

ORGANIZADORES

Ovidio Orlando Filho

Nilma Gonçalves Cavalcante

ESTUDOS META-AVALIATIVOS

diretrizes,
categorias,
padrões
e critérios

ORGANIZADORES

Ovidio Orlando Filho

Nilma Gonçalves Cavalcante

ESTUDOS META-AVALIATIVOS

diretrizes,
categorias,
padrões
e critérios

FUNDAÇÃO
◀ **cesgranrio**

FACULDADE
◀ **cesgranrio**

 **pimenta**
2024
São Paulo

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP)

E82

Estudos Meta-Avaliativos: diretrizes, categorias, padrões e critérios / Organização Ovidio Orlando Filho, Nilma Gonçalves Cavalcante. – São Paulo: Pimenta Cultural, 2024.

Livro em PDF

ISBN 978-85-7221-334-9

DOI 10.31560/pimentacultural/978-85-7221-334-9

1. Meta-Avaliação. 2. Diretrizes. 3. Padrões. 4. Critérios.
I. Orlando Filho, Ovidio (Org.). II. Cavalcante, Nilma Gonçalves (Org.). III. Título.

CDD: 372.48

Índice para catálogo sistemático:

I. Educação - Avaliação

Simone Sales • Bibliotecária • CRB ES-000814/0

Copyright © Pimenta Cultural, alguns direitos reservados.

Copyright do texto © 2024 os autores e as autoras.

Copyright da edição © 2024 Pimenta Cultural.

Esta obra é licenciada por uma Licença Creative Commons:
Atribuição-NãoComercial-SemDerivações 4.0 Internacional - (CC BY-NC-ND 4.0).
Os termos desta licença estão disponíveis em:
<<https://creativecommons.org/licenses/>>.
Direitos para esta edição cedidos à Pimenta Cultural.
O conteúdo publicado não representa a posição oficial da Pimenta Cultural.

Direção editorial	Patricia Biegging Raul Inácio Busarello
Editora executiva	Patricia Biegging
Gerente editorial	Landressa Rita Schiefelbein
Assistente editorial	Júlia Marra Torres
Estagiária editorial	Ana Flávia Pivisan Kobata
Diretor de criação	Raul Inácio Busarello
Assistente de arte	Naiara Von Groll
Editoração eletrônica	Andressa Karina Voltolini
Estagiárias em editoração	Raquel de Paula Miranda Stela Tiemi Hashimoto Kanada
Imagens da capa	RAZUANA, pikisuperstar - Freepik.com
Tipografias	Acumin, Poster Gothic, Rockwell Nova
Revisão	Os autores e as autoras
Organizadores	Ovidio Orlando Filho Nilma Gonçalves Cavalcante

CONSELHO EDITORIAL CIENTÍFICO

Doutores e Doutoradas

Adilson Cristiano Habowski
Universidade La Salle, Brasil

Adriana Flávia Neu
Universidade Federal de Santa Maria, Brasil

Adriana Regina Vettorazzi Schmitt
Instituto Federal de Santa Catarina, Brasil

Aguimario Pimentel Silva
Instituto Federal de Alagoas, Brasil

Alaim Passos Bispo
Universidade Federal do Rio Grande do Norte, Brasil

Alaim Souza Neto
Universidade Federal de Santa Catarina, Brasil

Alessandra Knoll
Universidade Federal de Santa Catarina, Brasil

Alessandra Regina Müller Germani
Universidade Federal de Santa Maria, Brasil

Aline Corso
Universidade do Vale do Rio dos Sinos, Brasil

Aline Wendpap Nunes de Siqueira
Universidade Federal de Mato Grosso, Brasil

Ana Rosângela Colares Lavand
Universidade Federal do Pará, Brasil

André Gobbo
Universidade Federal da Paraíba, Brasil

Andressa Wiebusch
Universidade Federal de Santa Maria, Brasil

Andreza Regina Lopes da Silva
Universidade Federal de Santa Catarina, Brasil

Angela Maria Farah
Universidade de São Paulo, Brasil

Anísio Batista Pereira
Universidade do Estado do Amapá, Brasil

Antonio Edson Alves da Silva
Universidade Estadual do Ceará, Brasil

Antonio Henrique Coutelo de Moraes
Universidade Federal de Rondonópolis, Brasil

Arthur Vianna Ferreira
Universidade do Estado do Rio de Janeiro, Brasil

Ary Albuquerque Cavalcanti Junior
Universidade Federal de Mato Grosso, Brasil

Asterlindo Bandeira de Oliveira Júnior
Universidade Federal da Bahia, Brasil

Bárbara Amaral da Silva
Universidade Federal de Minas Gerais, Brasil

Bernadette Beber
Universidade Federal de Santa Catarina, Brasil

Bruna Carolina de Lima Siqueira dos Santos
Universidade do Vale do Itajaí, Brasil

Bruno Rafael Silva Nogueira Barbosa
Universidade Federal da Paraíba, Brasil

Caio Cesar Portella Santos
Instituto Municipal de Ensino Superior de São Manuel, Brasil

Carla Wanessa do Amaral Caffagni
Universidade de São Paulo, Brasil

Carlos Adriano Martins
Universidade Cruzeiro do Sul, Brasil

Carlos Jordan Lapa Alves
Universidade Estadual do Norte Fluminense Darcy Ribeiro, Brasil

Caroline Chioquetta Lorenset
Universidade Federal de Santa Catarina, Brasil

Cássio Michel dos Santos Camargo
Universidade Federal do Rio Grande do Sul-Faced, Brasil

Christiano Martino Otero Avila
Universidade Federal de Pelotas, Brasil

Cláudia Samuel Kessler
Universidade Federal do Rio Grande do Sul, Brasil

Cristiana Barcelos da Silva.
Universidade do Estado de Minas Gerais, Brasil

Cristiane Silva Fontes
Universidade Federal de Minas Gerais, Brasil

Daniela Susana Segre Guertzenstein
Universidade de São Paulo, Brasil

Daniele Cristine Rodrigues
Universidade de São Paulo, Brasil

Dayse Centurion da Silva
Universidade Anhanguera, Brasil

Dayse Sampaio Lopes Borges

Universidade Estadual do Norte Fluminense Darcy Ribeiro, Brasil

Diego Pizarro

Instituto Federal de Brasília, Brasil

Dorama de Miranda Carvalho

Escola Superior de Propaganda e Marketing, Brasil

Edson da Silva

Universidade Federal dos Vales do Jequitinhonha e Mucuri, Brasil

Elena Maria Mallmann

Universidade Federal de Santa Maria, Brasil

Eleonora das Neves Simões

Universidade Federal do Rio Grande do Sul, Brasil

Eliane Silva Souza

Universidade do Estado da Bahia, Brasil

Elvira Rodrigues de Santana

Universidade Federal da Bahia, Brasil

Éverly Pegoraro

Universidade Federal do Rio de Janeiro, Brasil

Fábio Santos de Andrade

Universidade Federal de Mato Grosso, Brasil

Fabrícia Lopes Pinheiro

Universidade Federal do Estado do Rio de Janeiro, Brasil

Felipe Henrique Monteiro Oliveira

Universidade Federal da Bahia, Brasil

Fernando Vieira da Cruz

Universidade Estadual de Campinas, Brasil

Gabriella Eldereti Machado

Universidade Federal de Santa Maria, Brasil

Germano Ehlert Pollnow

Universidade Federal de Pelotas, Brasil

Geymeesson Brito da Silva

Universidade Federal de Pernambuco, Brasil

Giovanna Ofretorio de Oliveira Martin Franchi

Universidade Federal de Santa Catarina, Brasil

Handherson Leylton Costa Damasceno

Universidade Federal da Bahia, Brasil

Hebert Elias Lobo Sosa

Universidad de Los Andes, Venezuela

Helciclever Barros da Silva Sales

Instituto Nacional de Estudos

e Pesquisas Educacionais Anísio Teixeira, Brasil

Helena Azevedo Paulo de Almeida

Universidade Federal de Ouro Preto, Brasil

Hendy Barbosa Santos

Faculdade de Artes do Paraná, Brasil

Humberto Costa

Universidade Federal do Paraná, Brasil

Igor Alexandre Barcelos Graciano Borges

Universidade de Brasília, Brasil

Inara Antunes Vieira Willerding

Universidade Federal de Santa Catarina, Brasil

Jaziel Vasconcelos Dorneles

Universidade de Coimbra, Portugal

Jean Carlos Gonçalves

Universidade Federal do Paraná, Brasil

Jocimara Rodrigues de Sousa

Universidade de São Paulo, Brasil

Joelson Alves Onofre

Universidade Estadual de Santa Cruz, Brasil

Jónata Ferreira de Moura

Universidade São Francisco, Brasil

Jorge Eschriqui Vieira Pinto

Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil

Jorge Luís de Oliveira Pinto Filho

Universidade Federal do Rio Grande do Norte, Brasil

Juliana de Oliveira Vicentini

Universidade de São Paulo, Brasil

Julierme Sebastião Moraes Souza

Universidade Federal de Uberlândia, Brasil

Junior César Ferreira de Castro

Universidade de Brasília, Brasil

Katia Bruginski Mulik

Universidade de São Paulo, Brasil

Laionel Vieira da Silva

Universidade Federal da Paraíba, Brasil

Leonardo Pinheiro Mozdzenski

Universidade Federal de Pernambuco, Brasil

Lucila Romano Tragtenberg

Pontifícia Universidade Católica de São Paulo, Brasil

Lucimara Rett

Universidade Metodista de São Paulo, Brasil

Manoel Augusto Polastreli Barbosa

Universidade Federal do Espírito Santo, Brasil

Marcelo Nicomedes dos Reis Silva Filho

Universidade Estadual do Oeste do Paraná, Brasil

Marcio Bernardino Sirino

Universidade Federal do Estado do Rio de Janeiro, Brasil

Marcos Pereira dos Santos

Universidad Internacional Iberoamericana del México, México

Marcos Uzel Pereira da Silva

Universidade Federal da Bahia, Brasil

Maria Aparecida da Silva Santandel
Universidade Federal de Mato Grosso do Sul, Brasil

Maria Cristina Giorgi
*Centro Federal de Educação Tecnológica
Celso Suckow da Fonseca, Brasil*

Maria Edith Maroca de Avelar
Universidade Federal de Ouro Preto, Brasil

Marina Bezerra da Silva
Instituto Federal do Piauí, Brasil

Mauricio José de Souza Neto
Universidade Federal da Bahia, Brasil

Michele Marcelo Silva Bortolai
Universidade de São Paulo, Brasil

Mônica Tavares Orsini
Universidade Federal do Rio de Janeiro, Brasil

Nara Oliveira Salles
Universidade do Estado do Rio de Janeiro, Brasil

Neli Maria Mengalli
Pontifícia Universidade Católica de São Paulo, Brasil

Patricia Biegging
Universidade de São Paulo, Brasil

Patricia Flavia Mota
Universidade do Estado do Rio de Janeiro, Brasil

Raul Inácio Busarello
Universidade Federal de Santa Catarina, Brasil

Raymundo Carlos Machado Ferreira Filho
Universidade Federal do Rio Grande do Sul, Brasil

Roberta Rodrigues Ponciano
Universidade Federal de Uberlândia, Brasil

Robson Teles Gomes
Universidade Católica de Pernambuco, Brasil

Rodiney Marcelo Braga dos Santos
Universidade Federal de Roraima, Brasil

Rodrigo Amancio de Assis
Universidade Federal de Mato Grosso, Brasil

Rodrigo Sarruge Molina
Universidade Federal do Espírito Santo, Brasil

Rogério Rauber
Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil

Rosane de Fatima Antunes Obregon
Universidade Federal do Maranhão, Brasil

Samuel André Pompeo
Universidade Estadual Paulista Júlio de Mesquita Filho, Brasil

Sebastião Silva Soares
Universidade Federal do Tocantins, Brasil

Silmar José Spinardi Franchi
Universidade Federal de Santa Catarina, Brasil

Simone Alves de Carvalho
Universidade de São Paulo, Brasil

Simoni Urnau Bonfiglio
Universidade Federal da Paraíba, Brasil

Stela Maris Vaucher Farias
Universidade Federal do Rio Grande do Sul, Brasil

Tadeu João Ribeiro Baptista
Universidade Federal do Rio Grande do Norte

Taiane Aparecida Ribeiro Nepomoceno
Universidade Estadual do Oeste do Paraná, Brasil

Taíza da Silva Gama
Universidade de São Paulo, Brasil

Tania Micheline Miorando
Universidade Federal de Santa Maria, Brasil

Tarcísio Vanzin
Universidade Federal de Santa Catarina, Brasil

Tascieli Feltrin
Universidade Federal de Santa Maria, Brasil

Tayson Ribeiro Teles
Universidade Federal do Acre, Brasil

Thiago Barbosa Soares
Universidade Federal do Tocantins, Brasil

Thiago Camargo Iwamoto
Universidade Estadual de Goiás, Brasil

Thiago Medeiros Barros
Universidade Federal do Rio Grande do Norte, Brasil

Tiago Mendes de Oliveira
Centro Federal de Educação Tecnológica de Minas Gerais, Brasil

Vanessa Elisabete Raue Rodrigues
Universidade Estadual de Ponta Grossa, Brasil

Vania Ribas Ulbricht
Universidade Federal de Santa Catarina, Brasil

Wellington Furtado Ramos
Universidade Federal de Mato Grosso do Sul, Brasil

Wellton da Silva de Fatima
Instituto Federal de Alagoas, Brasil

Yan Masetto Nicolai
Universidade Federal de São Carlos, Brasil

PARECERISTAS E REVISORES(AS) POR PARES

Avaliadores e avaliadoras Ad-Hoc

Alessandra Figueiró Thornton
Universidade Luterana do Brasil, Brasil

Alexandre João Appio
Universidade do Vale do Rio dos Sinos, Brasil

Bianka de Abreu Severo
Universidade Federal de Santa Maria, Brasil

Carlos Eduardo Damian Leite
Universidade de São Paulo, Brasil

Catarina Prestes de Carvalho
Instituto Federal Sul-Rio-Grandense, Brasil

Elisiene Borges Leal
Universidade Federal do Piauí, Brasil

Elizabeth de Paula Pacheco
Universidade Federal de Uberlândia, Brasil

Elton Simomukay
Universidade Estadual de Ponta Grossa, Brasil

Francisco Geová Goveia Silva Júnior
Universidade Potiguar, Brasil

Indiamaris Pereira
Universidade do Vale do Itajaí, Brasil

Jacqueline de Castro Rimá
Universidade Federal da Paraíba, Brasil

Lucimar Romeu Fernandes
Instituto Politécnico de Bragança, Brasil

Marcos de Souza Machado
Universidade Federal da Bahia, Brasil

Michele de Oliveira Sampaio
Universidade Federal do Espírito Santo, Brasil

Pedro Augusto Paula do Carmo
Universidade Paulista, Brasil

Samara Castro da Silva
Universidade de Caxias do Sul, Brasil

Thais Karina Souza do Nascimento
Instituto de Ciências das Artes, Brasil

Viviane Gil da Silva Oliveira
Universidade Federal do Amazonas, Brasil

Weyber Rodrigues de Souza
Pontifícia Universidade Católica de Goiás, Brasil

William Roslindo Paranhos
Universidade Federal de Santa Catarina, Brasil

Parecer e revisão por pares

Os textos que compõem esta obra foram submetidos para avaliação do Conselho Editorial da Pimenta Cultural, bem como revisados por pares, sendo indicados para a publicação.

SUMÁRIO

Ovidio Orlando Filho

Nilma Gonçalves Cavalcante

Apresentação 11

CAPÍTULO 1

Alcineidy Nascimento Falcão

Maria das Graças Madeira

Ligia Gomes Elliot

**Meta-avaliação do Relatório de Avaliação
da implementação do Programa
de Estímulo ao Emprego Científico14**

CAPÍTULO 2

Dalva Carneiro

Albertina Thereza Pacheco Rodrigues

Ovidio Orlando Filho

**Avaliação do Relatório de Avaliação
sobre as informações de um
descumprimento de condicionalidades
do Programa *Bolsa Família*:
um estudo meta-avaliativo42**

CAPÍTULO 3

Karla Verônica Mendes da Silva

Michele Terra Duarte

Elaine Rodrigues Perdigão

**Meta-Avaliação do Relatório de Retorno
Social do Programa *A União faz a Vida*:
resultados e recomendações.....78**

CAPÍTULO 4

Patrícia Sousa de Souza

Lígia Silva Leite

Programa Melhoria da Educação:

a meta-avaliação colaborando para a qualidade de programas educacionais e seus processos avaliativos 118

CAPÍTULO 5

Hélio Huguenin Tavares Júnior

Virginia Virgínio de Lima

Glauco da Silva Aguiar

Estudo Meta-Avaliativo da Avaliação de Possíveis Impactos:

PROFMAT 160

CAPÍTULO 6

Cristina Borges Alves

Alessandra Baldner Pontes

Ovidio Orlando Filho

Meta-Avaliação da Avaliação Impacto Social do Programa *União Faz a Vida*, realizada pela Fundação SICRED.....

195

CAPÍTULO 7

Sheila Regina da Rocha de Magalhães

Luciana Vieira Santos

Lucí Hildenbrand

Meta-Avaliação da Avaliação de Impacto Social do Projeto *Palcos para a Inclusão* financiado pelo Programa Operacional para a Inclusão Social e Emprego de Portugal.....

248

CAPÍTULO 8

Flávia de Andrade Viveiros Quadrelli

Paula Magalhães da Silva Santos

Elaine Maria de Andrade Senra

Meta-Avaliação da Avaliação de Impacto Social dos Cursos Técnicos CEAP	275
---	------------

CAPÍTULO 9

Andréa Cristina Tavares Moreira

José Rafael Ayres da Motta

Mariana Sá Alcântara Gomes

Meta-Avaliação da Avaliação do Projeto de Promoção do Ecossistema de Investimento e Negócios de Impacto Socioambiental no Brasil	307
---	------------

Sobre os autores e as autoras	332
--	------------

Índice remissivo	340
-------------------------------	------------

APRESENTAÇÃO

É com grande satisfação que apresentamos este livro dedicado aos estudos meta-avaliativos, uma área essencial para a compreensão crítica e aprimoramento das avaliações. A meta-avaliação, entendida como a avaliação da avaliação (Scriven, 1991), busca garantir a qualidade, a utilidade e a credibilidade dos processos avaliativos em diferentes contextos.

A avaliação é uma ferramenta essencial da gestão, voltada para a tomada de decisões inteligentes em diversos campos, como educação, gestão pública, saúde e programas sociais, etc. No entanto, para que uma avaliação cumpra seu papel de forma adequada, é necessário avaliar sua qualidade, aferindo o seu mérito e o seu valor. Ou seja, sua eficiência, sua eficácia no atingimento de seus objetivos, além de sua utilidade, efetividade e legitimidade (Spinoza, 2012).

A presente publicação, produzida por docentes, egressos e discentes do Mestrado Profissional em Avaliação da Faculdade Cesgranrio, nas atividades didático-metodológicas da própria disciplina Meta-Avaliação, que discute a multidisciplinaridade tão necessária à área da avaliação, reúne diversas avaliações realizadas em programas destinados à execução de políticas públicas de áreas distintas e os submete à avaliação da avaliação. A intencionalidade em sua concepção residuiu no fato de se considerar relevante que tantos os programas de políticas públicas quanto as suas avaliações devam ter as suas qualidades aferidas por meio de diretrizes, padrões e critérios consagrados academicamente.

Ao longo da obra, foram abordadas diversas perspectivas teóricas e metodológicas, fornecendo um panorama abrangente sobre os princípios, desafios e processos de meta-avaliativos. O leitor encontrará reflexões aprofundadas sobre boas práticas para a

implementação de avaliações de qualidade e metodologias inovadoras (Worthen; Sanders; Fitzpatrick, 2004; Davidson, 2005; Scriven, 2015; Yarbrough; Shulla; Hopson; Caruthers, 2011; Elliot, 2011; Silva; Joppert; Gasparini, 2020; Ghiano; Wehrle; Fiscarelli; Valencia; Checa Gutiérrez; Amariles Erazo; Guzman; Soberón Álvarez, 2021).

O livro foi concebido para auxiliar pesquisadores, avaliadores, gestores e profissionais que estejam interessados em compreender como a meta-avaliação pode contribuir para o aperfeiçoamento dos processos avaliativos e para a tomada de decisões mais informadas e eficazes. Além disso, busca estimular o debate sobre a importância de estabelecer padrões rigorosos na avaliação e promover uma cultura de aprimoramento contínuo. Os estudos componentes de seus capítulos ilustram a aplicabilidade da meta-avaliação em diferentes domínios como, por exemplo, políticas públicas e gestão organizacional.

Esperamos que esta obra seja uma fonte valiosa de conhecimento e inspiração para todos aqueles comprometidos com a excelência na avaliação. Desejamos a todos uma leitura proveitosa e enriquecedora!

Ovidio Orlando Filho
Nilma Gonçalves Cavalcante

REFERÊNCIAS

DAVIDSON, E. Jane. **Evaluation methodology basics**: the nuts and bolts of sound evaluation. Thousands Oaks, CA: Sage, 2005.

ELLIOT, Ligia Gomes. Meta-avaliação: das abordagens às possibilidades de aplicação. **Ensaio**: aval. pol. públ. educ., Rio de Janeiro, v. 19, n. 73, p. 941-964, 2011.

GHIANO, Maria Celeste; WEHRLE, Andrea María; FISCARELLI, Andrea Peroni; VALENCIA, Sergio Martinic; CHECA GUTIÉRREZ, Marcia Itzel; AMARILES ERAZO, Fabiola; GUZMÁN, Ana Luiza; SOBERÓN ÁLVAREZ, Luis. **Diretrizes de avaliação para a América Latina e o Caribe**. 2. ed. [S. l.]: Relac, 2021. Disponível em: <https://www.deval.org/en/publications/diretrizes-de-avaliacao-para-a-america-latina-e-o-caribe-edicao-2021>. Acesso em: 5 set. 2024.

SCRIVEN, Michael. **Evaluation thesaurus**. Newbury Park: Sage Publications, 1991.

SCRIVEN, Michael. **Key evaluation checklist**. Kalamazoo, MI: Western Michigan University, 2015. Disponível em: <https://files.wmich.edu/s3fs-public/attachments/u1105/2023/kec-scriven.pdf>. Acesso em: 10 nov. 2024.

SILVA, Rogério; JOPERT, Márcia; GASPARINI, Max. Por que produzimos diretrizes para a prática de avaliação. **Nexo**, [S. l.], 25 ago. 2020. Disponível em: <https://www.nexojornal.com.br/ensaio/2020/Por-que-produzimos-diretrizes-para-a-pr%C3%AAtica-de-avalia%C3%A7%C3%A3o>. Acesso em: 16 jul. 2024.

SPINOZA, Roberto Moreno. Accountability. *In*: CASTRO, Carmem Lúcia Freitas de; GONTIJO, Cynthia Rúbia Braga; AMABILE, Antonio Eduardo de Noronha (org.). **Dicionário de políticas públicas**. Barbacena: EdUEMG, 2012.

WORTHEN, Blaine R.; SANDERS, James R.; FITZPATRICK, Jody L. **Avaliação de programas: concepções e práticas**. São Paulo: Editora Gente, 2004.

YARBROUGH, Donald B.; SHULLA, Lyn M.; HOPSON, Rodney K.; CARUTHERS, Flora A. **The program evaluation standards: a guide for evaluators and evaluation users**. 3. ed. Thousands Oaks, Califórnia: Sage, 2011.

1

*Alcineidy Nascimento Falcão
Maria das Graças Madeira
Lígia Gomes Elliot*

META-AVALIAÇÃO DO RELATÓRIO DE AVALIAÇÃO DA IMPLEMENTAÇÃO DO PROGRAMA DE ESTÍMULO AO EMPREGO CIENTÍFICO

INTRODUÇÃO

O histórico das avaliações no contexto de Portugal, considerando seus aspectos como prática extensiva e continuada, é relativamente recente, não sendo muito diferenciada, no campo temporal, do histórico latino-americano. Atualmente, os princípios de garantia e fomento da qualidade de objetos de distintas origens e naturezas estão desempenhando um papel estratégico em diferentes campos de atuação, tanto nos aspectos da produção de bens quanto naqueles relacionados à prestação de serviços. No contexto português, no que se refere ao campo educacional, é patente o pensamento de que os processos de melhoria da qualidade das ofertas dos serviços educacionais em todos os seus campos, tipos e modalidades se tornam, cada vez mais, fundamentais para o desenvolvimento do país.

No que se refere ao Ensino Superior, existe uma sinalização de que os sistemas de avaliação praticados em diferentes contextos do mundo globalizado, como na América Latina e Europa, por exemplo, passaram a adotar critérios, de certa forma, uniformizados (Freitas, 2012). Presume-se que essa padronização seja um efeito colateral tanto da globalização propriamente dita, principalmente em seu aspecto econômico, assim como de uma internacionalização que tende a padronizar essa oferta de ensino, não importando o contexto em que ela ocorra (Durham, 1996; Brennan; Shah, 2000; Mollis; Marginson, 2007; Amaral, 2005). Ou seja, no que se refere à Educação Superior, os sistemas educacionais se tornam, cada vez mais, “convergentes nas suas políticas e estruturas formais” (Freitas, 2012, p. 124).

Considerando a ciência como uma prática social e, geralmente, estando sujeita a realização de parcerias, interações e relações do conhecimento adquirido com a prática daquilo que se aprendeu e produziu em termos acadêmico (Batista, 2019), considera-se de fundamental relevância que a produção científica de maior nível, como

a que tem origem nos cursos *stricto sensu*, tenha seu emprego estimulado, de modo a gerar benefícios para a sociedade como um todo.

Segundo Batista (2019, p. 43):

É consenso que um dos ativos mais valiosos das Universidades são seus ex-alunos, visto que, conforme Teixeira & Maccari (2014), suas realizações garantem exposição e maior visibilidade da instituição de origem. Segundo os autores, a contribuição da educação para a sociedade é observada com base nas experiências dos Egressos durante o curso e sua concretização na trajetória escolhida.

Pautado nessa problemática, e considerando que produzir conhecimento científico e não o difundir ou praticá-lo é pura perda de tempo, este artigo tem o seu foco em um processo avaliativo realizado em Portugal que teve sua origem em um estudo avaliativo cujo objeto foi a implementação de um programa de estímulo ao emprego científico de formandos originários de cursos superiores realizados naquele país europeu.

Procurou-se, então, realizar uma meta-avaliação do estudo avaliativo posto em prática por uma equipe de especialistas nomeada pelo governo português, de modo a conhecer, mesmo que de forma superficial, as características do programa avaliado como, de maneira mais completa, a qualidade da avaliação realizada. Para isso, foram aplicados no processo de avaliação da implementação do Programa de Estímulo ao Emprego Científico, os critérios adotados por Scriven (2009) em sua Lista Chave de Verificação, no sentido de se aferir a qualidade da avaliação, em um processo denominado de meta-avaliação.

Por fim, a finalidade basilar deste artigo é a de contribuir com novas informações sobre a qualidade de avaliações, medidas por meio de processos meta-avaliativos, utilizando-se de critérios instituídos por Scriven (2009), possuidor de validação acadêmica.

Ressalta-se nesse aspecto que o contexto de Portugal, um importante país da comunidade lusófona, possui relevante influência na produção de conhecimento acadêmico na Europa.

REFLEXÕES SOBRE AVALIAÇÃO E META-AVALIAÇÃO

O termo avaliar vem no latim *avaliare*, que significa atribuir valor e mérito ao objeto em estudo. Avaliar é, portanto, atribuir juízo de valor sobre uma ação ou uma matéria. Já avaliação, para Worthen, Sanders e Fitzpatrick (2004, p. 35), é a “identificação, esclarecimento e aplicação de critérios defensáveis para determinar o valor, a qualidade, a utilidade, a eficácia ou a importância do objeto avaliado em relação a estes critérios”. Ou seja, sua definição está ligada a um processo que implica em escolhas norteadas por valores que almejam melhorar a qualidade daquilo que se avalia.

Logo, os resultados de uma avaliação devem ter significância, de forma a justificar os investimentos por parte de seus contratantes. Sob essa ótica, para se aferir se a avaliação atingiu os objetivos pelos quais foi implementada, considera-se que ela deva ser também avaliada. A ação de avaliar uma avaliação, além de garantir uma maior credibilidade ao avaliador original, possui a força de evitar possíveis problemas existentes nos processos originalmente implementados, o que pode causar mais malefícios que benefícios, considerando a utilização de seus resultados.

