores das escolas e dos professores das disciplinas e anos escolares avaliados, o que comprova o estreitamento do currículo e a ênfase em algumas disciplinas com relação a outras, em decorrência da avaliação externa, o que se aproxima dos efeitos descritos na literatura e pesquisas (Chamberlin et al., 2002; Ravitch, 2009).

Os gestores políticos foram os intervenientes que se considera que mais realizam análise dos resultados das avaliações. Porém, também se compreende que tal fato ocorre somente se houver uma política de indução forte. A utilização geralmente ocorre no sentido de obter informações para realizar um planejamento de ações para a rede de ensino.

No entanto, a constante mudança de gestores políticos foi apontada como um ponto negativo, no sentido de serem realizados planejamento e ações com base na análise dos resultados de uma avaliação, principalmente, no que concerne aos níveis estadual e municipal, visto que há uma descontinuidade da política na mudança de um gestor para outro (Gatti, 2012), o que gera desconfianças e resistências. Isso foi referido como “efeito zigue-zague” por Cunha (1995). Num momento, uma avaliação é implementada e, ao mudar o gestor, a avaliação acaba ou muda toda a sua estrutura.

. . . o problema é que os gestores mudam muito, uma certa inconstância política, de partido político. Mas os professores são mais constantes, eles se articulam melhor, se eles criam uma rede de comunicação é difícil que quem chegue mude tudo (DP-B).

Como sintetizado na Figura 26, fica claro que os decisores políticos entendem que para se falar em apropriação de resultados, deve-se considerar fatores externos, como política de incentivo ou direcionamento para a apropriação, seja ela uma política de estado, da rede, seja da escola. Esse grupo considera que os gestores políticos são os intervenientes que mais realizam análise de resultados, porém para que essa análise ocorra, deve haver uma política de indução forte.

Resultado da avaliação externa	
DP-A	A apropriação dos diferentes intervenientes depende da política de cada estado. Se houver um movimento para utilização dos resultados, a apropriação ocorre para melhorar a qualidade da educação, pensando em diferentes ações. Os docente só vão conhecer e analisar os resultados da avaliação externa se houver uma política para isso no estado. O gestor só vai analisar se ele tiver uma política educacional forte, ou seja, se ele usa os resultados para fazer uma política pública voltada para equidade e qualidade, formação de gestor escolar. Agora, se não tiver, ele não realizará a análise dos resultado.
DP-B	Os professores e os gestores têm um interesse na apropriação dos resultados, a partir de dados e análises divulgadas. Quanto aos alunos, acredito que não, até porque nem sei se eles têm maturidade para isso. As avaliações externas do governo federal têm um alcance, as da redes estaduais e municipais têm outro. A gente tem visto que as redes municipais e estaduais tendem a fazer relatórios pedagógicos com os resultados e passam para as escolas, para os professores, aí eu vejo que os professores analisam. Quando as AE são das redes estaduais, vejo muitos seminários, os resultados indo para o professor interpretar junto com o relatório pedagógico, o desempenho dos alunos, a explicação do desempenho, os exemplos de itens que eles sabem resolver. Isso eu acho que causa impacto nos professores. É diferente do resultado nacional, que é mais abrangente. A nível estadual e municipal, eu tenho absoluta certeza de que os gestores analisam os resultados, porque as secretarias chamam todos os diretores, apresentam os resultados. Já no âmbito nacional eu não sei.
DP-C	Com relação ao aluno, como os resultados não são individuais, o máximo que podemos associar a eles é o resultado de suas escolas, e o replanejamento do trabalho que terá foco neste resultado, o que os impacta diretamente. Os professores acabam por se apropriarem dos resultados para repensar e planejar sua prática. Já os gestores se apropriam dos resultados para se informar e planejar ações para a sua rede. Não acredito que todos os professores conheçam os resultados da avaliação externa, mas acredito que a maioria conheça. No entanto, penso que tenhamos um percentual menor de professores que analisam ou fazem uso dos resultados. Não acredito que haja uma equipe gestora que não faça uso ou análise destes resultados.
DP-D	Os diferentes intervenientes se apropriam das avaliações externas de formas diferentes. Com relação aos alunos e família, acredito que não há apropriação. Os professores e os gestores acabam por se apropriar dos resultados para pensar em ações para as suas práticas. No âmbito da rede de ensino, acredito que conhecem e analisam de uma maneira geral, tanto os professores quanto os gestores.
DP-E	O conhecimento e a análise de resultados por parte dos professores depende muito da escola e da rede. Algumas redes dão um grande valor ao resultado da avaliação, discutem esses resultados com os docentes, outras não. Eu acredito que seria importante que as redes analisassem os resultados das avaliações externas, discutissem isso com os professores, identificassem os problemas, fizessem as margens dos erros e pudessem, a partir daí, melhorar a formação dos professores e o processo de ensino e aprendizagem. Em geral, os gestores escolares não analisam os resultados das avaliações. Com exceção de algumas redes que têm coordenadores pedagógicos, que valorizam os resultados da avaliação para melhorar a formação dos professores, mas em geral, eles não analisam.
DP-F	Eu acho que é muito diferente a apropriação. Os gestores, nas redes em que é dada muita visibilidade ao desempenho, colocam resultado do Ideb em lugares a que os pais têm acesso. Os professores menos, sendo importante lembrar que, como o Ideb mede mais Português e Matemática, os professores de outras disciplinas às vezes não prestam a mesma atenção. Mas quando a escola tem um resultado muito bom, isso mobiliza a comunidade e tem alguns diretores que impõem a meta a toda a escola. Os pais sabiam a meta de melhoria do Ideb, os alunos eventualmente também sabiam, mas isso é raro. Então eu acho que impacta muito os diretores, os professores daquela série avaliada e impacta menos os professores de outras séries e outras disciplinas. Os alunos ainda não. Os pais ficam sabendo que aquela escola é boa. Com relação aos gestores, depende da liderança da rede.

Figura 26

Dimensão “Resultado da avaliação externa”: decisores políticos

Essas foram as percepções apresentadas pelos decisores políticos participantes sobre a avaliação da aprendizagem e sua relação com a qualidade do ensino, estando dividida nas dimensões desses objetos de pesquisa. Na parte a seguir, será discutida a percepção dos professores da educação básica.

Narrativa dos Professores da Educação Básica

Nesta seção, a avaliação externa da aprendizagem dos alunos será vista a partir das narrativas dos professores participantes da pesquisa. Foram selecionados professores de português e matemática do ensino fundamental II, de escolas públicas do município do Rio de Janeiro.

Avaliação externa da aprendizagem.

Os professores da educação básica identificaram uma maior variedade de propósitos associados às avaliações externas, tais como avaliar a qualidade da educação; realizar um diagnóstico do processo de ensino e aprendizagem; apoiar as políticas públicas de educação; avaliar o trabalho docente; e contribuir para o desenvolvimento e melhoria dos processos pedagógicos. Assim, os propósitos que são trazidos pelos professores revelam que as avaliações externas têm a função política e pedagógica, desde uma característica somativa até a formativa.

Alguns mecanismos de processamentos foram considerados pelos professores participantes, possibilitando conhecer mais sobre a natureza da avaliação externa. Os mecanismos evidenciados foram: índices maquiados, periodicidade não contínua, avaliações mal organizadas, critérios de correção desconhecidos pelos docentes, provas descontextualizadas e desconsideração das particularidades das escolas. Há uma preocupação com o aspecto técnico da avaliação, como se pode ver na fala dos docentes, indicando que esses não a consideram confiável, com qualidade e válida:

. . . eu acho que a característica predominante desse tipo de avaliação é enganação. Eu sei que é bem sério isso que eu estou falando, mas a verdade é que essa avaliação serve meramente pra instituição de índice, que muitas vezes é maquiado. . . (P-B).

Acredito que a avaliação externa é mal organizada e descontextualizada (P-G).

Para além de descreverem alguns pontos negativos que compreendem os mecanismos de processamento, os professores também sublinharam alguns efeitos negativos, como contribuir para a padronização do ensino e para a responsabilização do professor.

A preocupação com os dados levantados pelos professores participantes já foi pesquisada por Ingram, Louis e Schroeder (2004). Os autores afirmam que há uma disponibilidade dos professores em utilizar dados para a melhoria do ensino; por outro lado, há preocupações consideráveis no que diz respeito às informações disponíveis e como estas estão sendo utilizadas com o intuito de avaliar os docentes. A pesquisa americana evidenciou que, para os professores entrevistados para o estudo, há possibilidade de utilização de dados das avaliações externas, no entanto, nas condições presentes no contexto estudado, seria irreal.

Observa-se que, para os professores entrevistados, há uma função pedagógica na avaliação externa. No entanto, quando discorrem sobre sua natureza, os aspectos negativos, tanto dos mecanismos de processamento quanto dos efeitos, são os que sobressaem, o que acaba por obscurecer essa função.

As perspectivas defendidas pelos professores da educação básica participantes na pesquisa estão resumidas na Figura 27. A análise da Figura permite destacar que para esse grupo de participantes, as avaliações externas são de natureza tanto somativa quanto formativa. Os propósitos apresentados confirmam a presença das duas naturezas, já que para os professores a avaliação externa apresenta como propósitos: avaliar a qualidade da educação; realizar um diagnóstico do processo de ensino e aprendizagem; apoiar as políticas públicas de educação; avaliar o trabalho docente; e contribuir para o desenvolvimento e melhoria dos processos pedagógicos.

Natureza e propósitos	
P-A	Uma proposta de avaliação do trabalho docente.
P-B	Eu acho que a característica predominante desse tipo de avaliação é enganação, e eu sei que é bem sério isso que eu estou falando, mas a verdade é que essa avaliação serve meramente pra se instituir índices que muitas vezes são maquiados. Se joga muitas vezes na conta do professor preparar o aluno pra esse tipo de coisa. Na maioria das vezes, esta nota não é divulgada, não pontualmente, você acaba tendo um índice da sua escola, mas pontualmente, aluno por aluno, você não sabe quanto tirou, como foi o critério de correção daquilo. E existe um problema também com relação à ementa, ela não reflete muitas vezes os problemas que a gente encara dentro da unidade escolar. E aí isso reflete no índice com certeza, chega-se a resultados que não são concretos, e a gente acaba que tendo que aturar consequências de uma coisa mal feita.
P-C	Identificar, aferir, analisar o processo de ensino-aprendizagem e dar <i>feedback</i> para os professores sempre reformularem suas práticas. Além de dar uma ideia abrangente das condições em que se encontra a educação escolar dos alunos no Brasil.
P-D	A avaliação externa é diagnóstica e visa à padronização do ensino.
P-E	São necessárias, mas, no Brasil, são descontínuas e muitas vezes “maquiadas”.
P-F	Avalia a qualidade do ensino e fornece informações importantes a fim de promover políticas públicas decorrentes da aprendizagem dos alunos, identificando falhas e lançando mão de estratégias eficazes, visando melhorar o ensino.
P-G	Acredito que a avaliação externa é mal organizada e descontextualizada.
P-H	Diagnóstico dos educandos, dos professores, porém nem sempre é real.
P-I	Uma prova enviada pelo governo para ter dados de quanto o aluno aprendeu.

Figura 27

Dimensão “Natureza e propósitos”: professores

Com base nos dados quantitativos, mostrados na Tabela 5, observa-se que os professores concordam que o principal propósito da avaliação externa é avaliar a qualidade da educação, o que evidencia ser mais uma função política somativa. No entanto, nas entrevistas, esse propósito apareceu apenas em duas respostas. Isso pode ter ocorrido pelo fato de os professores, na sua maioria, destacarem mais aspectos negativos da natureza das avaliações externas, o que acabou por apagar outras características. Nas duas respostas em que aparece o propósito de avaliar a qualidade da educação (P-C e P-F), os professores consideram apenas pontos positivos. Com isso, evidencia-se que, quando a percepção dos professores está, sobretudo, nos aspectos negativos da avaliação externa, ocorre uma forte tendência de contaminação, ou seja, os aspectos negativos invalidam todo o processo.

Os professores discordam que as avaliações no âmbito nacional contribuem para a regulação do currículo. Sobre isso, o único ponto abordado na entrevista foi a “padronização do ensino” (P-D), sendo a avaliação externa um instrumento que possibilita regular o que é ensinado pelo professor, o que acaba de alguma forma

impactando o currículo, seja impulsionando mudanças em currículos inadequados seja promovendo o afinamento curricular (Au, 1997; Esteban & Fetzner, 2015; Madaus et al., 2009; Sousa, 2003; Stecher, 2012).

Sobre a avaliação docente, os professores não consideram que a avaliação externa possibilite isso. Tal aspecto pode ser explicado pelos pontos negativos dos mecanismos de processamento trazidos nas entrevistas, pelas limitações da avaliação ou por não ser possível realizar uma avaliação do professor por meio do desempenho do aluno. Porém, apesar de não acreditarem nessa possibilidade, há a percepção de que o Estado busque avaliar o trabalho do professor por meio da avaliação externa, como afirmado pelos professores nas entrevistas.

. . . eu sempre considereí como uma proposta de avaliação externa do trabalho que eu desenvolvo (P-A).

Diagnóstico dos educandos, dos professores, porém nem sempre é real (P-H).

Tabela 5
Natureza e propósitos

Questões	Concordo	Discordo
As avaliações externas têm como principal propósito avaliar a qualidade da educação	53,2%	46,8%
As avaliações externas contribuem para que, em nível nacional, todos os alunos aprendam os aspectos fundamentais do currículo	44,4%	55,7%
As avaliações externas permitem avaliar o trabalho pedagógico dos docentes	26,6%	73,4%

Alguns professores consideram que a avaliação externa pode ser relevante e importante para o processo de ensino e aprendizagem, o que reforça a presença de sua natureza formativa como evidencia a Figura 27, pois “permite ao professor reavaliar suas práticas e, dessa forma, replanejar constantemente” (P-D); “os envolvidos (educador e educando) conseguem mensurar a efetividade do ensino” (P-E); “Para sinalizar o nível intelectual dos alunos e efetuar algumas mudanças se o resultado não for o esperado” (P-H); e “observando os acertos e os descritores em que houve mais dificuldade” (P-I).

Portanto, um número de professores parece perceber que a avaliação externa é relevante para a organização das suas ações pedagógicas, tendo em vista que fornece dados que podem ser usados como apoio pedagógico, contribuindo com pistas sobre o que e quando ensinar aos alunos, evitando, assim, perda de tempo didático (Evers & Walberg, 2002; Madaus et al., 2009). Isso a revela não como uma atividade isolada, mas como um processo que traz a sua contribuição na ação pedagógica (Nevo, 1997), o que destaca o seu caráter formativo.

Há também um grupo de docentes participantes que reconhece que as avaliações externas têm potencial para serem relevantes, com influência nas suas próprias práticas. Entretanto, devido aos seus mecanismos de processamento, isso não se verifica. As afirmações seguintes ilustram a percepção desse grupo: “O professor deve receber o *feedback* da avaliação de seus alunos e isso nem sempre acontece.” (P-F) e “Uma avaliação externa, se bem organizada, pode revelar mais da instituição avaliada.” (P-G). De acordo com Marchesi (2002, p. 33), “poucos são os professores que duvidam da sua importância e da sua utilidade [avaliação externa]. E contudo, muitos desconfiam da forma como está sendo posta em prática.”, o que fica evidente nas respostas apresentadas nas entrevistas.

Porém, outros professores têm percepções opostas acerca da relevância das avaliações externas no que se refere à sua ação pedagógica, ou seja, consideram que tal relevância é nula (Stecher, 2002), não influenciando o seu trabalho. Como afirmou P-A, “O nosso trabalho não muda em função dessas avaliações.”. Com isso, a avaliação externa estaria mais a serviço de uma função política.

A Figura 28 apresenta as perspectivas defendidas pelos professores, pode-se perceber que os professores participantes apresentam três posicionamentos: aqueles que consideram que a avaliação externa pode ser relevante e importante para o processo de ensino e aprendizagem; há os que reconhecem que as há um potencial para serem relevantes, no entanto os mecanismos de processamento anulam esse potencial; por fim, aqueles que não acreditam em tal relevância.

	Relevância e importância
P-A	A relevância da avaliação externa para quem está de fato lecionando é mínima. O nosso trabalho não muda em função dessas avaliações. No entanto, elas são importantes para os gestores e para os pais.
P-B	Quando a gente fala de processo de ensino e aprendizagem, a gente está falando da evolução do aluno ou da evolução da rede. Bom, para o aluno, eu acho que a relevância é zero. Essa nota existe para o governo apresentar um índice de quantidade de pessoas alfabetizadas ou quantidade de pessoas num determinado nível de inteligência, vamos colocar assim. Para o aluno, muito pouca, eu acho que seria mais interessante se a gente se preocupasse com esses números para avaliar a rede pública de ensino.
P-C	No sentido em que permite ao professor reavaliar suas práticas e dessa forma replanejar constantemente.
P-D	Permite fazer uma diagnose da aprendizagem e dar base para definir os rumos da educação no país.
P-E	Na medida em que os envolvidos (educador e educando) conseguem mensurar a efetividade do ensino.
P-F	O professor deve receber o <i>feedback</i> da avaliação de seus alunos e isso nem sempre acontece. Através dos descritores com maiores quantidades de erros, é possível identificar as habilidades não desenvolvidas por esses alunos, dando suporte ao professor ao replanejar suas práticas pedagógicas.
P-G	Uma avaliação externa, se bem organizada, pode revelar mais da instituição avaliada. Porém, o que normalmente a escola recebe são blocos de papéis para serem aplicados de qualquer forma aos alunos, dentro de um prazo outorgado sem explicações suficientes, mandam aplicadores despreparados para uma escola pública lotada. Sendo assim, as provas não são levadas a sério pelos alunos, que percebem a desorganização, a improvisação do evento "prova". Então as notas passam a não revelar nada, as avaliações ficam sem propósitos e acabam vistas como algo que atrapalha o andamento da escola. Um gasto de dinheiro público que se não organizar melhor, nunca vai medir nada.
P-H	Para sinalizar o nível intelectual dos alunos e efetuar algumas mudanças se o resultado não for o esperado.
P-I	Quando o professor utiliza essa avaliação num segundo momento, observando os acertos e os descritores em que houve mais dificuldade.

Figura 28

Dimensão "Relevância e importância": professores

Nas respostas ao inquérito por questionário, os professores discordaram de as avaliações externas serem relevantes para o processo de ensino e aprendizagem (55,7% dos respondentes), no entanto, o percentual de diferença é baixo, o que mostra um quadro equilibrado. Tal fato pode ocorrer devido à presença de três grupos, como verificado nas entrevistas. O primeiro seria composto pelos professores que concordam que as avaliações externas são relevantes para apoiar o processo de ensino e aprendizagem. Um segundo grupo compreende que há um potencial para tal relevância, porém há problemas em seus mecanismos de processamento que a limitam ou anulam. Por fim, o grupo que não considera que a avaliação externa seja relevante, não tendo, assim, uma função pedagógica. Com isso, alguns professores podem ter discordado da relevância dessa avaliação para o processo de ensino e aprendizagem devido a suas limitações e não por ela não apresentar potencial para ser relevante.

Os professores consideram que a credibilidade da avaliação externa é baixa, devido aos mecanismos de processamento, às políticas públicas que a embasam e, por conseguinte, a seus efeitos negativos.

Sobre os mecanismos de processamento, os professores destacam a padronização das avaliações, a forma como é feita a coleta dos dados, o fato de não se considerarem os fatores associados e de os governantes maquiarem os índices. Pode-se observar, no discurso dos professores, que há um entendimento de desarticulação entre o contexto escolar e a avaliação externa, não considerando as peculiaridades, as diversidades presentes nos espaços escolares, o que leva a um distanciamento entre os resultados obtidos e os sujeitos que os produzem (Esteban, 2009). Dessa forma, para o professor, não seria possível a utilização dos resultados.

Outro ponto que merece destaque na fala dos professores entrevistados é a manipulação dos resultados da avaliação externa, o que levaria a índices favoráveis para uma suposta qualidade da educação, para uma visão favorável de uma determinada política educacional.

. . . as pessoas que estão por trás disso, querendo gerar índices, acredito que elas tenham uma ótima ideia disso. Até porque, geram uns índices maravilhosos, que a gente sabe que não existem. De um milhão de pessoas alfabetizadas, a maioria mal e porcamente sabe escrever o próprio nome (P-B).

A ideia de que o governo maquia resultados leva a um quadro de dúvidas e incertezas plausíveis sobre os índices oficialmente divulgados e questionamentos sobre a credibilidade da avaliação externa. Os sujeitos escolares passam a ignorar tais resultados, pois entendem que não são fontes confiáveis de análise e apoio pedagógico, e nem para obtenção de informações provindas de uma função política somativa. Para além disso, gera um sentimento de raiva e desmotivação, tendo em vista que veem seu trabalho sendo manipulado.

A credibilidade relacionada ao efeito surge quando os professores destacam a pressão por resultado, a política de bonificação e a culpabilização dos professores e das escolas. Em geral, a resposta a essas ações governamentais seria a ação de preparação para o teste, podendo suceder num resultado que não representa a realidade acadêmica dos alunos (Esteban & Fetzner, 2015; Madaus et al., 2009).

Entretanto, em análise da literatura realizada por Bonamino e Sousa (2012), concluiu-se que, apesar de estudos indicarem poucos efeitos do SARESP no cotidiano escolar, havia reações de desconfiança e de resistência ao sistema por parte dos profissionais de educação. Tal fato pode ser explicado por haver uma percepção negativa

dos professores com relação à avaliação externa, que contamina todas as avaliações, sem realizar distinção entre suas funções e propósitos.

No que concerne às políticas públicas, o questionamento surge com relação a uma falta de confiabilidade nos objetivos que o governo tem, de fato, sobre os resultados das avaliações, pois estariam voltados à formulação de políticas públicas de acordo com seus interesses. Consideram que esses interesses não seriam necessariamente em prol de uma efetiva melhoria na qualidade da educação, haja vista que os resultados positivos da avaliação externa são considerados como uma chancela às iniciativas de política educacional (Sousa & Oliveira, 2010). Isso conduziria, assim, a uma manipulação dos resultados para fundamentação ou avaliação positiva de políticas já estabelecidas.

A Figura 29 apresenta as percepções mais relevantes que os professores participantes apresentaram no contexto desta pesquisa sobre a credibilidade. As narrativas descritas na Figura permitem verificar que, para esses professores, a credibilidade das avaliações externas para o seu grupo está associada a três fatores principais: a) aos mecanismos de processamento, b) às políticas públicas que a embasam e, por conseguinte, c) a seus efeitos negativos.