Scriven (1991) define a meta-avaliação como “avaliação da avaliação”. Segundo o autor é de fundamental importância que a própria avaliação seja avaliada, dando, assim, maior confiabilidade e garantia de uso efetivo dos seus resultados. Para Scriven (1991), a meta-avaliação tem duas principais finalidades que podem ser

implementadas juntas ou separadamente nos processos meta-avaliativos: a primeira delas está ligada a uma característica formativa, isto é, a que ocorre durante o desenvolvimento da avaliação de forma a avaliar seus processos de desenvolvimento na medida em que eles ocorrem. Esse tipo busca aprimorar a avaliação, antes de seu término, evitando, com isso, a geração de danos irreparáveis; em segundo lugar, a sua característica somativa, realizada depois que a avaliação termina e é implementada, geralmente, por meio do estudo de sua documentação. Esse tipo tem o propósito de aumentar a confiabilidade dos resultados alcançados.

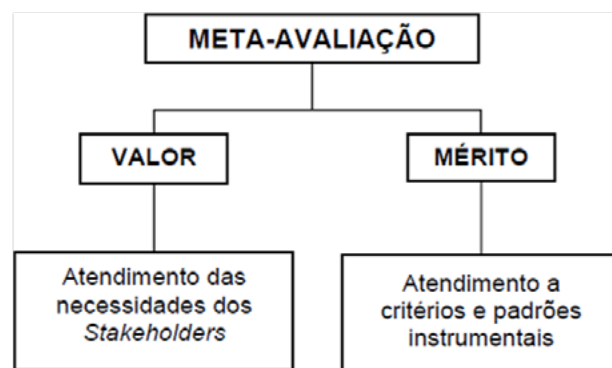
Nesse sentido, um estudo meta-avaliativo pode contribuir para que uma avaliação possua valor e seja realizada com o devido rigor, sustentando-se em metodologias consistentes, com critérios válidos, tendo como resultados recomendações qualificadas e apropriadas para serem utilizadas nos processos de melhoria do objeto avaliado.

Tanto a meta-avaliação formativa quanto a somativa podem ser implementadas por avaliadores internos, aqueles que integram a mesma organização em que a avaliação foi realizada, e avaliadores externos, aqueles que são exteriores à organização, não tendo qualquer contato com a avaliação realizada. No entanto, de acordo com Yarbrough, Shulha, Hopson e Caruthers (2011), elas se complementam, pois esses dois olhares diferentes contribuem para apresentar um panorama de maior amplitude da avaliação, tornando a meta-avaliação possuidora de uma feição mais sistemática. Essa colaboração entre os dois tipos de meta-avaliadores, internos e externos, tende a produzir resultados combinados, com uma perspectiva ao mesmo tempo interior e externo, introduzindo qualidade e credibilidade ao processo avaliativo meta-avaliado.

Os processos de coleta de dados de uma avaliação se diferenciam à medida em que ela tenha o caráter formativo, em que os dados são coletados durante a consecução de suas diferentes fases

de execução, ou somativo, em que os dados são coletados depois que a avaliação é encerrada, por meio dos documentos gerados pelo processo avaliativo como um todo. Normalmente o documento principal consultado neste segundo caso é o relatório final da avaliação. Em ambos os casos, esses dados devem ser observados de acordo com parâmetros que possibilitem conferir o valor produzido pela avaliação (mensuração das necessidades de informação dos *stakeholders* além do uso efetivo dos resultados gerados) e mérito (medida em que são atendidos os critérios e padrões instrumentais estabelecidos, fundamentados na eficiência e na eficácia de seus processos), conforme conceito exemplificado na Figura 1.

Figura 1 - Princípios de uma meta-avaliação



Fonte: Davok (2006, p. 87).

Torna-se importante ressaltar que os resultados da meta-avaliação devem servir de subsídios para a tomada de decisão no sentido de buscar melhoria para a qualidade da avaliação ou para a sua *accountability* (Davok, 2006). A meta-avaliação também deve servir para auxiliar os interessados a “perceber tanto os pontos fortes como as fragilidades da avaliação realizada” (Stufflebeam, 2001 *apud* Elliot, 2011, p. 944).

Conforme formulado por Elliot (2011, p. 960, grifo nosso), a meta-avaliação faz-se necessária porque “em diversas instâncias, a necessidade de meta-avaliar uma avaliação, seu **Relatório** ou o próprio processo avaliativo é a forma de atestar-lhes a qualidade ou de evidenciar pontos que ainda precisam ser burilados”.

Entre os caminhos estabelecidos como estratégia para se alcançar os objetivos traçados no *design* da meta-avaliação, no sentido de lhe conferir uma direção ou orientação, pautados em critérios definidos em seu planejamento, destacam-se os atributos do *Joint Committee on Standards for Educational Evaluations*, que abordam 30 padrões. No conjunto da obra, os padrões definem o que é uma avaliação de qualidade, servindo como guia tanto para a realização de avaliações quanto de meta-avaliações.

Elliot (2011, p. 943) afirma que, ao utilizar-se de sua própria definição de avaliação acrescida das dimensões descritas pelo *Joint Committee*, Stufflebeam (2001) definiu meta-avaliação como:

O processo de delinear, obter e aplicar informação descritiva e de julgamento sobre a utilidade, a viabilidade, adequação e precisão de uma avaliação e sua natureza sistemática, competente conduta, integridade/honestidade, respeitabilidade e responsabilidade social para orientar a avaliação e divulgar publicamente seus pontos fortes e fracos.

Para Yarbrough, Shulha, Hopson e Caruthers (2011), as meta-avaliações de diferentes objetos de estudo diferem em diferentes aspectos como suas dimensões, categorias e indicadores, suas abrangências, sua complexidade, duração e ciclos do processo, alocação de recursos, perspectivas de cenários e utilização de seus resultados. No entanto, todo processo ou estudo meta-avaliativo aplica evidências específicas ligadas a padrões, critérios ou diretrizes selecionadas, no sentido de alcançar conclusões sobre o nível de qualidade da avaliação e/ou de seus processos de implementação. Qualquer meta-avaliação pode ser orientada e fundamentada por diferentes *checklists*, formulários, tabelas e narrativas, desde que

possibilitem os usuários de seus resultados o acompanhamento de evidências, padrões, critérios, lógica e conclusões que os levem a inferências e decisões próprias e devidamente adequadas.

OBJETIVO E OBJETO DO ESTUDO

Considerando a definição de Scriven (1991) sobre meta-avaliação somativa, que tem o propósito de aumentar a confiabilidade dos resultados alcançados por uma avaliação já encerrada, o objetivo do presente estudo é meta-avaliar o Relatório de Avaliação da Implementação do Programa de Estímulo ao Emprego Científico. Como critério adotado para nortear o estudo foi utilizada a Lista Chave de Verificação da Avaliação de Scriven, adaptada em versão traduzida por Elliot (2011) do original *Key Evaluation Checklist* (Scriven, 1991).

O objeto do estudo é fruto de uma avaliação determinada pelo Ministro da Ciência, Tecnologia e Ensino Superior de Portugal que nomeou uma Comissão de Alto Nível para avaliar a implementação de um Programa de Estímulo ao Emprego Científico daquele país. A finalidade do Programa era proporcionar “incentivo à contratação de novos doutorados e ao desenvolvimento de planos de emprego científico e de carreiras científicas pelas instituições científicas e acadêmicas, públicas ou privadas” (Nazaré; Rendas; Cunha; Santos, 2020, p. 5).

O Programa, instituído em 2020, envolveu integrantes representantes de quatro importantes universidades portuguesas (Universidade de Aveiro, Universidade do Porto, Universidade Nova de Lisboa e Universidade do Minho).

Na Introdução do Relatório, a Comissão informa que começou a abordagem ao problema com uma breve descrição do contexto, no que concerne à evolução do número e áreas de doutoramento

e, genericamente, à questão dos pós-doutorados. Na sequência, a Comissão fez uma análise global da implementação do Programa de Estímulo ao Emprego Científico. As análises conduziram às recomendações relativas à evolução futura do Programa.

Dessa forma, este estudo pretende responder à seguinte questão meta avaliativa:

Até que ponto a avaliação expressa no Relatório de Avaliação da Implementação do Programa de Estímulo ao Emprego Científico atende aos critérios expostos na Lista Chave de Verificação da Avaliação de Scriven (*Key Evaluation Checklist*) elaborada por Scriven (1991)?

PROCEDIMENTOS METODOLÓGICOS

O estudo foi iniciado por meio de uma leitura detalhada do conteúdo do objeto de estudo, o Relatório de Avaliação da Implementação do Programa de Estímulo ao Emprego Científico. Posteriormente, devido à abordagem somativa e ao caráter qualitativo do estudo, foram realizadas pesquisas em *sites* especializados, livros e instrumentos legais relacionados ao tema abordado no trabalho. Posteriormente foi escolhido o instrumento de coleta de informação a ser adotado no presente estudo. Nesse sentido, Elliot (2012, p. 13) esclarece:

Um instrumento de pesquisa ou avaliação consiste no recurso usado para coletar a informação de interesse sobre uma variável, característica, categoria ou dimensão do objeto, ou ainda evidências de indicadores. São os dados coletados pelo instrumento que auxiliam o avaliador ou pesquisador a acompanhar o desenvolvimento do objeto ou fenômeno focalizado, a obter informações mais precisas sobre ele, e a tirar conclusões sobre determinadas características suas.

Finalmente, para sua realização, optou-se como quadro de critérios da meta-avaliação a utilização da Lista Chave de Verificação da Avaliação, proposta por Scriven (2007). Para isso, utilizou-se os padrões e os critérios apresentados no Quadro 1, de modo a verificar quais foram identificados no Relatório.

Quadro 1 – Lista chave de verificação de Scriven

Pontos de Verificação		Julgamentos (A - E)
Preliminares	I. Sumário Executivo	
	II. Prefácio	
	III. Metodologia	
Fundamento	1. Antecedentes e Contexto	
	2. Descrições e Definições	
	3. Consumidores	
	4. Recursos	
	5. Valores	
Subavaliações	6. Processo	
	7. Resultados	
	8. Custos	
	9. Comparações	
	10. Exportabilidade	
Conclusões	11. Síntese	
	12. Recomendações e Explicações	
	13. Responsabilidade e Justificativa	
	14. Relatório e Apoio	
	15. Meta-avaliação	

Fonte: Scriven (2007 apud Elliot, 2011).

Os julgamentos de cada ponto de verificação obedeceram os seguintes critérios em relação aos seus níveis de atendimento: conceito A atinge todos os aspectos principais deste ponto e os expressam de forma clara e concisa; conceito B atinge a maioria dos aspectos relativos a este ponto, mas não cobre um ou dois deles, ou trata dos ingredientes principais, mas não são 100% claros; conceito C trata do ponto, mas perde alguns aspectos cruciais, ou perde vários aspectos importantes, ou é um tanto desorganizado ou pouco claro; conceito D tem um ou dois elementos que parecem tratar implicitamente do ponto, mas o tratamento é deficiente; e, conceito E perde o ponto totalmente (Davidson, 2005).

RESULTADOS

No primeiro nível, ou componente da Lista de Verificação de Scriven, denominado Preliminares, o primeiro Ponto de Verificação, (I) Sumário Executivo, deve fornecer, de forma sintética, uma visão geral dos resultados relativos aos últimos cinco pontos da Lista Chave de Verificação da Avaliação, que correspondem àqueles numerados entre 11 e 15.

Dessa forma, os autores que avaliaram a implementação do Programa de Estímulo ao Emprego Científico apresentaram, de forma resumida, entre outros dados, as contribuições trazidas pelo Programa, assim como alguns casos com repercussões significativas, tendo em vista transições na legislação trabalhista, o que não deve comprometer o objetivo de consolidar o Programa, conforme declarado na avaliação:

O Programa de Estímulo ao Emprego científico veio contribuir para estabilizar os vínculos constituídos com investigadores através de mecanismos vários, a maioria com sucesso (como documentam os números do

Observatório do Emprego Científico). No entanto, alguns casos, com repercussões significativas, resultaram em distorções do sistema, nomeadamente por combinação da norma transitória do novo regime jurídico com o PREVPAP (Programa de regularização extraordinária dos vínculos precários na Administração Pública). Estas situações, menos conseguidas, em grande parte resultantes de um sistema que tinha casos anacrónicos que se arrastavam há vários anos, em nada devem comprometer o objetivo de consolidar a situação laboral dos atores do sistema de I&D em Portugal (Nazaré; Rendas; Cunha; Santos, 2020, p. 3).

Portanto, o Ponto de Verificação (I) Sumário Executivo recebeu o conceito A, considerando que os autores apresentam, de forma sintética, uma visão geral dos resultados da avaliação do Programa.

Ainda no primeiro nível da Lista de Verificação, o segundo Ponto de Verificação, (II) Prefácio (Introdução), comunica “quem encomendou a avaliação e as razões principais, quais as questões avaliativas e quais os interessados ou públicos que aguardam os resultados da avaliação e/ou deles necessitam” (Elliot, 2011, p. 945). Nesse ponto, os autores indicaram que a avaliação foi encomendada pelo Ministério da Ciência, Tecnologia e Ensino Superior de Portugal, que nomeou uma comissão com o objetivo de avaliar a implementação do Programa de Estímulo ao Emprego Científico, conforme consta no trecho do Relatório meta-avaliado:

O Ministro da Ciência, Tecnologia e Ensino Superior nomeou uma Comissão de Alto Nível¹, tendo por missão avaliar a implementação do Programa de Estímulo ao Emprego Científico, como incentivo à contratação de novos doutorados e ao desenvolvimento de planos de emprego científico e de carreiras científicas pelas instituições científicas e académicas, públicas ou privadas (Nazaré; Rendas; Cunha; Santos, 2020, p. 5).

1

Despacho nº 349/2020, publicado em Diário da República, 2ª, série, de 10 de janeiro de 2020.

O Ponto de Verificação (II) Prefácio (Introdução), recebeu o conceito A. O terceiro e último Ponto de Verificação do nível das Preliminares, (III) Metodologia, apresenta os procedimentos utilizados (incluindo o desenho e abordagem adotados pela avaliação, entre outros procedimentos). Busca-se o entendimento do que é necessário para progredir nos procedimentos metodológicos da avaliação e como foram desenvolvidas, ou o que era preciso e não foi feito no processo avaliativo (Elliot, 2011).

O Ponto de Verificação (III) Metodologia recebeu o conceito D, considerando que o Relatório não apresenta um item para tratar especificamente da metodologia adotada para a avaliação, o que é feito apenas na Introdução do estudo, de forma sintética, conforme esclarecido no texto a seguir.

Entendeu-se começar a abordagem ao problema com uma breve descrição do contexto, em Portugal, no que concerne à evolução do número e áreas de doutoramento e genericamente à questão dos pós-doutorados. Seguidamente desenvolve-se uma análise global da implementação do Programa de Estímulo ao Emprego Científico. As análises desenvolvidas conduzem às recomendações relativas à evolução futura do Programa de Estímulo ao Emprego Científico (Nazaré; Rendas; Cunha; Santos, 2020, p. 5).

Assim, no segundo nível, Fundamentos, numerados de 1 até 5, são tratados alguns atributos da avaliação. O primeiro, (1) Antecedentes e Contexto, no qual é considerada a razão de existir do objeto, podendo ser um projeto, um programa, material, entre outros (Elliot, 2011), foi julgado satisfatório, sendo atribuído o conceito máximo A. O texto que fundamenta a avaliação pode ser encontrado no Relatório meta-avaliado:

Portugal tem um sistema diversificado de programas doutorais, ministrados num número muito alargado de instituições e cobrindo um espectro muito diverso de áreas científicas. Em comparação com outros países europeus

e a partir de 2009, o número de doutores em ciências da educação aumentou, sendo de 8,5%, em 2015, um valor muito elevado em comparação com os 4% identificados, nesse mesmo ano, na Irlanda, Espanha e Suécia (Nazaré; Rendas; Cunha; Santos, 2020, p. 6).

O segundo Ponto de Verificação, (2) Descrições e Definições do objeto, também foi julgado procedente, recebendo conceito A. O trecho extraído do documento corrobora a análise realizada:

A definição mais consensual de pós-doutoramento é a de um período (normalmente de 3 anos) de trabalho em investigação científica e inovação “ainda supervisionado”, durante o qual um jovem doutor aprende a ganhar independência no trabalho de investigação, e adquire competências para prosseguir uma carreira profissional no âmbito da investigação ou noutra setor de atividade público ou privado. É durante este período que é avaliado e avalia a possibilidade de poder vir a seguir uma carreira científica/académica numa instituição de ensino superior ou em instituições de I&D ou uma carreira não científica/académica pública ou privada. Apesar desta amplitude, a expectativa para a generalidade dos que iniciam este percurso, tem sido, até aqui, de aceder a uma carreira académica ou de investigação (Nazaré; Rendas; Cunha; Santos, 2020, p. 8).

O terceiro Ponto de Verificação, (3) Consumidores, que são os principais interessados, neste caso, os investigadores e as instituições científicas e académicas que oferecem os cursos, atinge todos os aspectos apreciados e foi considerado A, como mostrado na página cinco “como incentivo à contratação de novos doutorados e ao desenvolvimento de planos de emprego científico e de carreiras científicas pelas instituições científicas e académicas, públicas ou privadas” (Nazaré; Rendas; Cunha; Santos, 2020, p. 5).

O conceito atribuído ao Ponto de Verificação (4), Recursos, que se caracteriza como aquele disponível para manter o objeto e ajudá-lo a ser eficaz em seus propósitos, foi D. A definição desse

conceito é: “Tem um ou dois elementos que parecem tratar implicitamente do ponto, mas o tratamento é deficiente” (Davidson, 2005 *apud* Elliot, 2011, p. 945). Os autores não mencionam diretamente os recursos (humanos, monetários utilizados), fazendo apenas uma menção à Fundação para a Ciência e a Tecnologia (FCT), principal responsável pelo financiamento da atividade de investigação em Portugal, de acordo com o que se apreende no texto:

Em Portugal, o financiamento da atividade de investigação é responsabilidade principalmente assumida pela Fundação para a Ciência e a Tecnologia (FCT), que goza de estatuto de Instituto Público, recebe financiamento público, para além de outros, mas não tem o mesmo grau de autonomia das suas congéneres europeias, com responsabilidades similares (Nazaré; Rendas; Cunha; Santos, 2020, p. 5).

O último Ponto de Verificação desse nível, (5) Valores, está relacionado diretamente ao padrão adotado pela avaliação: valores relativos a requisitos legais, éticos, profissionais, culturais, históricos, tradicionais, científicos, tecnológicos, políticos, econômicos, o que não esgota a lista de possibilidades (Scriven, 2007).

Na redação do estudo, os autores expõem a importância do objeto estudado, mostrando os embasamentos legais, as circunstâncias econômicas e demográficas. Logo, depreende-se do relatório que a Comissão de avaliadores cumpriu o objetivo para o qual foi nomeada, respeitando os requisitos legais, éticos e demais valores da lista de responsabilidades de Scriven. Assim, recebeu no julgamento o conceito A, que indica o atingimento pleno do critério mais completo de Davidson (2005).

O conjunto de circunstâncias destacadas pelos avaliadores e seu reflexo no Programa avaliado, pode ser visto no texto a seguir:

O conjunto de circunstâncias econômicas, demográficas e os estímulos dos mecanismos de financiamento fez avolumar o número de doutorados a trabalhar em projetos de I&D, predominantemente em universidades e

unidades de I&D, sem grande probabilidade de contratação permanente, dada a exiguidade de lugares de carreira disponíveis, quando existiam. Essa disponibilidade era controlada, dependendo dos países, diretamente pelo Estado, pelas próprias instituições com regime de instituto público com gestão de direito privado ou com regime de instituição sem fins lucrativos ou ainda parcerias público/privado (Nazaré, Rendas; Cunha; Santos, 2020, p. 9).

O terceiro nível, Subavaliação, é onde, segundo Elliot (2011, p. 946):

São dirigidas ao processo e aos resultados ou impactos da avaliação, aos custos (considerados pelo autor como elemento-chave para determinar valor), às comparações de um programa com outro mais efetivo e, também, à capacidade de os componentes da avaliação terem valia para uma futura utilização em outro local. A relevância geral da avaliação também faz parte desse nível, incluindo o que está indo bem, o que está faltando, qual o melhor uso dos recursos, respeitando-se a relação custo-efetividade.

Nesse nível, os Pontos de Verificação estão numerados de seis até 10, onde o primeiro, (6) Processo, não foi demonstrado totalmente, recebendo, com isso, o conceito B, pois trata dos ingredientes principais, mas não são 100% claros. Conforme citado no texto:

Entendeu-se começar a abordagem ao problema com uma breve descrição do contexto, em Portugal, no que concerne à evolução do número e áreas de doutoramento e genericamente à questão dos pós-doutorados. Seguidamente desenvolve-se uma análise global da implementação do Programa de Estímulo ao Emprego Científico. As análises desenvolvidas conduzem às recomendações relativas à evolução futura do Programa de Estímulo ao Emprego Científico (Nazaré; Rendas; Cunha; Santos, 2020, p. 5).

O Ponto de Verificação (7), Resultados ou Impactos da avaliação, foi julgado coerente e recebeu o conceito A, pois atinge todos

os aspectos principais deste ponto e os expressam de forma clara e concisa. O Relatório trata de alguns itens pertinentes:

Os dados disponíveis indicam que o PEEC tem contribuído para o aumento da contratação de investigadores doutorados, nacionais e estrangeiros, pelas instituições públicas e privadas. De referir ainda que muitos institutos politécnicos afirmaram que o PEEC lhes conferiu maior estabilidade ao permitir consolidar algumas equipas de investigação (Nazaré, Rendas; Cunha; Santos, 2020, p. 14).

No Ponto de Verificação (8), Custos, o Relatório também recebeu o conceito A, tendo em vista que, sendo considerado elemento-chave na determinação de valor, os autores apontaram de forma adequada:

Em Portugal, o financiamento da atividade de investigação é responsabilidade principalmente assumida pela Fundação para a Ciência e a Tecnologia (FCT), que goza de estatuto de Instituto Público, recebe financiamento público, para além de outros, mas não tem o mesmo grau de autonomia das suas congéneres europeias, com responsabilidades similares (Nazaré; Rendas; Cunha; Santos, 2020, p. 6).

O Ponto de Verificação (9), Comparações, recebeu o julgamento A, dado que se refere à comparação entre programas vizinhos dos melhores da Europa, conforme citado no texto: “Como já descrito, reconhece-se que o PEEC contribuiu para o aumento da maturidade e qualidade do sistema científico em Portugal, aproximando-o dos melhores na Europa (Nazaré; Rendas; Cunha; Santos, 2020, p. 16). Já o Ponto de Verificação (10), Exportabilidade, não se aplica a esse estudo, considerando ser um programa próprio do governo de Portugal, possuindo características impostas pela legislação local. Por isso, não deve ser considerada sua exportabilidade para outros contextos que apresentem diferenças substantivas em relação ao contexto português, tendo em vista aspectos como leis, culturas e padrões éticos e específicos.

O quarto e último nível da Lista de Verificação de Scriven (2015) é denominado Conclusões. Está numerado de 11 até 15 e diz respeito às recomendações e explanações, que são opcionais, e a responsabilidade dos resultados e justificativas, mas nem sempre as avaliações aportam informações dessas características. Os relatórios trazem a divulgação dos resultados, os formatos e os destinatários.

Finalmente, o último Ponto de Verificação trata, em especial, da meta-avaliação, zelando pelos pontos fortes e fracos da avaliação e as conclusões a respeito de padrões de avaliação: utilidade, viabilidade, adequação e precisão (Elliot, 2011). O primeiro Ponto de Verificação desse nível, Síntese (11), recebeu conceito B, pois atinge a maioria dos aspectos relativos a este ponto, mas não cobre um ou dois deles. No entanto, o Relatório trata de alguns itens pertinentes:

Este conjunto de mecanismos lançados no decorrer da anterior legislatura tem sido objeto de uma monitorização rigorosa e divulgação sistemática, cujos resultados têm sido publicados pelo Observatório do Emprego Científico, revelando de forma visível o esforço realizado de apoio à Ciência e ao Ensino Superior em Portugal, o qual deve ser inequivocamente reconhecido (Nazaré; Rendas; Cunha; Santos, 2020, p. 13).

O Relatório apresenta recomendações na página 16:

Seleção de prioridades dirigida às áreas com mais impacto no desenvolvimento económico e social do País, o que poderá conduzir ao aumento do número de doutorados que opta por trabalhar no sistema produtivo e empresarial e na administração pública em outras carreiras para além das carreiras docentes do ensino superior e de investigação. O desenvolvimento económico e social do País requer reconhecimento explícito da necessidade de priorização de determinadas áreas, quer a nível do governo quer ao nível das instituições científicas e de ensino superior. [...] Esta escolha não será uma medida permanente, mas corresponde unicamente à antecipação possível do futuro (Nazaré; Rendas; Cunha; Santos, 2020, p. 16).

O PEEC é um programa de excelência. No entanto, foi sinalizada a necessidade de adoção de políticas de índole geral cuja concretização requer um esforço exigente, em particular nas circunstâncias, entretanto surgidas. Logo, foi atribuído ao Ponto de Verificação (12), Recomendações e Explicações, por sua abrangência e pertinência, o conceito A.

O Ponto de Verificação (13), Responsabilidade e Justificativa, enfatiza, ao mesmo tempo, a grande importância de métodos sólidos em contextos específicos. Assim, recebeu o conceito A, pois atingiu todos os aspectos principais deste ponto e os expressam de forma clara e concisa. Conforme cita a página 18:

(vi) lançar um programa para incentivar as entidades do sistema científico e tecnológico a contratar doutorados, numa parceria com empresas, ficando um determinado tempo na entidade em regime de aprendizagem que integrará também a empresa. No final desse período de aprendizagem o doutorado seria integrado na empresa. (Nazaré, Rendas; Cunha; Santos, 2020, p. 18).

O Ponto de Verificação (14) recebeu conceito A, pois na meta-avaliação realizada Relatório e Apoio, foi atingido integralmente, trazendo todos os requisitos pertinentes, cumprindo informar apenas que o Relatório e os resultados foram satisfatoriamente contemplados, conforme o texto:

Profissionalizar a atividade de I&D em instituições académicas e científicas. A sustentabilidade do sistema de ciência em Portugal e, portanto, a consolidação do emprego científico requer que as instituições de ensino superior e as de investigação 17 “profissionalizem” a atividade de I&D para os investigadores, isto é, que adotem percursos profissionais integrando posições remuneradas claramente previstas em carreiras/mapas de pessoal, uma parte substancial das quais ocupadas por detentores de contratos de trabalho por tempo indeterminado. Os critérios de progressão e avaliação devem ser transparentes, exigentes e consequentes e serem definidos «por forma

a consolidar na sociedade portuguesa o entendimento de que o conteúdo das publicações científicas e a sua apropriação académica, científica, social ou económica, é muito mais importante do que as métricas de publicação ou a sua apreciação em função das entidades» (in preâmbulo do Decreto-Lei nº 57/2016, de 29 de agosto) (Nazaré; Rendas; Cunha; Santos, 2020, p. 16-17).

No quinto e último Ponto de Verificação desse nível, (15) Meta-avaliação, foram encontrados sinais de atendimento aos padrões do *Joint Committee* (Utilidade, Viabilidade, Adequação e Precisão). No entanto, alguns pontos não foram contemplados plenamente, considerando-se que a meta-avaliação foi realizada de forma satisfatória. Julgado procedente, recebeu o conceito B.

O Quadro 2 expõe os resultados do estudo meta-avaliativo, apresentando os conceitos recebidos em cada ponto de verificação da Lista Chave de Verificação de Scriven.

Quadro 2 – Resultados do estudo meta-avaliativo

Pontos de Verificação		Julgamentos (A - E)
Preliminares	I. Sumário Executivo	A
	II. Prefácio	A
	III. Metodologia	D
Fundamento	1. Antecedentes e Contexto	A
	2. Descrições e Definições	A
	3. Consumidores	A
	4. Recursos	D
	5. Valores	A

Pontos de Verificação		Julgamentos (A - E)
Subavaliações	6. Processo	B
	7. Resultados	A
	8. Custos	A
	9. Comparações	A
	10. Exportabilidade	—
Conclusões	11. Síntese	B
	12. Recomendações e Explicações	A
	13. Responsabilidade e Justificativa	A
	14. Relatório e Apoio	A
	15. Meta-avaliação	B

Fonte: Os autores (2023, adaptado de Scriven, 2007).

CONSIDERAÇÕES FINAIS E RECOMENDAÇÕES

A meta-avaliação se caracteriza por trazer a avaliação como objeto de interesse. É de natureza qualitativa e pretende estabelecer um julgamento de valor, a partir de critérios pré-definidos, com base em informações existentes ou construídas sobre uma determinada avaliação para subsidiar eventuais tomadas de decisão. Assim, tem-se os componentes essenciais da avaliação – informação, estabelecimento de critérios, julgamento e tomada de decisão – aplicados sobre a própria avaliação.

Considerando o Relatório de Avaliação da Implantação do Programa de Estímulo ao Emprego Científico, fruto de um estudo

avaliativo, pode-se concluir que é um Programa de qualidade, que agrega valor aos interessados em ingressar no mercado científico. Conforme pode-se perceber no julgamento aqui sintetizado.

Foram investigados 14 Pontos de Verificação. Apenas um, que faz parte do nível Subavaliações, foi descartado por não se ajustar ao estudo meta-avaliativo proposto. O quantitativo de conceitos que cada Ponto de Verificação recebeu consta a seguir.

Como resposta à Questão meta-avaliativa, pôde-se verificar, considerando os três parâmetros compostos nas Preliminares e mais os 15 pontos de verificação do Fundamento, Subavaliação e Conclusões, um total de 12 deles obteve conceito 'A' Também na meta-avaliação realizada ficou constatado que nenhuma das quatro categorias foi totalmente contemplada com o conceito máximo. Nas Preliminares, o ponto de verificação "III. Metodologia" obteve conceito 'D'. O mesmo aconteceu com "4. Recursos" da categoria "Os Fundamentos". Nas "Subavaliações", o ponto de verificação '6, Processo' obteve conceito 'B', igualmente aos pontos de verificação "11. Síntese" e "15. Meta-avaliação" da Categoria Conclusões.

Pautado nesses resultados considera-se que a avaliação expressa no Relatório meta-avaliado, em que pese sua boa qualidade, considerando os critérios da Lista Chave de Verificação de Scriven, ainda possui pontos de fragilidades.

Em complemento, ressalta-se que os processos de avaliação, e, por conseguinte de meta-avaliação, possuem caracteres reflexivos e críticos. O Programa de Estímulo ao Emprego Científico, por sua vez, incontestavelmente possui qualidade e desempenha papel importante na absorção de formandos oriundos do ensino superior. No entanto, ficou demonstrado que a metodologia não foi explicitada no Relatório, tornando ausente um aspecto relevante ao avaliador. No ponto de verificação Fundamento, no que se refere aos recursos, foi observado também que ficou omissa. Sabendo-se que o

Programa é financiado pelo FCT, o Relatório simplesmente não cita nenhum suporte financeiro.

Entretanto, embora atinjam um conceito satisfatório, há dimensões que ficam abaixo do esperado. Por isso, acredita-se que, de alguma forma, a melhoria desses conceitos deva ser pactuada com o governo português, visando o aperfeiçoamento e a melhoria da qualidade das ofertas de emprego científico para formandos originários de ensino superior.

Face ao exposto, são apresentadas algumas recomendações, visando a implementação do processo avaliativo meta avaliado neste estudo:

1. Discorrer sobre a metodologia utilizada durante a Avaliação, bem como a abordagem, ausentes no Relatório.
2. Identificar os custos empregados na Avaliação, objetivando-se mais transparência nos gastos de recursos públicos, seguindo o princípio de clareza e equidade.
3. Explanar de forma clara e objetiva, como transcorreu o processo, a fim de contribuir para o aumento da maturidade e qualidade do sistema científico em Portugal.
4. Descrever os critérios utilizados do Programa, bem como o método científico adotado, buscando revelar o alcance da qualidade desejada.
5. Atender aos padrões do *Joint Committee* (Utilidade, Viabilidade, Adequação e Precisão), com o objetivo de suscitar maior sustentação e clareza nas dimensões que se deseja atingir, garantindo maior qualidade aos estudos avaliativos a serem realizados.

MÉRITO DO ESTUDO

Ligia Gomes Elliot

Como anunciado, um parecer a cada capítulo do livro *Estudos Meta avaliativos*: diretrizes, categorias, padrões e critérios é acrescentado após o texto meta avaliativo sobre o relatório analisado.

Todos os capítulos são fruto do estudo de um relatório de avaliação sobre programas diversos e foram produzidos por dois mestrandos da disciplina Meta-Avaliação, integrante do currículo da Pós-Graduação em Avaliação, Mestrado, oferecida pela Faculdade Cesgranrio. As duplas de mestrandos são acompanhadas por um orientador.

Em avaliação, interessa e muito conhecer a opinião de especialistas sobre estudos realizados, de modo que, ao ter o conhecimento do todo, se possa aquilatar o mérito do estudo. Segundo o emblemático filósofo da avaliação, Michael Scriven, que lamentável e recentemente nos deixou, mérito diz respeito a características intrínsecas do trabalho avaliativo.

Mas que características são essas? Valem inicialmente para nortear a meta-avaliação realizada do relatório e mais ainda para verificar sua qualidade por meio desta segunda análise, desta feita sobre o que foi apresentado. Assim, torna-se necessário explicitar aspectos importantes ou indispensáveis que tragam, ao leitor, justamente aqueles componentes utilizados pelos autores e que vão responder pelo mérito do trabalho.

No presente capítulo, Meta-Avaliação do Relatório de Avaliação da Implementação do Programa de Estímulo ao Emprego Científico, de autoria das mestrandas Alcineidy Nascimento Falcão e Maria das Graças Madeira e sua orientadora, uma leitura preliminar nos leva a aspectos que compõem o texto meta-avaliativo. Serão,

portanto, enunciados e comentados a partir de evidências contidas no próprio capítulo em análise.

Ao buscar a relevância da meta avaliação, a própria temática do relatório justifica o estudo. O relatório focaliza a Avaliação da Implementação do Programa de Estímulo ao Emprego Científico em 2020 e consta no Relatório de formandos oriundos de cursos superiores em Portugal. O Processo avaliativo foi conduzido por uma comissão instituída pelo Ministério de Ciência, Tecnologia e Ensino Superior de Portugal, composta por representantes de conceituadas Universidades, de Aveiro, do Porto, Nova de Lisboa e do Minho (Falcão, Madeira e Elliot, 2024). O trabalho da comissão reflete o grande interesse desses *stakeholders* no produto entregue por seus cursos de formação para empregos na área científica.

Como afirmam Falcão, Madeira e Elliot (2024, p. 24),

No contexto português, no que se refere ao campo educacional, é patente o pensamento de que os processos de melhoria da qualidade das ofertas dos serviços educacionais em todos os seus campos, tipos e modalidades se tornam, cada vez mais, fundamentais para o desenvolvimento do país.

O relatório representa contribuição significativa para a área educacional, sendo atual pois foi publicado em 2020, além de resguardar originalidade, uma vez que foi a primeira avaliação da implementação do programa. Interessava inicial e primordialmente aos *stakeholders* das instituições de ensino superior portuguesas, como resposta à sociedade ao novo programa de formação tecnológica.

Em relação à organização lógica do que é apresentado, percebe-se claramente o que pretendeu o estudo, em linguagem bastante clara, informativa e organizada, o que leva a assumir um trabalho cuidadosamente elaborado e orientado. Neste aspecto, o respeito às normas ortográficas foi seguido, não sendo encontradas incorreções em relação à escrita da norma culta.

O objeto da meta-avaliação teve sua relevância destacada e reconhecida.

Ao informar os procedimentos metodológicos aplicados à meta-avaliação, as autoras optaram pela escolha da Lista Chave de Verificação da Avaliação, proposta por Scriven (2007). A Lista Chave tem sido amplamente utilizada em avaliações e meta-avaliações, merecendo estudos posteriores e atualizações (Davidson, 2005). Sua utilização no estudo do Relatório foi pertinente, permitindo certificar-lhe as características de mérito.

A estrutura da Lista Chave e o uso de pontos de verificação embasaram a atribuição de conceitos e a emissão de juízos de valor sobre o Relatório. Há clareza e precisão na indicação deste julgamento, o que possibilita conhecer melhor o Relatório, em aspectos positivos, plenamente alcançados, e naqueles que apresentaram inconsistências.

A síntese dos resultados e as recomendações feitas são úteis e bastante pertinentes. A leitura atenta do capítulo e, se possível, do próprio Relatório meta-avaliado representam oportunidade de aprendizagem sobre a metodologia adotada.

REFERÊNCIAS

AMARAL, Alberto. Bolonha, o ensino superior e a competitividade econômica. *In*: SERRALHEIRO, Jose P. **O processo de Bolonha e a formação dos educadores e professores portugueses**. Porto: Profedições, 2005.

BATISTA, Maria Clarice Lima. **Produção científica dos egressos de pós-graduação em Psicologia**: redes de colaboração e domínios científico. Orientador: Sérgio Dias Cirino. Coorientador: Thiago Magela Rodrigues Dias. 2019. 158 f. Dissertação (Mestrado em Psicologia) - Universidade Federal de Minas Gerais, Faculdade de Filosofia e Ciências Humanas, Belo Horizonte, 2019. Disponível em: <https://repositorio.ufmg.br/handle/1843/30140>. Acesso em: 28 jun. 2023.

BRENNAN, John; SHAH, Tarla. **Quality assessment and institutional change:** experiences from 14 countries. London: Centre for higher Education Research and Information, 2000.

DAVIDSON, Jane E. **Evaluation methodology basics.** Thousands Oaks, CA: Sage, 2005.

DAVOK, Delsi F. **Modelo de meta-avaliação de processos de avaliação da qualidade de cursos de graduação.** Orientador: Jair dos Santos Lapa. 2006. 274 f. Tese (Doutorado) – Programa de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2006. Disponível em: <https://repositorio.ufsc.br/xmlui/bitstream/handle/123456789/88899/231045.pdf?sequence=1&isAllowed=y>. acesso em 28 jun. 2023.

ELLIOT, Ligia Gomes. Meta-avaliação: das abordagens às possibilidades de aplicação. **Ensaio:** aval. pol. públ. educ., Rio de Janeiro, v. 19, n. 73, p. 941-964, out./dez. 2011.

ELLIOT, Ligia Gomes. FILHO, Ovídio Orlando. **Meta avaliações de políticas públicas do governo:** o que os relatórios relevam os relatórios de avaliação. Fundação Cesgranrio, Rio de Janeiro, 2019.

FREITAS, Antônio Alberto da Silva Monteiro de. Avaliação da educação superior no Brasil e Portugal homogeneização ou diferenciação?. **Avaliação.** Campinas; Sorocaba, SP, v. 17, n. 1, p. 119-136, 2012. Disponível em: <https://doi.org/10.1590/S1414-40772012000100007>. Acesso em: 27 jun. 2023.

MOLLIS, Marcela; MARGINSON, Armon. The assessment of universities in Argentina and Austrália: between autonomy and heteronomy. **Higher Education.** [S.l.], v. 43, p. 311-330, 2002. Disponível em: https://www.researchgate.net/publication/226556782_The_assessment_of_universities_in_Argentina_and_Australia_Between_autonomy_and_heteronomy. Acesso em: 27 jun. 2023.

NAZARÉ, Maria Helena; RENDAS, António; CUNHA, António M.; SANTOS, José Marques dos. **Relatório de avaliação da implementação do Programa de Estímulo ao Emprego Científico:** Comissão de avaliação. Portugal, 2020. Disponível em: <https://former.fct.pt/apoios/contratacaodoutorados/empregocientifico/docs/RelatorioProgramaEstimuloEmpregoCientifico.pdf>. Acesso em: 27 jun. 2023.

SCRIVEN, Michael. **Evaluation thesaurus.** 4. ed. Newbury Park, CA: Sage, 1991.

SCRIVEN, Michael. **Key evaluation checklist**. Kalamazoo, MI: Western Michigan University, 2015. Disponível em: <https://files.wmich.edu/s3fs-public/attachments/u1105/2023/kec-scriven.pdf>. Acesso em: 10 nov. 2024.

SCRIVEN, Michael. **Standards for evaluations of educational programs, projects, and materials**. Hightstown, NJ: McGraw-Hill Book Co., 1981.

WORTHEN, Blaine R.; SANDERS, James R.; FITZPATRICK, Jody L. **Avaliação de programas: concepções e práticas**. São Paulo: Editora Gente, 2004.

YARBROUGH, Donald B., SHULLA, Lyn M.; HOPSON, Rodney K.; CARUTHERS, Flora A. **The program evaluation standards: a guide for evaluators and evaluation users**. 3. ed. Thousands Oaks, Califórnia: Sage, 2011.

2

Dalva Carneiro
Albertina Thereza Pacheco Rodrigues
Ovidio Orlando Filho

AVALIAÇÃO DO RELATÓRIO DE AVALIAÇÃO SOBRE AS INFORMAÇÕES DE UM DESCUMPRIMENTO DE CONDICIONALIDADES DO PROGRAMA BOLSA FAMÍLIA:

UM ESTUDO META-AVALIATIVO

INTRODUÇÃO

Usualmente, estamos sempre avaliando alguma coisa ou fato ocorrido e emitindo algum juízo de valor. Segundo Worthen, Sanders e Fitzpatrick (2004), a prática da avaliação precede a sua definição e remonta aos primórdios da história humana, podendo mesmo ser considerada como característica fundamental das pessoas. Considerando o campo das organizações, tanto públicas quanto privadas, atualmente a realização de avaliações passou a fazer parte de suas rotinas, visando adequar seus processos de gestão às inovações trazidas pelo veloz desenvolvimento dos processos de produção de bens e serviços.

No campo da administração pública, Carvalho (1999) considera que emissão de juízos de valor sobre as ações governamentais são revestidas de elevado grau de complexidade. Esse fato aumenta os níveis de dificuldades na realização processos avaliativos, pois eles podem assumir diversas feições como, avaliações de pessoas, de projetos e programas, por exemplo, devendo ser considerados seus objetivos, motivos, consequências e desfechos. Assim, ao avaliar políticas públicas de modo geral, face à importância e impactos que elas acarretam na vida das pessoas e das organizações, é necessário que sejam utilizados critérios e padrões bem definidos e testados, no sentido de se realizar um processo avaliativo possuidor de um modelo formal e adequadamente estruturado. Ao avaliar políticas públicas, face à importância e ao impacto que elas acarretam na vida das pessoas, faz-se necessário trabalhar com uma avaliação mais formal e estruturada.

Comumente as pessoas, de modo informal, avaliam os serviços públicos oferecidos à medida que suas expectativas e necessidades são ou não atendidas. Segundo Worthen, Sanders e Fitzpatrick (2004), a avaliação é uma forma básica de comportamento humano e, no modo informal, a pessoa faz escolhas baseadas em percepções subjetivas.

Com o avanço tecnológico e com as informações circulando de forma rápida em todas as camadas sociais, os indivíduos começam a tomar consciência dos seus direitos e deveres em relação à sociedade e, como cidadãos, do estado em relação a eles. Assim, buscam de forma mais organizada acompanhar os programas de políticas públicas dos governos destinados a atender suas necessidades, considerando as diferentes instâncias administrativas do país. De tal modo, a implementação de avaliações dos programas e projetos implementados pela administração pública se torna fundamental, pois possibilita que haja transparência entre as intervenções promovidas pelo Estado e a população, a qual sofre os impactos dos serviços públicos, sejam eles bem prestados ou não (Elliot, 2019).

Nesse campo de ação governamental, a Controladoria-Geral da União (CGU) se apresenta com grande relevância, pois, “é o órgão do governo federal responsável pela defesa do patrimônio público e pelo incremento da transparência na gestão, por meio de ações de controle interno, auditoria pública, correição, ouvidoria e prevenção e combate à corrupção” (Controladoria Geral da União, 2023).

A metodologia utilizada pela CGU visa avaliar a execução dos programas federais, o alcance dos objetivos e a adequação do gerenciamento, tendo como primeiro processo a Tomada de Decisão com as seguintes fases: mapeamento e conhecimento das políticas públicas; hierarquização das políticas públicas; e priorização de ações de governo para avaliação (Controladoria Geral da União, 2015).

Como exemplo de programas avaliados pela CGU, pode-se citar o Programa Bolsa Família, possuidor de grande relevância no Governo brasileiro. Criado em 2003 e regulamentado pelo decreto nº 5.209/2004 (Brasil, 2004), o Bolsa Família se caracteriza como um programa de transferência direta e condicionada de renda, objetivando diminuir a desigualdade social e a pobreza do país. Para ser beneficiário do Programa, além de estar em condição de pobreza, é necessário cumprir certas condicionalidades relacionadas à educação

e à saúde. As condicionalidades abrangem a realização de exames pré-natais, o acompanhamento nutricional, a vacinação e o cumprimento da frequência escolar mínima (Martins; Ruckert, 2022).

O Programa é considerado e reconhecido internacionalmente por já ter tirado milhões de pessoas da fome (Brasil, 2023). A esse respeito, o Bolsa Família visa integrar políticas públicas assegurando direitos básicos como saúde, educação e assistência social. Alguns requisitos básicos são solicitados às famílias para terem direito aos seus benefícios. Para recebê-los, as famílias, entre outros requisitos, têm o compromisso de cumprir obrigações relacionadas à frequência escolar, à saúde de crianças e gestantes, caracterizadas como “condicionalidades”. Especificamente no campo da Educação, o Programa Bolsa Família aparece com o intuito de solucionar diversos problemas, como, por exemplo, a evasão escolar, cujas variáveis de resultados mais utilizadas são: frequência escolar, evasão escolar e taxa de matrícula (Araujo, Frio; Alves, 2021).

Por se tratar de um programa de grande escala e impacto, a CGU emitiu em 2021 o Relatório de Avaliação, como parte do Ciclo 2020 de avaliações do Conselho de Monitoramento e Avaliação de Políticas Públicas (CMAP), sobre a política de transferência de renda diretamente às famílias em condição de pobreza e extrema pobreza - o Bolsa Família -. Trata-se do maior programa de transferência condicionada de renda do mundo em números absolutos, beneficiando uma quantidade superior a 13 milhões de famílias – referência de 01/2020, com orçamento anual regular aproximado de R\$ 32 bilhões nos exercícios de 2019 e 2020 (Controladoria Geral da União, 2021).

A metodologia utilizada para nortear as avaliações da CGU é disponibilizada na internet por meio do Manual da Metodologia para Execução de Programas de Governo (AEPG). A publicação tem por objetivo orientar servidores e outros órgãos públicos na função constitucional de avaliar a execução de políticas públicas. De acordo com esse Manual:

A avaliação da execução de programas de governo caracteriza-se por ter como diretriz principal a obtenção de diagnósticos acerca da eficácia, eficiência e economicidade, assim como, quando possível, da efetividade das políticas públicas, diferenciando-se fortemente, portanto, das auditorias de conformidade (compliance) (Controladoria-Geral da União, 2015).

Dessa forma, as atividades desenvolvidas pela CGU visam garantir a defesa do patrimônio público, considerando seu patrimônio e incrementando a transparência da gestão, por meio de realizações de auditoria, correição, prevenção, combate à corrupção e ouvidoria (Controladoria Geral da União, 2023).

OBJETIVO, OBJETO DO ESTUDO E QUESTÃO META-AVALIATIVA

O presente artigo possui como objetivo realizar uma meta-avaliação da avaliação realizada pela CGU sobre a política Transferência de Renda Diretamente às Famílias em Condição de Pobreza e Extrema Pobreza - Bolsa Família, tendo como objeto de análise o Relatório Final da referida avaliação e que foi elaborado pela citada instituição.

No sentido de responder o objetivo no presente estudo, procurou-se responder a seguinte questão meta-avaliativa:

Até que ponto a avaliação da política Transferência de Renda Diretamente às Famílias em Condição de Pobreza e Extrema Pobreza - Bolsa Família -, realizada pela CGU, cumpriu os padrões de qualidade definidos por Yarbrough, Shulha, Hopson e Caruthers (2011) na obra *The program evaluation standards: a guide for evaluators and evaluation users*.

A AVALIAÇÃO DA CGU: APRESENTAÇÃO E CARACTERÍSTICAS

A avaliação realizada pela CGU focalizou a análise das estruturas de governança do acompanhamento das condicionalidades previstas na legislação que implementou o Programa, especialmente no que tange a informação sobre possíveis descumprimentos daquilo que foi previamente instituído.

Assim, a avaliação se concentrou no acompanhamento das referidas condicionalidades, especialmente no modo com que informações sobre seus possíveis descumprimentos ocorreram, objetivando gerar recomendações no sentido de superar as vulnerabilidades encontradas. Com isso, procurou-se auxiliar os Ministérios da Cidadania (MC), da Educação (MEC) e da Saúde (MS) para a implementação de melhorias nas políticas públicas implementadas pelos citados órgãos da federação, sempre orientadas para a superação das vulnerabilidades que pudessem impactar negativamente no auxílio às famílias beneficiadas pelo Programa.

Ressalta-se que o processo de acompanhamento de condicionalidade do Programa, como elemento fundamental para a melhoria ou a introdução de novas políticas públicas, proporciona informações úteis para a tomada de decisão na vida dessas famílias. Por sua vez, destaca-se que as famílias que não cumprem as condicionalidades ficam em situação de maior vulnerabilidade, sendo privadas de seus direitos básicos, pois elas possuem prioridade no atendimento pela assistência social em suas localidades, o que pode prejudicar seus atendimentos pelos direitos que possuem em relação ao Programa. Dessa forma, essas famílias necessitam ter conhecimento desses registros no sentido de terem condições de enfrentamento a essas vulnerabilidades.

Uma questão principal norteou o estudo avaliativo: “De que forma as informações de descumprimento de condicionalidades são utilizadas para subsidiar políticas e orientar ações direcionadas à superar as vulnerabilidades das famílias?” (Controladoria Geral da União, 2021 p. 7).

Derivada da questão principal, foram elaboradas algumas questões específicas sobre normatização do monitoramento e utilização de informações para a tomada de decisões, procurando-se conhecer: se existiram responsabilidades sobre o tratamento das informações; se as informações sobre o acompanhamento das famílias que descumprem condicionalidades foram apresentadas aos que tomam as decisões e se são avaliadas periodicamente; se as informações possibilitaram identificar as famílias com maior vulnerabilidade; se houve controle na manutenção de informações atualizadas no Sistema de Condicionalidades (Sicon) e se os ministérios da Cidadania, Saúde e da Educação utilizaram tais informações de condicionalidades para promover o aprimoramento das políticas.

O trabalho de campo ocorreu entre os dias 22/09/2020 a 15/01/2021 por meio de testes, análises, indagações e consolidação de informações coletadas pela equipe de auditoria que atendeu rigorosamente às normas de auditoria aplicáveis ao serviço público federal.

FUNDAMENTOS METODOLÓGICOS ADOTADOS NO ESTUDO META-AVALIATIVO

De acordo com Scriven (2018), a meta-avaliação pode ser formativa, somativa ou ambas. Quando cumprem o papel formativo podem sugerir o aprimoramento do processo de avaliação antes que seja tarde para corrigi-lo, mas quando assumem o caráter somativo, avaliando após o término da avaliação, pode aumentar a confiabilidade dos resultados alcançados em uma avaliação já finalizada.

Neste trabalho adotou-se uma abordagem somativa, isto é, o objeto de estudo, o Relatório final do processo avaliativo posto em prática pela CGU e elaborado depois da avaliação encerrada, foi o documento principal sobre o qual procurou-se debruçar, analisar e emitir juízos de valor sobre qualidade da avaliação expressa no documento. Também foram realizadas análises bibliográfica e documental em livros, periódicos e instrumentos legais relacionados ao tema do estudo, procedimento necessário e adequado ao entendimento do que se pretendeu analisar no relatório meta-avaliado. Por isso, devido ao subjetivismo, sempre presente em avaliações dessa natureza, o conteúdo do estudo está sujeito a pontos e contrapontos relacionados a possíveis pensamentos divergentes em relação aos conceitos emitidos em seu escopo.

Cabe observar que o estudo teve caráter qualitativo, sendo conduzido por meio de interpretações do conteúdo do Relatório, tendo como referência os 30 padrões do *Joint Committee on Standards for Educational Evaluation* (JCSEE), cujas normas orientam avaliadores sobre as características que devem possuir as boas avaliações. Os padrões são categorizados em cinco diferentes atributos com o objetivo de aferir a qualidade de avaliações em aspectos relacionados a sua utilidade, exequibilidade, adequação, precisão e responsabilização (American National Standards Institute, 2023).

De acordo com Yarbrough, Shulha, Hopson e Caruthers (2011), o primeiro atributo do *Joint Committee* é o de Utilidade de uma avaliação (Utility Standards). Ele representa um importante componente do valor da avaliação e é aferido por meio do nível de satisfação das necessidades dos *stakeholders*, considerando a relação aos seus processos e produtos. Ainda segundo os autores, “um bom ponto de partida para o entendimento do proveito da avaliação é examinar a variedade de utilizações possíveis de seus processos, resultados e produtos”. Esse atributo é composto por oito padrões: Credibilidade do Avaliador, Atenção aos Interessados (*stakeholders*), Propósitos Negociados, Explicitação de Valores, Informação Relevante, Produtos

e Processos Significativos, Comunicação e Relatórios Apropriados e no Prazo e Preocupação com Consequências e Influências (Yarbrough; Shulha; Hopson; Caruthers, 2011, p. 4).

O segundo atributo do *Joint Committee* é o da Exequibilidade da avaliação (Feasibility Standards). Uma avaliação exequível é aquela cuja implementação possui potencialidades de alcançar um nível adequado tanto no que se refere à sua efetividade quanto aos aspectos relacionados à sua eficiência. Trabalhar esses dois parâmetros, efetividade e eficiência, está diretamente relacionado à necessidade de se agregar valor aos processos avaliativos (Yarbrough; Shulha; Hopson; Caruthers, 2011, p. 4). De acordo com Yarbrough, Shulha, Hopson e Caruthers (2011), são quatro os padrões deste atributo: Gerenciamento do Projeto, Procedimentos Práticos, Viabilidade do Contexto e Uso dos Recursos. Para os autores a atenção à exequibilidade acrescenta valor às avaliações, em três modos:

1. Em primeiro lugar, a atenção à exequibilidade destaca os requisitos administrativos e logísticos da avaliação que precisam ser administrados. A boa administração de um projeto de avaliação melhora as probabilidades de sucesso.
2. Em segundo lugar, a atenção à exequibilidade une os procedimentos possíveis aos procedimentos práticos para uma avaliação específica. Os avaliadores devem perguntar: "Posso adotar estes procedimentos, neste momento, para este programa?"
3. Em terceiro lugar, uma exequibilidade adequada serve como pré-condição para outros atributos da qualidade. [...] A maior eficiência pode liberar recursos a serem canalizados para necessidades adicionais, tais como mais competência cultural ou mais informações. Assim, o uso efetivo e eficiente de recursos permite melhoria do impacto e da qualidade da avaliação (Yarbrough; Shulha; Hopson; Caruthers, 2011, p. 72, tradução livre).

O terceiro Atributo do Joint Committee é o da Adequação (Propriety) da avaliação. É composto por sete padrões: Orientação Responsiva e Inclusiva, Acordos Formais, Respeito e Direitos Humanos, Clareza e Equidade, Transparência e Abertura, Conflitos de Interesses e Responsabilidade Fiscal (Yarbrough; Shulha; Hopson; Caruthers, 2011). Esse conjunto de padrões faz referência ao que é apropriado, verdadeiro, legal, correto, oportuno e justo na avaliação. A adequação abarca três campos que se sobrepõem: 1) Direitos, deveres e responsabilidades dos avaliadores; 2) Sistemas de leis, regulamentos e regras; 3) Funções e obrigações inerentes à prática profissional da avaliação (Yarbrough; Shulha; Hopson; Caruthers, 2011, p. 106, tradução nossa).

Para Yarbrough, Shulha, Hopson e Caruthers, (2011), o quarto atributo do *Joint Committee* se refere à Precisão da avaliação (Accuracy Standards) e é composto de oito padrões: Conclusões e Decisões Justificadas, Informação Válida, Informação Fidedigna, Explicitação das Descrições do Programa e do Contexto, Gerenciamento da Informação, Análises e Delineamentos (designs) Sólidos, Explicitação do Raciocínio da Avaliação e Comunicação e Relatório. Para os autores, os padrões de Precisão da avaliação resultam da:

Exatidão das representações, proposições e resultados, em especial aqueles que fundamentam julgamentos acerca da qualidade dos programas ou de seus componentes. Em geral, alcança-se a precisão por meio de teorias, justificativas, métodos e delineamentos solidamente fundamentados. A avaliação deve buscar o máximo de precisão que for exequível, adequada e útil ao embasamento de conclusões e decisões sólidas em uma situação específica (Yarbrough; Shulha; Hopson; Caruthers, 2011, p. 158, tradução nossa).

Por fim o quinto atributo do Joint Committee é Responsabilização da avaliação (Evaluation Accountability), composto pelos padrões Documentação da avaliação, Meta-avaliação Interna e Meta-avaliação Externa, se refere à utilização responsável

de todos os recursos da (ou relacionados à) avaliação, com a finalidade de produzir valor. Segundo Yarbrough, Shulha, Hopson e Caruthers (2011, p. 226), a fundamentação da Responsabilização da avaliação está abalizada nas respostas das seguintes questões:

1. Como os programas são implementados?
2. Como os programas podem ser melhorados?
3. Qual o proveito, o significado ou a importância que os *stakeholders* conferem aos programas; e
4. Qual a relação entre os custos (inclusive em oportunidades) e os benefícios do programa?