	Credibilidade
P-A	Os professores acreditam que essas avaliações não avaliam nada. Se há algum empenho da classe em fazer com que seus alunos se saiam bem na avaliação externa, esse empenho deve-se a alguma forma de bonificação (normalmente financeira) relacionada aos resultados dos alunos nas avaliações. O público em geral só vê o resultado final. Os únicos grupos que vi até hoje interessados nessas avaliações são os pais que desejam colocar seus filhos em escolas públicas, pesquisadores e jornalistas. Nem os alunos se mostram muito interessados pelo resultado final dessas avaliações. Sou professora de língua portuguesa, meus alunos fazem uma prova de múltipla escolha. Na minha opinião, a prova não avalia o que é mais importante: a capacidade de escrita dos alunos. Acho uma prova mal feita, com questões muito simples.
P-B	Muito pouca a credibilidade dessa avaliação para os professores, e eu tiro pelos meus colegas, com quem eu converso bastante. É uma prova que não se preocupa com o perfil do aluno, com as nuances do que está acontecendo aqui dentro. Avaliam-se as pessoas de uma maneira única, como se todo mundo fosse uma máquina e pudesse reproduzir as mesmas coisas. Os governantes, as pessoas que estão por trás disso, querendo gerar índices, acredito que eles tenham uma ótima ideia disso, até porque geram-se uns índices maravilhosos, que a gente sabe que não existem. Os alunos, os pais, pessoas que não estão nesse meio, eu acho que, por uma questão de falta de informação, muitos deles acreditam piamente no índice que é mostrado. O conhecimento é construído através de um raciocínio, quando você faz uma prova de múltipla escolha você só quer saber a resposta final, e às vezes o aluno teve todo o raciocínio, todo um trabalho, errou em algum lugar, isso acontece, ou não soube explicar bem um determinado ponto, e aquele erro acabou sendo acumulado, fazendo ele chegar na resposta final que não era correta. Mas você consegue perceber que ele teve um princípio de raciocínio correto, que ele tentou. Acho que uma avaliação discursiva seria melhor.
P-C	Os professores dão pouco crédito à avaliação externa, dadas as condições em que os dados são coletados e aferidos e os objetivos que os governos têm de fato com estes resultados, para estabelecerem políticas públicas de acordo com seus interesses. Sendo dessa forma, agentes ativos fazem pressão sobre escolas por resultados, sem levar em consideração opiniões dos atores verdadeiros do processo, ou seja, os professores. O público em geral informa-se pela mídia, que veicula os resultados apurados e manipulados pelos governos, segundo seus interesses. E não conhecem a realidade do chão da escola. No geral, as avaliações externas são bem formuladas. Entretanto, se o interesse é ter resultados favoráveis, a tendência é formular avaliações muito básicas e pouco aprofundadas para não dar margem a resultados indesejados.
P-D	A avaliação externa consiste num instrumento de diagnóstico, mas não se pode deixar de lado as especificidades de cada local, de cada público. Por isso é extremamente relevante a avaliação do professor. Então, muitos profissionais não dão muito crédito à avaliação externa. Acho que quanto maior o nível de instrução das pessoas, [estas] tendem a valorizar mais a avaliação dos professores do que a externa. A avaliação externa pode ser bem elaborada, mas muitos alunos a fazem de qualquer jeito e esse não pode ser o único parâmetro.
P-E	Eu as acho importantíssimas, embora saiba que muitos colegas professores as desaprovem. A credibilidade depende do grau de conhecimento que se tenha sobre a aplicação dos exames. Quando se sabe que o resultado pode ser fraudado para que a escola obtenha uma boa classificação no Ideb, questionam-se seus fins. Mas essa é uma percepção [de] que apenas os envolvidos dispõem. Alunos, pais e grande público não conseguem ligar a Prova Brasil que fazem na escola, por exemplo, aos índices que veem no telejornal. Acho que são bons testes.
P-F	Muitos professores não veem a avaliação externa como uma ferramenta importante para rever seu trabalho com os alunos. Penso totalmente o contrário, acredito que a avaliação externa se torna um termômetro pedagógico da aprendizagem do aluno e das práticas do professor. No entanto, há fatores como a falta de estruturas física e pedagógica, e problemas sociais que influenciam negativamente nesses indicadores. Acredito que uma pessoa que não vivencia o dia a dia da educação pública acredita na avaliação externa bem mais do que aquela que possui certa intimidade com esse meio. Acho que a avaliação externa deve melhorar muito, principalmente em Matemática. O aluno terá muito mais facilidade para resolver problemas que fazem parte de sua realidade.
P-G	Acho que os professores não dão crédito a essas provas, apenas rezam para que seus alunos se saiam bem, para que depois nem a escola nem eles sejam culpados por um mau desempenho dos alunos em provas descontextualizadas, desorganizadas e mal elaboradas. O público externo não sabe das condições em que essas provas são aplicadas aos alunos, por isso acredita. Tanto que os resultados viram estatísticas e gráficos e vão parar no jornal e em planos de governo irresponsáveis. Só quem conhece bem os alunos são seus professores. Mesmo assim, uma avaliação externa seria possível se mais bem pensada.
P-H	Muitos professores acham que não funcionam. Mas o público em geral sim, pois são números, dados concretos. Porém, a educação não pode ter um único instrumento para avaliação. Temos que vê-la de uma forma holística. A avaliação externa é bem formulada, porém é impossível saber a fundo a capacidade do aluno através de uma prova.
P-I	Os professores dão pouco crédito à avaliação externa. O público não considera a avaliação externa séria e dá mais valor à avaliação realizada pelo professor. Acho a avaliação externa muito fácil.

Figura 29

Dimensão "Credibilidade": professores

Relativamente ao público em geral, os professores consideram que as avaliações externas seriam mais credíveis devido à falta de informação, tanto do contexto escolar quanto do processamento das avaliações, o que é ratificado nos dados quantitativos apresentados na Tabela 6.

Sobre a qualidade das questões, o grupo de professores participantes ficou dividido, pois alguns entrevistados afirmam que não consideram as provas com qualidade, por serem fáceis e terem apenas questões de múltipla escolha. Alguns acreditam que as provas precisam melhorar, mas há uma qualidade mínima. Por fim, também houve aqueles que julgaram que a avaliação externa tem qualidade, sem ponderar pontos negativos do instrumento. No entanto, no questionário, os professores apontaram uma falta de confiança no sistema de avaliação.

Tabela 6
Credibilidade

Questões	Concordo	Discordo
Para a maioria dos docentes, as avaliações externas são credíveis	21%	79%
Para a maioria dos cidadãos, as avaliações externas são consideradas mais credíveis do que para os atores educacionais	61,3%	38,7%
De modo geral, as avaliações externas são construídas com qualidade, permitindo, no que se refere às questões formuladas, conhecer bem o que os alunos sabem e são capazes de fazer	33,1%	66,9%

Com base na percepção da credibilidade dos professores com relação à avaliação externa, é importante entender qual é a utilidade dessa avaliação, para verificar se ela está sendo útil para os professores e outros intervenientes. Os professores participantes consideram que a avaliação externa possui utilidade, tanto para o sistema, ao gerar índices para o governo, subsidiar políticas públicas, contribuir para a tomada de decisão para investimentos nas escolas, como também para a escola, pois seria útil para aprimorar o trabalho docente, gerar reflexões de professores e diretores e o monitoramento da evolução do aluno.

No entanto, cabe destacar os verbos na condicional em algumas falas como “Seria útil sendo um dos parâmetros...” (P-D), “Creio que deveria servir para investimentos em escolas...” (P-G), ou seja, a avaliação externa só pode ser útil se houver uma condição,

que é expressa em seguida por P-G: “se os índices pudessem traduzir minimamente a realidade, o que não fazem”. Isso pode ocorrer por uma inconsistência entre a avaliação e a realidade escolar, pelo fato de se maquiarem os resultados – o que provoca falta de confiança na avaliação externa –, ou pelos mecanismos de processamento. Com isso, fica evidente que os professores consideram não apenas a função política da avaliação externa, mas também a pedagógica, porém não veem formas de esta função ser considerada devido aos problemas levantados.

Outro ponto é a falta de conhecimento sobre o processo, expresso pelo professor P-A, que entende haver outros critérios, porém não os conhece. Afirma, por conseguinte, que a utilidade estaria vinculada a uma avaliação que tenha “critérios que incluam todos os aspectos relacionados à formação humana escolar.”, o que aponta para uma falta de conhecimento dos limites dessa avaliação. É perceptível que o professor P-A entende que a avaliação externa deveria ser capaz de dar conta de todas as dimensões avaliativas possíveis no contexto escolar, o que não cabe à avaliação externa. Para que a avaliação externa tenha utilidade para seus usuários, eles devem compreender os meandros da avaliação, as suas limitações e as alternativas de seus usos (Alavarse et al., 2017; Nevo, 1997).

Como se pode ser visto na Figura 30, os professores consideram que a avaliação externa pode ser útil, porém pontuam que, com o atual desenho e processos, ela não apresenta utilidade para os professores ou para outros intervenientes. Também demonstram pouco conhecimento sobre suas potencialidades e limitações, ou seja, o que ela efetivamente pode fazer e qual é o seu limite com relação ao contexto escolar.

	Utilidade
P-A	Não faço ideia dos motivos que estimulam uma avaliação externa nesses moldes. Mas repare que as avaliações externas não são compostas só de provas. Há outros critérios. Quanto a esses outros critérios, eu nem sei quais são. No entanto, acho importante haver uma avaliação externa ampla das escolas, uma avaliação que seja feita a partir de critérios que incluam todos os aspectos relacionados à formação humana escolar.
P-B	Meramente para gerar índice para o governo, para mais nada que não seja isso.
P-C	Mostrar o grau de alfabetização, capacidade de interpretação e nível de conhecimento da própria língua dos alunos em âmbito geral.
P-D	Diagnóstico do nível de aprendizagem dos alunos. Seria útil sendo um dos parâmetros a serem analisados para dar base a novas políticas públicas na área de educação.
P-E	Avaliar resultados. Para mim, a utilidade principal é permitir que eu aprimore meu trabalho.
P-F	A avaliação externa proporciona uma visão horizontal da qualidade do ensino. Através dos indicadores que ela nos fornece, é possível criar ações e estratégias nos casos de baixo desempenho, tanto para melhorar a aprendizagem do aluno quanto como elemento de reflexão para professores e diretores.
P-G	Creio que deveria servir para investimentos em escolas que precisassem melhorar os índices, isso se os índices pudessem traduzir minimamente a realidade, o que não fazem.
P-H	Acho que para ver o nível intelectual de um determinado grupo: país, estado, município... e tentar melhorar, corrigir os erros.
P-I	Observar a evolução gradativa do aluno.

Figura 30

Dimensão "Utilidade": professores

De acordo com os dados quantitativos, como mostra a Tabela 7, os professores não consideram que a avaliação externa seja útil para que os atores educacionais ou o público em geral conheçam a qualidade do ensino, o que destaca a baixa credibilidade com relação à função pedagógica somativa e no que tange ao conhecimento da população sobre esses dados. Relativamente às políticas públicas, o quadro ficou equilibrado. Considerando à luz das entrevistas, pode-se afirmar que há um entendimento de função política formativa, porém, com o quadro de limitações da avaliação externa, há uma tendência de se desconsiderar um real papel dessa função.

Tabela 7
Utilidade

Questões	Concordo	Discordo
As avaliações externas são úteis para que os atores educativos conheçam a qualidade da educação que estão proporcionando	36,3%	63,7%
As avaliações externas são úteis para que a opinião pública em geral conheça a qualidade da educação que está sendo proporcionada aos seus cidadãos	31,5%	68,7%
As avaliações externas são úteis para corrigir e melhorar as políticas públicas de educação	49,2%	50,8%

Assim, os professores não negam a importância e a utilidade da avaliação externa, contudo, evidenciam as limitações intrínsecas a ela. Por isso, cabe ressaltar a emergente necessidade de se pensar a projeção da avaliação externa de forma cuidadosa, de modo que tudo que é feito, do começo ao fim, afetará o seu uso (Patton, 2005). Deve-se, também, reconhecer suas limitações e verificar o quanto elas impactam na utilização dos resultados da avaliação pelos diferentes intervenientes (Almeida & Freitas, 2013; Bauer, 2012; Gatti, 2012, 2014).

A baixa credibilidade atribuída à avaliação externa pelos professores participantes resulta numa baixa utilização. Isso se deve ao fato de que, ao entender que a utilidade é vista com base na necessidade em relação aos resultados da avaliação de um determinado grupo de interessados (Patton, 2005), se esses resultados não são críveis, não há motivo para usá-los, buscando-se, assim, outras formas de informação e apoio pedagógico.

Os professores consideram que, para os seus pares, não há utilização dos resultados das avaliações externas. Mesmo os que responderam que dão algum tipo de utilização, afirmaram, em seguida, que isso não corresponde ao movimento da maioria. Alguns exemplos de utilização foram citados, como: replanejar o trabalho pedagógico, ações com objetivo de melhorar os resultados, análise no final do ano letivo, observar as necessidades dos alunos. De forma muito tímida, contudo, houve um movimento no sentido de reconhecer sua pontualidade, visto que não há confiança para fundamentar ações com base nesses resultados.

Alguns professores destacaram a falta de relação entre os resultados e a realidade da sala de aula, o que seria um impeditivo para a utilização, como fica evidente em “. . .

acho que os professores apenas lamentam, porque [o resultado] é normalmente tão comprometido que não dá pra fazer nada com ele.” (P-G).

Para os gestores educacionais, os professores consideram três possibilidades. A primeira consiste em ter ações de cobranças direcionadas aos professores para melhores resultados, pois são cobrados por instâncias superiores, o que impulsiona uma pressão verticalizada (Freitas, 2007; Moe, 2003). O segundo movimento é a ação voltada apenas para a divulgação, ou seja, disponibilizam os resultados em suas escolas, porém sem ação concreta com eles. Por fim, o de análise e discussão dos resultados, para o desenvolvimento de estratégias com o objetivo de alcançar melhores resultados, uma ação que busca fazer a mediação entre os resultados da avaliação externa e os professores, principalmente em reflexões nas reuniões pedagógicas. Em pesquisa com gestores escolares do estado do Rio de Janeiro, Rosistolato e Viana (2014) indicam que os gestores compreendem que a avaliação externa da aprendizagem dos alunos já está inserida no cotidiano escolar, não ignorando seus reflexos em suas práticas profissionais. No estudo, mesmo os que questionaram a avaliação externa, destacaram que é preciso utilizar os dados da avaliação, uma vez que há a necessidade de mediar os diálogos entre as unidades escolares que administram e a gestão central.

No que se refere aos decisores políticos, os professores consideram que a utilização maior é para campanha política, chegando a maquiagem os resultados com o objetivo de ganhar popularidade, o que aparece recorrentemente na fala dos professores. Esse fato dialoga com a baixa credibilidade da avaliação externa apresentada, já que os resultados não seriam reais.

Tanto com relação aos gestores escolares quanto aos decisores políticos, alguns professores colocaram que não tinham ideia de como são utilizados os resultados das avaliações externas por eles, mostrando falta de conhecimento e envolvimento com a avaliação externa ou falta de diálogo entre as diferentes instâncias.

Os professores entendem que a avaliação externa pode ser um recurso para regular e reorientar o trabalho docente, mas não pode ser considerada a única ferramenta, devendo, assim, ser conjugada com outras ações. Um exemplo trazido pelos professores nesse sentido foi identificar erros e acertos para corrigir falhas no trabalho pedagógico. Porém, dois professores foram contrários à avaliação externa regular e a reorientar o trabalho docente. O professor P-A assinalou que isso não pode acontecer, considerando ser até mesmo antiético. Já o professor P-G afirmou que as avaliações externas não podem ser utilizadas para esse fim por não traduzirem a realidade.

Sobre o currículo, os professores afirmaram que, se há um controle sobre o cumprimento do currículo, é algo muito básico, elementar. Isso aponta para a construção, no estado do Rio de Janeiro, do currículo mínimo, cuja finalidade foi definida como

Orientar, de forma clara e objetiva, os itens que não podem faltar no processo de ensino-aprendizagem, em cada disciplina, ano de escolaridade e bimestre. Com isso pode-se garantir uma essência básica comum a todos e que esteja alinhada com as atuais necessidades de ensino, identificadas não apenas nas legislações vigentes, Diretrizes e Parâmetros Curriculares nacionais, mas também nas matrizes de referência dos principais exames nacionais e estaduais (Seeduc, 2013, p. 2).

Assim, o currículo mínimo tem relação direta com as matrizes de referência das avaliações externas geridas pela federação e pelo estado. Esse é considerado um uso, a partir do processo ou desenho da avaliação, pela instância de gestão educacional central, ou seja, secretaria educacional do estado ou do município.

A utilização dos resultados para a elaboração de políticas públicas foi considerada algo possível, mas que não acontece na prática. São poucas as pesquisas que objetivam estudar essa relação. Pode-se concluir, com base em estudos que tomam como referência o PISA, que a relação entre políticas públicas e avaliação externa depende de fatores para além do propósito da avaliação, uma vez que há governos que utilizam os resultados para formular suas políticas educacionais, e outros não (c.f. Araújo & Tenório, 2017; Bieber & Martens, 2011; Lingard, 2016; Michel, 2017; Niemann, Martens & Teltemann, 2017).

Na Figura 31, pode-se verificar os pontos de vista defendidos pelos professores participantes na pesquisa. De acordo com esse grupo, há uma baixa utilização pelos professores dos resultados da avaliação, muito em decorrência da falta de relação entre os resultados e a realidade da sala de aula, o que seria um impeditivo para a utilização. Com relação aos gestores, podem ocorrer três movimentos: cobrança, ações de divulgação e discussão dos resultados. No que tange aos decisores políticos, os professores consideram que há mais utilização voltada para campanha política.

	Utilização
P-A	Não sei qual a utilização que os professores e os gestores escolares dão à avaliação externa, eu, por exemplo, nunca dei nenhuma utilidade aos resultados das avaliações. Os resultados da avaliação externa servem como <i>ranking</i> e para que os pais conheçam o desempenho das escolas. O resultado das provas feitas diretamente pelos alunos serve para medir o grau de letramento e raciocínio matemático dos alunos, acredito que os decisores políticos utilizam os resultados para terem estes parâmetros. O trabalho pedagógico nunca pode ser definido estruturado ou pensado a partir da avaliação externa. Acredito até que seja antiético fazer isso. A prova aplicada aos alunos trazia algumas competências e habilidades relacionadas com o currículo mínimo. No entanto, pelo menos a de Língua Portuguesa, era muito elementar. Não entendo como essas provas poderiam ser um regulador do cumprimento do currículo mínimo. Acredito que a avaliação como um todo, aquela que inclui todos os critérios avaliativos, tenha bastante utilidade para se formularem políticas públicas.
P-B	A melhor coisa que a gente pode filtrar disso aí é pegar os índices que são dados, assumir obviamente que eles não estão de acordo com nossa realidade, ou seja, ignorar todos os fatores que a prova ignora, e através disso tentar formas de obter um índice melhor. Não sei como os gestores utilizam estes resultados, estou na rede desde 2015 e nunca vi nenhum gestor usando. Para o governo, é basicamente campanha, e quando eu digo campanha eu falo, na maioria das vezes, de enganação. É o político que muda um parâmetrozinho, por exemplo. Não tenho dúvidas de que podem regular e reorientar o trabalho docente, desde que as avaliações sejam aplicadas levando em consideração todas as nuances daquela turma, da escola já seria uma coisa muito geral, e pontualmente de cada aluno, o que seria o ideal. A utilização para políticas públicas não ocorre de maneira nenhuma. Eu acho até que deveriam, talvez a ideia principal por trás dela seja essa.
P-C	Da minha parte, utilizo a avaliação externa para replanejar meu trabalho pedagógico. Já as equipes gestoras das escolas se sentem pressionadas. Quase sempre pressionam os docentes também por resultados, que muitas vezes não correspondem à realidade, para agradarem às secretarias de educação, que por sua vez querem agradar os governantes do momento. Os políticos as utilizam para realizarem políticas públicas e ganharem assim popularidade e permanecerem no poder. Sim, a avaliação externa pode reorientar e regular o trabalho docente, sempre. Toda vez que um determinado item tem uma avaliação ruim pela maioria dos alunos, buscamos mecanismos para voltarmos a ele e reforçarmos, assim, sua compreensão e fixação pelos alunos. Sempre que uma avaliação é aplicada, nela constam os conteúdos mais básicos que devem ser alcançados por todos os alunos brasileiros, fazendo com que todos os professores da rede pública sejam forçados a cumprir esse currículo mínimo. Como são niveladas por baixo, essas avaliações acabam por mascarar a realidade e não obrigam os governos a investirem mais nas escolas. Porque estes atingem seus objetivos e fazem parecer que está tudo bem, assim é que se dá a relação da avaliação externa com políticas públicas.
P-D	Os professores utilizam a avaliação externa para melhorar os resultados, mas muitos não se importam com essas avaliações. Os gestores procuram melhorar os resultados. Acredito que os decisores políticos as utilizem para definir políticas públicas para a área de educação. O trabalho docente pode ser sim muito reorientado e regulado. Principalmente se o professor tiver acesso aos erros e acertos dos alunos. Por exemplo, se muitos alunos erram uma determinada questão de interpretação, o professor pode perceber se o problema é de figuração, implícitos, inferência etc. O governo tenta utilizar a avaliação externa para padronizar o cumprimento do currículo, é uma forma de formulação de política pública.
P-E	Eu faço uma avaliação dos resultados ano a ano, mas acho que muitos colegas professores nem querem saber dos resultados. Muitos não gostam de ser avaliados. Os gestores da escola imprimem e publicam num mural. Nunca soube que as utilizassem para desenvolver medidas de combate ao mau desempenho escolar. Já os políticos, não faço ideia de como eles utilizam a avaliação externa. Sobre o trabalho docente, a avaliação externa não pode ser a única ferramenta, mas ajuda bastante quando quero saber como meu aluno está em relação aos alunos de outras escolas, municípios e estados. Controlar o currículo, aqui no Brasil? Ninguém fiscaliza nada, em contrapartida ninguém se responsabiliza por nada também. Por isso, também não vejo ligação entre avaliação externa e as políticas públicas.
P-F	Os professores não dão nenhuma utilização à avaliação externa. Os gestores geralmente cobram quando há pedido de explicações por chefia superior, mas não buscam junto com o professor ações que possam melhorar os resultados. Pelo menos é o que eu vivo com frequência. Acho que os decisores políticos utilizam apenas para disputa de poder. Também acredito que o controle do currículo se dá somente para disputa de poder, tendo em vista que currículo vai muito além de habilidades e competências. Se as políticas públicas vissem a educação pelo aluno real e não ideal, acredito que teríamos alunos com melhores desempenhos. E nas políticas públicas, as estratégias planejadas devem ser executadas. A meritocracia por si só não resolve as mazelas da educação.
P-G	Com o resultado mesmo das avaliações, acho que os professores apenas lamentam, porque é normalmente tão comprometido que não dá pra fazer nada com ele. Já tive aluno que não dominava a escrita (dito analfabeto) e uma prova objetiva o conduziu ao projeto Aceleração, porque avaliou o aluno como bom leitor. Ele mesmo me confessou que deu sorte, marcou qualquer coisa. Complicado usar pedagogicamente esses resultados que traduzem um imaginário e não a realidade. Os gestores e os decisores políticos conversam sobre o resultado nas reuniões e tentam elaborar estratégias no Plano Político-Pedagógico da escola para tentar alcançar resultados melhores. Acredito que, se as avaliações externas não traduzem a realidade, não podem servir para regular ou reorientar o trabalho do professor. As avaliações externas são utilizadas pelo governo para desvalorizar moralmente a educação pública e o professor desse lugar. Fazem uma prova ruim, com aplicação desorganizada e, portanto, resultados comprometidos, mas os utilizam mesmo assim para avaliar negativamente o trabalho das escolas públicas em sua maioria. Com relação às políticas públicas, acredito que sim, pois normalmente um empresário usa os dados comprometidos e lança aberrações como meritocracia etc. na educação pública.
P-H	A partir da avaliação externa, os professores observam as necessidades dos alunos e tentam trabalhar esses pontos. Nas escolas, os gestores geralmente buscam identificar as falhas e corrigi-las. Os políticos, por sua vez, utilizam para cobrar resultados. Para o trabalho docente, acredito que pode ser um dos sinalizadores para atingirmos as metas, mas não o único. Sobre o controle do currículo, penso que sim, é uma das formas de controlar. Acredito que os dados das avaliações externas deveriam ser utilizados para formular políticas públicas, funciona como um sinalizador, mas não é o que acontece na prática. Bom seria se investissem mais em educação.
P-I	As avaliações externas viram somente mais uma nota para os professores, os gestores as utilizam apenas para cobrar metas dos professores e os políticos como meio de propaganda política. Com relação a regular e reorientar o trabalho do professor, acho que não podem ser o instrumento principal nem o único, porque não leva em conta o processo de aprendizagem do aluno, mas apenas o raciocínio final. A grande serventia é percebermos a aprendizagem de determinados descritores. As avaliações externas são utilizadas como instrumento de meritocracia, muitas vezes. Não são utilizadas para formular políticas públicas de educação. Viraram apenas instrumento de controle.

Figura 31

Dimensão "Utilização": professores

Com base no questionário evidenciado na Tabela 8, observa-se que os professores estão divididos com relação à utilização da avaliação externa, o que pode ser devido a uma falta de conhecimento sobre seu uso ou por ações diferentes no âmbito das escolas, ou seja, escolas que realmente utilizam os resultados da avaliação externa e aquelas que não o fazem. Com relação à utilização governamental, os professores concordaram que o governo utiliza os resultados das avaliações externas com o intuito de avaliar escolas e professores, e também como uma forma de controle do currículo.

Tabela 8
Utilização

Questões	Concordo	Discordo
As equipes gestoras das escolas utilizam os resultados das avaliações externas para melhorarem a organização e o funcionamento pedagógico das suas instituições	54%	46%
Os resultados das avaliações externas são utilizados pelos docentes para melhorarem a qualidade do seu ensino	41,9%	58,1%
Os resultados das avaliações externas são utilizados pelo Governo para avaliar as escolas e/ou os professores	79%	21%
Os resultados das avaliações externas são utilizados pelas escolas para regular e reorientar o seu trabalho pedagógico	52,4%	47,3%
Os governos utilizam as avaliações externas como forma de controlar se os professores cumprem o que está previsto no currículo	69,4%	30,7%
Os resultados das avaliações externas são utilizados para formular políticas públicas de educação	47,6%	52,4%

Os professores entrevistados, na sua maioria, consideram que a categoria não está satisfeita com as avaliações, e algumas justificativas dizem respeito à manipulação de resultados pelo governo, ou seja, devido à falta de credibilidade da avaliação externa, por ações fraudulentas realizadas pelos políticos com os resultados da avaliação. Há críticas também à política de premiação, com critérios considerados injustos pelos professores, como pode ser visto na Figura 32.