Similarmente, Yarbrough, Shulha, Hopson e Caruthers (2011, p. 226) destacam que a análise desse atributo procura investigar:

1. Como as avaliações são implementadas;
2. Como as avaliações podem ser melhoradas;
3. Qual o proveito, o significado ou a importância que os *stakeholders* conferem às avaliações; e
4. Qual a relação entre os custos (inclusive em oportunidades) e os benefícios da avaliação.

Baseado na definição de cada atributo, e seus conjuntos de padrões, ficam estabelecidos os critérios constantes do Quadro 1 para valorar cada um deles, considerando as análises realizadas no conteúdo do Relatório avaliado.

Quadro 1 – Critérios de atendimento dos atributos do *Joint Committee*

Atributos	Critérios de atendimento
Utilidade e Precisão	Foram considerados Totalmente Atendidos (TA) quando no mínimo sete de seus padrões foram atendidos; foram considerados como Parcialmente Atendido (PA) quando pelo menos quatro dos seus padrões foram atendidos; os atributos foram considerados como Não Atendidos (NA) quando menos de quatro de seus padrões não foram atendidos.
Exequibilidade	Foi considerado Totalmente Atendido (TA) quando os seus quatro padrões foram atendidos; foi considerado como Parcialmente Atendido (PA) quando dois ou três de seus padrões foram atendidos; o atributo foi considerado como Não Atendido (NA) quando menos dois de seus padrões não foram atendidos.
Adequação	Foi considerado Totalmente Atendido (TA) quando no mínimo seis de seus padrões foram atendidos; quando pelo menos quatro dos seus padrões foram atendidos; os atributos foram considerados como Não Atendidos (NA) quando menos de quatro de seus padrões não foram atendidos.
Responsabilização	Foi considerado Totalmente Atendido (TA) quando os seus os três padrões foram atendidos; foi considerado como Parcialmente Atendido (PA) quando dois de seus padrões foram atendidos; o atributo foi considerado como Não Atendido (NA) quando menos de dois de seus padrões não foram atendidos.

Fonte: os autores (2023).

ANÁLISES E RESULTADOS DA META-AVALIAÇÃO

Após detalhada leitura realizada no Relatório de Avaliação sobre a política Transferência de Renda Diretamente às Famílias em Condição de Pobreza e Extrema Pobreza - Bolsa Família, foram realizadas as seguintes análises de seu conteúdo, considerando cada atributo e suas fundamentações de acordo com o estipulado por Yarbrough, Shulha, Hopson e Caruthers, (2011).

O ATRIBUTO DE UTILIDADE

Yarbrough, Shulha, Hopson e Caruthers, (2011, p. 3) afirma que para haver Credibilidade do Avaliador “as avaliações devem ser conduzidas por pessoal qualificado que estabeleça e mantenha credibilidade no contexto da avaliação”. Nesse sentido, o relatório estudado foi elaborado por um órgão da administração pública federal que possui grande expertise na condução de avaliações e monitoramento de programas sociais implementados pelo governo brasileiro, a CGU. Por si só, esse aspecto classifica este padrão como atendido, considerando os critérios estabelecidos no estudo.

Segundo Yarbrough, Shulha, Hopson e Caruthers, (2011, p. 3), para que haja Atenção aos Interessados (*stakeholders*) “as avaliações devem dedicar atenção a todos os indivíduos e grupos envolvidos no programa e por elas afetados”. Logo, em relação a este padrão, apesar de corretamente sintetizados e explicitados, os resultados da avaliação não apresentam indicações de que tenham chegado a todos os indivíduos e grupos envolvidos no programa, ou seja, não há indicação se os resultados dessa avaliação foram conhecidos, discutidos e criticados pelos envolvidos nas instâncias mais afetadas.

Segundo Yarbrough, Shulha, Hopson e Caruthers, (2011, p. 3), para que os propósitos sejam negociados eles devem ser identificados e continuamente negociados com base nas necessidades dos *stakeholders*. Nesse sentido, o relatório do estudo não evidencia esse aspecto de negociação. O Relatório avaliado não contempla esse padrão, pois, o aspecto considerado não é expresso em sua redação.

Para Yarbrough, Shulha, Hopson e Caruthers, (2011), no padrão relacionado à explicitação de valores, uma avaliação deve especificar os valores culturais e individuais que fundamentam objetivos, processos e julgamentos. O relatório em questão, deixa perceber essa importância em relação aos valores culturais no que se refere à necessidade de combater a pobreza, a extrema pobreza

e as preocupações com à saúde e à cidadania, conforme pode ser destacado no trecho a seguir:

O processo de acompanhamento de condicionalidades do PBF (Anexo II) gera um grande volume de informações sobre a frequência escolar e sobre as condições de saúde das famílias beneficiárias. Essas informações podem ser utilizadas para subsidiar a tomada de decisão sobre melhorias das condições de vida dessas famílias (Controladoria-Geral da União, 2021 p. 7).

Segundo Yarbrough, Shulha, Hopson e Caruthers, (2011), a Informação deve ser relevante no sentido de servir às necessidades identificadas e emergentes dos *stakeholders*. Nesse sentido, a avaliação realizada soube identificar as necessidades emergentes dos *stakeholders* por meio de informações obtidas e constantes do Anexo VI do Relatório. Entretanto, a utilidade só será garantida se houver monitoramento do seu aproveitamento, o que somente poderá ser aferido a *posteriori*, considerando a futura utilização dos resultados da avaliação realizada.

Segundo Yarbrough, Shulha, Hopson e Caruthers, (2011), a apresentação de produtos e processos significativos devem construir atividades, descrições e julgamentos de modo que os participantes se sintam incentivados a redescobrir, reinterpretar sua compreensão e seus comportamentos. No presente estudo, o relatório oferece recomendações estimulantes aos interessados que poderão ser úteis aos *stakeholders*, como a expressa no trecho a seguir:

Achado nº 2 – Desenvolver, em conjunto com os Ministérios da Educação e da Saúde, novas estratégias de abordagem que facilitem a identificação e a obtenção de informações sobre os motivos de descumprimento das condicionalidades pelas famílias beneficiárias (Controladoria Geral da União, 2021 p. 39).

De acordo com Yarbrough, Shulha, Hopson e Caruthers, (2011), a Comunicação e os relatórios apropriados e no prazo são

fundamentais para que as avaliações atendam às constantes necessidades de informação de seus vários públicos. A esse respeito, o relatório recomenda continuidade no acompanhamento, como a seguir exposto:

Achado nº 3 - Implementar melhorias do Sistema de Condicionalidades (Sicon) de forma que seja possível identificar e caracterizar todas as famílias beneficiárias com acompanhamento de condicionalidades, principalmente aquelas em descumprimento, acessando, por meio do Sistema, informações com adequado nível de detalhamento, úteis ao gestor no mapeamento das dificuldades existentes, incluindo os seguintes aspectos, não exaustivos: (...) iv. Incluir no relatório de acompanhamento familiar extraído do Sistema, informações sobre a situação em que a família se encontrava quando a assistência social iniciou seu acompanhamento (família beneficiária, com condicionalidades, em descumprimento, o motivo da intervenção da assistência social) (Controladoria Geral da União, 2021 p. 39).

Quanto ao prazo do relatório, é possível dizer que o Relatório apresenta em seu escopo a intencionalidade de atender com rapidez e adequação às necessidades detectadas, considerando o volume e a complexidade dos dados. Contudo, não está evidenciado a estratégia utilizada para a devolução destes resultados o que é sumamente importante para a sua melhor utilização.

Yarbrough, Shulha, Hopson e Caruthers (2011) indicam que nos aspectos relacionados à preocupação com consequências e influências, a avaliação deve promover o uso responsável e adequado de seus resultados além de prevenir consequências negativas e indesejáveis. Nesse ponto, o relatório prevê o acompanhamento das consequências, relatando que procurará identificar as necessidades identificadas pela avaliação no sentido de se tomar as providências cabíveis para cada caso, conforme pode ser observado no trecho a seguir:

Para os Ministérios setoriais envolvidos, normatizar o processo de acompanhamento das condicionalidades do Programa Bolsa Família – PBF, a fim de que sejam incluídos procedimentos de avaliação regular dos dados relativos ao descumprimento de condicionalidades pelas famílias beneficiárias. Esses ajustes, em conformidade com as melhores práticas de governança, deverão prever, pelo menos, a realização de reportes regulares para a alta administração (Controladoria-Geral da União, 2021 p. 39).

Em consonância com os critérios de valoração dos atributos (quadros 1 e 2), o relatório da CGU atendeu parcialmente aos padrões estabelecidos pelo *Joint Committee* no atributo Utilidade. Em síntese, o Padrão Utilidade está presente na avaliação em questão na medida em que ela foi conduzida por uma organização competente em seus profissionais e focalizou sua atenção nos interessados envolvidos.

Nesse sentido, ficou evidenciado a atenção prioritária aos valores culturais dos beneficiários do programa bolsa família, no que diz respeito, principalmente, às suas condições de pobreza e extrema pobreza. Sua relevância se destaca na atenção às vulnerabilidades às quais os beneficiários do programa bolsa família estão sujeitos.

O relatório da avaliação expressa também recomendações estimulantes aos beneficiários envolvidos no sentido de darem prosseguimento ao cumprimento das condicionalidades. A continuidade do processo avaliativo e as consequentes modificações necessárias que surgirem são recomendadas na conclusão da avaliação.

Embora se trate de um estudo profissionalmente conduzido, evidenciando a qualidade aqui apontada, vale fazer algumas recomendações pertinentes ao Padrão da Utilidade, ou seja, introduzir o processo de negociação com os interessados na tomada de decisões em todo o processo avaliativo desde a sua concepção até seus resultados e negociações. Outro aspecto que merece ser implementado para reforçar a utilização dos resultados é o monitoramento continuado desse aproveitamento. Nesse acompanhamento será

possível identificar a preocupação com as consequências e influências dos resultados.

Quadro 2 – Resultados da meta-avaliação em relação ao Atributo de Utilidade

Padrões	Classificação do Atendimento
Credibilidade do Avaliador	A
Atenção aos Interessados (<i>stakeholders</i>)	NA
Propósitos Negociados	NA
Explicitação de Valores	A
Informação Relevante	PA
Produtos e Processos - Significativos e comportamentos	A
Comunicação e Relatórios Apropriados e no Prazo	PA
Preocupação com Consequências e Influências	A

Legenda: A - atendido; PA - parcialmente atendido; NA - não atendido.

Fonte: Os autores (2023).

O ATRIBUTO DE EXEQUIBILIDADE

De acordo com Yarbrough, Shulha, Hopson e Caruthers (2011), o Gerenciamento do projeto exige que as avaliações devam adotar estratégias que conduzam a um gerenciamento eficaz. Nesse sentido, o gerenciamento inclui o planejamento, a programação, a seleção de pessoal, o monitoramento, a comunicação, a aquisição e a análise de riscos. De um modo geral, essas etapas foram cumpridas pela densidade de suas informações e detalhamento de seus resultados.

De acordo com Yarbrough, Shulha, Hopson e Caruthers (2011), os Procedimentos Práticos sugerem que o relatório da avaliação seja detalhado e extenso descrevendo funções, articulações, papéis e

responsabilidades dos integrantes responsáveis pelo processo avaliativo. Entende-se que o relatório em questão tenha cumprido esses requisitos na sua concepção, principalmente considerando as informações constantes de seus anexos. Entretanto, não sintetiza uma estruturação dos procedimentos de coleta de informação incluindo os responsáveis para essa condução, o que seria uma informação útil para os responsáveis pelo programa avaliado.

De acordo com Yarbrough, Shulha, Hopson e Caruthers (2011), a Viabilidade do Contexto significa que as avaliações devem identificar, monitorar e equilibrar interesses políticos e culturais e necessidades individuais e de grupos. No relatório, é visível nas recomendações o cumprimento dessas exigências (Controladoria Geral da União, 2021 p. 39).

De acordo com Yarbrough, Shulha, Hopson e Caruthers (2011), o Uso dos recursos significa que as avaliações devem utilizar os recursos com efetividade e eficiência. No Relatório da avaliação realizada não é mencionada uma análise a respeito do aproveitamento de recursos para a realização da avaliação conduzida pela CGU.

As análises realizadas no Relatório avaliado, em relação ao Atributo da Exequibilidade da Avaliação apresentaram os seguintes resultados, expressos no Quadro 3 a seguir.

Quadro 3 – Resultados da meta-avaliação em relação ao Atributo de Exequibilidade

Padrões	Classificação do Atendimento
E1 Gerenciamento do Projeto	A
E2 Procedimentos Práticos	PA
E3 Viabilidade do Contexto	A
E4 Uso dos Recursos	NA

Legenda: A - atendido; PA - parcialmente atendido; NA - não atendido.

Fonte: Os autores (2023).

De acordo com os critérios de atendimento dos atributos (quadros 1 e 3), o relatório da CGU atendeu parcialmente aos padrões estabelecidos pelo *Joint Committee on Standards for Educational Evaluation* no atributo Exequibilidade. Em síntese, a presente avaliação atendeu, ainda que parcialmente, ao Padrão de Exequibilidade no que diz respeito à condução de um Programa de grandes proporções, com eficiência e efetividade. Mais especificamente, o Gerenciamento do Projeto é percebido no relato minucioso das diferentes fases do seu desenvolvimento. Com relação aos Procedimentos Práticos a quantidade e a qualidade das informações obtidas deixam perceber a boa prática no atendimento a esse atributo. Sobre a Viabilidade do Contexto, vale dizer que é visível a intenção de um equilíbrio entre valores culturais e políticos e necessidades individuais e de grupo. Contudo, com o propósito de oferecer uma contribuição construtiva a essa avaliação é recomendável que os procedimentos práticos na metodologia de coleta de informações sejam explicitados para melhor compreensão do estudo na sua totalidade. Similarmente, o Uso de Recursos na condução e concepção da avaliação, embora subentendido, convém que ele seja explicitado no seu planejamento e no seu conteúdo para se obter informação sobre custo/efetividade.

O ATRIBUTO DE ADEQUAÇÃO

De acordo com Yarbrough, Shulha, Hopson e Caruthers (2011, p. 105), para que haja uma Orientação Responsiva e Inclusiva “as avaliações devem ser responsivas aos *stakeholders* e as suas comunidades”. Assim, as avaliações responsivas devem ficar atentas às necessidades e características dos *stakeholders* e de seus contextos e uma avaliação inclusiva complementa a responsividade e reconhece todos os *stakeholders*, não somente parte deles, procurando conhecer, também, seus contextos, necessidades e experiências. Nesse sentido, nesse sentido, conforme o trecho extraído do Relatório, a CGU destaca o papel do Conselho Gestor Interministerial

do Programa Bolsa Família (CGIPBF) composto pelos titulares dos Ministérios da Cidadania (MC), da Educação (MEC), da Saúde (MS), do MEC, do Ministério da Economia (ME), da Casa Civil da Presidência da República e da Caixa Econômica Federal, cujas deliberações teriam caráter prioritário quando da avaliação do PBF e das políticas públicas complementares. Porém, deve-se ressaltar que para cumprir sua finalidade faz-se necessário a regulamentação de suas competências e a forma de funcionamento.

Embora a finalidade do Conselho esteja prevista na lei que estabelece o PBF, encontra-se inoperante desde a sua criação por falta de regulamentação sobre suas competências, composição e funcionamento. Cabe ressaltar que a Senarc/MC, que cumpre o papel da extinta Secretaria-Executiva do CGIPBF¹⁰, tem atuação limitada com os demais órgãos específicos singulares da saúde (SAPS/MS) e educação (Semesp/MEC), por falta de uma coordenação em nível hierárquico superior (Controladoria Geral da União, 2021, p. 12).

Para Yarbrough, Shulha, Hopson e Caruthers (2011, p. 105), os Acordos Formais da avaliação “devem ser negociados de modo que as obrigações fiquem explícitas, e sejam considerados os contextos culturais, as necessidades e as expectativas dos clientes e de outros *stakeholders*”. Na avaliação, os acordos formais são contratos vinculantes entre avaliadores e as partes por/para quem conduz o processo. Na análise da equipe de auditoria da CGU, a Secretaria Nacional de Renda de Cidadania- Senarc/MC confirma que incorporou as atribuições da extinta Secretaria-Executiva do CGPBF. “Por outro lado, não foi identificada, no nível estratégico, estrutura intersetorial instituída, formal ou informalmente, na alta administração e capaz de assumir as funções deliberativas do Conselho” (Controladoria Geral da União, 2021, p. 45).

Dessa forma, sendo o CGPBF o órgão colegiado que faz parte da estrutura de governança do PBF, observa-se que a Lei nº 10.836, de 09.01.2004, que o criou, não foi revogada ou alterada de forma que seja extinta por vontade

alheia ao legislador, portanto, cabe acompanhamento da recomendação pela CGU.

Segundo Yarbrough, Shulha, Hopson e Caruthers (2011, p. 105), para que haja Respeito e Direitos Humanos “as avaliações devem ser planejadas e conduzidas com a preocupação de proteger os direitos humanos e a legalidade, e manter a dignidade de participantes e outros *stakeholders*.” Este padrão eleva os avaliadores a um grau de compromisso com o respeito e com os direitos humanos, superior ao que é exigido pelos legisladores. Os avaliadores da CGU destacam que o maior detalhamento dos motivos apresentados pelas famílias referentes ao descumprimento das condicionalidades auxiliaria no aprimoramento das ações governamentais direcionadas às famílias mais vulneráveis. De acordo com o Órgão:

A adequada especificação dos motivos facilita a atuação dos gestores públicos, tanto pela identificação dessas famílias e encaminhamento às áreas competentes, buscando livrá-las da situação de vulnerabilidade, como, num sentido mais amplo, pela identificação dos principais riscos que acometem as famílias descumpridoras, para que as políticas sejam direcionadas e focalizadas nas maiores necessidades dos beneficiários, um dos objetivos da gestão de condicionalidades (Controladoria Geral da União, 2021, p. 19).

Conforme Yarbrough, Shulha, Hopson e Caruthers (2011), neste padrão Clareza e Equidade, as avaliações devem ser inteligíveis e justas, ao tratar dos propósitos e necessidades dos *stakeholders*. Os avaliadores demonstraram um monitoramento cuidadoso e uma comunicação eficiente ao avaliar uma política pública de grande importância como o Bolsa família, destacando os pontos a serem desenvolvidos nas situações apresentadas de maneira clara e objetiva.

Já no Padrão Transparência e Abertura, de acordo com Yarbrough, Shulha, Hopson e Caruthers (2011, p. 105), “as avaliações devem fornecer a todos os *stakeholders* descrições completas das constatações, limitações e conclusões, a não ser que isso

represente violação das leis ou das convenções sociais". As análises foram focadas na avaliação da estrutura normativa que dá suporte ao processo, da qualidade das informações disponíveis no Sistema de Condicionalidades (Sicon) e dos casos concretos existentes de utilização dessas informações para aprimoramento de políticas públicas. E o relatório de avaliação apresentado pela CGU, apresenta os dados coletados por meio de tabelas, gráficos e achados com manifestação de cada unidade examinada, que serviram de subsídios para a sua conclusão e recomendações.

Para Yarbrough, Shulha, Hopson e Caruthers (2011, p. 105), no padrão Conflitos de Interesses "as avaliações devem identificar franca e honestamente e trabalhar conflitos de interesses, sejam reais ou disfarçados, que possam comprometer a avaliação." Os avaliadores registraram e apontaram o conflito existente, visto que:

Os Ministérios da Cidadania, Educação e Saúde vêm desenvolvendo ações com base nas informações sobre o descumprimento de condicionalidades. Essas ações possuem diferentes graus de complexidade e alcance, mas trazem em comum à ausência de formalização das iniciativas e a precariedade de sua continuidade em função da falta de apoio da alta administração, que se encontra afastada dos processos decisórios (Controladoria Geral da União, 2021, p. 40).

Segundo Yarbrough, Shulha, Hopson e Caruthers (2011, p. 105), para que haja Responsabilidade Fiscal "as avaliações devem prestar contas de todos os recursos despendidos e cumprir processos e procedimentos fiscais determinados pela lei". A responsabilidade fiscal na avaliação se refere aos deveres e obrigações para um bom gerenciamento de todos os recursos da avaliação, de modo que sejam efetiva e adequadamente utilizados, sem desperdício.

Os avaliadores registraram a expansão do PBF de 2004 a 2014, e o aumento do valor médio do benefício desde abril de 2019. Isso, aliado às restrições orçamentárias, fez com que, a partir de maio,

a concessão de benefícios do Programa fosse realizada somente para famílias dos grupos prioritários. Focalizaram na redução no número de famílias beneficiárias e no valor total repassado pelo Programa. Os avaliadores não abordaram a questão fiscal (Controladoria Geral da União, 2021, p. 50).

No Quadro 4 é apresentado um resumo das classificações obtidas após avaliação de cada um dos padrões do atributo Adequação de acordo com Yarbrough, Shulha, Hopson e Caruthers (2011). De acordo com os critérios de valoração dos atributos (Quadro 1), o Relatório de avaliação sobre as informações de descumprimento de condicionalidades do PBF no que se refere a esse atributo atendeu totalmente aos padrões estabelecidos.

Pode-se afirmar que O Relatório apresentou uma avaliação responsiva observando que neste grande programa, como um quebra-cabeças, faltava uma peça principal, o Conselho Gestor Interministerial do Programa Bolsa Família (CGIPBF), que inoperante, por falta de regulamentação, limitava, de certa forma, a atuação dos Ministérios MC, MS e ME. Como órgão legitimado pela Constituição Federal para avaliar a execução dos programas de governo, a CGU destaca o seu papel, principalmente pela ausência do Conselho e por não haver estrutura Intersetorial instituída, formal ou informalmente, na alta administração, capaz de assumir as funções deliberativas do Conselho. Também foi apresentado no Relatório uma preocupação com os direitos e vulnerabilidades do público do PBF ao ressaltar a necessidade de aprofundar os motivos dos descumprimentos das condicionalidades a fim de direcionar e focalizar as políticas públicas para as maiores necessidades dos beneficiários.

Apesar de atendido a maioria dos padrões estabelecidos pelo *Joint Committee* neste atributo, cabe ressaltar que o padrão Responsabilidade Fiscal não foi atendido por não apresentar o monitoramento da utilização dada aos recursos.

Quadro 4 – Resultados da meta-avaliação em relação ao Atributo de Adequação

Padrões	Classificação do Atendimento
Orientação Responsiva e Inclusiva	A
Acordos Formais	A
Respeito e Direitos Humanos	A
Clareza e Equidade	A
Transparência e Abertura	A
Conflitos de Interesses	A
Responsabilidade Fiscal	NA

Legenda: A – atendido; PA – parcialmente atendido; NA – não atendido.

Fonte: os autores (2023).

O ATRIBUTO DE PRECISÃO

Conforme Yarbrough, Shulha, Hopson e Caruthers (2011, p. 157), nesse padrão Conclusões e Decisões Justificadas, “as conclusões e decisões da avaliação devem ser explicitamente justificadas nas culturas e nos contextos em que têm consequências”. Os avaliadores concluíram sinalizando para a estrutura de governança das decisões que precisa ser aprimorada, assim como, o aperfeiçoamento dos dados extraídos do Sicon - Sistema de Condicionalidades, o estabelecimento de reportes regulares e a atribuição de responsabilidade para a tomada de decisão. Nesta conclusão, sugerem as melhorias que precisam ser implementadas, e justificam detalhadamente as lacunas encontradas que dificultam a utilização das informações de descumprimento de condicionalidades do programa bolsa família.

De acordo com Yarbrough, Shulha, Hopson e Caruthers (2011, p. 157), para que haja Informação Válida, “a informação da avaliação deve servir aos propósitos pretendidos e sustentar interpretações

válidas". Este padrão enfoca também a abrangência e a suficiência das informações. A abrangência se refere ainda às várias fontes de informação e aos métodos para a coleta dessas informações, com base em perspectivas e opiniões diferentes, ligadas aos objetivos da avaliação. A CGU (2021, p. 7) direcionou o seu trabalho a partir da seguinte questão avaliativa "De que forma as informações de descumprimento de condicionalidades são utilizadas para subsidiar políticas e orientar ações direcionadas a superar as vulnerabilidades das famílias?"; e apontou a metodologia utilizada, como pode ser observado no trecho a seguir:

Os trabalhos de campo foram realizados no período compreendido entre os dias 22.09.2020 e 15.01.2021, por meio de testes, análises, indagações e consolidação de informações coletadas pela equipe de auditoria, em estrita observância às normas de auditoria aplicáveis ao Serviço Público Federal (Controladoria Geral da União, 2021, p. 8).

Neste aspecto, o relatório apresenta informações consistentes baseadas nos dados obtidos por meio do Sicon, no arcabouço normativo e na análise de casos concretos, e a observação das normas da auditoria facilita a interpretação das informações em contextos específicos.

Para Yarbrough, Shulha, Hopson e Caruthers (2011, p. 157), para que ocorra a Informação Fidedigna "os procedimentos da avaliação devem gerar informações suficientemente confiáveis e consistentes, de modo que sirvam aos usos pretendidos". Nesse sentido, este padrão mostra claramente como melhorar e documentar a confiabilidade, não apenas em avaliações de resultados, como em todos os procedimentos ligados à seleção, coleta e análise das informações da avaliação. Os avaliadores da CGU além de obter os dados por meio de diversas fontes de informação, elaboraram gráficos e tabelas com diferentes indicadores do programa Bolsa Família, gerando informações consistentes e fidedignas.

Segundo Yarbrough, Shulha, Hopson e Caruthers (2011, p. 157), no padrão Explicitação das Descrições do Programa e do Contexto, “as avaliações devem documentar os programas e seus contextos com apropriado detalhe e escopo para os propósitos estabelecidos”. Além da questão avaliativa, os avaliadores da CGU formularam seis subquestões que nortearam todo o trabalho avaliativo:

- “Há normatização sobre o monitoramento e a utilização de informações na tomada de decisões?”;
- “Existem responsabilidades, papéis, instâncias de avaliação, modelos de reporte e fluxos de tratamento das informações definidos?”;
- “As informações relativas ao acompanhamento das famílias que descumprem condicionalidades são apresentadas aos tomadores de decisão e avaliadas periodicamente?”;
- “As informações disponíveis possibilitam a identificação e a caracterização das famílias em situação de maior vulnerabilidade?”;
- “Existem controles voltados para a manutenção de informações atualizadas e suficientes no Sicon?”; e
- “Os Ministérios da Cidadania, da Saúde e da Educação utilizam as informações de condicionalidades para a definição de estratégias e aprimoramento de suas políticas?”

(Controladoria Geral da União, 2021, p. 7-8).

Dessa forma, os avaliadores da CGU descreveram de forma clara as especificidades, inoperâncias e melhorias do Programa Bolsa Família no que tange as informações de descumprimento das condicionalidades do referido programa.

No padrão Gerenciamento da Informação, Yarbrough, Shulha, Hopson e Caruthers (2011, p. 157) expressam que “a avaliação deve empregar métodos sistemáticos de coleta, revisão, verificação e

armazenamento das informações.” Conforme relatado nos padrões Informação Válida e Informação Fidedigna, os avaliadores da CGU obtiveram os dados por meio de diversas fontes de informação, e elaboraram gráficos e tabelas com diferentes indicadores do programa Bolsa Família, utilizando métodos quantitativos e qualitativos. Essa combinação de abordagem enriqueceu a precisão da avaliação.

Yarbrough, Shulha, Hopson e Caruthers (2011, p. 157) no padrão Análises e Delineamentos (Designs) Sólidos declaram que “a avaliação deve empregar delineamentos tecnicamente adequados e análises apropriadas aos propósitos estabelecidos”. Trata da importância das análises e dos delineamentos nas interpretações válidas. Os avaliadores da CGU ao formular as subquestões, conforme explicitado no padrão Explicitação das Descrições do Programa e do Contexto, delinearam uma estrutura de raciocínio, para desenvolver análises consistentes e conclusões que respondam à questão avaliativa “De que forma as informações de descumprimento de condicionalidades são utilizadas para subsidiar políticas e orientar ações direcionadas a superar as vulnerabilidades das famílias?” mencionadas no padrão Informação Válida.

Conforme Yarbrough, Shulha, Hopson e Caruthers (2011, p. 157), para que ocorra a Explicitação do Raciocínio da Avaliação “a fundamentação da avaliação, desde informações e análises até resultados, interpretações, conclusões e julgamentos, deve ser clara e completamente documentada”. A avaliação da CGU foi direcionada para a análise do processo de acompanhamento de condicionalidades do Programa Bolsa Família, e nas informações de como o descumprimento das condicionalidades vêm sendo utilizadas pelos Ministérios da Cidadania (MC), da Educação (MEC) e da Saúde (MS). Acompanharam e analisaram o fluxo das condicionalidades no MEC e no MS e concluíram baseados em evidências ou lacunas encontradas nas informações e ações praticadas pelos Ministérios relativas ao descumprimento das condicionalidades. O Relatório está documentado conforme explicitado nos padrões Informação Válida e Informação Fidedigna.

A Comunicação e Relatório, Yarbrough, Shulha, Hopson e Caruthers (2011, p. 157) expressa que nesse padrão “as comunicações da avaliação devem ter escopo adequado, evitando concepções errôneas, vieses, distorções e erros”. De acordo com o relatório da CGU:

Quanto à qualidade das informações disponíveis para a tomada de decisão, o acompanhamento das condicionalidades realizado no âmbito da educação e da saúde é amplo e possibilita a identificação daquelas famílias em descumprimento. No entanto, ainda é preciso maior detalhamento das informações coletadas sobre essas famílias, especificamente no que se refere às razões que as levam ao descumprimento (Controladoria Geral da União, 2021, p. 40).

Dessa forma, os avaliadores destacam que as informações coletadas pelos Ministérios sobre o descumprimento das condicionalidades precisam ser mais bem detalhadas. Apresentam uma avaliação imparcial e equilibrada, quando enfatizam os aspectos positivos, visto que o acompanhamento das condicionalidades do PBF vem produzindo regularmente dados sobre o número de famílias acompanhadas pelas pastas de Educação, Saúde e Cidadania, mas alertam da carência de uma qualificação das informações – e das ações – relativas aos motivos que levam as unidades familiares a descumprirem essas obrigações.

Em consonância com os critérios de valoração dos atributos (quadros 1 e 5), o relatório da CGU atendeu totalmente aos padrões estabelecidos pelo *Joint Committee* no atributo Precisão. Foram apresentadas informações consistentes e fidedignas obtidas de diversas fontes, descrevendo de forma clara as especificidades, inoperâncias e melhorias do Programa Bolsa Família. Delinearam uma estrutura de raciocínio, analisando o fluxo das condicionalidades nos Ministérios de Educação e Saúde e concluíram baseados em evidências ou lacunas encontradas nas informações e ações praticadas por esses Ministérios.

Quadro 5 – Resultados da meta-avaliação em relação ao Atributo de Precisão

Padrões	Classificação do Atendimento
Conclusões e Decisões Justificadas	A
Informação Válida	A
Informação Fidedigna	A
Explicitação das Descrições do Programa e do Contexto	A
Gerenciamento da Informação	A
Análises e Delineamentos (designs) Sólidos	A
Explicitação do Raciocínio da Avaliação	A
Comunicação e Relatório	A

Legenda: A – atendido; PA – parcialmente atendido; NA – não atendido.

Fonte: Os autores (2023).

O ATRIBUTO DE RESPONSABILIZAÇÃO

Para Yarbrough, Shulha, Hopson e Caruthers (2011, p. 225), no padrão Documentação da Avaliação “as avaliações devem documentar por completo seus propósitos negociados, designs implementados, procedimentos, dados e resultados”. O relatório de avaliação está bem documentado e atento aos detalhes, explicitando por meio de gráficos, tabelas e achados, as respostas à questão avaliativa e as subquestões que nortearam a avaliação.

No padrão Meta-avaliação interna, segundo Yarbrough, Shulha, Hopson e Caruthers (2011, p. 225), os “avaliadores devem empregar estes e outros padrões aplicáveis para examinar a responsabilização dos delineamentos da avaliação, os procedimentos adotados, as informações coletadas e os resultados”. A CGU direcionou o seu trabalho a partir da seguinte questão avaliativa “De que

forma as informações de descumprimento de condicionalidades são utilizadas para subsidiar políticas e orientar ações direcionadas a superar as vulnerabilidades das famílias?” e de outras subquestões delineando uma estrutura de raciocínio, para desenvolver análises e conclusões consistentes. Obtiveram dados por meio de diversas fontes de informação, e elaboraram gráficos e tabelas com diferentes indicadores do programa Bolsa Família, utilizando métodos quantitativos e qualitativos, gerando informações consistentes e fidedignas, mas sem abordar os recursos despendidos nessa avaliação.

Conforme Yarbrough, Shulha, Hopson e Caruthers (2011, p. 225), no padrão Meta-avaliação Externa, “patrocinadores, clientes, avaliadores e outros *stakeholders* devem incentivar a realização de meta-avaliações externas, utilizando estes padrões e outros que forem aplicáveis”. Os avaliadores da CGU observaram que no âmbito da saúde, educação e assistência social, as normas de gestão das condicionalidades do PBF definem as responsabilidades e competências sobre o processo de acompanhamento nos níveis federal, estadual, distrital e municipal. Mas registraram que não estão formalmente institucionalizados os fluxos, instâncias decisórias, meios e a periodicidade com que as informações sobre o descumprimento das condicionalidades devam ser avaliadas. Essa ausência normativa impossibilita que tais informações sejam tratadas e encaminhadas, continuamente, no âmbito decisório da alta administração (Controladoria-Geral da União, 2021). Apesar das excelentes observações, não foram mencionados os valores despendidos na avaliação.

Em conformidade com os critérios de valoração dos atributos (quadro 1), e em consonância com os padrões estabelecidos pelo *Joint Committee*, o atributo Responsabilização não foi atendido. Este atributo se refere à utilização responsável de todos os recursos da (ou relacionados à) avaliação, com a finalidade de produzir valor. Segundo Yarbrough, Shulha, Hopson e Caruthers (2011, p. 226), “as avaliações devem prestar contas de todos os recursos despendidos e cumprir processos e procedimentos fiscais determinados pela lei”.

Quadro 6 – Resultados da meta-avaliação em relação
ao Atributo de Responsabilização

Padrões	Classificação do Atendimento
Documentação da Avaliação	A
Meta-Avaliação Interna	PA
Meta-Avaliação Externa	PA

Legenda: A – atendido; PA – parcialmente atendido; NA – não atendido.

Fonte: Os autores (2023).

No sentido de sintetizar os resultados alcançados por este estudo meta-avaliativo, são apresentados no quadro 7 os resultados finais da meta-avaliação realizada.

Quadro 7 - Resultado da Meta-Avaliação

Atributos do Joint Committee	Classificação do Atendimento
Utilidade	PA
Exequibilidade	PA
Adequação	TA
Precisão	TA
Responsabilização	NA

Legenda: TA = Totalmente Atendido; PA = Parcialmente Atendido; NA = Não Atendido.

Fonte: Os autores (2023).

Desse modo, como resposta à questão meta-avaliativa, pautado nos cinco atributos de qualidade da avaliação, considerando seus 30 padrões de qualidade, o estudo aferiu que os atributos de Utilidade e de Exequibilidade foram atendidos parcialmente; os atributos de Adequação e Precisão foram atendidos plenamente; e, o Atributo de Responsabilização não foi atendido.

CONSIDERAÇÕES FINAIS

Os resultados aqui sintetizados, relacionados à meta-avaliação do conteúdo do Relatório de Auditoria nº 830251 - Ministério da Cidadania - Avaliação das ações adotadas pelos Ministérios da Cidadania, Saúde e Educação sobre as informações de descumprimento das condicionalidades do Programa Bolsa Família, elaborado pela CGU, se referem:

- a. ao processo de avaliação das informações sobre descumprimento das condicionalidades julgando-o como não formalmente institucionalizado o que fragiliza o aprimoramento de políticas públicas;
- b. a inoperância do conselho gestor interministerial do programa;
- c. e, ao acompanhamento das condicionalidades de educação e saúde, indicando que necessita de aprimoramento.

Concentrando-se no descumprimento das condicionalidades do referido Programa com a intenção de oferecer melhorias direcionadas a superar as vulnerabilidades das famílias beneficiárias, o presente trabalho estabeleceu como referência teórica os 30 padrões do *Joint Committee*, cujas normas orientam avaliadores sobre as características que devem possuir as boas avaliações.

Nesse sentido, embora mais de 95% beneficiários cumpram as condicionalidades, o motivo de descumprimento mais utilizado não permite a caracterização das vulnerabilidades. Cabe se observar que o conteúdo do Relatório meta-avaliado indica que os recursos apresentados pelas famílias beneficiadas podem auxiliar na identificação das vulnerabilidades que as levam a esses descumprimentos.

Nesse aspecto, cabe citar que o Sicon permite a identificação das famílias mais vulneráveis, mas elas não têm sido priorizadas no

acompanhamento familiar pelos serviços socioassistenciais. É relevante se observar também que os dados disponíveis no Sicon não permitem identificar com clareza e exatidão as situações vivenciadas pelas famílias, havendo um distanciamento da alta administração quanto às decisões voltadas para o aprimoramento e a implantação de políticas públicas. Finalmente, ressalta-se que a CGU alertou para o risco de descontinuidade de boas práticas na utilização das informações sobre o descumprimento de condicionalidades.

Aos resultados deste estudo são agregadas as seguintes recomendações:

- a. Em relação ao Atributo de Utilidade: garantir a participação dos interessados no processo (*stakeholder*), avaliadores e avaliados, desde a sua concepção até seus resultados e sua aplicação; assegurar que os resultados da avaliação sejam utilizados na prática; e, monitorar o aproveitamento dos resultados reconhecendo avanços e corrigindo as falhas sem depreciações desnecessárias.
- b. Em relação ao Atributo de Exequibilidade: é importante explicitar a metodologia utilizada e os respectivos instrumentos para busca de informações.
- c. Em relação ao Atributo de Adequação: Incluir nos relatórios de avaliação os procedimentos de aplicação dos recursos.
- d. Em relação ao Atributo de Precisão, recomenda-se sempre discutir com os *stakeholders* as dificuldades e soluções referentes aos procedimentos na busca de qualidade.
- e. Em relação ao Atributo de responsabilização: desenvolver a prática da meta-avaliação interna, com a participação pessoal especializado da própria CGU, de forma a conferir a qualidade do processo avaliativo realizado; e, externa, com meta-avaliações conduzidas por avaliadores externos à

CGU, de forma a contribuir para a melhorias dos processos avaliativos do órgão, considerando um olhar independente e visando conferir-lhe maior credibilidade.

Em síntese, a avaliação aqui contemplada, é digna de apreciação tanto pela sua relevância social e cultural como pelo mérito de sua ação profissionalmente conduzida. Mais ainda, vale destacar a importância de sua contribuição a um programa social pertencente à uma política de estado tão crucial para a população brasileira de baixa renda.

MÉRITO DO ESTUDO

Ovidio Orlando Filho

Os estudos meta-avaliativos somativos analisam a qualidade de avaliações já realizadas, verificando sua validade, confiabilidade e impacto. Ou seja, são avaliações das avaliações, feitas com critérios rigorosos para identificar pontos fortes e limitações.

O caminho metodológico seguido foi centrado nas boas práticas avaliativas disseminadas pelo *Joint Committee on Standards for Educational Evaluation*, uma organização dos Estados Unidos que estabelece critérios para avaliar programas educacionais, instituições e políticas públicas. Esses padrões são amplamente utilizados para garantir avaliações justas, rigorosas e éticas.

O trabalho apresentou uma definição clara do tema abordado, apontando com clareza o que é meta-avaliação, explicando a importância desse tipo de estudo para melhorar metodologias avaliativas e aperfeiçoar políticas e práticas institucionais. O texto foi elaborado de modo organizado e com uma estrutura lógica, apresentando uma boa problematização onde foram considerados a conceituação de meta-avaliação, sua relevância e os objetivos do estudo.

Seu desenvolvimento foi realizado com profundidade de conceitos relevantes ao tema estudado e o caso prático, exposto pela avaliação meta-avaliada foi apresentada com referências sólidas e boa fundamentação teórica.

REFERÊNCIAS

AMERICAN NATIONAL STANDARDS INSTITUTE. **Joint Committee on Standards for Educational Evaluation**. Washington: ANSI Headquarters, 2023. Disponível em: https://webstore.ansi.org/sdo/JCSEE?gclid=CjwKCAjw70IBhB8EiwAnoOEK3LV2WBZjrc_PhVSBRCmuCr53WfmaOU7J8xZtDS05vf0uLczqKHBoCJicQAvD_BwE. Acesso em: 11 jul. 2023.

ARAUJO, Jevuks Matheus; FRIO, Gustavo Saraiva; ALVES, Pedro Jorge Holanda. O efeito do Bolsa Família sobre a distorção idade-série. **Estud. Econ.**, São Paulo, v. 51 n. 2, abr./jun. p. 343-371, 2021. Disponível em: <https://www.scielo.br/j/ee/a/64SfhkWgjZbVcCkb8gq45bq/?lang=pt>. Acesso em: 3 jul. 2023.

ARAUJO, Nayanna M. SANTOS, Olívia R.; AGUIAR, Glauco. Relatório de avaliação de impactos do Programa de Desenvolvimento e Crescimento Integrado com Sustentabilidade: um estudo meta avaliativo. *In*: ELLIOT, L. G., ORLANDO FILHO, O. (org.). **Estudos meta avaliativos**: aplicações metodológicas. São Paulo: Pimenta Cultural, 2020.

BRASIL. Ministério do Desenvolvimento e Assistência Social, Família e Combate à Fome. Brasília, DF, 2023. Disponível em: <https://www.gov.br/mds/pt-br/acoes-e-programas/bolsa-familia> Acesso em: 4 jul. 2023.

CARVALHO, Maria do Carmo Brant. Avaliação participativa: uma escolha metodológica. *In*: RICO, Elizabeth Melo (org.). **Avaliação de políticas sociais**: uma questão de debate. 2. ed. São Paulo: Cortez, 1999.

CONTROLADORIA GERAL DA UNIÃO. Ministério da Cidadania. Relatório de Auditoria nº 830251. **Avaliação das ações adotadas pelos Ministérios da Cidadania, Saúde e Educação sobre as informações de descumprimento das condicionalidades do Programa Bolsa Família**. Brasília, DF, 2021. Disponível em: <https://eaud.cgu.gov.br/relatorios>. Acesso em: 4 jul. 2023.

CONTROLADORIA GERAL DA UNIÃO. **Manual da metodologia para avaliação da execução de programas de governo.** Brasília, DF: Secretaria Federal de Controle Interno, 2015. Disponível em: https://repositorio.cgu.gov.br/bitstream/1/44969/5/Manual_AEPG_2015.pdf. Acesso em: 4 jul. 2023.

CONTROLADORIA GERAL DA UNIÃO. **Institucional.** Brasília, DF, 2023. Disponível em: <https://www.gov.br/cgu/pt-br/acesso-a-informacao/institucional>. Acesso em: 4 jul. 2023.

ELLIOT, Ligia Gomes; ORLANDO FILHO, Ovídio (org.). **Meta avaliações de políticas públicas do governo:** o que revelam os relatórios de avaliação. São Paulo: Pimenta Cultural, 2019.

MARTINS, Barbara Amaral; RUCKERT, Fabiano Quadros. O Programa Bolsa Família e a condicionalidade educacional: uma análise do desempenho escolar de estudantes em situação de pobreza. **Revista Brasileira de Educação.** Mato Grosso do Sul, v. 27, e270101. 2022. Disponível em: <https://www.scielo.br/j/rbedu/a/y3VsyT9jLBTbXmghfWhyt8r/?lang=pt#>. Acesso em: 3 jul. 2023.

OLIVEIRA, Lilian Ribeiro de; PASSADOR, Claudia Souza. Ensaio teórico sobre as avaliações de políticas públicas. **Cad. EBAPE.BR.** São Paulo, v. 17, n. 2, abr./jun., 2019. Disponível em: <https://www.scielo.br/j/cebape/a/svZxsKnLTZ4RWnLGG93bYfH/?format=pdf&lang=pt>. Acesso em: 10 jul. 2023.

SCRIVEN, Michael. **Avaliação:** um guia de conceitos. Rio de Janeiro: Paz e Terra, 2018.

WORTHEN, Blaine R.; SANDERS, James R.; FITZPATRICK, Jody L. **Avaliação de programas:** concepções e práticas. São Paulo: Editora Gente, 2004.

YARBROUGH, Donald B.; SHULLA, Lyn M.; HOPSON, Rodney K.; CARUTHERS, Flora A. **The program evaluation standards:** a guide for evaluators and evaluation users. 3. ed. Thousands Oaks, Califórnia: Sage, 2011.

3

*Karla Verônica Mendes da Silva
Michele Terra Duarte
Elaine Rodrigues Perdigão*

META-AVALIAÇÃO DO RELATÓRIO DE RETORNO SOCIAL DO PROGRAMA *A UNIÃO FAZ A VIDA:* RESULTADOS E RECOMENDAÇÕES

INTRODUÇÃO

Um dos maiores desafios na realização de avaliações de impacto em programas e projetos sociais no Brasil é a alta complexidade e os custos envolvidos (Fabiani; Rebehy; Camelo; Vicente; Mosaner, 2018). Nesse contexto, destaca-se a relevância de investir em avaliações rigorosas, capazes de aprimorar as intervenções sociais e promover a equidade e a sustentabilidade nas sociedades atuais.

Em diversos países, as avaliações têm se tornado essenciais para a tomada de decisões. No Brasil, ainda não são amplamente adotadas, o que torna necessário promover sua disseminação e discutir seus diferentes tipos. A avaliação de impacto, por exemplo, busca estabelecer uma relação de causa e efeito entre uma intervenção e seus resultados, com foco na atribuição desses resultados.

Segundo Gertler, Martínez, Premand, Rawlings e Vermeersch, (2018, p. 45), “a avaliação de impacto fornece informações sobre se o programa causou as mudanças desejadas nos resultados”. No contexto das intervenções sociais, o impacto refere-se à mudança social gerada por um programa ou projeto, e sua mensuração ajuda a refletir sobre estratégias que impulsionam a transformação social (Instituto para o Desenvolvimento do Investimento Social, 2022c, 2021).

O Instituto para o Desenvolvimento do Investimento Social já conduziu avaliações de dezenas de projetos para organizações como Amigos do Bem, Gerando Falcões, Parceiros da Educação, Petrobras, Sesc e Vale. No contexto da avaliação de impacto social do Programa *A União Faz a Vida*, objeto deste estudo meta-avaliativo, o Instituto para o Desenvolvimento do Investimento Social desenvolveu a Teoria da Mudança (Gertler; Martínez; Premand; Rawlings; Vermeersch, 2018) do programa com base na interação com a equipe da Fundação Sicredi, Centrais e Cooperativas Sicredi, bem como na coleta de dados qualitativos e quantitativos junto a alunos, familiares de alunos e professores.

Independentemente das técnicas e abordagens utilizadas, é indiscutível a imprescindibilidade de se realizar avaliações em programas e projetos sociais, que visam gerar benefícios para a sociedade e promover uma mudança social para as populações mais vulneráveis, a fim de mensurar se os objetivos traçados e os resultados almejados estão realmente sendo atingidos. Paralelamente ao crescimento da importância das avaliações no Brasil, cresce a necessidade de que elas sejam conduzidas de forma ética e que tenham qualidade técnica. De acordo com Silva, Joppert e Gasparini (2020), a qualidade das avaliações, devido aos seus significados, tamanhos potenciais e limites, tem sido debatida mundialmente. Uma das formas de verificar a qualidade das avaliações é realizando meta-avaliações.

META-AVALIAÇÃO: ATESTANDO A QUALIDADE DE AVALIAÇÕES

“Meta-avaliação é a avaliação de avaliações”, assim definiu Scriven (2018, p. 361). Worthen, Sanders e Fitzpatrick (2004) apontam que a meta-avaliação tem como finalidade ajudar a avaliação a realizar seu potencial. Elliot (2011, p. 960) justifica a necessidade de meta-avaliar, como uma busca pela qualificação das avaliações: “é a forma de atestar-lhes a qualidade ou de evidenciar pontos que ainda precisam ser burilados”.

Chianca (2001, p. 83) assevera que há quatro momentos cruciais em que uma meta-avaliação deve ser realizada para que mudanças qualitativas sejam introduzidas na avaliação: (i) após a finalização do plano/desenho da avaliação; (ii) após o desenvolvimento dos instrumentos e das estratégias para a coleta de dados; após as primeiras análises dos dados; e, (iii) o final do processo avaliativo.

Para julgar o valor e o mérito de uma avaliação é preciso que sejam estabelecidos critérios, diretrizes e padrões de julgamento para aferir sua qualidade e evidenciar o que precisa ser melhorado. Por isso, padrões e diretrizes como as Diretrizes para a prática de avaliação no Brasil (Silva; Joppert; Gasparini, 2020), norteadoras do presente estudo, são essenciais para a condução de uma boa meta-avaliação.

OBJETIVO DO ESTUDO

O presente estudo teve como objetivo avaliar o Relatório da Avaliação do Retorno Social do Investimento do Programa A União Faz a Vida da Fundação Sicredi, estabelecendo como base metodológica as Diretrizes para a prática de avaliação no Brasil. Trata-se de uma meta-avaliação somativa. De acordo com Stufflebeam (2011), as meta-avaliações somativas, ou seja, aquelas conduzidas após uma avaliação, auxiliam os interessados a perceber os pontos fortes e limitações de uma avaliação, além de julgar seu mérito e valor.

De fato, a meta-avaliação somativa possibilita avaliar não só a qualidade da avaliação, na busca em aumentar a sua credibilidade, mas também permite aferir a competência do avaliador.

OBJETO DO ESTUDO

A Fundação de Desenvolvimento Educacional e Cultural do Sistema de Crédito Cooperativo (Fundação Sicredi) é uma Organização da Sociedade Civil de Interesse Público (OSCIP) responsável por iniciativas educacionais, culturais, sociais e sustentáveis alinhadas aos valores do Sicredi. Seu foco está na preservação

da essência do cooperativismo, promovendo ações voltadas à educação e cultura.

Entre suas atividades, a Fundação Sicredi apoia diversos programas por meio de cooperativas locais e centrais regionais, sendo o Programa *A União Faz a Vida* sua principal iniciativa de responsabilidade social. Criado em 1995 e implementado nacionalmente, o programa baseia-se nos princípios de cooperação e cidadania para fomentar o desenvolvimento social das regiões onde atua e conscientizar a população sobre seu papel como agentes transformadores da realidade.

O principal objetivo do programa é “transformar crianças e jovens em protagonistas de seu processo de aprendizagem e de sua atuação cidadã, para que promovam impactos positivos em suas comunidades” (Instituto para o Desenvolvimento do Investimento Social, 2022c, p. 5).

A metodologia educacional implementada pelo programa é baseada no desenvolvimento de projetos junto aos alunos de Educação Infantil, Educação Fundamental e Ensino Médio, além de organizações da sociedade civil que atuam com crianças e jovens. A proposta metodológica utiliza “a pedagogia ativa e participativa, com ênfase nas interações, interesses e curiosidades dos alunos, além da contextualização das áreas de conhecimento aos elementos da realidade local de cada município e região” (Instituto para o Desenvolvimento do Investimento Social, 2022c, p. 12).

Segundo o relatório, o Programa contribui diretamente para o alcance de dois dos dezessete Objetivos para o Desenvolvimento Sustentável (ODS) promovidos pela Organização das Nações Unidas (ONU) como parte da agenda 2030: ODS 4 Educação de Qualidade e ODS 17 Parcerias e Meios de Implementação.

Considerando a importância, o tempo de existência, a capilaridade, os propósitos de longo prazo e a necessidade de estimar

o retorno social dos recursos investidos e se estes estão realmente trazendo benefícios para a sociedade e mudando positivamente realidades, evidencia-se apropriada a condução de uma avaliação.

Nessa perspectiva, a Fundação Sicredi encomendou ao Instituto para o Desenvolvimento do Investimento Social uma avaliação de impacto do Programa A União Faz a Vida. A equipe de avaliadoras do Instituto para o Desenvolvimento do Investimento Social composta pela Diretora-Presidente, pela Gerente de Financeira-Administrativa, pela Gerente de Projetos e pela Analista de Projetos do instituto elaboraram o relatório final de apresentação dos resultados da avaliação: O Relatório da Avaliação do Retorno Social do Investimento do Programa A União Faz a Vida da Fundação Sicredi, objeto deste estudo meta-avaliativo.

Segundo o relatório, a avaliação utilizou como metodologia o protocolo SROI (*Social Return On Investment*), desenvolvido pela organização *Social Value UK*, a fim de mensurar o impacto social do Programa.

O protocolo SROI apoia-se em sete princípios: (1) envolver os *stakeholders*; (2) entender o que muda; (3) valorizar as coisas que importam; (4) incluir somente o que for material (ou seja, o que é relevante); (5) não reivindicar em excesso; (6) ser transparente; e (7) verificar o resultado.

Para realizar a avaliação SROI do Programa A União Faz a Vida, a equipe de avaliadoras do Instituto para o Desenvolvimento do Investimento Social, além da construção de uma Teoria da Mudança (mapa de como a realidade pode ser alterada), seguiram os seis estágios preconizados pelo protocolo: (1) estabelecer o escopo e identificar os *stakeholders*-chave; (2) mapear resultados; (3) evidenciar resultados e atribuir-lhes um valor; (4) estabelecer o impacto; (5) calcular o SROI; e (6) relatar, usar e incorporar as conclusões.

O relatório especifica que, como delimitação do estudo avaliativo, optou-se por mensurar apenas os impactos percebidos por alunos e professores do Ensino Fundamental I, que participaram do Programa durante o ano de 2019, nas quatro Centrais nas quais a iniciativa é desenvolvida há pelo menos 1 ano: I) Central Sul Sudeste, II) Central PR/SP/RJ; III) Central Brasil Central e IV) Central Centro Norte (estas são nomenclaturas utilizadas pela organização) (Instituto para o Desenvolvimento do Investimento Social, 2022c).

No que se refere à sua estrutura, o relatório é composto pelo sumário executivo, sete capítulos e sete apêndices: o sumário executivo traz um resumo panorâmico do estudo; o primeiro capítulo introduz o Programa e sua metodologia e os objetivos da avaliação; o Capítulo 2 aborda o protocolo SROI; o Capítulo 3 estabelece o escopo do estudo e identifica os *stakeholders*; o Capítulo 4 aborda a construção e testagem da Teoria da Mudança; o Capítulo 5 expõe o engajamento dos *stakeholders* e a coleta de dados quantitativos e qualitativos; o Capítulo 6 trata da construção do modelo SROI e o Capítulo 7 apresenta os resultados da avaliação, incluindo conclusões e recomendações. Os sete apêndices apresentam os roteiros e questionário utilizados, os cálculos das *proxies* financeiras, bibliografia e glossário.

PROCEDIMENTOS METODOLÓGICOS

Para a realização da meta-avaliação do *Relatório da Avaliação do Retorno Social do Investimento do Programa A União Faz a Vida* da Fundação Sicredi, optou-se por adotar as *Diretrizes para a Prática de Avaliação no Brasil*, publicadas em 2020 pela Rede Brasileira de Monitoramento e Avaliação (RBMA) em colaboração com a Agenda de Avaliação do Grupo de Institutos, Fundações e Empresas (Gife).

As diretrizes brasileiras foram formuladas com base no documento produzido pela ReLAC (Rede de Monitoramento, Avaliação e Sistematização da América Latina e do Caribe), intitulado *Diretrizes para Avaliação para a América Latina e o Caribe*. A RBMA identificou a necessidade de criar diretrizes adaptadas ao contexto brasileiro, com o objetivo de refletir a prática e a cultura de avaliação no Brasil. Essas diretrizes servem como um instrumento para promover o diálogo e a reflexão entre os diversos atores dos setores público e privado, que podem se beneficiar de seu uso (Silva; Joppert; Gasparini, 2020).

Silva, Joppert e Gasparini (2020, p. 7) afirmam que “a RBMA compreende que avaliação é o exercício sistemático de compreender e valorar um objeto qualquer, política, programa, projeto etc., segundo critérios explícitos e consistentes para um dado conjunto de atores”. Os autores também destacam que as Diretrizes podem fundamentar processos de análise crítica das avaliações realizadas, ou seja, as meta-avaliações, oferecendo conteúdo para avaliar o sentido, a qualidade e a relevância dessas avaliações. Além disso, as diretrizes podem contribuir para qualificar avaliações de impacto social, especialmente aquelas voltadas para temas de responsabilidade social, como a avaliação meta-avaliada neste estudo.

Apesar das limitações deste estudo, centradas na utilização do conteúdo do relatório disponibilizado, optou-se por avaliar todas as 26 diretrizes propostas pela RBMA, com o intuito de contribuir para sua disseminação no contexto acadêmico. As diretrizes estão distribuídas em quatro dimensões mutuamente dependentes, que tratam das características desejáveis das avaliações, conforme o quadro a seguir:

Quadro 1 – Dimensões e diretrizes a serem meta-avaliadas

Dimensões	Diretrizes
Aprendizagem e Utilização (composta por sete diretrizes)	1. Níveis de participação adequados aos interessados
	2. Propósitos acordados mutuamente
	3. Perguntas relevantes
	4. Resultados úteis
	5. Conclusões fidedignas
	6. Recomendações contextualizadas
	7. Comunicação cuidadosa
Direitos e Integridade	1. Respeito aos direitos das pessoas
	2. Diversidade e equidade
	3. Identidades culturais
	4. Reciprocidade nas trocas
	5. Explicitação de interesses
	6. Transparência
	7. Legalidade
Contextualização e Valoração	1. Avaliações contextualizadas
	2. Descrição do objeto
	3. Critérios válidos
	4. Critérios explícitos

Dimensões	Diretrizes
Método e Viabilidade	1. Viabilidade
	2. Tempo oportuno
	3. Metodologia consistente
	4. Técnicas e procedimentos válidos
	5. Percurso metodológico explícito
	6. Autonomia e independência
	7. Gestão responsável
	8. Meta-avaliação

Fonte: Os autores (2024).