Faz-se importante ressaltar o sentimento negativo do professor P-B com relação à avaliação externa, afirmando que se sente “menosprezado”, que o seu “trabalho não serve para nada”. Ball (2003), referindo-se à performatividade de professores em ambientes em que se trabalha em resposta a metas, indicadores e avaliações, entende que

tal fato representa uma esquizofrenia estrutural e individual de valores e propósitos, e uma inautenticidade e falta de sentido. Pode-se destacar um sentimento de frustração e desvalorização do trabalho como docente.

Com relação aos decisores políticos, os professores consideram que há uma satisfação maior, mas não seria porque a avaliação gera ações para uma melhoria na qualidade da educação, mas porque podem “maquiar os dados” ou “ganhar dinheiro”. Consideram que a avaliação, portanto, é um recurso de políticas fraudulentas.

Satisfação dos diferentes intervenientes	
P-A	Os professores não ficaram muito satisfeitos não. Inclusive porque a nota geral da avaliação, que incluía todos os critérios, fazia com que os docentes de algumas escolas recebessem 14º salário, enquanto outros de outras escolas, não. Como a gente não tem como interferir em todos os critérios (por exemplo, a escola ter quadra de esporte), nós achávamos injusto os critérios da premiação. Eu pelo menos achava. E olha que os professores da minha escola, inclusive eu, ganhamos todos os anos enquanto vigorou a premiação. Os outros intervenientes eu não saberia falar.
P-B	Eu acho que o único grupo que está insatisfeito mesmo é o dos docentes. Os políticos estão interessados nos índices e se eles precisarem maquiar, eles vão maquiar, nem que seja jogando um parâmetro lá em baixo. As pessoas em geral, se escutam que elas são lindas e maravilhosas, e que está todo mundo inteligente, todo mundo acha maravilhoso e ninguém se questiona sobre. Esse índice claramente não representa o que está acontecendo. Isso gera uma sensação de desconforto, principalmente porque a gente se sente menosprezado, na questão do nosso trabalho, a gente acha que o nosso trabalho não serve para nada.
P-C	Acredito que não, os professores mostram-se insatisfeitos, pois o governo manipula os dados segundo os seus interesses.
P-D	Acredito que os diferentes intervenientes estejam satisfeitos com a avaliação externa, pois ela auxilia na melhoria da qualidade da educação.
P-E	Sim, acredito que há um grau de satisfação de todos.
P-F	Jamais, não há satisfação.
P-G	O governo sim, pois ganha muito dinheiro com isso. Os professores não, pois não dão crédito a essa avaliação externa.
P-H	Acho que sobre a satisfação, fica dividido, temos os dois lados, um que acredita e o outro que não.
P-I	Não.

Figura 32

Dimensão "Satisfação dos diferentes intervenientes": professores

Os professores, no questionário, também se mostraram insatisfeitos com a avaliação externa, tanto com o número de avaliações propostas por ano quanto com o atual sistema de avaliação, como destaca a Tabela 9.

Tabela 9
Satisfação dos diferentes intervenientes

Questões	Concordo	Discordo
A maioria dos atores educativos está satisfeita com o número anual de avaliações externas que são administradas aos alunos.	23,4%	76,6%
De modo geral, os atores educativos estão satisfeitos com o atual sistema de avaliações externas.	10,5%	89,5%

Com base na análise no quadro de avaliação externa do Rio de Janeiro, fica claro o número excessivo de avaliações que, em algum momento, foram aplicadas. Além do sistema nacional de avaliação externa, há o sistema estadual, totalizando três tipos de avaliações diferentes ao longo de oito anos, e uma avaliação do município. Para além dessa alta demanda, há todos os outros pontos considerados pelos professores que embasam tamanha insatisfação, sendo os principais: as limitações e os problemas dos mecanismos de processamento, as políticas que sustentam as avaliações e os efeitos negativos que acarretam.

Esta primeira parte teve como objetivo trazer a percepção dos professores da educação básica sobre a avaliação externa da aprendizagem dos alunos. A seguir, a discussão aborda a qualidade da educação e a sua relação com a avaliação externa.

Qualidade da educação.

Relativamente ao processo de ensino e aprendizagem, os professores consideram que há uma relação, que é, todavia, permeada de questões que a dificultam. A falta de confiança na avaliação foi considerada, no sentido de não ter seriedade (P-B); de não estar em compasso com o trabalho dos professores e o desenvolvimento dos alunos (P-F); de ser uma relação pautada em pressão, produzindo efeito de preparação para o teste (P-G); e de o governo apresentar dificuldades em utilizar os resultados da avaliação para isso (P-E).

A relação da avaliação externa com o processo de ensino e aprendizagem também foi considerada com base no currículo, com os conteúdos ensinados no bimestre, como destacado pelo professor P-A, relacionando com as provas estaduais. Assim, a avaliação

externa possibilitaria conhecer mais sobre esse processo, conhecer o que está sendo ensinado e aprendido nas escolas. O Saerj e o Saerjinho foram desenhados com esse objetivo, ser um “instrumento útil no apoio às discussões pedagógicas, na elaboração de projetos educativos, no planejamento das aulas, na análise do material didático e na reflexão sobre a prática educativa.” (Rio de Janeiro, 2010, p. 5), o que o professor P-A conseguiu identificar.

Quanto ao fato de a avaliação externa permitir conhecer o que os alunos sabem e são capazes de fazer, os professores entrevistados apresentaram certa resistência, afirmando que é difícil fazê-lo com questões objetivas (P-B); o resultado é comprometido, pois não há envolvimento do aluno na realização da avaliação (P-C); não se pode generalizar tal afirmação (P-D); somente se houver acesso aos indicadores (P-F). Entretanto, mostram que haveria a possibilidade de ocorrer, desde que atendessem alguns critérios no desenho da avaliação e da comunicação dos resultados.

Desta forma, como se evidencia na Figura 33, pode-se concluir que os professores consideram que a avaliação externa pode ter função formativa, porém muitas questões acabam por criar obstáculos, vindo a apagá-la.

Avaliação externa e qualidade do processo de ensino e aprendizagem	
P-A	Em relação ao Saerj e ao Saerjinho, o que estava relacionado eram os conteúdos de cada bimestre, pois era uma avaliação com base nos conteúdos ensinados bimestralmente. Eu nunca avaliei as competências e habilidades dos alunos a partir das avaliações externas. Por isso, não posso responder essa pergunta.
P-B	Atualmente eu diria que de maneira nenhuma as avaliações externas estão relacionadas com o processo de ensino e aprendizagem, acho que elas deveriam. Se fossem índices feitos com um pouco mais de seriedade, seria um ótimo parâmetro para a gente avaliar o quanto estamos conseguindo caminhar, o quanto a gente de repente tem que mudar o nosso jeito de ensinar, a nossa forma de tocar aqueles alunos, mas não acontece isso, simplesmente não acontece. Conhecer o que os alunos sabem apenas com questões objetivas é difícil. Uma avaliação discursiva, pontuando até que momento foi o raciocínio do aluno, seria uma coisa mais interessante.
P-C	A qualidade do ensino vem caindo e se deteriorando a cada ano. Cada vez cobra-se menos dos alunos, com a desculpa de incluir a todos, e isso reflete no ensino médio e superior. Isso se reflete na avaliação externa, que cobra o básico. Para conhecer o que os alunos saberiam fazer, eles deveriam fazer as avaliações sozinhos. Em minha experiência, observo pouco comprometimento dos alunos com essas avaliações, muitos chutam as respostas para acabar mais rápido. Outros não leem os enunciados com atenção. Não estudam antes das avaliações e levam tudo na brincadeira, por acreditarem que é fácil passar de ano, por não ter exemplo em casa ou quem os faça valorizar a escola. Não acreditam de verdade numa perspectiva de vida através do estudo.
P-D	Em parte dá para relacionar a avaliação externa com o processo de ensino e aprendizagem. Há muitos alunos que estão em determinada série sem ter proficiência para estar naquela série. Há muitos alunos analfabetos funcionais no ensino fundamental II. Logo, sua aprendizagem dos conteúdos da série será prejudicada. Não se pode medir a qualidade do ensino de um professor que dá aula para uma turma com alunos com essas características. A avaliação externa é apenas um parâmetro. Não se pode generalizar os níveis de proficiência e aprendizagem dos alunos apenas pela avaliação externa.
P-E	Acho que elas influenciam no processo de ensino e aprendizagem, mas não sei se o governo sabe tirar o melhor partido dos resultados que obtém. Bem, conhecer o que os alunos sabem as avaliações internas também permitem, o diferencial das avaliações externas é que avaliam sobretudo os docentes.
P-F	Muito difícil relacionar a avaliação externa com o ensino e a aprendizagem, pois existem avaliações que não têm o mínimo de compasso com o professor e com o aluno. Somente conseguiremos conhecer o que os alunos sabem à medida que tivermos acesso aos indicadores dessa avaliação, pois nem sempre chegam em nossas mãos.
P-G	Vejo que a avaliação externa está pouco relacionada com o ensino do professor e a aprendizagem dos alunos. Na realidade, só no que tange à preparação para aquele momento prova. Com relação a saber sobre os alunos, penso que em pouca medida. Ela não avalia inteligências múltiplas, por exemplo.
P-H	Não soube responder.
P-I	Quando os resultados são calmamente analisados, há uma relação entre avaliação externa e o processo de ensino e aprendizagem. Da mesma forma, só conseguimos verificar o que os alunos conseguem fazer pela avaliação externa observando as porcentagens de erros e acertos, pois percebemos a evolução da turma.

Figura 33

Dimensão “Avaliação externa e a qualidade do processo de ensino e aprendizagem”: professores

Fundamentado nos dados quantitativos, com base na Tabela 10, pode-se concluir que os professores não consideram que a avaliação externa apresenta relação com o processo de ensino e aprendizagem. Contudo, ao realizar uma análise a partir das

entrevistas, ficou evidenciado que os professores concordam com a função formativa da avaliação externa, porém apresentam certa resistência em decorrência de suas limitações.

Tabela 10

Avaliação externa e a qualidade do processo de ensino e aprendizagem

Questões	Concordo	Discordo
As avaliações externas geram reflexões de natureza pedagógica que contribuem para melhorar a qualidade do ensino e das aprendizagens	42,74%	57,30%
As avaliações externas dão indicações relevantes acerca do que é importante aprender e ensinar	38,70%	61,30%
As avaliações externas são o meio mais eficaz para melhorar o ensino dos professores e as aprendizagens dos alunos	15,30%	84,70%
As avaliações externas medem com rigor o que os alunos sabem e são capazes de fazer	12,10%	87,9%

Os professores têm em conta que a avaliação externa influencia as suas práticas de ensino. Isso ocorre, segundo eles, quando realizam reflexões a partir dos resultados, acabando, assim, por reformular suas práticas se necessário. Algumas falas marcaram um caráter de uma ação do próprio professor nesse movimento de análise dos resultados – “Basta o professor perceber onde os alunos estão apresentando dificuldades, dar uma remodelada nas aulas e ver se os resultados melhoram, costumam melhorar.” (P-E).

No entanto, é oportuno frisar que “dar uma remodelada na aula” não mostra necessariamente uma mudança na prática de ensino, podendo ser apenas mudança em ações em sala de aula, como fazer revisão e aplicar mais atividades de um determinado conteúdo cobrado no teste ou fazer simulados. Isso vai ao encontro da teoria defendida por Luxia (2007) e Firestone et al. (1998) de que a avaliação externa estaria promovendo algumas mudanças, no entanto, não apresenta sucesso nas estratégias de instruções básicas, ou seja, na prática docente.

Outra fala marca a influência da avaliação externa na prática pedagógica a partir da imposição da utilização dos resultados. Contudo, o professor considera os dados como algo não correspondente com a realidade da sala de aula, o que leva a crer que ele não os usaria de forma voluntária, a partir de uma reflexão sobre os resultados da avaliação e

uma ação consciente. Há também a resistência por utilizar informações provindas de avaliadores externos.

As avaliações externas influenciam na prática porque obrigam o professor a dar conta desses dados, ou seja, o professor é obrigado a usar dados comprometidos com sua turma, que, na realidade, ninguém melhor do que ele conhece (P-G).

Yerly (2014) realizou uma pesquisa com professores suíços para analisar os efeitos da avaliação externa sobre a prática docente e os fatores que os influenciam. Os resultados mostraram que a avaliação externa é, geralmente, considerada pelos professores entrevistados como um “mal necessário”, sendo considerada pedagogicamente útil. Há ainda aqueles que destacam os efeitos positivos, como a contribuição para a prática de avaliação somativa. Por fim, alguns evidenciam os efeitos nefastos, como ensinar para o teste e o estreitamento do currículo ensinado. Essas informações se assemelham muito aos dados levantados por esta tese.

Com relação às preocupações num contexto de avaliação, os professores trouxeram preocupações voltadas para os resultados e suas consequências, como: melhorar os resultados da aprendizagem dos alunos; grau de seriedade por parte do aluno no momento da realização da prova; como os resultados voltam para a escola; ser culpabilizado ou ter a sua reputação colocada em dúvida por resultados não satisfatórios. São preocupações, em grande parte, com relação a ações externas, provindas dos órgãos governamentais e direcionadas aos professores, apoiadas nos resultados da avaliação externa. Essas preocupações podem gerar efeitos negativos, como preparação para o teste, fraude e segregação dos alunos (Haney & Madaus, 1986; Madaus et al., 2009; Stecher, 2002). Mesmo quando afirmam ter preocupação com melhorar o rendimento do aluno na avaliação e o grau de seriedade do aluno no momento da prova, são reflexos de cobranças por resultados (Freitas, 2007; Moe, 2003).

Os professores tiveram dificuldades em responder sobre a consistência das avaliações. No entanto, os que responderam, em sua maioria, entenderam que a avaliação interna é mais consistente que a externa, tendo em vista que é feita pelo professor, sendo, assim, mais profunda. Eles consideram que, por conhecerem melhor a realidade de seus alunos e o contexto de sua escola, as suas avaliações seriam mais consistentes. Não pontuaram aspectos relacionados a questões técnicas em nenhuma das duas avaliações. Dessa maneira, o que importa para o professor na hora de avaliar é o conhecimento sobre o objeto que se deseja avaliar.

Tal fato está relacionado a certo grau de desconfiança por parte dos professores com relação a avaliadores externos. Isso não provém de um aspecto técnico, mas sim de o professor entender que os avaliadores externos, por não terem relação com o contexto escolar, não estariam capacitados para realizar a avaliação, pois desconhecem a realidade de uma sala de aula.

Eu tendo a achar que a avaliação interna tem um pouco mais de valor, no sentido de que ela é preparada pelo professor da turma. Ele conhece bem os alunos, sabe até onde eles conseguiram caminhar, ele tem ideia de até aonde vai o raciocínio daquele grupo. . . (P-B).

Para esse grupo, não se considera que a avaliação externa influencia suas práticas avaliativas. No entanto, alguns pontos foram levantados, no sentido de realização de algumas ações em decorrência de uma pressão para melhores resultados: a utilização da nota da avaliação estadual para compor a nota final do aluno no bimestre, referente à obrigatoriedade do professor de utilizar a nota do Saerjinho para compor a nota bimestral do aluno (Portaria Seeduc/Sugen nº 419 de setembro de 2013); treino para o teste, utilizando itens nas provas internas, possivelmente provenientes de banco de itens disponibilizado pelo governo; redução do processo avaliativo à aplicação de prova, sem uma reflexão sobre o sistema avaliativo, seja ele interno ou externo, ou a utilização de outros meios de avaliação.

Não apareceu, entre as respostas, o conhecimento da avaliação externa como um ganho de *know-how* em avaliação, o que era o esperado pela proposta de oficinas de elaboração de itens e a oferta aos professores de um banco de itens *on-line*, para que assim pudessem pensar em suas avaliações e melhorar sua metodologia avaliativa. Com as ações apresentadas, o que se destaca é apenas a reprodução de certos processos da avaliação externa na avaliação interna.

Da mesma forma, os professores também não consideram que as avaliações externas influenciem o estudo dos alunos, seja por falta de preocupação ou atenção do aluno para com esse tipo de avaliação, seja pelo fato de o resultado não chegar até ele. Mesmo o resultado chegando, o professor P-B considera que o aluno não se reconhece nesses resultados, evidenciando, assim, que não haveria um trabalho dos professores ou gestores escolares em realizar uma reflexão junto aos alunos com relação ao desempenho da escola, da turma ou destes na avaliação externa.

Eles [alunos] não se preocupam com se preparar, e nunca vi também esse resultado chegar até eles. Até porque, se chegar, vai chegar de uma maneira muito geral, então o aluno não vai se reconhecer ali. Ele não vai ver que, de repente, o 9º ano teve só 5% dentro de uma média aceitável. Ele não vai entender que, de repente, ele faz parte dos 95%, então acho que isso não condiciona o estudo dele em nada (P-B).

Em resumo, os especialistas participantes consideraram que a avaliação externa influencia as suas práticas de ensino, no entanto não ficou claro se realmente ocorre uma mudança na prática de ensino ou apenas em algumas ações realizadas em sala de aula. Com relação aos alunos, os professores entendem que não há influência da avaliação em seus estudos, como fica evidente na Figura 34.

Avaliação externa, prática de ensino e avaliação da aprendizagem	
P-A	Acho que a avaliação externa não influencia as práticas dos professores. Imagino que os professores se preocupem com a culpabilização do professor se os alunos não atingirem o resultado esperado. A avaliação externa seria consistente se fosse feita a partir de outros moldes e se fosse uma política pública de melhoria da qualidade da educação. A avaliação interna é mais consistente, pois são feitas pelo professor. Não sei com outros professores, no meu caso somente com aproveitamento da nota das provas para compor a nota do bimestre. Não acho que as AE condicionam o estudo dos alunos.
P-B	Acredito que não influenciam as práticas docentes. Pode até ser o objetivo fundamental por trás disso, embora não seja alcançado da maneira que a gente aplica. Acho que no mundo ideal onde a coisa fosse feita da forma correta, seria muito impactante. Eu acho que a principal preocupação é com o nível de seriedade com que é feita a prova. É mais uma perda de tempo do que uma coisa que o aluno leve a sério, e por conta disso, dessa não seriedade, os índices acabam sendo diferentes do que deveria ser, eles não refletem nossa realidade. Então isso também é muito preocupante. Eu tendo a achar que a avaliação interna tem um pouco mais de valor, no sentido de que ela é preparada pelo professor da turma, ele conhece bem os alunos, eles sabe até onde ele conseguiu caminhar, ele tem ideia de até onde vai o raciocínio daquele grupo. A prova em larga escala coloca um padrão, e a gente sabe que num país muito vasto como o Brasil, uma população com mais de duzentos milhões de pessoas, isso é difícil. O norte para o professor em sua avaliação interna é o trabalho que ele tem com o grupo, não a avaliação externa. Sobre o estudo dos alunos, eu diria que na rede municipal não condiciona, porque não tem um incentivo, então é feita por fazer na maioria das vezes. Eles não se preocupam em se preparar. Nunca vi também esse resultado chegar até eles, até porque, se chegar, vai chegar de uma maneira muito geral, então o aluno não vai se reconhecer ali.
P-C	A influência da avaliação externa na prática dos professores ocorre à medida que, observando os resultados, os professores reformulem suas práticas. Acredito que a maior preocupação dos professores é em aprofundarem os conhecimentos e tornarem os alunos cidadãos atores ativos de seu destino. As internas são sempre mais profundas e menos superficiais que as avaliações externas, mas seguem ambas os currículos básicos para as séries em questão. A avaliação externa nem sempre condiciona as práticas avaliativas dos professores, pois muitos optam por seguir caminhos próprios, mais tradicionais. De quantidade de informação em detrimento da qualidade. Já no estudo dos alunos, de nenhuma forma, pois eles têm pouca responsabilidade no processo.
P-D	As AE influenciam muito pouco na prática do professor, eles ficam preocupados em melhorar os resultados da aprendizagem dos alunos, acaba por ser algo apenas em busca de resultados. Considero que ambas as avaliações são consistentes, porém cabe dizer que uma não pode excluir ou sobrepor a outra. Percebo que a avaliação externa condiciona as práticas de avaliação dos professores no sentido de eles começarem a treinar para os teste, utilizando itens em suas provas. Da mesma forma acontece com os alunos, eles acabam por direcionar o estudo para fazer a prova.
P-E	A influência da avaliação externa nas práticas docentes é total. Basta o professor perceber onde os alunos estão apresentando dificuldades, dar uma remodelada nas aulas e ver se os resultados melhoram, costumam melhorar. A maior preocupação dos professores no contexto de avaliação externa é que os alunos as façam como uma loteria, ou seja, chutando as respostas. A avaliação externa não condiciona as práticas de avaliação dos docentes, mas sim direciona para a busca de melhora de resultados. Não acredito que a avaliação externa condiciona o estudo dos alunos.
P-F	Quando tenho retorno das avaliações, sempre interpreto os indicadores e, caso veja algum resultado discrepante, reflito acerca das minhas práticas, tentando elaborar aulas mais interessantes. A maior preocupação do professor é a de ter sua reputação posta em dúvida (o que acho que nada tem a ver) e teme pela cobrança. Não tenho medo de que me cobrem desde que me ofereçam suporte para qualquer proposta. Acredito que a avaliação externa pode condicionar a avaliação do professor, mas penso que isso vai depender muito da escola, de qual aluno se quer formar. Condicionar o estudo dos alunos só se for no sentido de processo totalmente mecanizado, onde a aprendizagem não se desenvolve.
P-G	A avaliação externa influencia na prática porque obriga o professor a dar conta desses dados, ou seja, o professor é obrigado a usar dados comprometidos para sua turma, que, na realidade, ninguém melhor que ele conhece. Os professores se preocupam com o modo como os alunos vão encarar aquele momento e como esse resultado vai voltar para a escola. Na avaliação interna é onde o professor pode atuar, portanto, de maneira mais consistente que a externa. Inclusive, nas internas, o professor pode atuar de forma não tradicional, avaliando as inteligências múltiplas dos alunos e seus potenciais diversos. Acredito que de nenhuma forma influencia o estudo dos alunos. Eles sabem que não vale ponto e não ligam. Quando vale ponto, alguns estudam.
P-H	No meu caso, influenciam na minha prática, pois passo a refletir sobre os resultados. Acredito que a maior preocupação dos professores é querer preparar o aluno para fazer a avaliação externa de forma consciente e segura. Não acredito que a AE condicione a avaliação dos professores ou o estudo dos alunos. O que pode acontecer é o aluno estudar somente o que vai cair na prova.
P-I	A avaliação externa influencia pouco as práticas, pois os professores não têm esse instrumento como algo prático. A preocupação do professor é com o resultado final do aluno, se ele conseguirá um bom resultado. Não acho nenhuma das duas avaliações consistentes, porque não avaliam o processo de aprendizagem, só o resultado final. A avaliação externa condiciona a avaliação dos professores, pois esses passam a olhar a avaliação como prova, um mero instrumento e com formalidades. Não condicionam o estudo dos alunos, pois eles não se importam com essas provas.

Figura 34

Dimensão: “Avaliação externa, prática de ensino e avaliação da aprendizagem”: professores

Sobre a influência da avaliação na prática de ensino, os dados quantitativos mostram que os professores ficam divididos sobre o tema, como destaca a Tabela 11. Isso aparece no momento em que não há diferença quantitativa nas respostas sobre o condicionamento das práticas pedagógicas, a promoção dos docentes em ações destinadas ao preparo do aluno para os testes e a influência no processo de ensino.

No entanto, os professores tendem a concordar quando são apresentadas ações pedagógicas específicas sob influência da avaliação externa, ou seja, concordam que, em um contexto de avaliação externa, o ensino é, sobretudo, para a realização de testes. Além disso, consideram que há uma maior ênfase em práticas de avaliação com o objetivo de classificar em detrimento de uma avaliação para melhorar a aprendizagem, o que apareceu nos exemplos trazidos pelos professores entrevistados. Os professores discordaram da afirmativa de que as práticas avaliativas dos docentes e os estudos dos alunos são condicionados pela avaliação externa.