O julgamento do atendimento da avaliação às diretrizes brasileiras foi realizado com base na análise subjetiva das autoras a partir da leitura apurada do relatório meta-avaliado. De acordo com Araújo, Santos e Aguiar (2020, p. 152), “todo juízo de valor formulado na meta-avaliação possui caráter qualitativo e subjetivo, tendo sido estabelecido de acordo com a interpretação da realidade, baseada no olhar crítico de seus autores”.

Para julgar os resultados das diretrizes avaliadas foi estabelecido um critério composto por uma escala de cinco níveis: Atendida (A); Atendida Parcialmente (AP); Não Atendida (NA); e Não se Aplica (NSA). Cada dimensão meta-avaliada será classificada de acordo com o seguinte quadro:

Quadro 2 – Classificação da meta-avaliação por dimensão

Dimensões	Número de Atendimentos de diretrizes por Dimensão	Classificação final da Dimensão
Aprendizagem e Utilização Direitos e Integridade	7	Excelente
	6 a 5	Boa
	4	Aceitável
	Menor que 4	Deficiente
Contextualização e valoração	4	Excelente
	3	Boa
	2	Aceitável
	Menor que 2	Deficiente
Método e Viabilidade	8	Excelente
	7 a 6	Boa
	5 a 4	Aceitável
	Menor que 4	Deficiente

Fonte: Os autores (2024) adaptado de Araújo, Santos e Aguiar (2020).

No que se refere à questão avaliativa, este estudo pretende responder: Em que medida o Relatório da Avaliação do Retorno Social do Investimento do Programa A União Faz a Vida, da Fundação Sicredi, atende às Diretrizes para a prática de avaliação no Brasil?

Observa-se que uma das etapas de grande importância numa avaliação é a definição das perguntas avaliativas que irão nortear o estudo, pois uma das finalidades das avaliações é justamente responder perguntas. Elas servem para direcionar e dar alicerce à avaliação. “Sem elas, a avaliação não teria foco e o avaliador teria considerável dificuldade para explicar o que vai ser examinado, como e por quê” (Worthen; Sanders; Fitzpatrick, 2004, p. 341).

ANÁLISE E RESULTADOS DA META-AVALIAÇÃO

Nesta seção são apresentados os resultados e a análise do Relatório da Avaliação do Retorno Social do Investimento do Programa A União Faz a Vida da Fundação Sicredi à luz das 26 Diretrizes para a prática de avaliação no Brasil. Cada diretriz, que orientou a meta-avaliação, será definida e avaliada individualmente. Ao final, os quadros três e quatro mostram os resultados gerais referentes às dimensões e diretrizes meta-avaliadas.

DIMENSÃO 1 – APRENDIZAGEM E UTILIZAÇÃO

A primeira dimensão enfatiza como a apreensão do conhecimento nas avaliações pode contribuir para sua utilização. De acordo com Silva, Joppert e Gasparini (2020), essa dimensão é desdobrada em sete diretrizes:

1. Nível de participação adequado aos interessados

Essa diretriz considera a necessidade de participação, envolvimento e engajamento dos interessados nas avaliações, pois “Os processos de concepção e realização das avaliações devem incluir as perspectivas dos interessados e acolher suas demandas, perguntas e preocupações” (Silva; Joppert; Gasparini, 2020, p. 14).

Um dos princípios do protocolo SROI, aplicado à avaliação, é o envolvimento dos *stakeholders*. Os capítulos 3 (p. 17 a 20) e 5 (p. 31 a 35) do Relatório tratam detalhadamente da identificação e do engajamento dos *stakeholders*, que puderam compartilhar suas percepções acerca de impactos concretos que decorreram do seu envolvimento com a iniciativa.

Para identificá-los foram realizadas entrevistas exploratórias e análise da documentação do programa. Foram entrevistadas 13 pessoas da equipe da Fundação Sicredi, das Cooperativas Sicredi e dos Municípios Parceiros, abrangendo indivíduos com diferentes visões do Programa. A partir das entrevistas e documentação analisada foram identificados 10 grupos de *stakeholders* influenciados pelo Programa, porém na análise da intervenção somente dois grupos foram envolvidos: alunos e professores, ou seja, as partes interessadas mais impactadas, beneficiárias diretas da iniciativa.

O processo de coleta de dados dos grupos priorizados foi feito através da realização de grupos focais (dados qualitativos) e aplicação de questionários (dados quantitativos). Os 16 grupos focais realizados no mês de outubro de 2019, em oito municípios das quatro centrais participantes do programa foram compostos por pais, alunos e professores.

Os relatos dos grupos focais contribuíram para validação da lista dos *stakeholders* diretamente impactados pelo Programa e identificação dos professores como público apto a responder os questionários. A princípio seriam realizadas uma coleta piloto e uma coleta oficial de dados junto aos alunos por meio da aplicação de questionários, porém devido à pandemia da COVID-19, tornou-se inviável a coleta de questionários com os alunos e o estudo teve prosseguimento com os resultados obtidos na aplicação dos questionários aos professores.

Devido ao detalhamento apresentado no relatório acerca do nível de envolvimento dos *stakeholders* no processo avaliativo, julgou-se a diretriz como “Atendida”.

2. Propósitos acordados mutuamente

O caráter multilateral necessário à definição dos objetivos de avaliação é enfatizado nessa dimensão.

Os principais objetivos da avaliação são explicitados no sumário executivo e na introdução do Relatório, quais sejam:

- a. Compreender o impacto do Programa A União Faz a Vida, por meio de um estudo avaliativo que demonstre a efetividade do investimento do Sicredi, Parceiros e Apoiadores;
- b. Apoiar o planejamento estratégico e o processo de tomada de decisão da Fundação Sicredi, quanto ao Programa A União Faz a Vida;
- c. Auxiliar a Fundação Sicredi a identificar potenciais pontos de aprimoramento e otimização do impacto gerado pelo Programa A União Faz a Vida.

Para atingir os objetivos da avaliação a equipe de avaliação do Instituto para o Desenvolvimento do Investimento Social aplicou o protocolo SROI e construiu a Teoria da Mudança a partir do envolvimento com os *stakeholders* e os impactos nela descritos foram validados, detalhados e mensurados junto aos beneficiários do Programa.

Outros dois aspectos do escopo do estudo que foram definidos a partir da interação com os *stakeholders* foram o recorte geográfico [Centrais que, em 2019, possuíam, no mínimo, um ano de participação no Programa, sendo estas as centrais: Sul Sudeste; PR/SP/RJ; Brasil Central (BRC) e Centro Norte (nomenclaturas utilizadas pela Fundação Sicredi)] e o recorte temporal (período de 2019), conforme explicitado no item 3.1 do capítulo 3, p. 18.

Logo, considerou-se que a segunda diretriz meta-avaliada foi “Atendida”, devido à descrição detalhada dos objetivos acordados.

3. Perguntas relevantes

As perguntas avaliativas concorrem para que os objetivos traçados sejam atingidos. A principal questão avaliativa foi citada no

relatório de forma indireta em sua página 13: “Os resultados da avaliação SROI retratam em que medida as intervenções são eficientes e de que forma os resultados são percebidos pelos grupos de interesse (também chamados de *stakeholders* ou públicos-alvo)”.

Observou-se, ainda, que o título do capítulo 4 traz uma pergunta relevante - Como o Programa A União Faz a Vida gera mudanças? – a qual é respondida ao longo do mesmo. E o item 6.2 do capítulo 6 buscou responder a seguinte pergunta: Incidência dos resultados: o que mudou depois do Programa A União Faz a Vida?

Tendo em vista que a pergunta avaliativa norteadora do relatório não foi destacada prévia e explicitamente, conclui-se que a diretriz Perguntas Relevantes foi “Atendida Parcialmente”.

4. Resultados úteis

A utilidade é o principal atributo que deve estar presente em uma avaliação. A diretriz, ora analisada, enfoca os usos dos resultados das avaliações.

O relatório, em seu Sumário Executivo e no capítulo 7 (p. 63 a 75), apresenta minuciosamente os resultados da avaliação SROI do Programa A União Faz a Vida, aderentes aos principais objetivos traçados pelo estudo, com informações úteis que poderão contribuir para a tomada de decisão dos gestores do Programa.

Os resultados obtidos evidenciaram que o Programa traz benefícios relevantes para seus participantes e públicos envolvidos, conforme pode ser exemplificado no trecho a seguir:

Os resultados da avaliação SROI indicam que, a cada R\$ 1 investido no Programa A União Faz a Vida, são criados R\$ 4,07 em valor social, isto é, o impacto social gerado é quatro vezes o valor investido, demonstrando a grande contribuição do Programa na vida de alunos e professores (Fabiani; Rebehy; Camelo; Vicente; Mosaner, 2020, p. 63).

Cabe mencionar que para o cálculo do SROI levou-se em conta, além dos valores investidos no Programa pelo Sicredi, os aplicados pelos Apoiaadores e Prefeituras Municipais no ano de 2019, período correspondente à avaliação. De acordo com o relatório, o valor social gerado pelo Programa é cerca de R\$ 465 milhões. 88,5% deste valor corresponde às transformações percebidas nos alunos participantes da intervenção e 11,5% corresponde às mudanças percebidas nos professores que aderiram ao Programa.

Pela demonstração rica e aprofundada dos resultados por meio de textos explicativos, tabelas e gráficos, analisando cada componente do protocolo SROI, julgou-se que a diretriz Resultados úteis foi “Atendida”.

5. Conclusões fidedignas

As conclusões geradas pelas avaliações devem ser consistentes e coerentes com os objetivos avaliativos.

O relatório dedica o item 7.4 do capítulo 7 às conclusões e recomendações resultantes do estudo avaliativo. As conclusões foram expostas de maneira clara, detalhada e seguiram uma ordenação lógica que respondeu à mensuração de cada um dos cinco eixos de mudança propostos na Teoria da Mudança para os alunos e professores. De modo geral, a avaliação comprovou que o Programa União Faz a Vida atinge seu objetivo principal de transformar crianças e jovens em protagonistas de seu processo de aprendizagem e sua atuação cidadã, para que promovam impactos positivos em suas comunidades. Outra conclusão relevante, alcançada a partir das informações coletadas durante o estudo foi:

O Programa A União Faz a Vida gera impactos sociais relevantes, não apenas para alunos e professores, mas também para as comunidades da região. Um dos fatores que contribui de forma muito significativa para o expressivo índice de retorno do Programa é sua capilaridade e

capacidade de alcance: em 2019, mais de 327 mil alunos e mais de 31 mil professores se envolveram com o Programa. Muito além do resultado positivo, o estudo evidencia o caráter transformador da iniciativa e seu potencial de intensificar ainda mais as transformações sociais que busca alcançar. (Instituto para o Desenvolvimento do Investimento Social, 2022c, p. 9).

Assim, a diretriz Conclusões Fidedignas foi julgada como “Atendida”.

6. Recomendações contextualizadas

As recomendações das avaliações devem ser fundamentadas em suas investigações, achados e análises, alinhadas às perguntas e objetivos da avaliação. Devem considerar lições aprendidas, endereçar atores e processos específicos e respeitar as condições de governabilidade e viabilidade (Silva, Joppert, Gasparini, 2020).

No relatório, as recomendações foram devidamente contextualizadas e explicitadas de forma clara e suficiente no item 7.4 do capítulo 7 (p. 72 a 74) e são, resumidamente:

- a. Priorização das variáveis de impacto;
- b. Livre Adesão ao Programa;
- c. Suporte pedagógico e financeiro;
- d. Ações de engajamento e reconhecimento dos Professores;
- e. Maior participação dos pais e familiares de alunos nas iniciativas do Programa.

Cada item recomendado foi amplamente descrito e justificado, sendo considerados como aspectos orientadores para manutenção, melhoria e ampliação do Programa. Destarte, considerando os critérios adotados no estudo, considerou-se que a diretriz em questão foi “Atendida”.

7. Comunicação cuidadosa

A comunicação das avaliações deve ser clara e pontual, utilizando diversos canais e linguagens para facilitar a compreensão pelos interessados. Deve apresentar os resultados, seus limites e evidências, contextualizando sua aplicação.

De acordo com o relatório do Programa A União Faz a Vida, uma das etapas que envolvem uma análise SROI é relatar, usar e incorporar as conclusões. Essa etapa preconiza que é crucial compartilhar os resultados da avaliação com os *stakeholders*, e reagir a eles, incorporando processos com bons resultados. O relatório final com os resultados da avaliação se encontra disponível para livre consulta no sítio eletrônico do Sicredi e do Instituto para o Desenvolvimento do Investimento Social, acessível a qualquer interessado que tenha acesso à internet. Em relação aos processos avaliativos, o texto do relatório não informa como e se foram sendo oportunamente comunicados aos públicos interessados durante o processo de desenvolvimento da avaliação e nem se foram elaborados e comunicados relatórios parciais. Assim sendo, a diretriz Comunicação Cuidadosa foi considerada “Atendida Parcialmente”.

DIMENSÃO 2 – DIREITOS E INTEGRIDADE

Em linhas gerais, essa dimensão trata do respeito aos aspectos sociais, culturais e políticos dos *stakeholders* e aos parâmetros legais e normativos a serem observados. De acordo com Silva, Joppert e Gasparini (2020), essa dimensão é desdobrada em sete diretrizes:

1. Respeito aos direitos das pessoas

Durante toda a condução do processo avaliativo, deve-se proteger os direitos dos participantes, assegurando consentimento

informado, confidencialidade das informações e respeito às diversidades culturais, sociais e institucionais, bem como ao cumprimento das leis e normas aplicáveis.

Em relação a esta diretriz, pôde-se extrair do relatório que os questionários foram aplicados aos professores, garantindo-lhes o anonimato. Na introdução do instrumento (Apêndice 4, p. 87), é informado aos respondentes sobre o sigilo do questionário e a desnecessidade de identificação pessoal. Quanto aos dados qualitativos coletados nos grupos focais, consta no roteiro (Apêndice 3, p. 78) que todas as informações coletadas seriam utilizadas de forma anônima.

Não foram identificados registros a respeito de questões culturais, políticas, etc., relacionadas aos contextos locais dos municípios onde foram coletados os dados qualitativos. Dessa forma, julgou-se que a diretriz foi “Atendida Parcialmente”.

2. Diversidade e equidade

As avaliações devem garantir igualdade de tratamento aos envolvidos, considerando suas diversidades e necessidades, e adotar critérios que exponham desigualdades, buscando reduzi-las e enfrentando práticas discriminatórias, como racismo e sexismo.

A avaliação envolveu diversos atores sociais: os gestores do Programa, os apoiadores, os municípios parceiros, alunos, familiares de alunos e professores, sendo que a análise priorizou os beneficiários diretos, visto que seu objetivo principal era medir o impacto social do Programa. No relatório de avaliação não são apontadas informações que pudessem levar a inferir que houve uma preocupação e adequação no manejo de questões culturais, étnico-raciais, socioeconômicas e outras relativas aos grupos envolvidos na avaliação.

Pela ausência de registros no relatório meta-avaliado que pudessem assegurar que houve um tratamento equânime aos *stakeholders*, considerou-se que a diretriz Diversidade e equidade é qualificada como “Não Atendida”.

3. Identidades culturais

Os traços culturais dos grupos envolvidos e afetados pelas avaliações devem ser considerados e respeitados. Depreende-se que uma avaliação que envolve um programa nacional que está presente em quatro regiões do país deveria explicitar em seu relatório os aspectos relevantes às identidades culturais dos envolvidos, porém não foram identificados registros que evidenciassem o cumprimento da diretriz. Assim sendo, esta diretriz foi julgada como “Não Atendida”.

4. Reciprocidade nas trocas

As avaliações devem fomentar confiança e equilíbrio entre os grupos envolvidos, evitando a supremacia de interesses de grupos privilegiados. Devem reconhecer desigualdades, como o racismo estrutural, e evidenciá-las para contribuir com sua eliminação.

Desse modo, considerando-se a ausência de informações no relatório meta-avaliado que pudessem assegurar que houve reciprocidade nas trocas entre os interessados, a diretriz é classificada como “Não Atendida”.

5. Explicitação de interesses

Deve ser dada especial atenção à existência de conflitos de interesses que possam comprometer a avaliação de modo significativo.

A avaliação foi encomendada pela fundação gestora do programa à uma equipe externa de avaliadores. O Instituto para o Desenvolvimento do Investimento Social como organização da sociedade civil de interesse público (OSCIP) que presta apoio técnico, consultoria e realização de avaliações de impacto a investidores sociais como fundações, empresas, famílias, etc., não demonstrou haver conflitos de interesses com o objeto avaliado, ou seja, o

Programa A União Faz a Vida. Ou seja, o Programa e seus processos encaminharam dentro da normalidade em relação a esses aspectos.

Ressalta-se que o próprio Instituto para o Desenvolvimento do Investimento Social, 2022a, não paginado) menciona em seu sítio eletrônico: “trabalhamos para aumentar o impacto do investimento social privado, construindo parcerias e projetos de impacto, oferecendo consultoria e produzindo e compartilhando conhecimento. Assim, melhoramos a vida das pessoas”. Além disso, na análise do relatório, pode-se observar no roteiro dos grupos focais (apêndice 3, p. 78) a preocupação da equipe de avaliadoras em afirmar sua autonomia e independência, como visto nos registros a seguir:

Importante deixar claro que não fazemos parte da Sicredi. Nosso intuito é entender as mudanças geradas pelo Programa. [...] Trata-se de um trabalho solicitado pela Sicredi para compreender se o Programa alcança o seu objetivo e de que forma pode ser melhorado.

Ante a percepção da ausência de conflitos de interesses e a clareza do papel a ser desempenhado pela equipe de avaliação, considerou-se que a diretriz Explicitação de Interesses encontra-se classificada como “Atendida”.

6. Transparência

A transparência deve permear todos os processos e procedimentos de desenvolvimento e execução das avaliações.

De acordo com o relatório, um dos princípios nos quais se fundamenta o protocolo SROI de avaliação aplicado pela equipe Instituto para o Desenvolvimento do Investimento Social é ser transparente. A análise do relatório permitiu identificar que a transparência permeou todos os procedimentos e processos conduzidos pela avaliação. Além disso, a transparência é mencionada em mais duas seções do documento: (i) nos agradecimentos à abertura, transparência, receptividade e colaboração da Fundação Sicredi ao longo do

estudo; e, (ii) na coleta de dados qualitativos, indicando a transparência dos resultados dos questionários aplicados aos docentes sem identificação pessoal, a fim de se evitar vieses e garantir a credibilidade das respostas coletadas (Instituto para o Desenvolvimento do Investimento Social, 2022c).

Em relação à transparência na comunicação dos resultados, se pode notar, já na fase inicial de realização das entrevistas exploratórias, por meio do roteiro (Apêndice 2, p. 77), a preocupação da equipe avaliadora em apresentar os objetivos para o uso e divulgação dos resultados. Por fim, acrescenta-se, também, que o Relatório se encontra disponibilizado, acessível a quaisquer interessados, nos sítios eletrônicos da organização avaliadora, o Instituto para o Desenvolvimento do Investimento Social, além da Fundação Sicredi, contratante da avaliação.

Dessa forma, como todos os aspectos recomendados pela Diretriz mostraram-se presentes no desenvolvimento da avaliação, julgou-se a Transparência como “Atendida”.

7. Legalidade

Nessa diretriz são ressaltados os aspectos legais nos quais as avaliações devem estar embasadas. Não foi identificado no relatório nenhuma menção aos documentos legais firmados, observação de legislações e normas pertinentes ao objeto avaliado, além de outros elementos necessários ao cumprimento da diretriz Legalidade, que foi, desse modo, considerada como “Não Atendida”.

DIMENSÃO 3 – CONTEXTUALIZAÇÃO E VALORAÇÃO

A terceira dimensão enfatiza a importância de se explicitar critérios tecnicamente consistentes e que respondam de forma adequada às questões avaliativas. Essa dimensão é desdobrada em

quatro diretrizes: Avaliações contextualizadas, Descrição do objeto, Critérios válidos e Critérios Explícitos.

1. Avaliações Contextualizadas

Nessa diretriz pretende-se avaliar o contexto cultural no qual o objeto está inserido. Cabe destacar que um preceito importante do projeto é a valorização da comunidade e ativos locais. Dentro dessa perspectiva, no relatório é afirmado que os professores possuíam maior flexibilidade para adaptar conteúdos programáticos ao contexto social e cultural no qual o aluno vivia, vinculando temas diversificados e tornando assim a aprendizagem mais interessante, pois ela estava adaptada ao seu dia a dia.

Entretanto não foram encontradas informações que comprovassem que o trabalho dos professores se tornou flexível ou trabalhou com conteúdo que se relacione ao contexto social dos beneficiados, portanto, as meta-avaliadoras observaram que os aspectos recomendados pela diretriz não se mostraram presentes na avaliação. Dessa forma, julgou-se que a diretriz Avaliações Contextualizadas foi classificada como “Não Atendida”.

2. Descrição do Objeto

Esta diretriz aborda a descrição do objeto bem como suas principais dimensões teóricas. Por meio da leitura das páginas 5, 6 e 7 do relatório de avaliação se pôde observar uma descrição detalhada do objeto avaliado. Nessa parte do instrumento foram enfatizados seus principais objetivos, estipulados no sentido de se compreender os impactos do Programa A União Faz a Vida. Procurou-se conhecer, por meio de um estudo avaliativo: a efetividade dos investimentos oriundos do Sicredi, apoiadores e parceiros; apoiar o planejamento estratégico e o processo de tomada de decisão da Fundação Sicredi quanto ao Programa; e, auxiliar a Fundação Sicredi a identificar

potenciais pontos de aprimoramento e otimização dos impactos por ele gerados (Instituto para o Desenvolvimento do Investimento Social, 2022c, p. 22).

O relatório aborda que o objetivo proposto foi o de aumentar o interesse, envolvimento e protagonismo no processo de aprendizagem; criar um maior senso de cidadania e coletividade; fortalecer os relacionamentos interpessoais; e, desenvolver habilidades socio-emocionais de indivíduos participantes do Programa. Já o objetivo de longo prazo foi o de “transformar crianças e jovens em protagonistas de seu processo de aprendizagem e sua atuação cidadã, para que promovam impactos positivos em suas comunidades” (Instituto para o Desenvolvimento do Investimento Social, 2022c, p. 23).

A partir desse objetivo, é possível encontrar as dimensões teóricas, lógicas, estratégias e técnicas operacionais para atingi-lo como: Atividades de articulação e mobilização; Formação de assessores pedagógicos; Formação de educadores; Desenvolvimento do Programa dentro das escolas.

Desta forma, de acordo com o observado no escopo do relatório meta-avaliado, foi verificado que a diretriz Descrição do Objeto foi totalmente contemplada, julgando-se, com isso, que a diretriz foi “Atendida”.

3. Critérios Válidos

Para Silva, Joppert e Gasparini (2020), essa diretriz aponta a importância de que os critérios tecnicamente defensáveis sejam válidos tanto para o objeto quanto para os atores interessados.

De acordo com o Instituto para o Desenvolvimento do Investimento Social (2022c) os critérios utilizados seguiram os preceitos do protocolo avaliativo SROI – *Social Return On Investment*, com os principais objetivos:

- a. Compreender o impacto do Programa por meio de um estudo avaliativo que demonstre a efetividade do investimento do Sicredi, apoiadores e parceiros;
- b. Apoiar o planejamento estratégico e o processo de tomada de decisão da Fundação Sicredi;
- c. Auxiliar a Fundação Sicredi a identificar potenciais pontos de aprimoramento e otimização dos impactos gerados pelo Programa.

Portanto, a avaliação formulada utilizou técnicas e critérios tecnicamente defensáveis e os aspectos recomendados pela diretriz se mostraram presente na avaliação, julgando-se, portanto, que a diretriz relativa aos Critérios Válidos foi “Atendida”.

4. Critérios explícitos

Nesta diretriz enfatiza-se que os critérios de valoração devem ser diretos e evidentes, ou seja, explícitos. Para Silva, Jopert e Gasparini (2020, p. 21) “além de consistentes e válidos, os critérios nos quais as avaliações se baseiam devem ser explícitos, de modo a favorecer que os interessados compreendam, apreciem, critiquem e validem, ou não, os mecanismos de valoração utilizados”.

Os critérios são explicitados por meio da aplicação do modelo *SROI* e foram divididos nas seguintes etapas: Mensuração da intensidade com que cada mudança ocorreu; ajuste da incidência do resultado que foi verificada no passo anterior; definição e atribuição de valores Proxy; definição do período do benefício; e, desconto dos benefícios e custos são financeiramente para que representem o valor presente (Instituto para o Desenvolvimento do Investimento Social, 2022c).

Dessa forma, pôde-se observar que no processo avaliativo foram utilizados mecanismos de valoração consistentes e válidos.

Portanto, os aspectos recomendados pela diretriz se mostraram presentes na avaliação, julgando-se que a diretriz Critérios Explícitos foi “Atendida”.

DIMENSÃO 4 – MÉTODO E VIABILIDADE

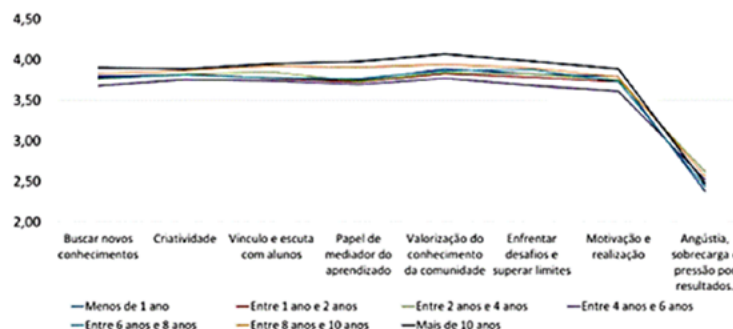
A quarta dimensão aborda o gerenciamento de metodologias consistentes que respondam ao objeto avaliado e às finalidades da avaliação. Essa dimensão é composta por oito diretrizes:

1. Viabilidade

Essa diretriz se baseia na análise do contexto social, regional, político, institucional e outros que auxiliem a construir com os atores interessados metodologias viáveis.

De acordo com Silva, Jopert e Gasparini (2020, p. 23), “As avaliações devem examinar periodicamente os contextos sociais, culturais, regionais, políticos, institucionais e subjetivos em que são realizadas e construir, com os atores interessados, estratégias metodológicas viáveis a tais contextos”.

Foi examinado por meio da análise do tempo de programa apresentado no relatório que o projeto realizou diversas avaliações com os atores interessados e seus contextos sociais, mantendo assim, a viabilidade do programa (Gráfico 1).

Gráfico 1 – Tempo de Programa

Fonte: Instituto para o Desenvolvimento do Investimento Social (2022c).

Desta forma, pode-se afirmar que os aspectos recomendados pela diretriz se mostraram presente na avaliação, sendo a diretriz da Viabilidade classificada como “Atendida”.

2. Tempo oportuno

Nesta diretriz observa-se a importância de se adequar a avaliação ao tempo político e institucional em que se pautará para a sua realização.

A avaliação SROI estabelece um período de benefício, que explicita o tempo pelo qual os projetos e seus benefícios estarão em funcionamento. No caso do Programa A União Faz a Vida, o relatório aponta que o tempo de duração do projeto é de sete anos, mas que se trata de um período influenciado pela duração das atividades ou por outros fatores externos (Instituto para o Desenvolvimento do Investimento Social, 2022c).

Conforme a avaliação, a Fundação Sicredi demonstra, em dados, os resultados da utilização do programa no tempo oportuno, visto que do ponto de vista dos professores, o programa continuará tendo efeitos na vida dos alunos por, aproximadamente, seis anos

após seu término. E, em suas vidas, por quase sete anos após seu encerramento, de acordo com os Gráficos 2 e 3:

Gráfico 2 - Período de benefício para alunos do Programa A União Faz a Vida (anos)



Fonte: Instituto para o Desenvolvimento do Investimento Social (2022c).

Gráfico 3 - Período de benefício para professores do Programa A união Faz a Vida



Fonte: Instituto para o Desenvolvimento do Investimento Social (2022c).

Logo, as meta-avaliadoras afirmam que a avaliação realizada pela Fundação SICRED avaliou os aspectos recomendados pela diretriz, julgando que a diretriz Tempo Oportuno foi "Atendida".