Tabela 11

Avaliação externa, prática de ensino e avaliação da aprendizagem

Questões	Concordo	Discordo
As avaliações externas condicionam as práticas pedagógicas dos professores	45,2%	54,8%
A maioria dos docentes promove ações específicas de ensino destinadas a preparar os seus alunos para as avaliações externas	51,6%	48,4%
As avaliações externas influenciam os processos de ensino utilizados pelos professores	58,9%	41,1%
Em geral, em contextos de avaliações externas, o ensino é orientado sobretudo para a realização de testes	74,2%	25,8%
As práticas de avaliação dos professores são sempre muito condicionadas pelas avaliações externas	39,5%	60,5%
Num contexto de avaliações externas, as práticas de avaliação para classificar tendem a sobrepor-se às práticas para melhorar as aprendizagens	67,7%	32,3%
As avaliações externas influenciam as formas como os alunos estudam	36,3%	63,7%

Mesmo entendendo que há uma função pedagógica formativa e que a avaliação apresenta influência nas práticas docentes, os professores consideraram que a apropriação

dos resultados é baixa, como mostra a Figura 35. Isso ocorre devido à tendência de ignorar os resultados, por não verem o contexto escolar nesse resultado e/ou por não conseguirem compreendê-lo. Ou, ainda, ocorre porque a apropriação se dá apenas para quem está no topo da gestão, não tendo utilização, por conseguinte, para os outros intervenientes, principalmente os atores escolares.

No que tange aos professores, a apropriação dos resultados foi considerada para melhorar a escola e a prática do professor; para definir os caminhos a serem tomados na educação, tanto no nível do aluno quanto do Estado. Esses propósitos coadunam com a função formativa da avaliação (Wiliam, 2006). Em geral, os professores consideram que têm acesso à divulgação dos resultados das avaliações externas, seja pela escola seja pela mídia, o que confirma que os resultados estão sendo disponibilizados, porém, fica evidente que devolver o resultado à escola não é o mesmo que incorporá-lo.

Os resultados continuam a não provocar os impactos esperados na qualidade da educação. A literatura afirma que tal fato pode suceder por que as tentativas de simplificação dessas informações complexas não têm apresentado êxito ou não faz parte da dinâmica escolar pautar seu planejamento em resultados de avaliação externa (Fontanive, 2013; Sousa & Oliveira, 2010; Fontanive, Elliot & Klein, 2007). No entanto, deve-se analisar também como esses resultados chegam à escola, pois o que fica evidente é que a divulgação não é sempre vista como algo positivo, ou seja, como uma devolutiva para fins pedagógicos sobre as informações recolhidas, mas sim como uma ação de divulgação do fracasso dos agentes escolares, tendo em vista que não se faz nada com a informação.

Os professores e os gestores das escolas conhecem muito os resultados das avaliações externas, pois estão nos murais da escola, nas reuniões, na TV, em todo lugar, popularizando nosso dito fracasso e não fazendo nada para mudar isso (P-G).

Dessa forma, os agentes escolares veem comprovados os problemas da qualidade da educação, são culpados por eles, porém não há uma reflexão sobre esses dados, servindo apenas como um discurso coercitivo, a partir de uma pressão (Freitas, 2007; Moe, 2003; Nevo, 1997). Um sistema utilizado não como apoio ao professor, mas sim como forma de cobrança por resultado.

Assim, concorda-se que os professores conhecem os resultados das avaliações, no entanto não os analisam, seja por falta de conhecimento para a incorporação de tais dados, seja por resistência a esse sistema avaliativo.

Ter o conhecimento eu acho que vai muito da proatividade de cada um. Em geral, não são resultados que ficam escondidos, isso sim é uma coisa que eu sempre vi fazerem, que é expor esses resultados, que fica na sala dos professores para quem quiser ver. Agora, analisar esses resultados, eu acho complicado, eu acho que a maioria deles não perde muito tempo, uma vez que eles acreditam que isso não está refletindo a realidade. Valeria mais a pena a gente fazer uma análise dos resultados das nossas provas internas (P-B).

Com relação aos gestores, os professores entendem que eles, em geral, analisam, seja por interesse nos dados, para passarem para os professores nas reuniões pedagógicas, para entenderem dados discrepantes, seja pela pressão de órgãos superiores.

Resultado da avaliação externa	
P-A	Acredito que os professores não se apropriam do resultado das avaliações externas. Quanto aos outros personagens envolvidos, imagino que tivessem a intenção de usar com fins de melhorar as escolas. No entanto, essa política não foi implementada. Os resultados das avaliações externas são divulgados, porém acredito que os professores podem conhecer as notas, mas não as analisam. Acho que, como para os gestores há um peso diferente, eles devem analisar, mas só acho, não sei.
P-B	Eu nunca vi os alunos tomarem propriedade dos resultados. Com relação aos gestores políticos, sempre com relação a índices. Os professores, pelos menos os que eu tenho contato, têm uma tendência muito forte a ignorar. Ter o conhecimento eu acho que vai muito da proatividade de cada um. Em geral, não são resultados que ficam escondidos. Uma coisa que eu sempre vi fazerem é expor esse resultado, que fica na sala dos professores para quem quiser ver. Agora analisar esses resultados, eu acho complicado, eu acho que a maioria dos professores não perde muito tempo, uma vez que eles acreditam que isso não está refletindo a realidade. Valeria mais a pena a gente fazer uma análise dos resultados das nossas provas internas. Quanto aos gestores das escolas, também nunca vi ninguém usar. Mas depende de que tipo de gestor escolar a gente está falando, se a gente tem um gestor escolar que realmente se empenha na questão da educação pública, ele tem dois caminhos a tomar. Ele vai procurar de alguma maneira incentivar os alunos a fazerem aquela prova da melhor forma possível, para buscar melhorar a instituição durante a sua gestão; ou ele vai entender que isso não reflete a realidade e não vai nem perder o tempo olhando para isso. E a gente tem o gestor escolar que faz uma gestão não pela educação ou pelo aluno, faz uma gestão por quem está acima, seja político, seja secretaria municipal. Esse gestor vai sempre tentar melhorar o índice, por mais que ele não acredite em nada do que está ali – não importa a que custo, eu tenho que melhorar. Nem que tenha que maquiagem notas, eu acho que são essas as posturas que estão mais presentes atualmente.
P-C	Os resultados das avaliações externas estão sempre disponíveis, assim a apropriação vai do interesse do professor e do gestor. Também acredito que os professores são obrigados a analisar os resultados, dada a pressão da secretaria. Os gestores são os maiores interessados, por isso sempre analisam. Principalmente porque bons resultados implicam recursos financeiros para as escolas e, muitas vezes, bonificações para todos os profissionais de educação.
P-D	Os resultados das avaliações externas são utilizados como um dos parâmetros para definir os caminhos a serem tomados na educação daquele aluno, daquela turma, daquela unidade, daquele município, estado. Porém, acredito que os professores não conhecem esses resultados e não analisam. Os gestores escolares procuram analisar e passar para os professores.
P-E	Os resultados são divulgados pelo governo, mas ninguém sabe o que fazer com os resultados, pois há uma distância abissal entre os decisores políticos e a realidade nas escolas. Poucos professores conhecem e analisam esses resultados, pois eles são muito resistentes às avaliações. Os gestores, se analisam, não tomam providências.
P-F	Vejo que os diferentes intervenientes se apropriam dos resultados somente quando estes estão no topo da lista, com resultados favoráveis. Os professores só conhecem os resultados quando os gestores transferem as informações por meio de reuniões pedagógicas, o que quase nunca acontece. E a análise só é feita se o resultado for muito discrepante. Os gestores escolares a mesma coisa, só analisam resultados discrepantes.
P-G	Os professores e os gestores das escolas conhecem muito os resultados das avaliações externas, pois estão nos murais da escola, nas reuniões, na TV, em todo lugar, popularizando nosso dito fracasso e não fazendo nada para mudar isso. Os gestores pressionam o professor para dar uma aula melhor, entupindo cada vez mais as salas de alunos.
P-H	Sempre vemos um gráfico das nossas turmas para tentarmos melhorar, por isso acredito que os professores conhecem o resultado. Na minha unidade escolar, vemos juntos os resultados e discutimos sobre o assunto. Porém, acredito que a minoria dos professores analisa os resultados. Já os gestores, penso que a maioria faz essa análise, por haver muita cobrança em cima deles.
P-I	Ninguém nem vê o resultado dessas provas. Elas são uma nota no sistema. Os professores só conhecem os resultados quando vão muito atrás, e não chegam a analisar. Os gestores da escola só analisam se baterem a meta estabelecida.

Figura 35

Dimensão: "Resultado da avaliação externa": professores

No levantamento quantitativo, Tabela 12, os professores consideraram que a maioria dos docentes tem acesso aos resultados da avaliação externa, porém não analisa esses resultados, o que foi confirmado nas entrevistas. Não há um consenso entre os professores no sentido de consultarem e conhecerem esses resultados, o que pode ser por causa de diferentes ações nas escolas, realizadas pelos gestores escolares.

Em sua maioria, os professores discordam do fato de receberem formação específica para a compreensão dos resultados. Tal fato aponta para uma falta de formação

para o conhecimento da avaliação externa, dificultando, desse modo, a compreensão dos professores para realizar uma relação entre essa avaliação e o trabalho pedagógico (Alavarse et al., 2017).

Tabela 12

Resultado da Avaliação Externa

Questões	Concordo	Discordo
A maioria dos docentes tem acesso aos resultados das avaliações externas obtidos pelos alunos da sua escola	60,5%	39,5%
A maioria dos docentes consulta os resultados obtidos pelos alunos nas avaliações externas da sua escola	54%	46%
A maioria dos docentes analisa os resultados das avaliações externas	39,5%	60,5%
Os docentes têm acesso à formação específica que lhes permite compreender os resultados dos alunos nas avaliações externas	28,2%	71,8%
A maioria dos docentes conhece bem o desempenho da sua escola nas avaliações externas	48,4%	51,6%
Os resultados das avaliações externas apresentam uma crescente melhoria ao longo dos anos	67,7%	32,3%

Estas foram as percepções apresentadas pelos professores da educação básica que foram entrevistados, com o objetivo de saber mais sobre a avaliação da aprendizagem e sua relação com a qualidade do ensino, estando dividida nas dimensões desses objetos de pesquisa. Na próxima seção, será apresentada a interpretação dos dados da pesquisa, com base na comparação da narrativa dos diferentes grupos entrevistados, dos dados quantitativos e da análise e interpretação dos documentos dos diferentes sistemas de avaliação contemplados neste trabalho.

Comparação entre os Grupos e Interpretação dos Resultados

Após descrever e analisar a narrativa dos diferentes intervenientes com relação às dimensões dos objetos de estudo desta tese, busca-se realizar uma síntese das evidências mais importantes encontradas em cada grupo participante e um cruzamento de

informações entre grupos, para que assim se possa delinear, de forma interpretativa, os objetos desta pesquisa, e responder às questões.

Os dados qualitativos e quantitativos mostram que, à avaliação externa da aprendizagem dos alunos, são atribuídas funções políticas, isto é, contribuir para políticas públicas e gerenciamento dos sistemas educacionais; bem como pedagógicas, quando é um instrumento que visa colaborar com o processo de ensino e aprendizagem. Além disso, apresenta tanto a característica somativa quanto formativa para as duas funções, ou seja: função pedagógica somativa, função pedagógica formativa, função política somativa e função política formativa.

Os três grupos de intervenientes entrevistados, ao longo da análise das diferentes dimensões, levantaram aspectos para uma dupla função da avaliação externa, como pode ser visto na Tabela 13.

Tabela 13

Comparação dos propósitos da avaliação externa entre os diferentes grupos

Propósitos	Especialistas	Decisores políticos	Professores
Avaliar/monitorar a qualidade da educação	X	X	X
Avaliar/subsidiar políticas educacionais	X	X	X
Contribuir com o processo de ensino e aprendizagem	X	X	X
Avaliar o trabalho do professor	X		X
Medir o desempenho do aluno	X		
Contribuir com a promoção da equidade		X	

No entanto, entender que a avaliação externa pode ter dupla função não garante que ela consiga desenvolvê-las. Pelo contrário, fica evidente que há uma baixa utilização no contexto brasileiro, principalmente no que compreende às ações mais formativas.

Em cada um dos três grupos, um número de participantes considerou que as avaliações externas são relevantes para apoiar o processo de ensino e aprendizagem, o que confirma sua função pedagógica formativa. No entanto, a maior parte pontuou os mecanismos de processamento das avaliações externas e as políticas que a embasam como fatores limitadores. Isso permite destacar que a não utilização da avaliação externa

pelo professor, no sentido de melhorar a sua prática pedagógica, estaria relacionada com algumas preocupações, isto é, com as ações do governo que afetam os professores, culpando-os pelos resultados e atribuindo-lhes consequências; com a falta de credibilidade da avaliação; com o fato de o *feedback* não ocorrer de forma apropriada; e com a falta de relação entre a avaliação externa e a realidade da escola.

Em decorrência desses aspectos, poderia haver apenas uma função pedagógica somativa, no momento em que os atores escolares utilizariam a avaliação externa para um balanço do estado do aluno, embasamento para a tomada de decisão para o próximo ano letivo ou o estabelecimento de uma série histórica sobre o desenvolvimento do rendimento da escola. Do mesmo modo, ocorre a não utilização da avaliação externa no contexto escolar, devido à falta de acesso aos resultados, de uma falta de compreensão sobre esses resultados ou por uma resistência.

Tanto os especialistas quanto os professores entendem que as avaliações externas são úteis para o sistema educacional e para a escola de uma forma mais pontual. No entanto, para que isso aconteça, é imperativo o desenvolvimento de outros pontos, como formação de professores, apoio aos professores no que diz respeito à leitura e à análise de dados, maior credibilidade e relação entre a avaliação externa e a realidade escolar. Os decisores políticos também assinalam que pode haver uma utilidade da avaliação para a escola, no entanto, consideram que a principal estaria voltada para o sistema educacional, sendo útil para avaliar a qualidade da educação e subsidiar políticas públicas, o que significa uma inclinação para a função política somativa e formativa.

Apesar de, no discurso dos políticos entrevistados, ocorrer uma tendência maior para a função política, essa percepção vai de encontro aos movimentos dos estados e municípios de realizarem avaliações com maior direcionamento para a utilização dos professores. Percebe-se esse aspecto na análise dos materiais dos sistemas de avaliações do estado e do município do Rio de Janeiro, que sempre pontuam a dupla função dessas provas nos materiais de divulgação, até mesmo dando grande ênfase para a função pedagógica formativa, além de realizar avaliações com resultados por alunos, o que não seria necessário numa função política.

Pode-se concluir que os grupos consideraram que a avaliação externa influencia pouco o processo de ensino e aprendizagem. Os especialistas consideram que essa relação acontece apenas quando há uma cobrança do sistema, havendo um movimento dos professores ou gestores escolares de realizar uma preparação para o teste. Os decisores políticos consideram que ocorre baixa influência, pois os professores não sabem usar,

necessitando, assim, de maior investimento em formação inicial, capacitação continuada e treinamento com base nos processos da avaliação. Por fim, os professores consideraram que ocorre mais uma mudança de ações em sala de aula do que prática de ensino, em resposta a uma cobrança por resultados.

Relativamente à relação entre as avaliações externa e interna, parte dos grupos dos especialistas e decisores políticos consideraram que a avaliação externa é mais consistente, por conta de seus mecanismos de processamento. Contudo, entendem que possuem mecanismos diferentes e seria difícil realizar uma comparação. Ambos os grupos consideraram que o problema com a consistência da avaliação externa está na baixa formação do professor, que geraria uma falta de conhecimento no desenvolvimento da avaliação interna. Em contrapartida, os professores consideraram que a avaliação interna é mais consistente, pois são feitas pelos professores e esses conhecem mais a realidade de seus alunos, revelando uma resistência quanto ao avaliador externo.

A influência da avaliação externa sobre as práticas avaliativas dos professores não foi considerada pelos intervenientes. Alguns decisores políticos apenas evidenciaram o fato de o professor começar a refletir mais sobre suas avaliações. No entanto, todos os grupos pontuaram ações que podem ser desencadeadas, como maior frequência de itens fechados e simulados.

No que corresponde ao estudo dos alunos, fica claro que a avaliação externa não está sendo aliada, pois não proporciona indicações de como está o seu rendimento e do que pode ser feito para melhorá-lo. Nota-se, portanto, que os alunos estão muito longe de uma interação formativa com a avaliação externa.

Por outro lado, percebe-se um apoio ao processo de ensino de forma mais ampla, isto é, voltado também para as equipes de gestão do sistema educacional, não somente para o professor em sala de aula, segundo os especialistas e os decisores políticos. Nessa perspectiva, há um entendimento de que as ações orientadas para o processo de ensino compreendem não apenas as decisões dos professores ou gestores escolares, mas também diferentes esferas do sistema educacional. Tal fato é possível quando se podem realizar ações que venham a influenciar esse processo, como: formação de professores, contribuindo para uma mudança nas práticas de ensino; investimento em boas práticas pedagógicas, evidenciando-as e dando incentivo para a sua multiplicação; e melhoria do material didático e de apoio ao professor, fortalecendo os pontos que necessitam de maior intervenção pedagógica. Isso levaria a avaliação externa a realizar uma função política

formativa, identificando as dificuldades no processo de ensino e podendo, assim, sugerir soluções ou realizar ações para superá-las, apoiando esse processo.

Mesmo havendo a possibilidade da função pedagógica somativa e formativa, parece claro que a avaliação externa serve mais como referência quando empregada na função política, seja ela para formulação, avaliação e reformulação de políticas educacionais; seja para a regulação do trabalho docente; seja para pressão e controle verticalizados. A utilização da avaliação externa para a formulação de políticas públicas foi a única utilização considerada pelos três grupos de intervenientes, como apresentado na Tabela 13, entendendo que as políticas públicas estariam voltadas para a melhoria da qualidade da educação. Os decisores políticos e os especialistas afirmam tal relação, recebendo destaque a política voltada para formação de professores. Os professores entendem que a relação é possível, porém não aconteceria na prática, em grande parte pelos problemas encontrados nos mecanismos de processamento ou pelo fato de o governo não saber utilizar os dados gerados.

A Tabela 14 apresenta quais utilizações foram levantadas por cada grupo entrevistado, dividido por a quem caberia a utilização, ou seja, professores, gestores escolares e decisores políticos.

Tabela 14

Comparação das utilizações da avaliação externa entre os diferentes grupos

Utilização	Especialistas	Decisores políticos	Professores
Professores			
Falta de utilização dos resultados na dinâmica escolar	X	X	
Falta de relação entre os resultados e as realidades escolares			X
Gestores escolares			
Controle e cobrança verticalizada	X		X
Análise e discussão de resultados junto com os professores		X	X
Repensar as práticas pedagógicas		X	
Divulgação de resultados sem ação concreta			X
Decisores Políticos			
Formulação de políticas públicas	X	X	X
Pressão por resultados	X		X

Análise do desempenho das redes de ensino	X	
Formação de professores	X	
Campanhas políticas		X

Pode-se concluir, como aponta a Tabela 14, que a avaliação externa da aprendizagem dos alunos está sendo mais considerada no que corresponde à função política, ou seja, com gerenciamento do sistema educacional e/ou das escolas e formulação, avaliação e reformulação de políticas públicas. Mesmo que não fique claro para os professores o que está sendo feito e até mesmo que os decisores políticos não consigam precisar efetivamente como ocorre essa relação, as respostas dos grupos caminham para esse sentido.

Nesse contexto, a avaliação estaria sendo usada em sua função política somativa, isto é, medir o desempenho dos alunos, monitorar o sistema educacional, realizar uma série histórica para verificação da qualidade da educação, entre outros. Também está sendo usada em sua função formativa, tendo em vista que busca regular a implementação de práticas eficazes, corrigir currículos inadequados, contribuir com o aporte técnico avaliativo do professor, entre outros. No entanto, a função formativa ainda está caminhando de forma muito tímida, com pouco efeito positivo na prática docente, porém com ações estruturadas, como banco de itens, matrizes de referência baseando o currículo mínimo do estado, formação de professores com base nos resultados da avaliação externa, principalmente, nas avaliações estaduais e municipais, que se propõem a essas práticas.

Cabe ressaltar que há três movimentos que se destacam a partir da utilização dos resultados da avaliação externa na função política. O primeiro refere-se ao controle. Na gestão gerencial, que busca garantir o sucesso escolar com base na ênfase em resultados e indicadores de desempenho, o Estado implementa uma cultura organizacional apoiado em princípios de gestão estratégica e de controle de qualidade, realizando, assim, um controle por resultados, principalmente no resultado das avaliações externas da aprendizagem dos alunos, em detrimento de um controle rígido dos processos administrativos. Dessa forma, com dados geralmente estatísticos, fornece informações sobre alunos, turmas, escolas e sistemas educacionais, com o intuito de comparar, verificar e/ou monitorar a qualidade da educação.

Com base nesses resultados, ocorre uma transferência de responsabilidade por meio de um controle descendente, visto que as autoridades políticas e gestores do sistema

educacional desenvolvem metas para gestores escolares e professores alcançarem. O recurso mais empregado para o funcionamento desse gerenciamento é a pressão verticalizada, com ações que têm por finalidade cobrar resultados.

Por fim, uma forma de regulação das ações que acontecem em sala de aula, orientando os atores escolares a realizarem determinados padrões, seja de ações, como conteúdos a serem ensinados, tipos de provas utilizadas, seja de práticas de ensino a serem adotadas. Há um monitoramento do trabalho docente, controlando as ações no contexto escolar com o objetivo de se alcançarem as metas estabelecidas, o que levanta um questionamento sobre essa avaliação ser usada mais como uma avaliação do trabalho do professor do que da aprendizagem do aluno.

Visto que há um maior desenvolvimento da avaliação na função política, é necessário afirmar que há uma maior ênfase da função somativa, tendo em vista que se constatou que a realização da avaliação como medição no Brasil já está, de alguma forma, consolidada. Apesar de aparecerem críticas com relação ao antes e durante a aplicação da avaliação externa, o país vem apresentando cada vez mais fortalecimento da capacidade técnica do sistema de avaliação, muito em decorrência de uma política para a implementação de uma cultura avaliativa, buscando suporte em organizações internacionais. No entanto, o momento pós-avaliação, com a análise dos resultados e o seu retorno à escola com o objetivo de atrelá-lo ao processo de ensino e aprendizagem, ainda está fora da realidade do contexto escolar brasileiro. A avaliação externa, por ora, continua a realizar, em grande parte, o papel de medição, não evoluindo para um sistema de avaliação formativa efetiva, seja para as políticas públicas, seja para os professores e gestores escolares.

Se a intenção do caráter formativo é dirigir informações ou ações que possam auxiliar o desenvolvimento do professor em seu programa pedagógico, esse deve estar familiarizado e apresentar confiança nos resultados que servem como base para as informações ofertadas e ações sugeridas. Se, por outro lado, tiver como objetivo apresentar dados para a monitoração da qualidade da educação, assumindo assim um caráter somativo, deve-se comunicar o resultado de forma clara e credível para todos os intervenientes do processo educacional e para a sociedade como um todo. Para isso, outras características da natureza da avaliação externa devem ser consideradas.

Os mecanismos de processamento foram pontuados pelos três grupos ao descreverem a natureza da avaliação externa. Ao considerar esses mecanismos, mostrou-se a necessidade de haver um instrumento que tenha métodos confiáveis, com validade e

garantia de qualidade, havendo uma preocupação com o aspecto técnico da avaliação. Os especialistas e os decisores políticos abordaram as questões técnicas, como: padronização do instrumento, comparabilidade entre as informações coletadas, validade, avaliador externo etc. Essas características foram consideradas um ponto forte desse tipo de avaliação, porém com pouca discussão sobre os seus limites e pontos que devem ser mais bem analisados e desenvolvidos.

No entanto, os professores da educação básica destacam as características negativas dos mecanismos de processamento: a inconstância na aplicação; o ato de maquiagem os índices, com mudança de indicadores; a falta de organização na aplicação da prova; a descontextualização da prova, tendo em vista que não dialoga com a realidade escolar; o critério de correção ser desconhecido; e não se considerarem as particularidades das escolas, o que estaria mais relacionado à falta de correspondência entre a avaliação externa e o contexto escolar.

Os efeitos das avaliações externas foram mais evidenciados nos grupos dos especialistas e dos professores, do que entre os decisores políticos. Na realidade, ao se falar da natureza da avaliação, os decisores políticos não pontuaram nenhum efeito. Os citados foram, em sua maioria, efeitos negativos, como: estreitamento do currículo, padronização do ensino e responsabilização dos professores, sendo este último referido tanto pelos especialistas quanto pelos professores, o que realça uma política de transferência de responsabilidade. Muito aparece na narrativa dos intervenientes sobre os efeitos negativos no âmbito de sala de aula, mas pouco se discute sobre os efeitos positivos e os efeitos dos testes sobre os decisores políticos.

Os mecanismos de processamento e os efeitos atrelados à avaliação externa são os fatores que mais influenciam no grau de satisfação com relação a essa avaliação. Parece haver um consenso de maior satisfação para os decisores políticos, em virtude de serem apenas gerentes do processo e transferirem a responsabilidade para os atores escolares. Eles podem maquiagem resultados para que esses venham a corresponder às metas por eles antes projetadas. Além disso, veem uma evolução no sistema de avaliação do país e têm dados estatísticos para fundamentar suas políticas públicas.

No entanto, entende-se que há maior insatisfação com relação aos professores e especialistas – estes, em sua maioria, pela falta de confiança na credibilidade desse tipo de avaliação. Não acreditam que a avaliação externa esteja ajudando a melhorar a qualidade da educação, ou entendem que há relevância, no entanto, desconfiam da forma como está sendo posta em prática tanto nas políticas públicas quanto nos processos

pedagógicos. No que concerne aos professores, estão insatisfeitos devido à pressão exercida por órgãos superiores e, na maioria das vezes, externos ao contexto escolar; ao número de dados sem uso efetivo; e à desconfiança com relação aos resultados apresentados, mecanismos de processamento e as políticas que embasam as avaliações.

Entretanto, mesmo havendo uma parcela significativa de insatisfação com a avaliação externa, ficou evidente que há uma percepção de que deve haver um sistema de avaliação externa, sendo necessário tanto como função política quanto como função pedagógica. São necessárias, contudo, melhorias para que possam corresponder ao que objetivam, principalmente no que diz respeito à função política formativa e pedagógica formativa.

Por fim, é importante apontar que, como no Brasil ocorrem múltiplos sistemas de avaliação, visto que todos os níveis governamentais possuem um sistema de avaliação externa que, por vezes, tem mais de uma avaliação, houve pouca referência ao longo das entrevistas sobre a diferença entre elas. Apenas os especialistas pontuaram sobre esse fato, considerando que a ele também está atrelada a questão da validade, fidedignidade e consistência para com o fim proposto. Torna-se evidente que esse aspecto é visto mais em pesquisa do que compreendido por professores ou referenciado pelos decisores políticos. Em alguns momentos, havia o direcionamento da resposta para uma avaliação específica, geralmente as estaduais, tanto na fala dos professores quanto na dos decisores políticos. Porém, diferenciar os sistemas de avaliação brasileiros não foi um ponto importante no momento em que falaram sobre a avaliação externa.

Levando em consideração esse aspecto, pode-se concluir que as avaliações da aprendizagem dos alunos no Brasil vêm ocupando o mesmo espaço para a maioria dos intervenientes. Mesmo havendo diferenças entre as avaliações nacionais, estaduais e municipais, elas são vistas como um modelo único. Parece claro que há uma necessidade primeira em se falar do âmago dessa avaliação, de um sistema incorporado há anos no país, mas que ainda precisa de muito diálogo entre os diferentes intervenientes e o estabelecimento de objetivos claros, não apenas de indicadores com base estatística. A falta de clareza e discussões sobre as limitações presentes nas avaliações externas; a pretensão de utilização de dupla função na maioria dos casos, sem uma condução apropriada e informação sobre as diferentes finalidades; a dificuldade encontrada na comunicação dos resultados; e a falta de clareza dos atores escolares sobre a avaliação externa foram os obstáculos encontrados para a devida incorporação, principalmente, de seu caráter formativo.

Em observação dos aspectos analisados, evidencia-se a necessidade de promover discussões tanto na forma de pesquisas, quanto nas agendas políticas e nos espaços escolares, para além dos resultados da avaliação, também sobre seus fundamentos e objetivos. Uma discussão que vise uma interligação dos diferentes pontos de vista, para que possa haver uma percepção global da avaliação externa, com os diferentes intervenientes, e não uma visão fragmentada. Desse modo, será possível trabalhar a partir de uma demanda que seja plausível, não atribuindo à avaliação nem mais e nem menos do que ela efetivamente pode dar.

Com base na dupla função da avaliação externa presente em alguns sistemas de avaliação, é importante salientar que as diferentes funções encontram-se em *continuum* de desenvolvimento, ou seja, de um lado uma característica mais somativa, com base na medição e apresentação de dados, encontrando-se a parte mais desenvolvida da avaliação externa no contexto brasileiro; do outro lado, um polo mais formativo, com vistas a contribuir com o processo de ensino e aprendizagem, onde há pouco desenvolvimento, como mostra a Figura 36.

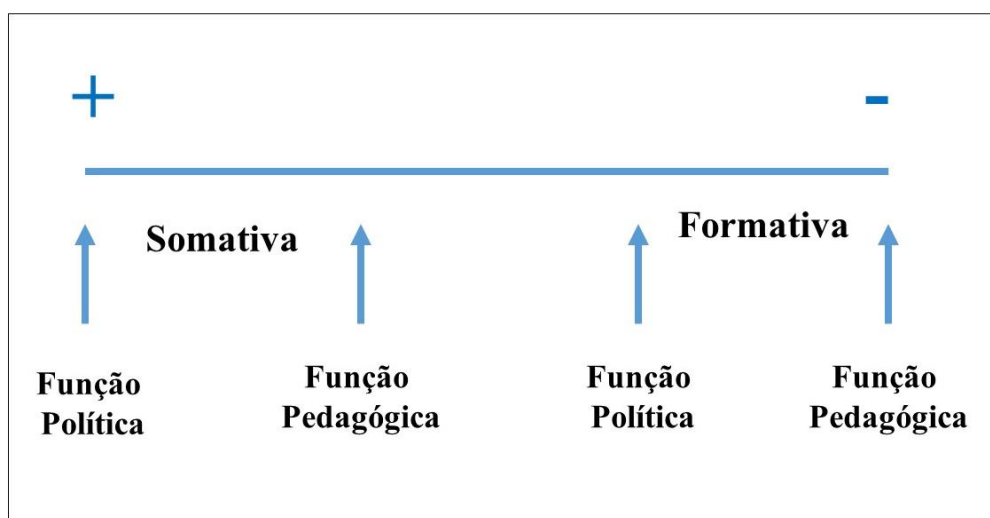


Figura 36

Continuum de desenvolvimento da dupla função da avaliação externa

Com isso, nessa perspectiva, a avaliação externa da aprendizagem dos alunos pode contribuir, em primeiro lugar, com a informação sobre a qualidade da educação, visto que gera, há anos e de forma constante, diferentes dados, o que corresponde à sua função política somativa. Em seguida, pode auxiliar os atores escolares, a partir de um olhar externo, a ter mais uma visão sobre o seu contexto, o que caracteriza a função pedagógica

somativa. Com menor intensidade, também permite a formulação de políticas públicas que possuem, como objetivo, a melhoria da qualidade da educação, o que configura a função política formativa. Por fim, com pouco desenvolvimento, a avaliação externa funciona como um instrumento de apoio no processo de ensino e aprendizagem, sendo utilizada por professores e alunos, o que corresponde à sua função pedagógica formativa.

O desenvolvimento desse *continuum* depende de uma gama de aspectos levantados ao longo da pesquisa, que necessitam receber atenção para que a avaliação externa desenvolva, de forma clara e credível, as diferentes funções e características que objetiva.

Capítulo VI

Conclusão

Este trabalho, de forma não generalista, buscou discutir a avaliação externa da aprendizagem dos alunos, com base na apreciação de três grupos de intervenientes relevantes no processo educativo, para que assim pudesse construir um quadro mais amplo, contemplando olhares de diferentes níveis, ou seja, do contexto escolar, dos gabinetes governamentais e dos centros de pesquisa.

Nesta seção, são apresentadas as conclusões consideradas mais relevantes tendo em conta, naturalmente, as questões de investigação que inicialmente se formularam e que, por uma questão de clareza, se formulam no início das referidas conclusões. A partir disso, são apontadas as limitações do trabalho, como também indicações para o desenvolvimento de estudos futuros.

1. Como se caracterizam e qual a natureza das avaliações externas da aprendizagem dos alunos?

A partir de uma análise descritiva das avaliações externas no Rio de Janeiro, realizada no capítulo IV – *Do internacional ao local: uma discussão acerca das avaliações externas internacionais, nacionais e do estado do Rio de Janeiro*, pode-se constatar um caráter de afunilamento do que se é avaliado. A avaliação externa, no município do Rio de Janeiro, apresenta avaliação do sistema educacional, nas avaliações internacional (PISA) e nacional (SAEB); avaliação da escola, na avaliação nacional (Prova Brasil); e avaliação do aluno, nas avaliações estadual (PNE, Saerj e Saerjinho) e municipal (Prova Rio).

Ficou evidente que as avaliações, em larga escala no Brasil, seguem diferentes desenhos e processos, dependendo do nível governamental que a administra e de seus objetivos, apresentando função política e/ou pedagógica. Para além dessas funções, fica claro a presença de um caráter somativo e formativo nessas avaliações, por exemplo, quando usada apenas com o propósito de medir a qualidade da educação, está diante de uma avaliação política somativa; ao servir como indicador para a formulação de políticas educacionais, pode-se considerar uma avaliação política formativa; ao realizar um

diagnóstico do desempenho dos alunos de uma determinada escola ou turma, há uma avaliação pedagógica somativa; por fim, quando contribui com o processo de ensino e aprendizagem, ocorre uma avaliação pedagógica formativa.

No entanto, na mesma avaliação, como evidenciado, há mais de uma função, podendo ter como objetivo, por exemplo, uma função política somativa e pedagógica formativa ao mesmo tempo, ao apresentar diferentes propósitos. As avaliações externas internacionais e nacionais tendem a apresentar a função política somativa e formativa. Já as avaliações locais, estadual e municipal, apresentam desde a função política somativa até a função pedagógica formativa, tendo assim muitos propósitos.

Segundo a apreciação dos intervenientes participantes da pesquisa, é possível que a avaliação externa desenvolva tanto a função política quanto a pedagógica. Contudo, observa-se, em suas narrativas, que há uma baixa utilização, principalmente no que se refere às ações mais formativas.

Outra forma de falar sobre a natureza da avaliação externa é referenciando seus mecanismos de processamento, entendendo a necessidade de se desenvolver metodologia que seja credível, com validade e garantia de qualidade, havendo uma preocupação com seu aspecto técnico. Para isso, foi citado: avaliador externo à escola, padronização dos instrumentos, comparabilidade de resultados, mecanismos de processamento recentes, se as etapas do exame são públicas, complexidade compatível com idade e série avaliadas, consistência dos resultados e análise dos resultados a partir de uma escala de proficiência.

Esses aspectos foram considerados como características importantes desse tipo de avaliação, principalmente pelos especialistas e decisores políticos, porém, com pouca discussão sobre os seus limites e pontos que devem ser melhor analisados e desenvolvidos. Por outro lado, os professores destacam as características negativas dos mecanismos de processamento, em grande parte, para justificar a falta de relação da avaliação externa com o contexto escolar, como: índices maquiados; falta de organização na aplicação das avaliações e periodicidade; critérios de correção não são públicos; e provas descontextualizadas.

Por fim, os efeitos resultantes da avaliação externa foram considerados ao se abordar a natureza dessa avaliação. É necessário destacar que muito se discorre sobre os efeitos negativos no âmbito de sala de aula, com pouca referência aos efeitos positivos e aos efeitos dos testes, com relação às políticas educacionais (Stecher, 2002).

O efeito mais considerado foi a responsabilização dos professores em decorrência dos baixos resultados, o que acaba por desencadear outros efeitos, como preparação para

o teste, trapaça, afunilamento do currículo entre outros (Darling-Hammond, 1994; Mons, 2009; Shepard, 2000).

2. Que relações se podem estabelecer entre as políticas públicas de educação e as avaliações externas da aprendizagem dos alunos?

A avaliação externa tornou-se uma ferramenta importante das políticas públicas educacionais, sendo utilizada por elas como instrumento para alcance de diversos propósitos, possibilitando cumprir funções em diferentes níveis, desde a sala de aula até o sistema educacional. A partir disso, a avaliação apresenta uma função política, ou seja, contribuir para a avaliação e formulação de políticas educacionais e para o gerenciamento dos sistemas educacionais. Essa função estaria no nível macro, nas avaliações externas internacionais e nacionais, e no nível meso, nas avaliações estaduais e municipais.

Nesta investigação, pode-se observar que há uma maior concordância na utilização da avaliação externa como referência para a função política, visto que, essa utilização foi a única citada pelos três grupos de mediadores. No entanto, não fica claro para os participantes da pesquisa como se dá essa relação, principalmente para os professores.

Isso pode ocorrer devido ao fato de, apesar de encontrar uma maior percepção de utilização, há pouca discussão na área que evidencie isso, principalmente, no Brasil. Há a necessidade de se estudar com profundidade suficiente e com rigor metodológico o impacto das recomendações ou efeitos das avaliações externas nas políticas públicas educacionais, verificando se há um aprendizado no nível governamental, que resulta em mudanças nas políticas educacionais; se ocorre apenas uma avaliação do desempenho dos sistemas educacionais, não causando nenhum impacto nas políticas educacionais; ou se os resultados da avaliação são utilizados apenas no sentido de legitimar reformas políticas que já são estruturadas por imperativos políticos internos e ideologias já enraizadas (Araújo & Tenório, 2017; Bieber & Martens, 2011; Lingard, 2016; Michel, 2017; Niemann, Martens & Teltemann, 2017).

Diante dessa realidade, pode-se concluir que a função política apresenta um forte movimento de controle de qualidade da educação, com base no controle de resultados das avaliações externas e de transferência de responsabilidade, dado que as autoridades políticas e gestores do sistema desenvolvem metas para gestores escolares e professores alcançarem, instituindo uma cultura de pressão verticalizada (Freitas, 2007; Mons, 2009).

Assim, parece haver uma maior inclinação à utilização da função política somativa, na qual objetiva medir o desempenho dos alunos, monitorar o sistema educacional, realizar uma série histórica para verificação da qualidade da educação e cobrar por resultados.

A função política formativa ainda encontra baixa utilização, uma das mais citadas, neste trabalho, foi a contribuição da avaliação externa para a formação de professores, principalmente no que concerne as avaliações estaduais e municipais, pois apresentam um olhar mais local. Isso confirma que a partir da ampliação de políticas de avaliações externas, houve a construção de novos debates sobre a formação docente, como afirmado por Alavarse et al. (2017).

Isso deve ser visto com cuidado, posto que é importante refletir sobre esse discurso que relaciona de forma direta a formação continuada do professor como melhoria da aprendizagem dos alunos, já que essa perspectiva infere que o problema da qualidade da educação se resume à formação e ao desempenho docente, desconsiderando outros fatores que influenciam no processo de ensino e aprendizagem (Alavarse et al., 2017; Bauer, 2012; Soares, 2007).

3. Que relações se podem estabelecer entre as avaliações externas da aprendizagem e a qualidade do sistema de ensino?

Ao longo dos anos, diferentes indicadores de qualidade da educação foram se estabelecendo no Brasil. Atualmente, esse indicador é representado pelos resultados dos testes padronizados da capacidade cognitiva dos estudantes, ou seja, um indicador com base num pressuposto epistemológico de qualidade como medida. É necessário ressaltar que a avaliação externa não representa a qualidade da educação, mas sim, deve ser considerada como um indicador, sendo utilizada para a garantia da qualidade, visto que, a partir de seus resultados, pode-se planejar e elaborar ações que tenham como objetivo melhorar a qualidade da educação.

É nesse contexto que se estabelece, muito em decorrência dos estudos comparativos internacionais, uma cultura avaliativa e o controle da qualidade dos sistemas educacionais, com o intuito de realizar um levantamento periódico de informações. Assim, inicia-se um movimento com o propósito de descrever a situação dos sistemas de ensino, fazer comparações, formular metas, fornecer explicações, diagnosticar problemas, tomar decisões, realizar avaliação e monitoramento.

Entretanto, observa-se que, em aproximadamente 30 anos de sistema de avaliação no Brasil, desenvolveu-se mais uma medição, vindo a descrever e comparar, do que uma avaliação do sistema de ensino, com o objetivo de contribuir com a qualidade da educação.

Esse contexto, vem ao encontro do PISA, que possui maior ênfase em gerar informações do desempenho dos alunos, com base na avaliação de conhecimento de habilidades, relacionando esse desempenho a temas de políticas públicas e realizando um monitoramento regular dos padrões de desempenho. Dessa forma, há ainda pouca evidência de movimentos que visem uma construção de conhecimento sobre os sistemas de ensino, para que assim possa contribuir com ações que possibilitem a melhoria na qualidade da educação.

O que vem se delineando é a transferência de responsabilidade, muito em resposta a um novo modelo de gestão gerencial da educação, na medida em que as autoridades políticas e gestores do sistema educacional acabam por desenvolver metas para gestores de escola e professores alcançarem, o que vem sendo considerado por Mons (2009) e Freitas (2007) como uma pressão verticalizada. Por isso, ocorre uma grande ênfase na responsabilização dos professores, tendo em vista que seriam eles os responsáveis por alcançar as metas estabelecidas nesse modelo de gestão dos sistemas de ensino.

Diante de tal situação, a avaliação externa apresenta uma ênfase maior em seu papel somativo, com maior objetivo em medir o desempenho dos sistemas de ensino. Como apresentado nas falas dos diferentes intervenientes, há muito mais um posicionamento de medição por parte do governo com um número considerável de dados, o que coaduna com o posicionamento de Bottani (1998) e Souza e Oliveira (2010). Quando se fala em ações para a melhoria da qualidade da educação desse sistema, há uma ideia maior de transferência de responsabilidade para os sujeitos escolares.

4. Em que medida as avaliações externas são relevantes para melhorar a qualidade da aprendizagem dos alunos e a qualidade dos processos de ensino?

A função formativa da avaliação apresenta evidências que podem ser usadas para adaptar o trabalho do professor às necessidades dos alunos, para que estes possam ter ganhos significativos na aprendizagem. Assim, os resultados da avaliação estariam a serviço de ajustes no ensino, envolvimento dos alunos com suas aprendizagens e desenvolvimento da escola, sendo uma aliada de todos. Uma avaliação que deve levar em

consideração a interatividade do processo de ensino e aprendizagem (Black & Wiliam, 1998, 2001; Villas-Boas, 2006).

A função política formativa busca ações que são orientadas para o processo de ensino, com isso, há uma compreensão de que essas ações não são apenas geridas pelos professores, mas também podem ocorrer em outros níveis do sistema educacional. Dessa forma, a função política formativa da avaliação externa é desenvolvida no sentido de identificar as dificuldades no processo de ensino e assim, sugerir soluções ou realizar ações para superá-las, apoiando esse processo. Sendo possível quando se realiza ações voltadas para influenciá-lo, como: formação de professores, investimento em boas práticas pedagógicas, melhoria do material didático e de apoio ao professor.

No que se refere à função pedagógica formativa, a avaliação externa teria como objetivo contribuir com informação sobre o aluno ou um grupo de alunos, seja para o professor realizar ajustes e refletir sobre o ensino, seja para o envolvimento do aluno com relação à sua aprendizagem. Assim, seria uma complementação da avaliação interna, um auxílio no processo de aprendizagem de cada aluno ou de um grupo de estudantes.

Contudo, apesar de se reconhecer a relevância da avaliação externa para o processo de ensino e aprendizagem, esta pesquisa constatou que há limitações para essa utilização, como os mecanismos de processamento e as políticas públicas que embasam a avaliação externa. Com isso, o caráter formativo, tanto da função política, quanto da pedagógica, está longe de uma verdadeira utilização, muito em decorrência da falta de envolvimento dos sujeitos escolares com essas avaliações, em grande parte por falta de credibilidade, tendo em vista que não consideram que a avaliação externa esteja ajudando as escolas em seus objetivos, não oferece *feedback* confiável e não está avaliando de forma adequada a aprendizagem dos alunos (Marchesi, 2002; Mons, 2009).

Como se pode ver, se há uma utilização da avaliação, em sala de aula, isso ocorre apenas quando há uma cobrança do sistema, havendo mais movimentos de preparação para os testes do que efetivamente em mudanças na prática de ensino (Darling-Hammond, 1994, Firestone et al., 1998; Mons, 2009; Shepard, 2000).

Outro posicionamento para a baixa utilização, seria a falta de conhecimento do professor em utilizar a avaliação externa a serviço do desenvolvimento de sua prática, havendo uma necessidade de investimento na formação inicial e capacitação continuada.

Independentemente do motivo que leva a baixa utilização da avaliação externa, no processo de ensino e aprendizagem, principalmente da função pedagógica formativa, o que fica evidente é a falta de envolvimento dos professores com esse processo

avaliativo. Os resultados apontam para a necessidade de se envolver os usuários no processo, para que assim ocorra um uso efetivo. Durante a pesquisa, ficou notória a falta de conhecimento dos professores sobre o real alcance da avaliação externa, sobre os seus mecanismos de processamento e seus propósitos, para além de demonstrarem baixa credibilidade pelo instrumento.

Parece haver um movimento contraditório dos professores participantes, pois ao mesmo tempo que entendem a relevância da avaliação externa, apresentam uma forte resistência em utilizar os seus dados. Entretanto, este estudo aponta que essa aparente contradição é na realidade formada por fatores que necessitam de maior desenvolvimento e estudo, para que dessa forma, ocorra um alinhamento entre a avaliação externa e os seus usuários, principalmente os professores. Especialmente, se há o desejo de solidificar a função pedagógica formativa da avaliação externa. Assim, será desejável que se desenvolvam algumas ações tais como:

- Envolvimento dos sujeitos escolares no processo - A avaliação participativa aumenta a probabilidade de utilização dos resultados, dando às partes interessadas maior confiança na qualidade das informações geradas e propriedade nos resultados da avaliação e sua aplicação (Nevo, 1997; Greene, 1988), assim como conhecimento dos mecanismos de processamento. Diante dessas considerações, é imperativo um maior investimento na participação dos professores no processo de avaliação externa para que possa alcançar uma verdadeira utilização dessa avaliação no processo de ensino e aprendizagem. Esse envolvimento ocorreu de forma mais acentuada no início do SAEB e depois não foi mais contemplado. As avaliações estaduais e municipais podem desenvolver essa característica, dado que apresentam uma abordagem mais local.

- Relação entre a avaliação externa e o contexto escolar – A avaliação externa deve ter um grau de congruência com relação às ações desenvolvidas na escola. A utilidade dessa avaliação para a prática docente depende, em grande medida, de um reconhecimento por parte do professor de seu trabalho nos resultados da avaliação. A desarticulação entre o contexto escolar e a avaliação externa acarretam em um distanciamento entre os resultados obtidos e os sujeitos que os produzem (Esteban, 2009).

- Política de responsabilização – Num contexto em que há uma forte responsabilização dos professores, com base nos resultados da avaliação externa, eles se preocupam mais com a forma como irão responder por esses resultados do que como utilizar essa avaliação como um instrumento relevante para o processo de ensino e aprendizagem. Nessa pesquisa, tornou-se evidente que a comunicação dos resultados está sendo vista com maior ênfase no sentido de expressar uma denúncia da baixa qualidade do ensino, chegando a escola de uma forma coercitiva, sem diálogo, apenas apontando os erros e pressionando mudanças. Muito mais realizando “descrições assertivas unilaterais e de julgamentos autoritários.” (Nevo, 1997, p. 94). Este trabalho mostra que os professores apresentam um certo grau de pressão verticalizada (Freitas, 1997; Mons, 2009), em decorrência dos resultados da avaliação. Vindo a desenvolver ações com o intuito de melhorar os resultados, ao invés de refletir sobre os modelos e efeitos da avaliação externa e lançar-se em ações que possam aprofundar o entendimento sobre o que ensinam e como ensinam. Corroborando com o posicionamento da literatura, a avaliação externa não tem efeito na prática do professor, mas sim, em ações em sala de aula, muito em resposta a uma tentativa de melhoria dos resultados (Shepard, 2000; Firestone et al., 1998; Darling-Hammond, 1994).