3. Metodologia consistente

Essa diretriz enfoca a necessidade de se utilizar diferentes métodos científicos para responder aos objetos avaliados e assegurar que as finalidades das avaliações sejam alcançadas. O relatório pretendeu estimar o retorno social dos recursos investidos e analisar sua contribuição para a sociedade. Ao utilizar o protocolo SROI –

Social Return On Investment, é mensurado com efetividade, por meio dos preceitos do SROI, o impacto social gerado pelo programa.

Portanto, as meta-avaliadoras observaram que os aspectos recomendados pela diretriz se mostraram presentes na avaliação, julgando-se que a diretriz Metodologia Consistente foi “Atendida”.

4. Técnicas e procedimentos válidos

Essa diretriz postula que sejam assegurados técnicas e procedimentos adequados para a compreensão da realidade.

A avaliação utilizou como técnica de investigação e análise a abordagem “Pré-Pós Design” (ou *Retrospective Pre Test*). Esse tipo de abordagem avaliativa trata de uma investigação que ocorre ao final do programa. Por meios dela é perguntado aos participantes como eles avaliam determinado assunto, comparando o antes (pré) com o depois (pós). Para sua realização foram aplicados questionários aos professores, alunos e familiares. Já no que tange à mensuração dos indicadores, foi utilizada uma escala de mensuração onde: 0 significa “nenhuma contribuição”; 1 significa “contribuição muito baixa”; 2 significa “contribuição baixa”; 3 significa “contribuição média”; 4 significa “contribuição alta”; e, 5 significa “contribuição muito alta” (Figura 1).

Figura 1 - Escala de intensidade



Fonte: Instituto para o Desenvolvimento do Investimento Social (2022c).

A Figura 2 apresenta um exemplo de mensuração de indicadores de impactos social nos alunos do Programa A União Faz a Vida.

Figura 2 - Eixo de impacto 1: sobre o aumento do interesse, envolvimento e protagonismo dos alunos no processo de aprendizagem

Aumento do interesse, envolvimento e protagonismo no processo de aprendizagem	Média Ponderada	Intensidade do Impacto
Maior interesse pelo espaço e atividades escolares	3,69	3,67
Aumento de interesses e conhecimentos que vão além do conteúdo formal abordado na escola	3,69	
Desenvolvimento de postura mais questionadora	3,67	
Maior autonomia e proatividade na condução e desenvolvimento das atividades escolares	3,64	

Fonte: Instituto para o Desenvolvimento do Investimento Social (2022c).

De acordo com o relatório (Instituto para o Desenvolvimento do Investimento Social, 2022c, p. 41) “A partir das respostas dos questionários, calcula-se a média ponderada de cada indicador, aplicando-se como peso o número de pessoas que selecionou cada uma das respostas da escala de intensidade”. Dessa forma, é possível averiguar que a técnica de investigação e análise dos dados reproduz, de forma segura, resultados válidos. Diante dessa afirmativa, as meta-avaliadoras consideram que os aspectos recomendados pela diretriz se mostraram presentes na avaliação, e julgou-se que, considerando as Técnicas e Procedimentos Válidos, a diretriz foi “Atendida”.

5. *Percurso metodológico explícito*

Nessa diretriz os métodos e técnicas utilizados nas avaliações devem estar explícitos. O relatório aponta que foi realizada a construção de uma rede de compromisso formada por gestores, parceiros, assessoria pedagógica e apoiadores. Também enfatiza que durante a realização do projeto foram realizados ajustes, dos quais os interessados recebiam informações e participavam.

Segundo o Instituto para o Desenvolvimento do Investimento Social (2022c, p. 12):

A metodologia do Programa A União Faz a Vida propõe o aprendizado baseado no desenvolvimento de projetos, que têm como característica a pedagogia ativa e participativa, com ênfase nas interações, interesses e curiosidades dos alunos, além da contextualização das áreas de conhecimento aos elementos da realidade local de cada município e região.

As etapas de aplicação foram as seguintes:

Índice Inicial: A partir do currículo escolar formal, o professor identifica os conteúdos que precisam ser desenvolvidos com os alunos e seleciona um território para a realização de uma expedição investigativa [...]; Índice Formativo: Mapeados os conhecimentos prévios, o professor e os alunos definem quais informações serão pesquisadas e quais conhecimentos pretendem construir a partir do projeto, ou seja, qual será o objetivo do estudo; Índice Final: O índice final consiste na etapa de fechamento do Projeto. Ele é realizado após todas as curiosidades dos alunos e todos os conteúdos de estudo terem sido trabalhados (Instituto para o Desenvolvimento do Investimento Social, 2022c, p. 12).

O processo de modelagem da avaliação foi pautado no Modelo SROI, adotando os seguintes passos: primeiro passo consistiu em medir a incidência dos resultados, ou seja, mensurar a intensidade com que cada mudança ocorreu; o segundo passo consistiu em ajustar a incidência do resultado que foi verificada no passo anterior (foram excluídos os fatores: contrafactual, atribuição e deslocamento); uma vez que a mudança efetiva (ou impacto) foi medida, o terceiro passo consistiu em definir e atribuir valores proxy; o quarto passo foi definir o período de benefício; e, finalmente, o quinto passo contemplou que os benefícios e custos foram financeiramente “descontados” para que representassem o valor presente (Instituto para o Desenvolvimento do Investimento Social, 2022c).

Sendo assim, o caminho metodológico traçado no relatório de avaliação atendeu ao Percurso Metodológico Explícito, tendo a diretriz sendo considerada como “Atendida”.

6. Autonomia e independência

Essa diretriz aborda a importância de se realizar uma avaliação autônoma, independente e responsável. O Instituto para o Desenvolvimento do Investimento Social é uma organização da sociedade civil que realiza avaliação de resultados e impactos alcançados por projetos de diversas organizações sociais. O instituto recebeu contribuições robustas de informações, dados, depoimentos, entre outros, da Fundação Sicredi, o que tornou as informações coletadas mais fáceis de serem analisadas e os resultados avaliativos explícitos e válidos.

Diante dessa afirmativa, as meta-avaliadoras consideram que os aspectos recomendados pela diretriz se mostraram presentes na avaliação, e julgou-se que a diretriz Autonomia e Independência foi “Atendida”.

7. Gestão responsável

Nessa diretriz aborda-se a importância do papel da gestão do avaliador. As avaliações devem ser conduzidas por equipes qualificadas, com boas condições de trabalho, cumprimento das leis, comunicação eficaz e respeito aos prazos e entregas. A gestão deve garantir o uso transparente e sustentável de recursos econômicos e naturais (Silva; Joppert; Gasparini, 2020).

O relatório demonstra que o projeto geriu de forma responsável os *stakeholders* envolvidos, a fim de assegurar de forma efetiva a realização do projeto. Cada grupo tem um papel específico a partir do uso de métodos variados para coletar informações e feedback.

1. **Alunos:** diretamente envolvidos com a metodologia do Programa, recebendo estímulos que promovem uma nova postura em relação ao ambiente escolar e ao aprendizado. O método de envolvimento utilizado para os alunos é o **grupo focal**, que permite uma interação direta e a coleta de percepções sobre suas experiências e aprendizados.
2. **Professores:** Os professores são formados na metodologia do Programa e incentivados a adotar novas abordagens no ensino. Para avaliar seu envolvimento, são utilizados **entrevistas, grupos focais e questionários quantitativos**.
3. **Família dos alunos:** As famílias têm um contato indireto com o Programa, interagindo principalmente através dos alunos. O envolvimento delas é avaliado por meio de **entrevistas e grupos focais**, que ajudam a entender como as mudanças na escola afetam a dinâmica familiar e a percepção sobre a educação.
4. **Comunidade local:** A comunidade desempenha um papel importante no compartilhamento de conhecimentos e experiências. O envolvimento da comunidade é mais pontual, e embora não haja um método específico mencionado, a interação com as atividades do Programa é essencial para enriquecer a experiência dos alunos.
5. **Assessoria Pedagógica:** Este grupo é capacitado na metodologia e é responsável pela formação dos educadores locais. Eles são considerados agentes viabilizadores da execução do Programa e participam do processo de avaliação através de **entrevistas**, que ajudam a entender como a formação e o acompanhamento impactam a implementação do Programa.
6. **Equipe de gestão das escolas, apoiadores e investidores locais, e poder público:** Esses grupos são descritos como agentes viabilizadores, e seu envolvimento é crucial para o

suporte e a execução do Programa. Embora não sejam beneficiários diretos, suas contribuições são fundamentais para o sucesso da iniciativa.

Em resumo, a diversidade de métodos de envolvimento, como entrevistas, grupos focais e questionários, permite uma avaliação rica e multifacetada do Programa. Em particular, com os grupos focais, o relatório descreve que foi permitido verificar a materialidade das hipóteses de impacto levantadas na Teoria de Mudança.

Desta forma, as meta-avaliadoras observaram que os aspectos recomendados pela diretriz se mostraram presentes na avaliação, e julgou-se que a Gestão Responsável foi “Atendida”.

8. Meta-avaliação

Nessa diretriz a avaliação, bem como o desenvolvimento da mesma deve passar por um processo avaliativo externo. A diretriz meta-avaliação “Não se Aplica”, pois não se enquadra no objetivo do relatório da avaliação. As sínteses dos resultados obtidos neste estudo meta-avaliativo são demonstradas nos Quadros 3 e 4 revelando os níveis de atendimento das diretrizes e a classificação final das quatro dimensões observadas:

Quadro 3 – Resultados geral da avaliação das diretrizes em relação aos critérios meta-avaliativos adotados

Dimensões e suas diretrizes	Classificação
Aprendizagem e Utilização	
Níveis de participação adequados aos interessados	A
Propósitos acordados mutuamente	A
Perguntas relevantes	AP
Resultados úteis	A
Conclusões fidedignas	A

Dimensões e suas diretrizes	Classificação
Recomendações contextualizadas	A
Comunicação cuidadosa	AP
Direitos e Integridades	
Respeito aos direitos das pessoas	AP
Diversidade e equidade	NA
Identidades culturais	NA
Reciprocidade nas trocas	NA
Explicitação de interesses	A
Transparência	A
Legalidade	NA
Contextualização e Valoração	
Avaliações contextualizadas	NA
Descrição do objeto	A
Crítérios válidos	A
Crítérios explícitos	A
Método e Viabilidade	
Viabilidade	A
Tempo oportuno	A
Metodologia consistente	A
Técnicas e procedimentos válidos	A
Percurso metodológico explícito	A
Autonomia e independência	A
Gestão responsável	A
Meta-avaliação	NSA

Legenda: A - atendida; AP - atendida parcialmente; NA - não atendida; NSA - não se aplica.
Fonte: Os autores (2022).

Quadro 4 – Resultados do julgamento das dimensões

Dimensões	Número de Atendimentos de diretrizes por Dimensão	Classificação final da Dimensão
Aprendizagem e Utilização	5	Boa
Direitos e Integridade	2	Deficiente
Contextualização e Valoração	3	Boa
Método e Viabilidade	7	Excelente

Fonte: Os autores (2022).

CONCLUSÕES E RECOMENDAÇÕES

O processo de meta-avaliação pautado na análise do relatório de avaliação do Programa A União Faz a Vida obteve como resultado o atendimento pleno de 17 das 26 diretrizes; três diretrizes foram atendidas parcialmente; cinco diretrizes não foram atendidas e uma diretriz foi considerada não aplicável ao contexto. Das quatro dimensões, uma foi classificada como excelente, duas foram classificadas como boas e uma deficiente, tendo atendido somente a duas das sete diretrizes integrantes. Cabe observar que a diretriz meta-avaliação não foi julgada por não se enquadrar no objetivo do relatório da avaliação, por isso foi desconsiderada no cômputo da dimensão Método e Viabilidade.

Como pontos fortes da avaliação realizada pelo Instituto para o Desenvolvimento do Investimento Social, as meta-avaliadoras destacam: objetivos bem definidos para condução da avaliação; o envolvimento e o engajamento dos *stakeholders* impactados pelo Programa; escolha de uma metodologia consagrada e adequada aos objetivos da avaliação; resultados e recomendações úteis para os públicos-alvo e para os tomadores de decisão.

Em relação à questão avaliativa norteadora do presente estudo: “Em que medida o Relatório da Avaliação do Retorno Social do Investimento do Programa A União Faz a Vida da Fundação Sicredi atende às Diretrizes para a prática de avaliação no Brasil?”, concluiu-se que o Relatório meta-avaliado alcançou um bom nível de atendimento das diretrizes para a prática de avaliação no Brasil.

Considerando que o objetivo de uma meta-avaliação é qualificar uma avaliação através da indicação de melhorias e da demonstração dos pontos positivos e limitações observadas, as meta-avaliadoras deste estudo apresentam como recomendações:

1. Indicar de forma direta e explícita as questões avaliativas, visto que elas funcionam como orientadoras do processo avaliativo;
2. Disponibilizar relatórios parciais aos interessados para que os resultados obtidos ao longo do processo avaliativo possam ser acompanhados e monitorados;
3. Incluir como anexo o Termo de Consentimento Livre e Esclarecido – TCLE, como documento garantidor do respeito aos direitos dos participantes.
4. Explicitar os critérios utilizados para lidar com as questões culturais, étnico-raciais e socioeconômicas dos públicos-alvo e os procedimentos éticos e legais empregados para contribuir com a redução das desigualdades e de qualquer tipo de discriminação;
5. Incluir os documentos legais pertinentes, citar as leis, normas e regulamentos de observância obrigatória;
6. Descrever o contexto do objeto avaliado, destacando os valores sociais e culturais das comunidades afetadas pelo programa e pela avaliação;

7. Contratar avaliadores externos para realização de uma meta-avaliação a fim de qualificar a avaliação realizada.

As recomendações têm a pretensão de contribuir para as reflexões sobre a avaliação conduzida e outras avaliações de impacto de programas e projetos sociais que venham a ser realizadas. Além disso, cabe enfatizar os benefícios potenciais de uma meta-avaliação e como é desejável que ela seja incluída como componente integrante e relevante já na fase de planejamento das avaliações.

MÉRITO DO ESTUDO

Elaine Rodrigues Perdigão

O artigo é bem estruturado, começando com um resumo claro que delinea os objetivos e a importância da meta-avaliação. As autoras utilizam uma abordagem sistemática para avaliar a conformidade do relatório com as diretrizes estabelecidas, o que é um ponto positivo, pois demonstra rigor metodológico. A análise é baseada em uma leitura apurada do relatório.

A meta-avaliação é relevante, especialmente em contextos de avaliação de impacto social, onde a credibilidade é crucial. As autoras destacam a importância de seguir diretrizes para garantir a qualidade das avaliações. A clareza e a relevância do estudo são pontos fortes.

O estudo apresenta a originalidade de ter como “fio condutor” as boas práticas avaliativas disseminadas por meio da publicação *Diretrizes para a Prática de Avaliação no Brasil* que orientam a aplicação de critérios claros e métodos diversificados que garantem uma análise completa de outras avaliações.

A sua aplicação, por meio de estudos meta-avaliativos somativos, proporciona o incentivo da utilização de técnicas avaliativas

que diagnosticam e acompanham o desenvolvimento de avaliações já encerradas, por meio do estudo de sua documentação.

REFERÊNCIAS

CHIANCA, Thomaz; MARINO, Eduardo; SCHIESARI, Laura. **Desenvolvendo a cultura da avaliação em organizações da sociedade civil**. São Paulo: Global, 2001.

FABIANI, Paula; REBEHY, Sofia; CAMELO, Rafael; VICENTE, Francisco José; MOSANER, Marcelo. **Avaliação de Impacto Social metodologias e reflexões**. São Paulo: Instituto para o Desenvolvimento do Investimento Social, [2020]. Disponível em: https://www.idis.org.br/wp-content/uploads/2018/05/Artigo_Avaliacao_Impacto_Social_06.pdf. Acesso em: 12 jun. 2022.

GERTLER, Paul J.; MARTÍNEZ, Sebastián; PREMAM, Patrick; RAWLINGS, Laura B.; VERMEERSCH, Christel M. J. **Avaliação de Impacto na Prática**. São Paulo: Banco Internacional para Reconstrução e Desenvolvimento; Banco Mundial, 2018.

INSTITUTO PARA O DESENVOLVIMENTO DO INVESTIMENTO SOCIAL. **Avaliação de Impacto e SROI**. Pinheiros, SP, 2022a. Disponível em: <https://www.idis.org.br/avaliacao-de-impacto-e-sroi/>. Acesso em: 12 jun. 2022.

INSTITUTO PARA O DESENVOLVIMENTO DO INVESTIMENTO SOCIAL. **O que é SROI e por que ele é útil?** Pinheiros, SP, 2022b. Disponível em: <https://www.idis.org.br/o-que-e-sroi-e-por-que-ele-e-util/>. Acesso em: 18 jun. 2022.

INSTITUTO PARA O DESENVOLVIMENTO DO INVESTIMENTO SOCIAL. **Relatório de Avaliação de Impacto**: Sicredi. Pinheiros, SP, 2022c. Disponível em: https://idis.org.br/files/RelatorioSROI_SICREDI.pdf. Acesso em: 1 jun. 2022.

ORLANDO FILHO, Ovídio; ELLIOT, Ligia G. (org.). **Estudos meta avaliativos**: aplicações metodológicas. São Paulo: Pimenta Cultural, 2020. 216 p. Disponível em: https://www.academia.edu/44513400/Estudos_meta_avaliativos_aplica%C3%A7%C3%B5es_metodol%C3%B3gicas. Acesso em: 1 jul. 2022.

SCRIVEN, Michael. **Avaliação**: um guia de conceitos. Rio de Janeiro: Paz e Terra, 2018.

SILVA, Rogério R.; JOPPERT, Márcia P.; GASPARINI, Max F. V. (org.). **Diretrizes para a prática de avaliação no Brasil**: livro eletrônico. Rio de Janeiro: Rede Brasileira de Monitoramento e Avaliação, 2020. Disponível em: https://rbma.site/book-category/biblioteca-digital/?wbg_author_s=Rog%C3%A9rio+Renato+Silva%3B+Marcia+Paterno+Joppert%3B+Max+Felipe+Vianna+Gasparini. Acesso em: 16 jul. 2022.

SILVA, Rogério; JOPPERT, Márcia; GASPARINI, Max. Por que produzimos diretrizes para a prática de avaliação. **Nexo**, [S. l.], 25 ago. 2020. Disponível em: <https://www.nexojournal.com.br/ensaio/2020/Por-que-produzimos-diretrizes-para-a-pr%C3%A1tica-de-avalia%C3%A7%C3%A3o>. Acesso em: 16 jul. 2022.

STUFFLEBEAM, Daniel L. The metaevaluation imperative. **American Journal of Evaluation**, Thousand Oaks, v. 22, n. 2, p. 183-209, 2001. DOI: <https://doi.org/10.1177/109821400102200204>. Disponível em: <https://journals.sagepub.com/doi/10.1177/109821400102200204>. Acesso em: 11 jun. 2022.

WORTHEN, Blaine R.; SANDERS, James R.; FITZPATRICK, Jody L. **Avaliação de programas: concepções e práticas**. São Paulo: Editora Gente, 2004.

4

Patrícia Sousa de Souza
Lígia Silva Leite

PROGRAMA MELHORIA DA EDUCAÇÃO:

A META-AVALIAÇÃO
COLABORANDO PARA A QUALIDADE
DE PROGRAMAS EDUCACIONAIS
E SEUS PROCESSOS AVALIATIVOS

INTRODUÇÃO

O termo “política pública”, largamente discutido na literatura brasileira e internacional, pode assumir diversas definições, podendo ser considerado um novo e conflituoso campo de estudo das ciências políticas. Na década de 1930 Harold Lasswell Introduziu a expressão *policy analysis* (análise de política pública), ocasião em esse novo campo do conhecimento passou a ser discutido nos ambientes acadêmicos, “como forma de conciliar conhecimento científico/acadêmico com a produção empírica dos governos e como forma de estabelecer o diálogo entre cientistas sociais, grupos de interesse e governo” (Souza, 2006, p. 23).

Normalmente, as políticas públicas são instituídas pela administração pública, em seus diferentes níveis, por meios de planos, projetos e programas e possuem como finalidade basilar o atendimento a uma necessidade da sociedade. A sua implementação, geralmente, obedece a um determinado ciclo com as seguintes fases: identificação do problema, formulação da agenda destinada a discutir o problema, formulação das alternativas possíveis para a sua solução, tomada de decisão de como isso será feito, implementação do programa e, por fim, avaliação dos resultados alcançados.

Na última etapa do ciclo, a avaliação dos resultados alcançados, torna-se necessário conhecer, entre outros, aspectos como:

Verificar se o programa foi conduzido e concluído com mérito, ou seja, se houve responsabilidade na utilização dos recursos públicos utilizados em sua implementação? [...] verificar se o programa possuiu valor, ou seja, se seus resultados foram, realmente, aproveitados e foram úteis para a sociedade? (Orlando Filho; Oliveira, 2019, p. 152).

Essas e outras importantes questões podem ser respondidas por meio dos resultados do processo avaliativo. Pois, esse é o objetivo fundamental da avaliação, de formular, ao seu final, recomendações

úteis que levem a administração pública à tomada de decisões inteligentes. O processo decisório, por sua vez, é conduzido no sentido de encontrar as melhores soluções dos aspectos relacionados à melhoria daquilo que foi avaliado. As soluções alcançadas podem ter diferentes significados como, por exemplo: reestruturar o programa de forma a atender mais plenamente as necessidades relacionadas à sua implementação, adequá-los a possíveis novas realidades contextuais, extinguir o programa ou mesmo criar outras políticas públicas.

Assim, avaliar, em seu sentido *stricto*, refere-se ao processo de determinar o mérito, relevância ou valor de algo. Ou, tomando por base termos de espectros de maior amplitude, porém, com a mesma referência ou proximidade, podem também significar: apreciar, analisar, examinar, criticar, graduar, revisar, interpretar, dentre outros. O processo de avaliar não é simples e envolve identificação de padrões, diretrizes ou critérios que sejam relevantes. Não se limita à simples medida. Avaliar de maneira criteriosa “é mais abrangente do que a maioria das pesquisas em ciências sociais” (Scriven, 2018, p. 185-186).

Essa complexidade conduz à reflexão de que os processos avaliativos também necessitam ser avaliados, de forma a se aferir aspectos fundamentais intrínsecos a toda avaliação que, segundo Yarbrough, Shulha, Hopson e Caruthers (2011), são: sua utilidade, exequibilidade, adequação, precisão e a responsabilização de seus executores em função dos resultados por ela alcançados.

Ao decidir pela avaliação de uma avaliação, processo que se denomina meta-avaliação (Scriven, 1967; Stufflebeam, 2001; Davidson, 2005), o avaliador se torna responsável por avaliar, examinar ou interpretar tanto um programa, quanto a si mesmo e as ações propostas; manifestando assim, a alta relevância da avaliação enquanto ciência, apresentando em ações o rigor ético a que está submetido voluntariamente e balizado por métodos e estratégias metodológicas próprias desta área. Portanto, o entendimento é que meta-avaliar é tornar a avaliação cada vez mais perfeita, aprimorada.

Ou seja, é estar na constante busca do aperfeiçoamento científico do próprio processo avaliativo originalmente posto em prática.

Este estudo meta-avaliativo carrega a intencionalidade de seus autores de desenvolverem seu estudo avaliativo de modo responsável e ético, buscando examinar e criticar um relatório desenvolvido por outrem, de modo a refazer um caminho na perspectiva de interpretar como programas podem servir de referência para outros programas de qualidade tendo em vista a respeitabilidade, a responsabilidade social e o rigor científico.

OBJETIVO DESTE ESTUDO META-AVALIATIVO

O objetivo deste estudo é meta-avaliar o processo avaliativo apresentado no Relatório de Avaliação Qualitativa e de Impacto do Programa Melhoria da Educação, elaborado pelo Itaú Social (2015 - 2017) e norteada por meio da utilização da metodologia pareamento por escore de propensão e regressão de diferenças em diferenças (avaliação de impacto) e análise documental, entrevistas e grupos focais (avaliação qualitativa). Como critérios estabelecidos para realizar a meta-avaliação foram utilizadas as Diretrizes da Avaliação para a América Latina e Caribe elaboradas por Rodríguez Bilella, Pablo D., Sergio Martinic Valencia; Luis Soberón Alvarez, Sarah D. Klier, Ana L. Guzmán Hernández e Esteban Tapella (2016).

OBJETO DO ESTUDO META-AVALIATIVO

O objeto deste estudo é o Relatório de Avaliação Qualitativa e de Impacto do Programa Melhoria da Educação, elaborado para

o Itaú Social, pelas empresas produtoras da avaliação: Instituto Datafolha, FRONTERES Pesquisa e Formação em Gestão e Políticas Públicas e Área de Avaliação de Impacto Itaú Unibanco. A avaliação realizada objetivou a geração de melhoria da gestão de educação de unidades educacionais dos municípios brasileiros, considerando as dimensões: rendimento escolar e gestão educacional.

DELINEANDO O OBJETO DA META-AVALIAÇÃO

O título do objeto avaliado é Avaliação Qualitativa e de Impacto do Programa Melhoria da Educação. Primeiro, cabe compreender o título a partir do foco de sua avaliação:

- Avaliação qualitativa, segundo a explicação de Scriven para este tipo de avaliação, se aplica muito bem ao desenho que fora feito neste relatório: “a avaliação rigorosa e a dicotomia entre qualitativo e quantitativo precisa ser definida claramente e colocada em perspectiva, caso contrário chega a confundir, mais do que explicar” (Scriven, 2018, p. 177)
- Avaliação de Impacto é a avaliação focada nos desfechos ou na recompensa, em vez de fazer a avaliação de processos, da entrega ou da implementação (Scriven, 2018, p. 166).
- A avaliação de programas é considerada por Scriven a maior área da avaliação que se tornou especialidade nos EUA na década de 1960 (Scriven, 2018, p. 153).

O sumário do relatório está organizado da seguinte forma: apresentação, o programa avaliado, avaliação de impacto (subdividido em: indicadores, metodologia e resultados); avaliação qualitativa (subdividido em: fonte de dados metodologia e resultados) e considerações finais.

É possível na leitura do relatório identificar os *stakeholders*: Oficina Municipal, Consórcio Intermunicipal do Vale do Paranapanema (CIVAP) e Consórcio Intermunicipal de Gestão Pública Integrada dos Municípios do Baixo Rio Paraíba (COGIVA). O público-alvo do programa consistiu nos gestores educacionais do município. Ao procurar compreender melhor esse público chegou-se à caracterização dos prefeitos, secretários, técnicos do Órgão Gestor da Educação Municipal, diretoria dos consórcios e equipe dos consórcios (Itaú Social, 2018).

Segundo a apresentação, o Programa Melhoria da Educação (edição 2015 – 2017) visa contribuir com as Secretarias Municipais de Educação para o “fortalecimento de sua capacidade em garantir acesso, permanência e aprendizagem com equidade para todos e todas na busca por uma educação inclusiva e de qualidade nas redes públicas municipais” (Itaú Social, 2018, p. 6).

A equipe técnica executora da avaliação (metodologia, instrumentos de coleta de dados) e produtora do relatório de avaliação de impacto e qualitativa foram: Instituto Datafolha, FRONTERES Pesquisa e Formação em Gestão e Políticas Públicas e a Área de Avaliação de Impacto Itaú Unibanco.

Para a avaliação de impacto a metodologia do relatório foi baseada “na comparação entre dados do grupo de tratamento e dados do controle, juntamente com a aplicação de técnicas estatísticas que buscam eliminar fatores que possam distorcer a percepção do pesquisador sobre os reais efeitos do programa avaliado” (Itaú Social, 2018, p. 12). Na subdivisão indicadores, está mencionado que recorreram a duas fontes de informação para saber sobre rendimento escolar (apontado como uma dimensão). O intuito era “avaliar o impacto do programa nos processos das secretarias municipais de educação e seu transbordamento para o ambiente escolar” (Itaú Social, 2018, p. 9). Para isto utilizaram uma abordagem quantitativa realizada em 2017.

A pergunta avaliativa que orientou a avaliação foi a seguinte: em que nível os objetivos almejados pelo Programa Melhoria da Educação foram atingidos já no curto prazo, logo após o término das atividades formativas junto aos consórcios COGIVA e CIVAP? Segundo a equipe avaliadora foi escolhida a avaliação de impacto como abordagem avaliativa. A mesma foi desenvolvida por meio da técnica de Pareamento por Escore de Propensão e a base de dados construída com indicadores de impacto antes e pós programa, e analisados por pesquisadores econométricos. Os resultados deveriam ser mapeados em três dimensões: rendimento escolar – desempenho escolar dos alunos e alunas ao nível das escolas; gestão-escolas – gestão educacional ao nível das escolas; gestão-secretarias municipais de educação – gestão educacional ao nível das secretarias de educação (Itaú Social, 2018, p. 8-12).