- Função formativa – Com base nas análises descritivas das avaliações externas e nas respostas dos participantes, merece ressaltar que, apesar de todo o esforço governamental, não houve um desenvolvimento significativo no campo formativo da avaliação externa, apesar dela possuir potencial para desenvolver essa função. Há ainda, uma maior ênfase nas práticas somativas, confirmando o fato de que a avaliação externa formativa se encontra ainda como um desafio que está longe de ser alcançado. Como afirma Pacheco (2012), há pontos importantes para o contexto escolar: políticas tendencialmente formativas; práticas predominantemente somativas; e entre o desejo e o realizado.

Limitações e Sugestões para Futuras da Pesquisas

Procurou-se neste trabalho, compreender a avaliação externa da aprendizagem dos alunos a partir das apreciações de diferentes intervenientes do processo educativo. Parece conveniente dizer que houve uma contribuição nesta pesquisa de participantes com grande prestígio no campo da avaliação externa da aprendizagem dos alunos e de professores, seja nos dados quantitativos ou nos qualitativos, para além de uma análise descritiva dos sistemas de avaliação, presentes no Rio de Janeiro, possibilitando a construção de um quadro relevante e significativo da avaliação externa a partir de diferentes perspectivas. Contudo, não se pode afirmar que abordamos de forma exaustiva os objetos de análise, longe de esgotar quaisquer pontos que se propuseram a trabalhar.

Dessa maneira, concluo esta tese, levantando alguns limites da pesquisa e indicando algumas sugestões para futuros trabalhos. Escolhi uma metodologia mista, porém com ênfase nos dados qualitativos, buscando usar os dados quantitativos como forma de complementação da parte qualitativa. Com isso, esse método apresenta uma lacuna no trabalho dos dados estatísticos, pois optou-se por discutir de forma mais acentuada as informações das entrevistas. Por isso, sugere-se que outras vias metodológicas, nomeadamente com uma maior ênfase na utilização da estatística inferencial dos dados obtidos através do inquérito por questionário, possam ser utilizadas para produzir outras leituras acerca das respostas dos professores.

Não houve análise de uma avaliação externa específica, mas sim, do contexto de avaliação externa, para que assim pudesse compreender a apreciação dos diferentes intervenientes sobre esse tipo de avaliação. Porém, entende-se que pode haver posicionamentos diferentes com relação às avaliações de diferentes instâncias (nacional, estadual, municipal), visto que apresentam desenho e processamentos diferentes. Para tal, recomenda-se estudos a partir da apreciação de diferentes intervenientes do processo educativo com relação às avaliações específicas de cada esfera governamental, para que possa haver maior análise das avaliações externas presentes no Brasil.

O recorte realizado, com os professores do município do Rio de Janeiro, de português e matemática, representa apenas uma parcela da apreciação dessa categoria. Cabe ressaltar, que o contexto no qual esses profissionais estão inseridos, apresenta desde o início, política de bonificação e diversas avaliações externas de diferentes níveis governamentais. Assim, faz-se necessário realizar uma análise com uma gama maior de professores de diferentes municípios e estados, para verificar como o contexto influencia

a percepção dos professores. Assim como, ouvir os professores das outras disciplinas não contempladas nesta tese.

Finalmente, optou-se por selecionar alguns atores do processo educativo para falar sobre a avaliação externa, sendo necessário considerar outras apreciações, como alunos, seus pais, gestores escolares, para que assim, haja estudos com todos os envolvidos no processo.

Ao realizar esta tese, ficou evidente para mim que há diversar pesquisas no campo da avaliação externa, principalmente apresentando relação dessa área com outras temáticas, como: currículo, gestão escolar, entre outros. No entanto, pouco se fala de alguns pontos, como: seu âmago, sua natureza, suas diferentes funções e propósitos. Ao longo do caminho que propus realizar, pude olhar para essa avaliação tomando como ponto de partida outros olhares, que muito conhecem desse processo, o que foi enriquecedor para o meu conhecimento. Pude entender que às vezes é necessário dar alguns passos para trás, com o objetivo de compreender para onde estamos indo. Identificar tensões e limites, com o intuito de clarificar o processo e não para desqualificá-lo. Hoje junto-me ao grupo que considera a utilidade da avaliação externa, tanto em sua função política quanto na pedagógica. Porém considero urgente a necessidade de clarificar mais este percurso para os diferentes intervenientes; proporcionar um processo avaliativo mais colaborativo, com mais diálogos entre as diferentes partes envolvidas; e realizar mudanças em alguns caminhos que optou-se por percorrer ao longo de sua implementação no território brasileiro.

Referências Bibliográficas

- Afonso, A. J. (2005). Políticas educativas e avaliação educacional: *Centro de estudos em educação*. Braga: Universidade do Minho.
- Afonso, A. J. (2009). Nem tudo o que conta em educação é mensurável ou comparável: Crítica à accountability baseada em testes estandardizados e rankings escolares. *Revista Lusófona de educação*, (13), 13-29.
- Afonso, A. J. (2012). Para uma concetualização alternativa de accountability em educação. *Educação & Sociedade*, 33(119), 471-484.
- Afonso, A. J. (2014). Questões, objetos e perspetivas em avaliação. *Avaliação: Revista da avaliação da educação superior*, 19(2).
- Ahn, T., & Vigdor, J. (2010). The Impact of Incentives on Effort: Teacher Bonuses in North Carolina. Program on Education Policy and Governance Working Papers Series. PEPG 10-06. Program on Education Policy and Governance, Harvard University.
- Ala-Harja, M., & Helgason, S. (2000). Em direção às melhores práticas de avaliação. *Revista do Serviço Público*, 51(4), 5-60.
- Alavarse, O. (2013). Desafios da avaliação educacional: ensino e aprendizagem como objetos de avaliação para a igualdade de resultados. *Cadernos Cenpec. Nova série*, 3(1).
- Alavarse, O., Bravo, M. H., & Machado, C. (2012). Avaliação como diretriz das políticas educacionais dos governos federal, estaduais e municipais: o caso brasileiro. In. III Congresso Ibero-americano de política e administração da educação. Zaragoza.
- Alavarse, O., Machado, C., & Arcas, P. H. (2017). Avaliação externa e qualidade da educação: formação docente em questão. *Revista Diálogo Educacional*, 17(54).
- Alderson, J., & Wall, D. (1993). Does Washback Exist?. *Applied Linguistics*, 14(2), 115-129.
- Alkin, M.C., & Taut, S.M. (2003). Unbundling evaluation use. *Studies in Educational Evaluation*, 29, 1-12.
- Alves, N., & Canário, R. (2002). The new magistracy of influence: changing governance of education in Portugal. *European Educational Research Journal*, 1(4), 656-666.
- Almeida, A. (2013). As políticas públicas de avaliação e a prática docente: percepções dos professores dos municípios do Rio de Janeiro e Duque de Caxias. Dissertação de Mestrado, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil.
- Almeida, I. C. (2008). A comparação internacional de indicadores de financiamento e gasto com educação. *Em Aberto*, 18(74).
- Almeida, L. C., Dalben, A., & Freitas, L. C. D. (2013). The Ideb: limits and illusions of an educational policy. *Educação & Sociedade*, 34(125), 1153-1174.
- Almeida, I. C., & Wolyne, E. (2007). A produção de indicadores educacionais no Brasil e a comparação internacional. *Revista Brasileira de Estudos Pedagógicos*, 79(193).
- André, M. (2015). Políticas de valorização do trabalho docente no Brasil: algumas questões. *Ensaio: Avaliação e Políticas Públicas em Educação*, 23(86), 213-230.
- Andrews, P., Ryve, A., Hemmi, K., & Sayers, J. (2014). PISA, TIMSS and Finnish mathematics teaching: An enigma in search of an explanation. *Educational Studies in Mathematics*, 87(1), 7-26.

- Antunes, M. P. L., & Galvão, C. (2015). Manuais escolares de Ciências Naturais de 8º ano em Portugal e estrutura conceptual do PISA 2006. *Revista Portuguesa de Educação*, 28(1), 139-169.
- Araújo, M. D. L. H. S., & Tenório, R. M. (2017). Resultados brasileiros no PISA e seus (des) usos. *Estudos em Avaliação Educacional*, 28(68), 344-380.
- Atkinson, A., Burgess, S., Croxson, B., Gregg, P., Propper, C., Slater, H., & Wilson, D. (2009). Evaluating the impact of performance-related pay for teachers in England. *Labour Economics*, 16(3), 251-261.
- Au, W. (2007). High-Stakes Testing and Curricular Control: A Qualitative Metasynthesis. *Educational Researcher*, 36(5), 258-267.
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of education policy*, 18(2), 215-228.
- Barroso, J. (2006). O Estado e a educação: a regulação transnacional, a regulação nacional e a regulação local. In Barroso, J. (Org.). *A regulação das políticas públicas de educação: espaços, dinâmicas e atores*. Lisboa: Educa, 41-70.
- Bauer, A. (2011). Formação continuada de professores e resultados dos alunos no SARESP: propostas e realizações. *Educação E Pesquisa*, 37(4), 809-824.
- Bauer, A. (2012). É possível relacionar avaliação discente e formação de professores? A experiência de São Paulo. *Educação em revista*, 28(2), 61-82.
- Bauer, A., Alavarse, O. M., & de Oliveira, R. P. (2015). Avaliações em larga escala: uma sistematização do debate. *Educação e Pesquisa*, 41, 1367-1384.
- Bauer, A., Sousa, S. M. Z. L., Neto, J. L. H., Valle, R. D. C., & Pimenta, C. O. (2017). Iniciativas de Avaliação do Ensino Fundamental em Municípios Brasileiros: mapeamento e tendências. *Revista Brasileira de Educação*, 22(71), 1-19.
- Bercier-Larivière, M., & Forgette-Giroux, R. (1999). L'évaluation des apprentissages scolaires: une question de justesse. *Canadian Journal of Education/Revue canadienne de l'éducation*, 169-182.
- Bieber, T., & Martens, K. (2011). The OECD PISA study as a soft power in education? Lessons from Switzerland and the US. *European Journal of Education*, 46(1), 101-116.
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74
- Black, P., & Wiliam, D. (2001) Inside the black box: raising standards through classroom assessment (London, King's College London School of Education).
- Bloom, B., Hastings, J., & Madaus, G. (1971). Handbook on formative and summative evaluation of student learning. *New York: McGraw-Hill*.
- Bonamino, A. (2013). Avaliação educacional no Brasil 25 anos depois: onde estamos. Vinte e cinco anos de avaliação de sistemas educacionais no Brasil: implicações nas redes de ensino, no currículo e na formação de professores. *Florianópolis: Insular*, 43-60.
- Bonamino, A., Coscarelli, C., & Franco, C. (2002). Avaliação e letramento: concepções de aluno letrado subjacentes ao SAEB e ao PISA. *Educação & Sociedade*, 23(81), 91-113.
- Bonamino, A., & Franco, C. (1999). Avaliação e política educacional: o processo de institucionalização do Saeb. *Cadernos de Pesquisa*, 108, 101-132.
- Bonamino, A., & Sousa, S. Z. (2012). Três gerações de avaliação da educação básica no. *Educação e Pesquisa*, 38(2), 373-388.
- Bondioli, A. (2004). O projeto pedagógico da creche e a sua avaliação: a qualidade negociada. *Campinas: Autores Associados*.

- Bonitatibus, S. G. (1989). Educação Comparada: conceito, evolução, métodos. *São Paulo: EPU*.
- Bottani, N. (1998). Os indicadores educacionais da OCDE: propósitos, limites e processos de produção. In *Seminário Internacional de Avaliação Educacional* (pp. 53-64). Brasília: Inep.
- Brasil. (1988). Constituição Federal de 1988. Brasília, Senado.
- Bresser-Pereira, L. C. (1996). Da administração burocrática à gerencial. *Revista do Serviço Público*, 47(1).
- Brooke, N. (2006). O futuro das políticas de responsabilização educacional no Brasil. *Cadernos de Pesquisa*, 36(128), 377-401.
- Brooke, N. (2013a). Controvérsias sobre políticas de alto impacto. *Cadernos De Pesquisa*, 43(148), 336-347.
- Brooke, N. (2013b). Sobre a equidade e outros impactos dos incentivos monetários para professores. *Estudos em Avaliação Educacional*, 24(55), 34-62.
- Brooke, N., & Cunha, M. A. D. A. (2011). A avaliação externa como instrumento da gestão educacional nos estados. *Estudos & Pesquisas Educacionais*, São Paulo, 2, 17-79.
- Brunner, J. J. (2003). Límites de la lectura periodística de resultados educacionales. *IIPÉ-UNESCO. Evaluar las Evaluaciones Buenos Aires*.
- Caed. (2015). Sistemas Estaduais de Avaliação da Educação Básica: Abrangência dos atuais sistemas. Recuperado em 30 fevereiro, 2016, de <http://www2.camara.leg.br/atividade-legislativa/comissoes/comissoes-permanentes/ce/audiencias-publicas-1/apresentacoes/apresentacao-lina-katia-ufjf-sistemas-de-avaliacao-da-educacao-basica>.
- Caed. (2008). Roteiro Básico para a Discussão e Apropriação dos Resultados. Recuperado em 30 fevereiro, 2016, de <http://www.avaliacaoexternasaerj.caedufjf.net/wp-content/uploads/2012/05/RoteiroBasicoDiscussaoApropriacaoResultadoSAERJ.pdf>
- Cardinet, J. (1986). A avaliação formativa, um problema atual. In L. Allal, J. Cardinet, & P. Perrenoud, A avaliação formativa num ensino diferenciado. *Coimbra: Livraria Almedina*.
- Cardinet, J. (1993). Avaliar é medir? *Rio Tinto: Edições ASA*.
- Carnoy, M., Khavenson, T., Fonseca, I., Costa, L., & Marotta, L. (2015). A educação brasileira está melhorando? Evidências do Pisa e do Saeb. *Cadernos de pesquisa*, 45(157), 450-485.
- Carnoy, M., & Loeb, S. (2002). Does External Accountability Affect Student Outcomes? A Cross-State Analysis. *Educational Evaluation and Policy Analysis*, 24(4), 305–331.
- Carnoy, M., Khavenson, T., Loyalka, P., Schmidt, W. H., & Zakharov, A. (2016). Revisiting the relationship between international assessment outcomes and educational production: Evidence from a longitudinal PISA-TIMSS sample. *American Educational Research Journal*, 53(4), 1054-1085.
- Carrasco, M. R., & Torrecilla, F. J. M. (2016). A avaliação das aprendizagens na América Latina. Comportamentos e tendências do desempenho escolar dos estudantes latino-americanos nos ensinos primário e secundário. *Sísifo*, (9), 31-46.
- Carvalho, C. P., Oliveira, A. C. P., & Lima, M. D. F. M. (2014). Avaliações externas: tensões e desafios para a gestão escolar. *Estudos em Avaliação Educacional*, 25(59), 50-76.
- Carvalho, E. (2014). Estudos comparados em educação: novos enfoques teórico-metodológicos. *Acta Scientiarum. Education*, 36(1), 129-141.
- Carvalho, L. M. (2006). Apontamentos sobre as relações entre conhecimento e política educativa. *Administração Educacional*, 6, pp. 36-45
- Castro, A. M. D. A. (2008). Administração gerencial: a nova configuração da gestão na América Latina. *RBPAAE*, 24(3), 289-406.

- Castro, M. H. G. (2009a). A consolidação da política de avaliação da educação básica no Brasil. *Meta: avaliação*, 1(3), 271-296.
- Castro, M. H. G. (2009b). Sistemas de avaliação da educação no Brasil: avanços e novos desafios. *São Paulo em Perspectiva*, 23(1), 5-18.
- Castro, M. H. G., Pestana, M. I. G.S, & Marte, M. A. S. (1995). O laboratório latino-americano de avaliação da qualidade da educação: um mecanismo de integração político-social. *Em aberto*, 68(15).
- Cardeira, D. G. S., & Almeida, B. A. (2012). A evolução das políticas de avaliação do RJ na visão dos professores: do Nova Escola ao SAERJ. *XVI ENDIPE - Encontro Nacional de Didática e Práticas de Ensino, Campinas*.
- Cardeira, D. G. S; Almeida, B. A, & Costa, M. (2014). Indicadores e avaliação educacional: percepções e reações a políticas de responsabilização. *Estudo em Avaliação Educacional*, 25(57), 198-225.
- Chamberlin, R., Wragg, T., Haynes, G., & Wragg, C. (2002). Performance-related pay and the teaching profession: A review of the literature. *Research Papers in Education*, 17(1), 31-49.
- Churchill, G. A. (2006). Marketing research: methodological foundations. *New York: Dryden Press*.
- Cimbricz, S. (2002). State-mandated testing and teachers' beliefs and practice. *Education policy analysis archives*, 10(2).
- Clímaco, M. C. (2005). Avaliação de Sistemas em Educação. *Lisboa: Universidade Aberta*.
- Coelho, P. S., & Esteves, S. P. (2007). The choice between a fivepoint and a ten-point scale in the framework of customer satisfaction measurement. *International Journal of Market Research*, 49(3), 313-339.
- Cohen, L., Manion, L., & Morrison, K. (2000). Research methods in education. *London/New York: Routledge/Falmer*.
- Comar, S. R. (2017). Avaliação no contexto do Laboratório Latino-Americano de Avaliação da Qualidade da Educação. *Estudos em Avaliação Educacional*, 28(68), 572-598.
- Costin, Cláudia. (2009, setembro 21). No Rio, professores e secretaria trabalham juntos. (Entrevista com Marcus Tavares). Revista Ponto Com. Recuperado em 03 julho, 2018, de <http://revistapontocom.org.br/entrevistas/%e2%80%9cprofessores-devem-ter-uma-critica-aqui-um-elogio-ali-mas-e-um-trabalho-conjunto%e2%80%9d>.
- Creswell, J. W. (2010). Projeto de pesquisa métodos qualitativo, quantitativo e misto. Porto Alegre: *Artmed*.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin*, 52(4), 281.
- Cunha, L. A. (1995). Os males do ziguezague. *Presença Pedagógica*, (6), 5-15.
- Dalben, A., & Almeida, L. C. (2015). Para uma avaliação de larga escala multidimensional. *Estudos em Avaliação Educacional*, 26(61), 12-28.
- Dantas, L. M. V. (2009). As contribuições das políticas de avaliação educacional em larga escala: o caso da avaliação de aprendizagem na Bahia. Tese de doutorado, Universidade Federal da Bahia, Bahia, BA, Brasil.
- Darling-Hammond, L. (1991). The implications of testing policy for quality and equality. *The Phi Delta Kappan*, 73(3), 220-225.
- Darling-Hammond, L. (1994). Performance-Based Assessment and Educational Equity. *Harvard Educational Review*, 64(1), 5-30.

- Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education policy analysis archives*, 8(1).
- Denzin, N. K. (1978). Triangulation: A case for methodological evaluation and combination. *Sociological methods*, 339-357.
- Denzin, N. K., Lincoln, Y. S., & Giardina, M. D. (2006). Disciplining qualitative research. *International journal of qualitative studies in education*, 19(6), 769-782.
- Depresbiteris, L. (2001). A avaliação na educação básica: ampliando a discussão. *Estudos em avaliação educacional*, (24), 137-146.
- Díaz-Barriga, A. (2005). TIMSS Y PISA. Dos proyectos internacionales de evaluación del aprendizaje escolar en ciencias. *Revista Eureka sobre enseñanza y divulgación de las ciencias*, 282-301.
- Díaz-Barriga, A. (2018). A prova PISA: idealização, cidadania global, imposição cultural e ausência de impacto pedagógico didático. In Ortigão, M.I.R. (Org.) Políticas de avaliação, currículo e qualidade: diálogos sobre o Pisa. *Curitiba: CRV*, 19-38.
- Dias Sobrinho, J. (2004). Avaliação ética e política em função da educação como direito público ou como mercadoria? *Revista Educação e Sociedade*, 25(88), 703-725.
- Dias Sobrinho, J. (2008). Qualidade, avaliação: do SINAES a índices. *Avaliação: Revista da Avaliação da Educação Superior*, 13(3).
- Dickel, A. (2010). O impacto do PISA na Produção Acadêmica Brasileira: contribuições para a discussão do currículo escolar. *Educação: teoria e prática*, 20(35), 201-201.
- Dourado, L. F. (2007). Políticas e gestão da educação básica no Brasil: limites e perspectivas. *Educação & Sociedade*, 28(100), 921-946.
- Dourado, L. F., & Oliveira, J. F. D. (2009). A qualidade da educação: perspectivas e desafios. *Cadernos Cedes*, 29(78), 201-215.
- Duarte, R. (2004). Entrevistas em pesquisas qualitativas. *Educar em Revista*, (24), 213-225.
- Duboc, A. P. M. (2007). Um olhar sobre o conceito de letramento do PISA e suas possíveis implicações para o ensino de línguas. *Revista CROP-Revista de Estudos Linguísticos e Literários em Inglês*, (12), 241-262.
- Eberts, R., Hollenbeck, K., & Stone, J. (2000). Teacher Performance Incentives and Student Outcomes. *Upjohn Institute Staff Working Paper*, 913-927.
- Enguita, M. F. (2001). O discurso da qualidade e a qualidade do discurso. In P. Gentili & T. T. da Silva (Org.). Neoliberalismo, Qualidade Total e Educação: visões. *Petrópolis: Vozes*, 10, 93-110.
- Esteban, M. T. (2009). Provinha Brasil: desempenho escolar e discursos normativos sobre a infância. *Sisifo*, (9), 47-56.
- Esteban, M. T., & Fetzner, A. R. (2015). A redução da escola: a avaliação externa e o aprisionamento curricular. *Educar em Revista*, (1), 75-92.
- Evers, W. M., & Walberg, H. J. (Eds.). (2002). School accountability. Stanford, CA: *Hoover Institution Press*.
- Fernandes, D. (2007). Percursos e desafios da avaliação contemporânea. Lição síntese apresentada nas provas de agregação. *Universidade de Lisboa. Faculdade de Psicologia e de Ciências da Educação*.
- Fernandes, D. (2008). Avaliação das Aprendizagens: Desafios às Teorias, Práticas e Políticas. *Lisboa: Texto Editores*.

- Fernandes, D. (2010a). Acerca da articulação de perspectivas e da construção teórica em avaliação educacional. In M. T. Esteban, & A. J. Afonso, *Olhares e interfaces: Reflexões críticas sobre a avaliação* (pp. 15-44). São Paulo: Cortez.
- Fernandes, D. (2010b). Reflexões acerca das relações entre os estudos internacionais de avaliação das aprendizagens e as políticas educativas. *Impacto das avaliações internacionais nos sistemas educativos*, 395-412.
- Fernandes, D. (2013). Avaliação em Educação: uma discussão de algumas questões críticas e desafios a enfrentar nos próximos anos. *Revista Ensaio*, 21(78), 9-32.
- Fernandes, D., Rodrigues, P. e Nunes, C. (2012). Uma investigação em ensino, avaliação e aprendizagens no ensino superior. In C. Leite e M. Zabalza (cords.), *Ensino superior: Inovação e qualidade na docência*, pp. 932 - 944. Porto: Centro de Investigação e Intervenção Educativas da Faculdade de Psicologia e de Ciências da Educação da Universidade do Porto.
- Firestone, W. A., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational evaluation and policy analysis*, 20(2), 95-113.
- Fletcher, P. R. (1995). Propósitos da Avaliação Educacional: uma análise das alternativas. *Estudos em avaliação educacional*, (11), 93-112.
- Fonseca, G. L. B. D. (2010). Qualidade dos indicadores educacionais para avaliação de escolas e redes públicas de ensino básico no Brasil. Tese de doutorado, Universidade Federal de Juiz de Fora, Minas Gerais, JF, Brasil.
- Fontanive, N. S. (2013). A divulgação dos resultados das avaliações dos sistemas escolares: limitações e perspectivas. *Ensaio: Avaliação e políticas públicas em Educação*, 21(78), 83-100.
- Fontanive, N. S., Elliot, L. G., & Klein, R. (2007). Os desafios da apresentação dos resultados da avaliação de sistemas escolares a diferentes públicos. REICE. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 5(2e), 262-273.
- Freitas, L. C. (2007). Eliminação adiada: o ocaso das classes populares no interior da escola e a ocultação da (má) qualidade do ensino. *Educação & Sociedade*, 28(100), 965-987.
- Freitas, L. C. (2016). A importância da avaliação e seus desafios: em defesa de uma responsabilização participativa. *Em Aberto*, 29(96).
- Gatti, B. A. (2008). Análise das políticas públicas para formação continuada no Brasil, na última década. *Revista Brasileira de educação*, 13(37), 57-70.
- Gatti, B. A. (2009). Avaliação de sistemas educacionais no Brasil. *Sísifo*, (9), 7-18.
- Gatti, B. A. (2013). Políticas de avaliação em larga escala e a questão da inovação educacional. *Série-Estudos-Periódico do Programa de Pós-Graduação em Educação da UCDB*, (33).
- Gatti, B. A. (2014). Avaliação: contexto, história e perspectivas. *Olhares: Revista do Departamento de Educação da Unifesp*, 2(1), 08-26.
- Gatti, B., Vianna, G., & Davis, C. (1991). Problemas e Impasses da Avaliação de Projetos e Sistemas Educacionais: Dois Casos Brasileiros. *Estudos Em Avaliação Educacional*, (4), 7-26.
- Gil, C. A. (2008). Métodos e Técnicas de Pesquisa Social. *São Paulo: Atlas*.
- Godinho, J. D., & Farias, M. E. (2013). Inserção de ciências visando a compatibilização do SAEB com o PISA e seus reflexos nas Políticas Educacionais. Atas do IX Encontro Nacional de Pesquisa em Educação em Ciências, 1-8.
- Goldenberg, M. (2011). A arte de pesquisar. *Rio de Janeiro: Record*.
- Greene, J. G. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review*, 12(2), 91-116.

- Guba, E. G. e Lincoln, Y. S. (1989). Fourth generation evaluation. *London: SAGE Publications*.
- Hadji, C. (2001). *Avaliação desmistificada*. Porto Alegre: Artmed Editora.
- Hanushek, E. A., & Raymond, M. E. (2005). Does school accountability lead to improved student performance?. *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management*, 24(2), 297-327.
- Harlen, W. (2001). The assessment of scientific literacy in the OECD/PISA project. In H. Behrendt, H. Dahncke, R. Duit, W. Graber, M. Komorek & A. Kross, Eds, *Research in Science Education – past, present and future* (pp. 49–60). Dordrecht: The Netherlands: Kluwer Academic Publishers.
- Harlen, W., & Gardner, J. (2010). Assessment to support learning. In J. Gardner, W. Harlen, L. Hayward, G. Stobart, & M. Montgomery (Eds.), *Developing teacher assessment* (pp. 15–28). New York, NY: Open University Press.
- Harlen, W., & James, M. (1997). Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education: Principles, Policy & Practice*, 4(3), 365-379.
- Harvey, L. (2006). Understanding quality. *EUA Bologna Handbook: Making Bologna work*, 1(1), 1-29.
- Haney, W., & Madaus, G. (1986). Effects of Standardized Testing and The Future of The National Assessment of Educational Progress. Working paper for the NAEP. Recuperado em 30 julho, 2018, de <https://eric.ed.gov/?id=ED279680>.
- Herman, J. L., & Golan, S. (1993). The effects of standardized testing on teaching and schools. *Educational measurement: Issues and practice*, 12(4), 20-25.
- Heubert, J. P., & Hauser, R. M. (1999). High stakes: Testing for tracking, promotion, and graduation. *A report of the National Research Council*. Washington, DC: National Academy Press.
- Hoffmann, J. (2006). Avaliação mediadora: uma prática em construção da pré-escola à universidade, 11. Porto Alegre: *Educação e Realidade*.
- Hong, W. P., & Youngs, P. (2008). Does high-stakes testing increase cultural capital among low-income and racial minority students?. *Education policy analysis archives*, 16(6).
- Horta Neto, J. L. (2007). Um olhar retrospectivo sobre a avaliação externa no Brasil: das primeiras medições em educação até o Saeb de 2005. *Revista Iberoamericana de Educación*, 42(5), 1-13.
- Horta Neto, J. L. (2010). Avaliação externa de escolas e sistemas: questões presentes no debate sobre o tema. *Revista Brasileira de Estudos Pedagógicos*, 91(227), 84-104.
- House, E. R., & Howe, K. R. (2000). Deliberative democratic evaluation in practice. In *Evaluation models* (pp. 409-421). Springer, Dordrecht.
- House, E. R., & Howe, K. R. (2003). Deliberative democratic evaluation. In *International handbook of educational evaluation* (pp. 79-100). Springer, Dordrecht.
- Hutchison, D., & Schagen, I. (2007). Comparisons between PISA and TIMSS: Are we the man with two watches. *Lessons learned: What international assessments tell us about math achievement*, 227-261.
- Iaies, G. (2003). Evaluar las evaluaciones. IIPE-UNESCO, Evaluar las evaluaciones. Una mirada política acerca de las evaluaciones de la calidad educativa (15-35). Buenos Aires: IIPEUNESCO.
- Inep/PISA. (2015). Brasil no PISA 2015. Brasília.
- Inep/PISA. (2001). PISA 2000: Relatório Nacional. Brasília.

- Inep/Pnud. (1992). Relatório do Sistema de Avaliação da Educação Básica – ciclo 1990. Brasília.
- Inep. (n.d.). Aneb e Anresc (Prova Brasil) - Inep. Portal.inep.gov.br. Brasília.
- Ingram, D., Louis, K. S., & Schroeder, R. G. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers college record*, 106(6), 1258-1287.
- Johns, R. (2005). One size doesn't fit all: Selecting response scales for attitude items. *Journal of elections, public opinion & parties*, 15(2), 237-264.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational researcher*, 33(7), 14-26.
- Johnson, K., Greenesid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on evaluation use: A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation*, 30(3), 377-410.
- Kamens, D. H., & McNeely, C. L. (2009). Globalization and the growth of international educational testing and national assessment. *Comparative education review*, 54(1), 5-25.
- Kazamias, A. M. (2012). Educação comparada: uma reflexão histórica. *Educação comparada: panorama internacional e perspectivas. Brasília: Unesco, Capes*, 1, 173-193.
- Kellaghan, T., & Madaus, G. (2003). External (public) examinations. In *International handbook of educational evaluation* (pp. 577-600). Springer, Dordrecht.
- Kellaghan, T. e Madaus, G. (2000). Outcome evaluations. In D. L. Stufflebeam; G. Madaus; & T. Kellaghan. (Eds), *Evaluation Models: viewpoints on educational and human services evaluation*. Dordrecht: Kluwer.
- Komorita, S. S. (1963). Attitude content, intensity, and the neutral point on a Likert scale. *Journal of Social Psychology*, 61(2), 327-334.
- Kooiman, J. (Ed.). (1993). Modern governance: new government-society interactions. *Calif: Sage*.
- Koretz, D. M., McCaffrey, D. F., & Hamilton, L. S. (2001). Toward a framework for validating gains under high-stakes conditions. Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education & Information Studies, University of California, Los Angeles.
- Lange, J. (2006). Mathematical literacy for living from OECD-PISA perspective. Recuperado em 30 julho, 2019, de <http://citeseerx.ist.psu.edu/viewdoc/similar?doi=10.1.1.500.5156&type=sc>
- Lee, J. (2008). Is test-driven external accountability effective? Synthesizing the evidence from cross-state causal-comparative and correlational studies. *Review of educational research*, 78(3), 608-644.
- Leviton, L.C., & Hughes, E.F.X. (1981). Research on the utilization of evaluations: A review and synthesis. *Evaluation Review*, 5, 525-549.
- Libânio, J. C. (2013). Didática. São Paulo: *Cortez Editora*.
- Lingard, Bob. (2016). PISA: Fundamentações para participar e acolhimento político. *Educação & Sociedade*, 37(136), 609-627.
- Linn, R. L. (2000). Assessments and accountability. *Educational researcher*, 29(2), 4-16.
- Looney, J. (2009). Assessment and Innovation in Education. *OECD-Education Working Papers*, (24), OECD Publishing, Paris.
- Löwy, M. (2013). As aventuras de Karl Marx contra o Barão de Münchhausen: marxismo e o positivismo na sociedade do conhecimento. *São Paulo: Cortez*.

- Luckesi, C. C. (1998). Verificação ou avaliação: o que pratica a escola. *Série Idéias*, (8), 71-80.
- Luckesi, C. C. (2014). *Avaliação da aprendizagem escolar: estudos e proposições*. São Paulo: Cortez editora.
- Lüdke, M., & André, M. E. (1986). Evolução da pesquisa em Educação. In. *Lüdke, M.; André, M. Pesquisa em Educação: abordagens qualitativas*. São Paulo, SP: EPU.
- Luxia, Q. (2007). Is testing an efficient agent for pedagogical change? Examining the intended washback of the writing task in a high-stakes English test in China. *Assessment in Education*, 14(1), 51-74.
- Madaus, G., & Russell, M. (2010). Paradoxes of high-stakes testing. *Journal of Education*, 190(1-2), 21-30.
- Madaus, G., Russell, M., & Higgins, J. (2009). The paradoxes of high stakes testing: How they affect students, their parents, teachers, principals, schools, and society. *Charlotte: Information Age*.
- Madaus, G. F.; Stufflebeam, D. L. (2000). Program Evaluation: a historical overview. In D. Stufflebeam, G. Madaus, & T. Kellaghan. (Eds). *Evaluation Models: viewpoints on educational and human services evaluation*. Dordrecht: Kluwer.
- Marchesi, A. (2002). Mudanças educativas e avaliação das escolas. In: Azevedo, J. (Org.). *Avaliação das escolas: consensos e divergências*. Porto: Edições ASA. pp. 33-49.
- McDonnell, L. M. (1994). Assessment Policy as Persuasion and Regulation. *American Journal of Education*, 102(4), 394-420.
- MEC. (2018). Ministro da Educação anuncia mudanças no sistema de avaliação a partir de 2019. Recuperado em 30 janeiro, 2019, de <http://portal.mec.gov.br/ultimas-noticias/211-218175739/65791-ministro-da-educacao-anuncia-mudancas-no-sistema-de-avaliacao-a-partir-de-2019>.
- MEC/Inep. (1994). Saeb: Sistema Nacional de Avaliação da Educação Básica – objetivos, diretrizes, produtos e resultados. Recuperado em 25 agosto, 2016 de <http://www.dominiopublico.gov.br/download/texto/me002195.pdf>
- Mendes, J. C. B., & Frangella, R. D. C. P. (2014). Avaliação: te olhei por um monóculo e enxerguei a prova. *Roteiro*, 39(2), 331-346.
- Menegão, R. D. C. S. G. (2015). Impactos da avaliação externa no currículo escolar: percepções de professores e gestores. Tese de doutorado, Universidade Estadual de Campinas, São Paulo, Brasil.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American psychologist*, 50(9), 741.
- Merriam, S. B. (1998). *Qualitative Research and Case Study Applications in Education. Revised and Expanded from "Case Study Research in Education"*. Jossey-Bass Publishers, 350 Sansome St, San Francisco, CA 94104.
- Michel, A. (2017). The contribution of PISA to the convergence of education policies in Europe. *European Journal of Education*, 52(2), 206-216.
- Miller, D. M., & Linn, R. L. (2000). Validation of performance-based assessments. *Applied Psychological Measurement*, 24(4), 367-378.
- Minayo, M. C. D. S. (1994). Pesquisa Social: Teoria, método e criatividade. *Vozes, Petrópolis/RJ*.
- Minayo, M. C. D. S., & Sanches, O. (1993). Quantitativo-qualitativo: oposição ou complementaridade?. *Cadernos de saúde pública*, 9, 237-248.

- Minayo, M. C. S., Assis, S. G., & Souza, E. R. (2005). *Avaliação por triangulação de métodos: abordagem de programas sociais*. SciELO-Editora FIOCRUZ.
- Moe, T. M. (2003). Politics, control, and the future of school accountability. In P. E. Peterson & M. R. West (Eds.). *No child left behind? The politics and practice of school accountability* (pp. 80-106). Washington, DC: Brookings Institution Press.
- Moita Lopes, L. P. D. (1994). Pesquisa interpretativista em Linguística Aplicada: a linguagem como condição e solução. *Delta*, 10(2), 329-338.
- Mons, N. (2009). Effects théoriques et réels des politiques d'évaluation standardisée. *Revue française de pédagogie* 169, 99-140.
- Moreno, A. P. (2016). O sistema permanente de apropriação dos resultados do SADEAM. Dissertação de mestrado, Universidade Federal de Juiz de Fora, Minas Gerais, JF, Brasil.
- Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist*, 22(2), 155-175.
- Nevo, D. (1998). Avaliação por diálogos: uma contribuição possível para o aprimoramento escolar. In: *Seminário internacional de avaliação educacional*, Rio de Janeiro. Anais... Brasília, DF: Inep, pp. 89-97.
- Nichols, S., Glass, G., & Berliner, D. (2012). High-stakes testing and student achievement: Updated analyses with NAEP data. *Education policy analysis archives*, 20(20).
- Niemann, D., Martens, K., & Teltemann, J. (2017). PISA and its consequences: Shaping education policies through international comparisons. *European Journal of Education*, 52(2), 175-183.
- Noizet, G., & Carveni, J-P. (1985). *Psicologia da avaliação escolar*. Coimbra: Coimbra Editora.
- Novaes, L. C. (2014). Os impactos da avaliação externa sobre o trabalho de professores na rede estadual paulista. *Roteiro*, 39(2), 283-310.
- Nóvoa, A. (2009). Modelos de análise em educação comparada: o campo e o mapa. In D. B. Souza, & S. A. Martínez. (Orgs), *Educação Comparada: Rotas de além-mar*. São Paulo: EJR Xamã Editora.
- Perez, J. R. R., & Passone, E. F. (2007). A perspectiva política da educação comparada e as avaliações internacionais da qualidade da educação. *Políticas Educativas*, 1(1).
- Oliveira, R. P., Bauer, A., Ferreira, M. P., Minuci, E. G., Lisauskas, F., Carvalho, M. X., ... & Caldeira, C. (2013). Análise das desigualdades intraescolares no Brasil. *Estudos e Pesquisas Educacionais*, 4, pp. 19-35.
- Oliveira, R. P., & de Araújo, G. C. (2005). Qualidade do ensino: uma nova dimensão da luta pelo direito à educação. *Red Revista Brasileira de Educação*. *Revista Brasileira de Educação*, 28, pp. 5-23.
- Orlandi, E. P. (1995). Texto e Discurso. *Revista Organon* 23, 9(23).
- Ortigão, X. (2018). Pisa 2012: uma análise dos itens de matemática. In Ortigão, M.I.R. (Org.) Políticas de avaliação, currículo e qualidade: diálogos sobre o Pisa. *Curitiba: CRV*, 19-38.
- Ozga, J. (2000). *Investigação sobre políticas educacionais: terreno de contestação*. Porto: Porto Editora.
- Pacheco, J. (1994). *A Avaliação dos Alunos na Perspectiva da Reforma*. Porto: Porto Editora.
- Pacheco, J. A. (2012). Avaliação das aprendizagens: políticas formativas e práticas sumativas. *Encontros de Educação*, 1-9.
- Pasquali, L. (2009). Psicometria. *Revista Escola de Enfermagem USP*, 43, 992-999.

- Patton, M.Q. (1997). *Utilization-focused evaluation: The New Century text*. Thousand Oaks, CA: Sage.
- Patton, M. Q. (2003). *Utilization-Focused Evaluation*. In T. Kellaghan, D. L. Stufflebeam, & Lori A. Wingate, (Eds.), *International handbook of educational evaluation*. Chicago: Kluwer Academic Publishers.
- Patton, M. Q. (2005). The challenges of making evaluation useful. *Ensaio: avaliação e políticas públicas em educação*, 13(46), 67-78.
- Perrenoud, P. (1999). Avaliação: da excelência à regulação das aprendizagens – entre duas lógicas. *Porto Alegre: Artes Médicas Sul*.
- Pestana, M. I. (1998). O sistema de avaliação brasileiro. In *Seminário Internacional de Avaliação Educacional* (pp. 15-30). Brasília: MEC/INEP/Unesco.
- Pilatti, L. A., Pedroso, B., & Gutierrez, G. L. (2010). Propriedades psicométricas de instrumentos de avaliação: um debate necessário. *Revista Brasileira de Ensino de Ciência e Tecnologia*, 3(1).
- Plank, D. N. (2001). Política Educacional no Brasil: Caminhos para a salvação. *Porto Alegre: Artmed Editora*.
- Plomp, T. (1998). O potencial dos estudos comparativos internacionais para o controle da qualidade da educação. In *Seminário Internacional de Avaliação Educacional* (pp. 41-52). Brasília: MEC/INEP/Unesco.
- Polato, A. (2014). A relação entre as avaliações em larga escala e a organização do trabalho escolar em duas escolas públicas estaduais do interior de São Paulo. Dissertação de mestrado, Universidade Estadual Paulista Júlio Mesquita Filho, São Paulo, SP, Brasil.
- Ravela, P., Arregui, P., Valverde, G., Wolfe, R., Ferrer, G., Martínez Rizo, F., ... & Wolff, L. (2008). Las evaluaciones educativas que América Latina necesita. *RIEE. Revista Iberoamericana de Evaluación Educativa*, 1(1), pp. 51-63.
- Ravitch, D. (2009). "Time to Kill" No Child Left Behind". *Education Digest: Essential Readings Condensed for Quick Review*, 75(1), 4-6.
- Reboloso, E. et al. (2002). Metaevaluation of a total quality management evaluation system. *Psychology in Spain*, 6(1), 12-25.
- Ribeiro, L.C. (1991). Avaliação da aprendizagem. *Lisboa: Texto Editora*.
- Richards, D., & Smith, M. J. (2002). Governance and public policy in the United Kingdom. New York: *Oxford University Press*.
- Rio de Janeiro. (2018). *Educação em Números - prefeitura.rio*. *Rio.rj.gov.br*. Recuperado em 20 dezembro, 2018, de <http://www.rio.rj.gov.br/web/sme/educacao-em-numeros>
- Rio de Janeiro. (2013). Portaria Seeduc/Sugen nº 419 de 27 de setembro. Recuperado em 13 maio, 2017, de <http://www.jusbrasil.com.br/diarios/59711674/doerj-poder-executivo-30-09-2013-pg-30/pdfView>
- Rio de Janeiro. (2012). Revista da Gestão Escolar – Sistema de Avaliação da Educação do Estado do Rio de Janeiro. Recuperado em 23 março, 2017, de http://www.avaliacaoexternasaerj.caedufjf.net/wp-content/uploads/2013/07/SAERJ_GESTAO_2012.pdf
- Rio de Janeiro. (2010). Revista do Sistema de Avaliação: Saerj – 2010. Recuperado em 23 março, 2017, de <http://www.avaliacaoexternasaerj.caedufjf.net/revistas-pedagogicas/2010-2/>.
- Rio de Janeiro. (2008) Revista do Professor de Avaliação da Educação: Saerj – 2008. Recuperado em 23 março, 2017, de <http://www.avaliacaoexternasaerj.caedufjf.net/revistas-pedagogicas/2008-3/>.

- Rio de Janeiro. (n.d.a). *Informações sobre a Avaliação Bimestral*. Recuperado em 29 dezembro, 2015, de <http://download.rj.gov.br/documentos/10112/451413/DLFE-39734.pdf/Saerjinhoinformacoes.pdf>
- Rio de Janeiro. (n.d.b). Jovens Turistas - o projeto. Jovens Turistas. Recuperado em 27 maio, 2016, de <http://www.jovensturistas.com.br/o-projeto/>
- Risopatrón, V. (2007). *El concepto de la Calidad de la Educación*. Santiago do Chile: Unesco; Orealc.
- Ristoff, D. I. (1995). Avaliação institucional: pensando princípios. *Avaliação institucional: teoria e experiências*. São Paulo: Cortez, 180, 37-51.
- Rodrigues, J. D. Z. (2018). Gerencialismo e responsabilização: repercussões para o trabalho docente nas escolas estaduais de ensino médio de Campinas/SP. Tese de doutorado, Universidade do Estado de Campinas, São Paulo, Brasil.
- Rodrigues, R. F. (2011). Usos e repercussões de resultados do SARESP na opinião de professores da rede estadual paulista. Dissertação de mestrado, Pontifícia Universidade Católica, São Paulo, SP, Brasil.
- Rojas, H. D. L. S. (2007). Formação do professor do Ensino Básico e a Avaliação educacional. *Estudos em Avaliação Educacional*, 18(37), 7-40.
- Rosistolato, R., & Prado, A. P. (2014). Os profissionais da educação e as avaliações externas de aprendizagem: uma comparação entre rio de janeiro e duque de Caxias. *Roteiro*, 39(2), 311-330.
- Rosistolato, R., & Viana, G. (2014). Os gestores educacionais e a recepção dos sistemas externos de avaliação no cotidiano escolar. *Educação e Pesquisa*, 40(1), 13-28.
- Rua, M. G. (2012). *Políticas Públicas*. Florianópolis. Recuperado em 15 fevereiro, 2016 de https://www.academia.edu/11259556/Políticas_Publicas_-_Maria_das_Gra%C3%A7as_Rua
- Sá, V. I. M. (2008). O discurso da qualidade no contexto da recomposição das desigualdades em educação. *Revista Brasileira de Estudos Pedagógicos*, 89(223).
- Sadler, D. R. (1989) Formative assessment and the design of instructional systems. *Springer*, 18(2), 119-144.
- Santana, A. D. C. M., & Rothen, J. C. (2015). A avaliação externa das escolas e a formação continuada de professores: o caso paulista. *Revista Diálogo Educacional*, 15(44), 89-110.
- Santos, A. F. (2015). Os Questionários Contextuais do SAEB 2013 e do PISA 2012: uma comparação dos elementos constituintes do Ensino. *Reuniões da ABAVE*, (8), 453-486.
- Santos, L. (2016). A articulação entre a avaliação somativa e a formativa, na prática pedagógica: uma impossibilidade ou um desafio?. *Ensaio: Avaliação e Políticas Públicas em Educação*, 24(92), 637-669.
- Saraiva, E. (2006). Introdução à teoria da política pública. In E. Saraiva; & E. Ferrareli. (Orgs.), *Políticas Públicas*. (pp. 21-42). Brasília: ENAP.
- Scaramucci, M. V.R. (2004). Efeito retroativo da avaliação no ensino/aprendizagem de línguas: o estado da arte. *Trabalho em Linguística Aplicada*, 43(2), 203-226.
- Schwandt T. A. (2006). Três posturas epistemológicas para a investigação qualitativa. In: Denzin N, Lincoln Y, organizadores. *O planejamento da pesquisa qualitativa - teorias e abordagens*. 2ª Ed. Porto Alegre: Artmed Bookman. pp. 193-218.
- Secchi, L. (2009). Modelos organizacionais e reformas da administração pública. *RAP - Revista de Administração Pública*, Rio de Janeiro, 43(2), 347-69.
- Sepe-Rio. (2019). Avaliação externa adapta escola à padronização e controle empresariais. Recuperado em 20 julho, 2019, de http://www.seperj.org.br/ver_noticia.php?cod_noticia=21447.

- Seeduc. (2013). Currículo Mínimo 2013. Recuperado em 4 agosto, 2019, de http://historiadabncc.mec.gov.br/documentos/CURRICULOS/Rio_de_Jneiro_Curriculo_Minimo_2013_Arte_Livro.pdf.
- Seeduc. (n.d.a) *Saerj/Saerjinho/Iderj - SEEDUC. Rj.gov.br*. Recuperado em 3 novembro, 2015, de <http://www.rj.gov.br/web/seeduc/exibeconteudo?article-id=843535>
- Seeduc. (n.d.b). *Secretaria de Estado de Educação - SEEDUC. Rj.gov.br*. Recuperado em 14 Outubro, 2015, de <http://www.rj.gov.br/web/seeduc/exibeconteudo?article-id=612060>
- Seeduc. (n.d.c). *Avaliação bimestral monitora o desempenho escolar dos alunos da rede estadual. Rj.gov.br*. Recuperado em 17 dezembro, 2015, de <http://www.rj.gov.br/web/seeduc/exibeconteudo?article-id=2397680>
- Seeduc. (n.d.d) *Docentes participam de oficina de elaboração de itens. Rj.gov.br*. Recuperado em 10 dezembro, 2015, de <http://www.rj.gov.br/web/seeduc/exibeconteudo?article-id=689382>
- Seeduc. (n.d.e). *Saerjinho conquista prêmio de gestão do BID. Rj.gov.br*. Recuperado em 19 dezembro, 2015, de <http://www.rj.gov.br/web/imprensa/exibeconteudo?article-id=2687289>
- Seeduc. (n.d.f). *Seeduc repudia atitudes que prejudicam alunos e professores. Rj.gov.br*. Recuperado em 14 dezembro, 2015, de <http://www.rj.gov.br/web/seeduc/exibeconteudo?article-id=519146>
- Seeduc. (n.d.g). *Seeduc apresenta ações já desenvolvida na rede, diante do resultado da pesquisa realizada pela Fundação Lemann. Rj.gov.br*. Recuperado em 14 dezembro, 2015, de <http://www.rj.gov.br/web/seeduc/exibeconteudo?article-id=2350652>
- Sellar, S., & Lingard, B. (2014). The OECD and the expansion of PISA: New global modes of governance in education. *British Educational Research Journal*, 40(6), 917-936.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 1-12.
- Shepard, L. A. (1990). Inflated test score gains: Is the problem old norms or teaching the test?. *Educational Measurement: Issues and Practice*, 9(3), 15-22.
- Shulha, L. M., & Cousins, J. B. (1997). Evaluation use: Theory, research, and practice since 1986. *Evaluation practice*, 18(3), 195-208.
- Silva, V. G., Gimenes, N. A., Moriconi, G. M., & Louzano, P. (2013). Uso da avaliação externa por equipes gestoras e profissionais docentes: um estudo em quatro redes de ensino público. *São Paulo: FCC/SEP*, 10-116.
- Stecher, B. M. (2002). Consequences of large-scale, high-stakes testing on school and classroom practice. *Making sense of test-based accountability in education*, 79-100.
- Silva, V. G. (2008). A narrativa instrumental da qualidade na educação. *Estudos em avaliação educacional*, 19(40), 191-221.
- Smith, M. L. (1991). Put to the Test: The Effects of External Testing on Teachers. *Educational Researcher*, 20(5), 8-11.
- Soares, J. F. (2007). Melhoria do desempenho cognitivo dos alunos do ensino fundamental. *Cadernos de pesquisa*, 37(130), 135-160.
- Soares, J. F., & Candian, J. F. (2007). O efeito da escola básica brasileira: as evidências do PISA e do SAEB. *Revista Contemporânea de Educação*, 2(4), 163-181.
- Soares, S. S. D., & Nascimento, P. A. M. M.. (2012). Evolução do desempenho cognitivo dos jovens brasileiros no Pisa. *Cadernos de Pesquisa*, 42(145), 68-87.
- Sousa, S. Z. (2003). Possíveis impactos das políticas de avaliação no currículo escolar. *Cadernos de Pesquisa*, (119), 175-190.

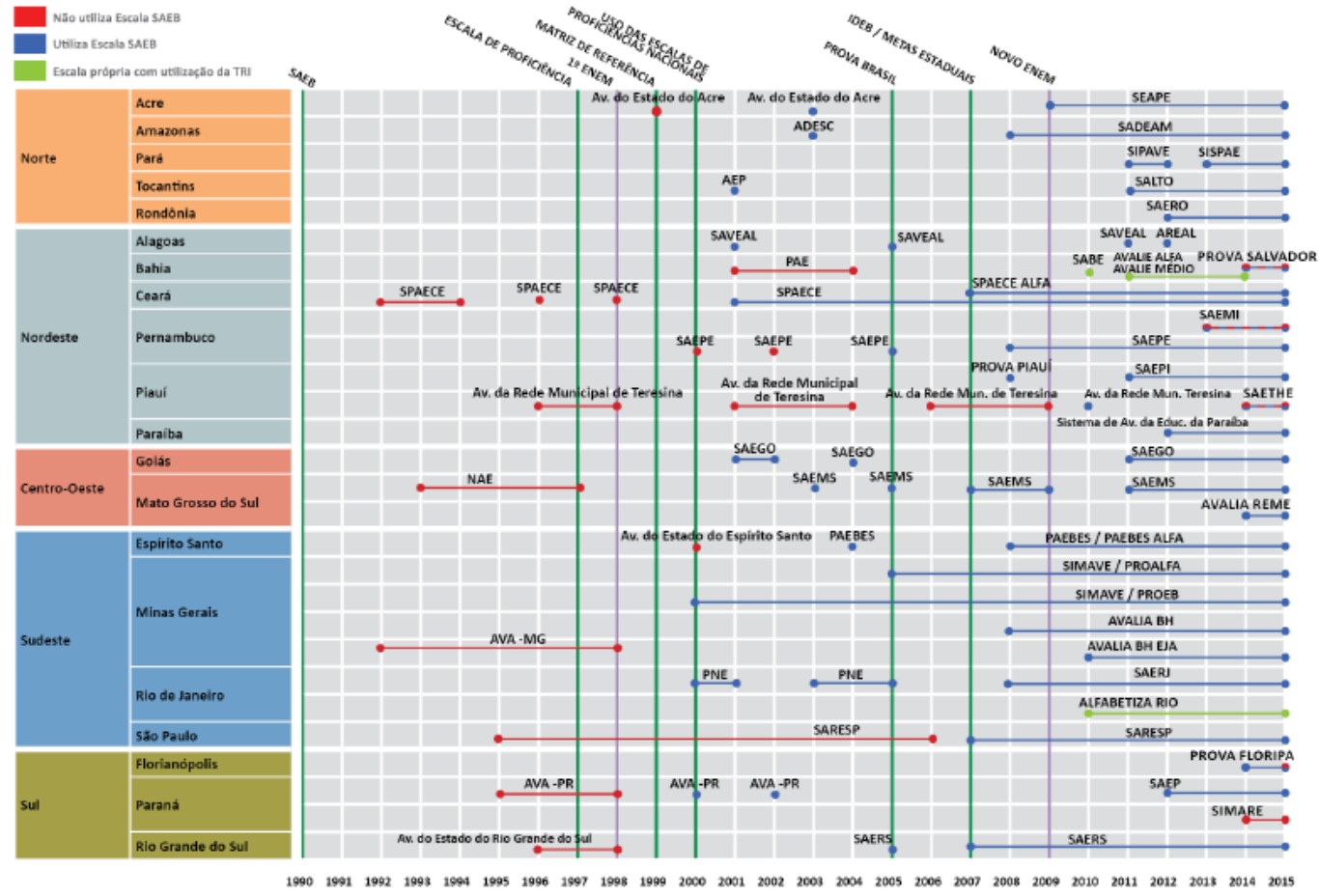
- Sousa, S. Z. (2008). Avaliação e carreira do magistério: premiar o mérito?. *Retratos da escola*, 2(2/3).
- Sousa, S. Z., & Arcas, P. H. (2010). Implicações da avaliação em larga escala no currículo: revelações de escolas estaduais de São Paulo. *Educação: teoria e prática*, 20(35), 181-181.
- Sousa, S. Z., & de Oliveira, R. P. (2010). Sistemas estaduais de avaliação: uso dos resultados, implicações e tendências. *Cadernos de Pesquisa*, 40(141), 793-822.
- Sousa, S. Z., Pimenta, C. O., & Machado, C. (2012). Avaliação e gestão municipal da educação. *Estudos em Avaliação Educacional*, 23(53), 14-36.
- Sousa, S. Z. (2014). Concepções de qualidade da educação básica forjadas por meio de avaliações em larga escala. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 19(2), 407-420.
- Souza, C. (2006). Políticas Públicas: uma revisão da literatura. *Sociologias*, (16), 20-45.
- Souza, E. R. D. (2009). Accountability de professores um estudo sobre o efeito da Prova Brasil em escolas de Brasília. Dissertação de mestrado, Universidade Estadual de Campinas, São Paulo, Brasil.
- Scriven, M. (2008). The Concept of a Transdiscipline: And of Evaluation as a Transdiscipline. *Journal of MultiDisciplinary Evaluation*, 5(10), 65-66.
- Spink, P. (1987). Continuidade e descontinuidade em organizações públicas: um paradoxo democrático. *Cadernos Fundap*, 7(13), 57-65.
- Stake, R. (2003) Responsive Evaluation. In: T. Kellaghan, & D. L. Stufflebeam. *International handbook of educational Evaluation*. (pp. 63-68). Boston: Kluwer Academic Publishers.
- Stufflebeam, D. (2000). Foundational Models for 21st Century Program Evaluation. In D. Stufflebeam, G. Madaus, & T. Kellaghan. (Eds). *Evaluation Models: viewpoints on educational and human services evaluation*. Dordrecht: Kluwer.
- Stufflebeam, D. L., & Shinkfield, A. J. (2007). Evaluation theory, models and applications. San Francisco, CA: *Jossey-Bass*.
- Tardif, Maurice. (2002). Saberes docentes e formação profissional. *Petrópolis*, RJ: Vozes.
- Tedesco, J. (2003). Evaluar las evaluaciones. In: *Evaluar las Evaluaciones: una mirada política acerca de las evaluaciones de la calidad educativa*. IPE – UNESCO Buenos Aires Argentina.
- Teodoro, A. (2001). Organizações internacionais e políticas educativas nacionais: A emergência de novas formas de regulação transnacional, ou uma globalização de baixa intensidade. In R. Stoer; L. Cortesão; J. Correia (Orgs), *Transnacionalização da educação: da crise da educação à "educação" da crise*. Porto: Edições Afrontamento.
- Tenório, R. M., & Araújo, M. D. (2016). Concepções de letramento matemático em avaliação educacional subjacentes ao PISA. *Revista Ciência (In) Cena*, (2), 66-82.
- Unesco-Santiago. (2008). SERCE: Segundo Estudio Regional Comparativo e Explicativo. Recuperado em 28 março, 2017, de <http://unesdoc.Unesco.org/images/0016/001609/160926POR.pdf>.
- Unesco-Santiago. (1998). Primer Estudio Internacional Comparativo sobre Lenguaje, Matemática y Factores Asociados en Tercero y Cuarto Grado. Recuperado em 28 março, 2017, de <http://unesdoc.Unesco.org/images/0012/001231/123130S.pdf>.
- Vasconcellos, C. S. (2010). Avaliação da aprendizagem: Práticas de mudança por uma práxis transformadora. *São Paulo: Libertad*.
- Vianna, H. M. (2002). Questões de avaliação educacional. *Avaliação: construindo o campo e a crítica*. Florianópolis: Insular.

- Vianna, H. M. (2005). Fundamentos de um programa de avaliação educacional. *Brasília: Líber Livro*.
- Vianna, H. M. (1983). Validade de construto em testes educacionais. *Educação e seleção*, (08), 35-44.
- Vieira, A. M. (2017). Acordes e dissonâncias do letramento científico proposto pelo PISA 2015. *Estudos em Avaliação Educacional*, 28(68), 478-510.
- Villas-Boas, A. (2006). Avaliação formativa e formação de professores: ainda um desafio. *Linhas Críticas*, 12(22), pp. 75-90.
- Villas-Boas, A. (2006). Avaliação formativa e formação de professores: ainda um desafio. *Linhas críticas*, 12(22), 159-179.
- Villas-Boas, A., Freitas, B. M., & Soares, S. L. (2016). O lugar da avaliação nos espaços de formação de professores. *Cadernos CEDES*, Campinas, 36(99), 239-254
- Vitelli, R. F., Fritsch, R., & Corsetti, B. (2018). Indicadores educacionais na avaliação da educação básica e possíveis impactos em escolas de Ensino Médio no município de Porto Alegre, Rio Grande do Sul. *Revista Brasileira de Educação*, 23.
- Waiselfisz, J. J. (2009). O ensino das ciências no Brasil e o PISA. São Paulo: *Sangari do Brasil*.
- William, D. (2006). Formative assessment: Getting the focus right. *Educational assessment*, 11(3-4), 283-289.
- Winters, M. A. R., & Sundararaman, V. (2008). The impact of performance pay for public school teachers: theory and evidence. *Harvard University*. (PEPG 08-15).
- Wolcott, H.F. (1994). Transforming qualitative data: Description, analysis, and interpretation. ThousandOaks, CA: *Sage*.
- Worthen B.R., Sanders J.R., Fitzpatrick J. (2004). Avaliação de programas: concepções e práticas. *São Paulo: Gente*.
- Wu, M. (2010). Comparing the similarities and differences of PISA 2003 and TIMSS. *OECD Education Working Papers*, 32, OECD Publishing.
- Yerly, G. (2014). Les effets de l'évaluation externe des acquis des élèves sur les pratiques des enseignants. Analyse du regard des enseignants du primaire. Tese de doutorado, Université de Fribourg, Fribourg, França.

Anexos

Anexo 1

Linha do Tempo | Avaliações Estaduais



Anexo 2



**INSTITUTO DE EDUCAÇÃO
DA UNIVERSIDADE DE LISBOA
COMISSÃO DE ÉTICA**

PARECER

A Comissão de Ética do Instituto de Educação da Universidade de Lisboa, tendo procedido à análise dos elementos relativos ao projeto de investigação de *Helen Vieira de Oliveira*, estudante do curso de Doutoramento em Educação, especialidade Avaliação em Educação, intitulado *Avaliações externas das aprendizagens e públicas: Uma discussão a partir das apreciações dos docentes*, considera que são respeitados os princípios éticos, bem como as orientações éticas para a investigação, expressos na Carta Ética para a Investigação em Educação e Formação do Instituto de Educação da Universidade de Lisboa.

Instituto de Educação da Universidade de Lisboa

28 de setembro de 2016

O Presidente,

Professor Feliciano H. Veiga

Alameda da Universidade
1649-023 Lisboa Portugal

T. +351 21 794 36 23
F. +351 21 793 34 09

geral@ieul.pt
www.ieul.pt



Anexo 3

Guião de Entrevista

Avaliação Externa das Aprendizagens

- 1- Em sua opinião, quais são as características que melhor definem avaliações externas?
- 2- Em que medida a avaliação externa pode ser relevante para o processo de ensino e aprendizagem?
- 3- Qual é, na sua opinião, a credibilidade que os políticos atribuem às avaliações externas?
- 4- Em que medida é que a credibilidade das avaliações externas varia se considerarmos diferentes intervenientes? (Por exemplo, será que o público em geral as considera mais credíveis do que os docentes ou os académicos?)
- 5- Em sua opinião, as avaliações externas são construídas com qualidade, permitindo, no que se refere às questões formuladas, conhecer bem o que os alunos sabem e são capazes de fazer?
- 6- Na sua opinião, para que servem as avaliações externas? Qual seria, dentre estas, sua utilidade principal?
- 7- Qual é, na sua opinião, a principal utilização que os docentes dão aos resultados das avaliações externas?
- 8- Como é que as equipas gestoras das unidades escolares utilizam os resultados das avaliações externas?
- 9- Como é que os decisores políticos e outros utilizam os resultados das avaliações externas?
- 10- Em sua opinião, as avaliações externas podem ser utilizadas para regular e reorientar o trabalho pedagógico dos professores e das escolas? De que forma? Pode dar um ou mais exemplos concretos?
- 11- Em que medida é que, na sua opinião, os resultados das avaliações externas são utilizados pelos governos para controlar o cumprimento do currículo e seus objetivos por parte dos professores e das escolas?
- 12- Em sua opinião, os resultados das avaliações externas são utilizados para formular políticas públicas de educação? De que forma? Pode dar um ou mais exemplos concretos?
- 13- Em sua opinião, os diferentes intervenientes estão satisfeitos com o processo de avaliação externa?

Qualidade do Ensino e das Aprendizagens

- 1- Em que medida é que, na sua opinião, as avaliações externas estão relacionadas com o ensino dos professores e com as aprendizagens dos alunos?
- 2- Em que medida é que, na sua opinião, as avaliações externas permitem conhecer o que os alunos sabem e são capazes de fazer?
- 3- Em que medida é que, na sua opinião, as avaliações externas influenciam as práticas de ensino dos professores?
- 4- Num contexto de avaliações externas, quais são, na sua opinião, as principais preocupações dos professores?
- 5- Até que ponto é que as avaliações internas e as avaliações externas são consistentes? Que explicações encontra para justificar a sua resposta?
- 6- Em sua opinião, as avaliações externas condicionam as práticas de avaliação dos professores nas salas de aula? De que formas? Pode dar um ou mais exemplos concretos?
- 7- De que formas é que, na sua opinião, as avaliações externas condicionam o estudo dos alunos?
- 8- Como vêm sendo apropriados os resultados das avaliações externas pelos diferentes intervenientes e interessados no processo? (Alunos, Professores, Gestores Educacionais, Decisores políticos)
- 9- Em que medida os docentes conhecem bem os resultados de suas escolas?
- 10- Até que ponto considera que os docentes analisam os resultados das avaliações externas?
- 11- Até que ponto considera que os gestores escolares analisam os resultados das avaliações externas?

Anexo 4

Questionário

Caro professor,

Sou aluna do curso de doutorado em Educação da Universidade de Lisboa. Minha pesquisa tem como foco o estudo das avaliações externas das aprendizagens dos alunos no Rio de Janeiro.

Venho solicitar a sua colaboração para o preenchimento de um pequeno questionário, que se destina a recolher informações acerca das seguintes áreas do conhecimento: Avaliações externas das aprendizagens e Qualidade do Ensino e das aprendizagens.

Suas respostas são muito importantes para o aporte de alguns dados sobre o alvo da pesquisa. Não há respostas certas ou erradas para cada um dos itens. Interessa, acima de tudo, que manifeste sua opinião.

Garantimos a confidencialidade das informações aqui coletadas, bem como a sua privacidade, assegurando que os dados serão apenas utilizados para fins de pesquisa. Para efeito de clareza e para evitar repetições, utilizou-se a expressão “avaliações externas” para se referir “avaliações externas da aprendizagem dos alunos”.

Desde já agradeço pela atenção e colaboração,

Helen Vieira de Oliveira

*Obrigatório

1. Declaro que fui informado(a) sobre a finalidade da pesquisa, que não há nenhuma identificação de quaisquer dados referentes a minha identidade e que todas as informações aqui coletadas serão utilizadas apenas para fins acadêmicos. Aceito, por conseguinte, participar da pesquisa. *

Marque todas que se aplicam.

Sim

Dados Pessoais

2. Sexo *

Marcar apenas uma oval.

Masculino

Feminino

3. Ano de nascimento (dia/mês/ano) *

4. Formação acadêmica mais elevada *

Marcar apenas uma oval.

- Graduação
- Especialização
- Mestrado
- Doutorado
- Pós-Doutorado
- Outros

5. Disciplina que leciona *

Marcar apenas uma oval.

- Língua Portuguesa
- Matemática

6. Anos de atuação como docente *

Marcar apenas uma oval.

- menos de 1 ano
- 1-5 anos
- 6-10
- 11-15
- 16-20
- 21-25
- 26-30
- mais de 30 anos

7. Anos de docência na Rede Pública do Município do Rio de Janeiro *

Marcar apenas uma oval.

- menos de 1 ano
- 1-5 anos
- 6-10
- 11-15
- 16-20
- 21-25
- 26-30
- mais de 30 anos

8. Ano em que leciona *

Marque todas que se aplicam.

- 6º ano
- 7º ano
- 8º ano
- 9º ano

9. Já fez parte da equipe de gestão escolar? *

Marcar apenas uma oval.

- Sim
- Não

Avaliação das Aprendizagens

Este bloco contempla as questões sobre Avaliação Externa das Aprendizagens. De acordo com a escala apresentada, por favor, manifeste a sua posição com relação à cada um dos itens que a seguir se apresentam.

1- As avaliações externas têm como principal propósito avaliar a qualidade da educação. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

2- As avaliações externas contribuem para que, a nível nacional, todos os alunos aprendam os aspectos fundamentais do currículo. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

3- As avaliações externas permitem avaliar o trabalho pedagógico dos docentes.

*

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

4- As avaliações externas são muito relevantes para apoiar o processo de ensino e aprendizagem. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

5- Para a maioria dos docentes as avaliações externas são credíveis. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

6- Para a maioria dos cidadãos as avaliações externas são consideradas mais credíveis do que para os atores educacionais. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

7- De modo geral, as avaliações externas são construídas com qualidade, permitindo, no que se refere às questões formuladas, conhecer bem o que os alunos sabem e são capazes de fazer. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

8- As avaliações externas são úteis para que os atores educativos conheçam a qualidade da educação que estão proporcionando. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

9- As avaliações externas são úteis para que a opinião pública em geral conheça a qualidade da educação que está sendo proporcionada aos seus cidadãos. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

10- As avaliações externas são úteis para corrigir e melhorar as políticas públicas de educação. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

11- As equipes gestoras das escolas utilizam os resultados das avaliações externas para melhorarem a organização e o funcionamento pedagógico das suas instituições. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

12- Os resultados das avaliações externas são utilizados pelos docentes para melhorarem a qualidade do seu ensino. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

13- Os resultados das avaliações externas são utilizados pelo governo para avaliar as escolas e/ou os professores. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

14- Os resultados das avaliações externas são utilizados pelas escolas para regular e reorientar o seu trabalho pedagógico. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

15- Os governos utilizam as avaliações externas como forma de controlar se os professores cumprem o que está previsto no currículo. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

16- Os resultados das avaliações externas são utilizados para formulação de públicas de educação. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

17- A maioria dos atores educativos está satisfeita com o número anual de avaliações externas que são administradas aos alunos. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

18- De modo geral, os atores educativos estão satisfeitos com o atual sistema de avaliações externas. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

Qualidade do Ensino e das Aprendizagens

Este bloco contempla as questões sobre Qualidade do Ensino e das Aprendizagens. De acordo com a escala apresentada, por favor, manifeste a sua posição com relação à cada um dos itens que a seguir se apresentam.

- 1- As avaliações externas geram reflexões de natureza pedagógica que contribuem para melhorar a qualidade do ensino e das aprendizagens. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

- 2- As avaliações externas dão indicações relevantes acerca do que é importante aprender e ensinar. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

- 3- As avaliações externas são o meio mais eficaz para melhorar o ensino dos professores e as aprendizagens dos alunos. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

4- As avaliações externas medem com rigor o que os alunos sabem e são capazes de fazer. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

5- As avaliações externas condicionam as práticas pedagógicas dos professores. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

6- A maioria dos docentes promove ações específicas de ensino destinadas a preparar os seus alunos para as avaliações externas. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

7- As avaliações externas influenciam os processos de ensino utilizados pelos professores. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

8- Em geral, em contextos de avaliações externas o ensino é orientado sobretudo para a realização de testes. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

9- As práticas de avaliação dos professores são sempre muito condicionadas pelas avaliações externas. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

10- Num contexto de avaliações externas, as práticas de avaliação para classificar, tendem a sobrepor-se às práticas de avaliação para melhorar as aprendizagens. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

11- As avaliações externas influenciam as formas como os alunos estudam. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

12- A maioria dos docentes tem acesso aos resultados das avaliações externas obtidos pelos alunos da sua escola. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

13- A maioria dos docentes consulta os resultados obtidos pelos alunos nas avaliações externas da sua escola. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

14- A maioria dos docentes analisa os resultados das avaliações externas. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

15- Os docentes têm acesso a formação específica que lhes permite compreender os resultados dos alunos nas avaliações externas. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

16- A maioria dos docentes conhece bem o desempenho da sua escola nas avaliações externas. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

17- Os resultados das avaliações externas apresentam uma crescente melhoria ao longo dos anos. *

Marcar apenas uma oval.

- Concordo
- Concordo Totalmente
- Discordo
- Discordo Totalmente

Anexo 5**Universidade de Lisboa – Instituto de Educação
Doutorado em Avaliação Educacional****Termo de Consentimento Livre e Esclarecido**

Pelo presente documento, eu _____

CPF: _____, aceito participar da pesquisa **Avaliações Externas da Aprendizagem: uma Discussão a partir das Apreciações dos Docentes e outros Intervenientes do Processo Educativo** da pesquisadora: **Helen Vieira de Oliveira**, CPF 105.565.977-37, aluna do Doutorado em Educação na Universidade de Lisboa.

Declaro que fui informado(a) que a pesquisa tem por finalidade compreender melhor a percepção dos principais intervenientes do processo educativo – docentes, educadores, pesquisadores e decisores políticos – acerca de um conjunto de aspectos da relação da avaliação externa, políticas públicas e o processo de ensino-aprendizagem no município do Rio de Janeiro.

Como participante da pesquisa, declaro que concordo em ser entrevistado uma ou mais vezes pela entrevistadora em local e duração previamente combinados, () permitindo a gravação das entrevistas / () não permitindo a gravação das entrevistas / () não se aplica a gravação das entrevistas.

Estou ciente que serei esclarecido(a) sobre a pesquisa em qualquer aspecto que desejar, estando livre para recusar a participar, retirar meu consentimento ou interromper a participação a qualquer momento.

() Autorizo / () Não autorizo que meu nome seja divulgado nos resultados da pesquisa, comprometendo-se, a pesquisadora, a utilizar as informações que prestarei somente para os propósitos da pesquisa.

Rio de Janeiro, de de 2018.

Assinatura do Entrevistado(a)

Atividade/Cargo/Função: _____

Contato do(a) Entrevistado(a): _____



Assinatura da Pesquisadora