O Capítulo 3 trata da avaliação qualitativa, explicitamente realizada em 2018, e conduzida pela FRONTERES Pesquisa e Formação em Gestão e Políticas Públicas. Segundo esta empresa, a avaliação buscou se aproximar do processo de implementação do Melhoria da Educação nos consórcios COGIVA e CIVAP registrando as ações desenvolvidas e as reações de curto prazo dos atores locais e foram divididas em três dimensões de análise, a saber: avaliação de desenho; avaliação de processo; avaliação de efeitos. Registra-se nesse Capítulo, ainda, a intenção desta metodologia responder às perguntas avaliativas do programa por meio dos quatro diferentes processos que foram foco da formação do Programa Melhoria da Educação pela Oficina Municipal: avaliação de desempenho, formação continuada, avaliação do ensino e aprendizagem e projeto político pedagógico.

A fonte de dados que alimentou esta avaliação qualitativa (após o término das intervenções do programa) foi a análise de documentos do programa e a coleta de dados primários. Foram abordados técnicos e gestores do Itaú Social, da Oficina Municipal e dos dois Consórcios, além de profissionais de nove municípios

que foram admitidos como estudos de caso (Itaú Social, 2018). Com vários objetivos diferentes foram adotadas duas técnicas na metodologia: análise documental e entrevista com roteiro semiestruturado, tendo como entrevistados técnicos e gestores do Itaú Social e da Oficina Municipal. Mencionam também ter utilizado entrevistas com atores-chave a fim de que relatassem a experiência dos municípios com o programa. Somado a isto, empreenderam grupos focais com base nos processos que compunham as formações. Ao todo foram 69 atores, dos dois consórcios envolvidos, participando de entrevistas e grupos focais (mencionados adiante no texto), ambos com roteiros semiestruturados comuns aos dois consórcios. Assim, foram alcançados os principais resultados do programa partindo da visão dos participantes.

No Capítulo 4, Considerações Finais, encontra-se a informação denominada de conjunto de avaliações realizadas (relatório), o período das avaliações entre 2017 e 2018 e o objetivo principal de investigar a eficácia da edição do Programa Melhoria da Educação no curto prazo junto aos dois consórcios CIVAP e COGIVA, e que para isto recorreu-se a múltiplos métodos de pesquisa. Cabe salientar que há uma informação em nota de rodapé na página 6 informando que em 2021 houve aumento no número de municípios no consórcio CIVAP. Continuando nas considerações, há informação de que os resultados obtidos se desdobraram em propostas de melhoria e visibilidade de resultados não previstos.

Foi identificada também uma declaração das limitações dos processos avaliativos: “o esforço dos pesquisadores em produzir metodologias sólidas e robustas” (Itaú Social, 2018, p. 20). Há a citação, como exemplo, de um questionário aplicado por telefone (para obtenção de dados para avaliação de impacto) como possível erro de medida. Ao final há uma afirmação de que as interpretações devem ser realizadas com cautela “por questões inerentes às pesquisas empíricas” (Itaú Social, 2018, p. 20), ainda mais com os resultados positivos encontrados no documento.

METODOLOGIA DESTA META-AVALIAÇÃO

O estudo meta-avaliativo utilizou análise documental e pesquisa bibliográfica. Foi norteado pelas Diretrizes de Avaliação para a América Latina e o Caribe em sua versão atualizada e consensualizada para meta-avaliar o relatório Avaliação Qualitativa e de Impacto do Programa Melhoria da Educação. Neste estudo foram utilizados as quatro dimensões e os vinte critérios destas diretrizes, cujas especificações estão maiores nesta versão 2021 buscando considerar a diversidade do interior dos países e de suas culturas (Ghiano; Wehrle; Fiscarelli; Martinic; Gutiérrez; Erazo; Guzmán; Álvarez, 2021).

Conforme Leite (2012, p. 101), “um processo avaliativo precisa percorrer numerosas etapas ou procedimentos para que a avaliação seja realizada com qualidade”. Com base nesta afirmação, esta meta-avaliação recorreu à lista de verificação para julgar o relatório segundo as Diretrizes de Avaliação para América Latina e o Caribe, a fim de obter resultados confiáveis ao final deste estudo; isto por ser a lista de verificação de critérios ou indicadores de mérito um instrumento adequado para a investigação da completude da implementação de um projeto da natureza deste que está sendo avaliado, identificando todas as dimensões de valor relevante e sem sobreposição. Além de ser um instrumento bastante versátil (Scriven, 2018, p. 87). Portanto, esta meta-avaliação está desenhada da seguinte forma:

Figura 1 - Desenho da metodologia do estudo meta-avaliativo

Fonte: Os autores (2023).

As Diretrizes de Avaliação para a América Latina e o Caribe foram as indicações de qualidade utilizadas neste estudo meta-avaliativo, compondo, com isso, os parâmetros representativos dos critérios adotados na meta-avaliação. O Quadro 1 apresenta as quatro dimensões e as vinte Diretrizes de Qualidade da Avaliação para a América Latina e o Caribe que norteou a verificação do nível de qualidade do relatório objeto deste estudo meta-avaliativo.

Quadro 1 - Dimensões e diretrizes de avaliação para a América Latina e o Caribe

Dimensões	Diretrizes
Rigorosidade	1.1. Avaliabilidade da Intervenção 1.2. Avaliação Contextualizada 1.3. Compreensão Detalhada da Intervenção 1.4. Perguntas Avaliativas Relevantes e Consensualizadas 1.5. Abordagens e Métodos Devidamente Explicados e Justificados
Ética e Princípios Jurídicos	1.6. Respeito aos Direitos das Pessoas 1.7. Transparência 1.8. Ética e Integridade Profissional 1.9. Legalidade 1.10. Autonomia
Compreensão Cultural	1.11. Direitos Culturais 1.12. Igualdade e Equidade 1.13. Reciprocidade e Interculturalidade
Relevância e Utilidade	1.14. Participação Efetiva e Consciente 1.15. Propósitos Acordados mutuamente 1.16. Valores Explícitos 1.17. Comunicação e Relatórios Específicos Apropriados 1.18. Interesse pelas Consequências e Incidências

Fonte: Os autores (2023) adaptado de Ghiano, Wehrle, Fiscarelli, Valencia, Checa Gutiérrez, Amariles Erazo, Guzmán e Soberón Álvarez (2021).

CRITÉRIOS DE ATENDIMENTO DAS DIMENSÕES

Os autores definiram os critérios de avaliação para cada dimensão das Diretrizes. Eles são válidos mesmo nos casos em que a diretriz não seja aplicável na avaliação do relatório; são eles: TA = Totalmente Atendida; PA =Parcialmente Atendida; NA = Não Atendida; na= Não Aplicável.

O atendimento das dimensões foi definido por meio dos seguintes critérios:

- a. A Dimensão Rigoriedade foi considerada 'Totalmente Atendida' quando todas as suas diretrizes foram atendidas; foram consideradas 'Parcialmente Atendida' se, pelo menos, três de suas diretrizes foram atendidas; e, 'Não Atendidas' no caso de menos de três de suas diretrizes não terem sido atendidas.
- b. A Dimensão Ética e Princípios Jurídicos foi considerada 'Totalmente Atendida' se suas cinco diretrizes tiverem sido atendidas; "Parcialmente atendida" quando três de suas diretrizes foram atendidas; e, 'Não Atendida' no caso de menos de três de suas diretrizes não serem atendidas.
- c. A Dimensão Compreensão Cultural foi considerada 'Totalmente Atendida' se todas as suas diretrizes foram atendidas; 'Parcialmente Atendida' se, pelo menos duas de suas diretrizes forem atendidas; e, 'Não Atendida' se menos de duas diretrizes tiverem sido atendidas.
- d. A Dimensão Relevância e Utilidade foi considerada 'Totalmente Atendida' se todas as suas diretrizes tiverem sido atendidas; 'Parcialmente Atendida' quando pelo menos seis de suas diretrizes tiverem sido atendidas; e, 'Não Atendida' se menos de seis diretrizes foram atendidas.

Para melhor compreensão dos critérios utilizados foi elaborado o Quadro 2 que indica a relação dimensão x conceito x número de diretrizes atendidas para cada dimensão, a saber:

Quadro 2 – Classificação das dimensões de acordo com os critérios de atendimento

Dimensão rigorosidade	
Classificações	Critérios de atendimento
Totalmente Atendida (TA)	= 5 diretrizes atendidas
Parcialmente Atendida (PA)	= 3 diretrizes atendidas
Não Atendidas (NA)	< 3 diretrizes atendidas
Dimensão ética e princípios jurídicos	
Totalmente Atendida (TA)	= 5 diretrizes atendidas
Parcialmente Atendida (PA)	= 3 diretrizes atendidas
Não Atendidas (NA)	< 3 diretrizes atendidas
Dimensão compreensão cultural	
Totalmente Atendida (TA)	= 3 diretrizes atendidas
Parcialmente Atendida (PA)	= 2 diretrizes atendidas
Não Atendidas (NA)	< 2 diretrizes atendidas
Dimensão relevância e utilidade	
Totalmente Atendida (TA)	= 7 diretrizes atendidas
Parcialmente Atendida (PA)	= 6 diretrizes atendidas
Não Atendidas (NA)	< 6 diretrizes atendidas

Fonte: Os autores (2023).

ENTENDENDO AS DIRETRIZES DA AVALIAÇÃO PARA A AMÉRICA LATINA E O CARIBE

Tomando por base as desigualdades nas regiões que compõem a América Latina e o Caribe, bem como, as peculiaridades desses lugares, foram criadas as Diretrizes de Qualidade de Avaliação para esse perfil geográfico, considerando as desigualdades externas e internas dos países dessas regiões. Dados do Banco Mundial apontam que ainda há desafios antigos que foram somados às novas demandas do século XXI. Nem se venceu obstáculos e há outros novos a serem entendidos e a serem ultrapassados (Ghiano; Wehrle; Fiscarelli; Martinic; Gutiérrez; Erazo; Guzmán; Álvarez, 2021, p. 21).

De acordo com estes mesmos autores, o propósito das Diretrizes da Avaliação para a América Latina e o Caribe é oferecer subsídios para a transformação das realidades, considerando a pobreza, a exclusão, as desigualdades e que intervenções democráticas e sérias que precisam levar em conta todos esses problemas; além do que a avaliação precisa ser constante: no planejamento, durante a intervenção e após estas. Isto, por ser a avaliação orientada por princípios e abordagens que podem apontar os rumos que essas transformações estão tomando. Este propósito é expresso por meio da figura proposta para uma avaliação apoiada nestas diretrizes:

Figura 1 – Representação dos principais parâmetros de uma avaliação



Fonte: Ghiano, Wehrle, Fiscarelli, Valencia, Checa Gutiérrez, Amariles Erazo, Guzmán e Soberón Álvarez (2021, p. 16).

Desenvolvedores de programas, que é foco de interesse deste trabalho meta-avaliativo, precisam ter em mente a necessidade de dar atenção a estas diretrizes, a fim de que percebam em suas intervenções as melhores práticas, aprendizados, pontos fortes e/ou fracos, e assim não os entendendo como erros no processo; mas como percurso processual que pode apontar novos caminhos, novas formas de fazer de maneira mais técnica, mais assertiva, aparelhando os atores envolvidos com informações para melhores decisões.

Para esta meta-avaliação utilizar as Diretrizes da Avaliação para a América Latina e o Caribe como critérios de qualidade para a avaliação do relatório de Avaliação Qualitativa e de Impacto do Programa Melhoria da Educação é trazer à tona uma das principais funções da avaliação: “a produção de informação e conhecimento através da avaliação é o caminho mais seguro para a melhoria dos resultados, impactos dos serviços públicos e para a qualidade da democracia na América Latina e o Caribe” (Ghiano; Wehrle; Fiscarelli; Valencia; Checa Gutiérrez; Amariles Erazo; Guzmán; Soberón Álvarez, 2021, p. 22).

ANÁLISE E RESULTADOS DA META-AVALIAÇÃO

A seguir é apresentada a análise do relatório, objeto desta meta-avaliação, a partir das diretrizes que compõem cada dimensão das Diretrizes da Avaliação para a América Latina e o Caribe.

DIMENSÃO 1 - RIGOROSIDADE

Diretriz condições de avaliabilidade

É relevante o entendimento do termo avaliabilidade. Para Scriven, “na etapa de planejamento, não se pode deixar que a preocupação com a avaliabilidade se torne um caso em que a avaliação orienta o projeto” (Scriven, 2018, p. 120). O entendimento da importância deste termo traz grandes melhorias para o programa porque refina o foco e traz uma ansiedade positiva, a ansiedade que deixa o avaliador sempre em estado de alerta.

Conforme as Diretrizes de Avaliação para a América Latina e o Caribe, em sua dimensão rigorosidade, o critério avaliabilidade está relacionado com a preocupação do avaliador de assegurar que a intervenção procure resolver necessidades e problemas de populações (Ghiano; Wehrle; Fiscarelli; Valencia; Checa Gutiérrez; Amariles Erazo; Guzmán; Soberón Álvarez, 2021).

No objeto meta-avaliado essas populações não estão totalmente registradas como atores avaliados, o que classifica esse critério como pouco contemplado na avaliação realizada. Isto é notório durante a leitura do documento, quando comparadas ao resultado. Nesse aspecto, deve-se destacar uma nota de rodapé no documento que menciona 40 municípios participantes, ao mesmo tempo em que indica somente nove municípios admitidos como estudos de caso (Itaú Social, 2018)

Outro aspecto identificado no resultado que diz: “o programa proporcionou melhor utilização dos recursos públicos, com ganhos em escala e menor sobreposição de esforços” (Itaú Social, 2018, p. 13-18), em contrapartida outra parte do relatório menciona que há piora no planejamento orçamentário dos municípios CIVAP.

Ao mesmo tempo não se identifica pergunta avaliativa que corresponda a este resultado a partir da tabela utilizada sobre avaliação de efeitos. Além disso, constam desta avaliação, mais de quatro tipos de processos diferentes avaliativos não esclarecidos (Itaú Social, 2018).

Diretriz: avaliação contextualizada

A contextualização da intervenção que é avaliada e da própria avaliação, segundo as Diretrizes de Avaliação para a América Latina e o Caribe, deve dar conta de que os problemas da população têm origens e fatores em processos antigos e atuais; e isto implica na compreensão de diferentes níveis de complexidade e consolidação estrutural. Não é possível avaliar sem que se observe todos esses processos. A linha temporal dos problemas localizados precisa estar bem dimensionada, pois ela é parte da avaliação.

Para entender ainda mais profundamente esse critério da avaliação contextualizada é possível, retomar Demo (2013), quando diz que a realidade social possui dimensões qualitativas e que, embora, a materialidade (perspectiva quantitativa) seja importante para os estudos (para as ciências de um modo geral) olhar em perspectiva para a dinâmica histórica é fundamental.

É preciso ressaltar também que os critérios de avaliabilidade e de avaliação contextualizada estão entrelaçados. Não por acaso um vem imediatamente seguido do outro no livro de diretrizes; e especifica que seja feita uma análise detalhada do contexto local onde é realizada a avaliação, assim como das partes envolvidas e suas condições, demandas e perspectivas, tendo em primeiro lugar feito uma previsão da possibilidade de fatores que possam se confrontar nos contextos limitando a realização da avaliação e seus resultados. (Ghiano; Wehrle; Fiscarelli; Valencia; Checa Gutiérrez; Amariles Erazo; Guzmán; Soberón Álvarez, 2021).

Levando em conta todos esses referenciais, o número de municípios participantes não estando claro, não declarando a história desses consórcios representantes de tais municípios e as informações sobre outros *stakeholders* se mantendo dispersas por todo o documento gerou o resultado de pouco suficiente em relação ao critério de avaliação contextualizada.

Diretriz compreensão detalhada da intervenção

Nesta diretriz há uma interessante recomendação:

A compreensão da intervenção deveria ser matéria do primeiro relatório apresentado e discutido com a equipe responsável da intervenção respondendo dúvidas, resolvendo possíveis desvios de interpretação e complementando pontos relevantes omitidos (Ghiano; Wehrle; Fiscarelli; Valencia; Checa Gutiérrez; Amariles Erazo; Guzmán; Soberón Álvarez, 2021, p. 32).

No entanto, encontra-se no documento meta-avaliado, a seguinte informação na apresentação (como uma síntese do que se encontra no documento):

Horizonte temporal: médio e longo prazo.

Resultado: efeitos significativos em todas as dimensões de análise, tanto qualitativas (avaliação de desenho, avaliação de resultado e avaliação de efeitos), quanto quantitativas (rendimento escolar, gestão educacional ao nível das escolas e gestão escolar ao nível das secretarias).

Conclusões: os resultados deram visibilidade a oportunidades de melhoria e resultados positivos inesperados. Também foram encontrados alguns resultados contraintuitivos, possivelmente decorrentes de imperfeições no processo de coleta e análise de dados (Itaú Social, 2021, p. 4).

Tomando a expressão horizonte temporal médio e longo prazo, contrapõe-se ao exposto no documento: no capítulo Avaliação de Impacto é afirmado que houve uma avaliação qualitativa de curto prazo em 2017 sem detalhamento. No entanto, há informação de que as avaliações foram realizadas entre 2017 e 2018. Retornando ao início do texto, para compreender o programa avaliado, há informação de nota de rodapé que data de 2021, ficando, assim, difícil identificar o período entre as avaliações e a publicação deste relatório, bem como definir o tempo à luz da menção ao horizonte temporal a curto prazo (Itaú Social, 2018).

São frágeis também algumas ideias apresentadas nas conclusões: “esse conjunto de avaliações investigou no curto prazo”; “valendo-se de múltiplos métodos de pesquisa”; “por maior que seja o esforço dos pesquisadores, os métodos utilizados são falíveis”, “devido a essas e outras questões inerentes às pesquisas empíricas é preciso que todo e qualquer resultado seja interpretado com cautela” (Itaú Social, 2018, p. 19-20).

Todos esses trechos retirados do documento apontam confusão entre os conceitos de pesquisa e de avaliação, revelando que não há compreensão detalhada do tipo de intervenção realizada. Não há consistência nem interligação lógico-causal em nenhuma etapa desses capítulos mencionados. Ratifica-se essa percepção com: “Avaliação é o processo de determinação do mérito, importância e valor das coisas, e as avaliações são produtos deste processo” (Scriven, 2018, p. 29). Em nenhum momento da conclusão pode-se perceber declaração sobre o mérito ou valor do Programa Melhoria da Educação.

E à propósito do que a dimensão rigorosidade diz em seu critério compreensão detalhada da intervenção “é necessário diferenciar entre a estratégia da intervenção e a forma que se organiza e são realizadas as operações da intervenção, assim como padrões de relação implícitos examinando sua interligação” (Itaú Social, 2018,

p. 34). Por todos esses motivos, é que se julgou como 'Não atendida' esta diretriz no documento meta-avaliado.

Diretriz perguntas avaliativas relevantes

Quando o relatório procura responder à pergunta avaliativa que era: em que nível os objetivos almejados pelo Programa Melhoria da Educação foram atingidos já no curto prazo, logo após o término das atividades formativas junto aos consórcios COGIVA e CIVAP? Nota-se a tentativa de responder esta pergunta a partir da explicação de três novas dimensões: rendimento escolar; gestão-escolas; gestão-secretarias municipais de educação, onde constam a tabela de indicadores adotados e principais informações associadas a eles; porém, os resultados elencados explicitam que "o programa aumentou a taxa de aprovação das escolas em 0,64 p.p na média CIVAP; e 6,04 p.p, na média COGIVA" (Itaú Social, 2018, p. 8). Quando se analisa o documento mais detalhadamente não há dados numéricos que possibilitem a visualização deste aumento. Isto torna fraca a pergunta e não traz uma resposta completa e crível.

Tomando como referencial as Diretrizes para a América Latina e o Caribe, este julgamento insuficiente como resultado da meta-avaliação torna-se mais significativo a partir da compreensão de que as perguntas avaliativas são coerentes com os propósitos e o contexto da avaliação, bem como sua estratégia e características da intervenção (Ghiano; Wehrle; Fiscarelli; Valencia; Checa Gutiérrez; Amariles Erazo; Guzmán; Soberón Álvarez, 2021, p. 35).

No documento meta-avaliado a quantidade de indicadores é bastante elevada em dimensões bastante distintas. Todas sendo avaliadas ao mesmo tempo. Pode-se notar isto no capítulo de indicadores da Avaliação de Impacto. E essa prática se repete na Avaliação Qualitativa quando se menciona que foi preciso fazer avaliação de desenho, de processo e de efeitos somadas a avaliações de desempenho, de formação continuada, do ensino e aprendizagem e do

projeto político pedagógico e que quando se tenta compreender, na leitura do texto, os indicadores associados a cada uma dessas avaliações mostradas, vê-se que estão de forma não plenamente transparentes (Itaú Social, 2018).

Diretriz abordagens e métodos devidamente explicados e justificados

O relatório meta-avaliado do Itaú Social, por vezes não faz a descrição de todas as abordagens e métodos utilizados. Por este motivo, esta diretriz foi considerada como “Não atendida”. É possível identificar, inclusive, muitas vezes em que o termo avaliação foi empregado, no entanto, não há explicação em algumas etapas sobre a abordagem e método utilizados, vide o capítulo avaliação de impacto:

Por meio de uma avaliação quantitativa realizada em 2017, logo após o término das atividades formativas junto aos consórcios” não explica como foi feita ou o que alcançou. Mas continua. “Para isso, desenhou-se uma avaliação de impacto capaz de mapear os resultados do programa em três dimensões: rendimento escolar: desempenho escolar dos alunos e alunas ao nível das escolas; gestão-escolas: gestão educacional ao nível das escolas; gestão-secretarias municipais de educação: gestão educacional ao nível das secretarias de educação (Itaú Social, 2018, p. 8).

Outro exemplo a partir do relatório em que o termo avaliação é utilizado e não há nenhuma descrição de abordagem: avaliação de desenho, avaliação de processo, avaliação de efeitos:

Avaliação de desenho sistematizou e analisou as estratégias, insumos e recursos disponibilizados pelo programa, bem como o encadeamento esperado de atividades até que os objetivos finais fossem alcançados; avaliação de processo investigou se e como as ações planejadas foram bem executadas, observando diferenças entre formulação e implementação; avaliação de efeitos registrou as mudanças na gestão financeira, administrativa e pedagógica que foram apontadas pelos profissionais envolvidos,

além das principais causas dos sucessos e insucessos verificados no período (Itaú Social, 2018, p. 15).

Talvez a avaliação de desenho tenha sido confundida com avaliação de pontos fortes em que é possível analisar os recursos disponíveis para um programa inclusive tempo, talentos, fundos e espaço (Scriven, 2018); ou mesmo com avaliação baseada em objetivos, cujo conceito diz que “mede o quando o programa atinge com clareza objetivos específicos” (Panton, 1990, *apud* Leite; Ferreira; Freitas, 2023, p. 145).

Há sério equívoco no conceito de avaliação de processo. O conceito de avaliação de processos se concentra inteiramente em:

Variáveis entre a entrada e saída (insumo e resultado), também pode se referir ao componente do processo de uma avaliação completa; a abordagem exclusivamente do processo é ilegítima a não ser que variáveis do processo e variáveis de resultado estejam disponíveis e a relação quantidade e qualidade dos resultados são relativamente fracas porque a razão de ser do programa é o resultado e não o processo (Scriven, 2018, p. 141-142).

Outras vezes o relatório utiliza o termo avaliação, explicando-o e o justificando a partir da perspectiva de pesquisadores e não de avaliadores vide: “uma vez construída a base de dados, os indicadores de impacto (antes de depois do programa) foram analisados por pesquisadores econométristas” (Itaú Social, 2018, p. 12); ou ainda, “por maior que seja o esforço dos pesquisadores em produzir metodologias sólidas e robustas os métodos utilizados são falíveis” (Itaú Social, 2018 p. 20).

As abordagens e metodologias empregadas no relatório do Programa Melhoria da Educação parecem sugerir que a linguagem e estrutura utilizadas no mesmo não garantem uma comunicação efetiva com seus líderes, uma vez que os relatórios avaliativos devem ser inteligíveis para outras pessoas que tenham interesse em reproduzir o método em seus contextos avaliativos.

Considerando o que no relatório é mencionado Francischini (2017), sobre saber mostrar os resultados, é possível ponderar as dificuldades vistas por essa meta-avaliação. Para este autor da área de planejamento estratégico, mostrar os resultados é fundamental; e a utilização de gráficos e diagramas podem auxiliar a comunicação do texto de forma mais fluida e transparente. No que tange a indicadores, por exemplo, estes devem ser simples e objetivos, além, do que a utilização de poucos indicadores a serem monitorados apresentam resultados mais eficientes para a tomada de decisões. Assim sendo, pode-se depreender que nesses aspectos o relatório não apresenta métodos devidamente explicados e justificados.

DIMENSÃO 2 - ÉTICA E PRINCÍPIOS JURÍDICOS

Diretriz Respeito dos direitos das pessoas

O resultado foi considerado insuficiente por alguns motivos relevantes. Não há menção no relatório meta-avaliado que explicita se houve consentimento dos respondentes devidamente registrado. No entanto, como foram coletados dados com base em procedimentos científicos, torna-se crível que, ao menos, a confidencialidade foi respeitada durante a coleta de informações.

Para as Diretrizes de Avaliação para a América Latina e o Caribe:

É necessário informar e explicar às pessoas o porquê e para que as informações estão sendo solicitadas, como isso pode afetá-los e como terão acesso aos resultados da avaliação. O procedimento é desenvolvido através de um protocolo anteriormente estabelecido de consentimento informado, devidamente registrado. É necessário contar com o consentimento explícito das pessoas no caso de desejar mencionar sua identidade ou utilizar sua imagem no resultado da avaliação ou outra forma de comunicação do processo e resultados avaliativos (Ghiano; Wehrle;

Fiscarelli; Valencia; Checa Gutiérrez; Amariles Erazo; Guzmán; Soberón Álvarez, 2021, p. 36).

Como não há no relatório informações ou explicações às pessoas do porquê e para que as informações foram solicitadas aos respondentes do processo avaliativo, considera-se que esta diretriz não foi devidamente atendida pela avaliação.

Diretriz Transparência

Não há critérios bem definidos no texto do relatório, nem diretrizes de avaliação na qual tenham se baseado como, por exemplo, as Diretrizes para a América Latina e o Caribe ou o *Joint Committee*. O que foi bastante destacado foram suas limitações como dito em: “ao longo do tempo não houve modificação da estratégia geral para atingi-lo” (Itaú Social, 2018, p.17); “relataram-se adversidades para implementar Projeto Político Pedagógico”; “a formação sobre o Projeto Político Pedagógico foi percebida como a de maior impacto”; as atividades sobre Formação Continuada parecem ter sido as mais deficitárias” (Itaú Social, 2018, 17-18). Por mais que essas frases revelem as dificuldades vividas, por traz delas há pouca transparência no que diz respeito a como se chegou a esses resultados. Elas são apenas descritivas. Por isso, o julgamento é não atendido nesta diretriz. O único aspecto positivo é que este relatório está disponível publicamente. Mas não é inteligível a toda e qualquer pessoa.

Segundo as Diretrizes em que se baseiam esta meta-avaliação:

A avaliação precisa ser guiada, do início ao fim, pelo princípio de transparência, o que significa destacar e dar conta dos propósitos, critérios, métodos e procedimentos utilizados, sobre seus alcances e limitações. Além disso, os termos de referência e os resultados da avaliação devem estar à disposição das partes envolvidas, a não ser que exista um critério de confidencialidade justificado. Os resultados da avaliação precisam ser documentos disponíveis publicamente, garantindo o princípio de

máxima publicidade (Ghiano; Wehrle; Fiscarelli; Valencia; Checa Gutiérrez; Amariles Erazo; Guzmán; Soberón Álvarez, 2021, p. 36).

Como não houve a apresentação de forma clara dos critérios adotados na avaliação, tanto no texto do relatório quanto nas diretrizes adotadas pela avaliação, considera-se que esta diretriz não foi atendida adequadamente.

Diretriz Ética e Integridade Profissional

Segundo Ghiano, Wehrle, Fiscarelli, Valencia, Checa Gutiérrez, Amariles Erazo, Guzmán e Soberón Álvarez (2021), esta diretriz preconiza que durante o percurso da avaliação se forem descobertas situações que afetam as pessoas envolvidas produzidas pela ação da intervenção ou pela avaliação é necessário que a equipe avaliadora assuma a responsabilidade ética abordando e também resolvendo o problema. Uma vez que o Projeto Político Pedagógico (PPP) é um documento que visa garantir a qualidade e a igualdade na educação, é importante que ele esteja adequadamente pronto.

Portanto, quando é dito que ocorreram adversidades para implementar as ações relacionadas ao PPP das escolas, evidencia-se a fragilidade do processo. É sabido que há dificuldade de monitoramento no apoio às escolas com maior dificuldade para finalizar este documento (PPP). No entanto, não há nenhum indício de que os interventores procuraram sanar essas dificuldades vistas nas escolas durante a execução do programa e nem após o programa. Então, esta diretriz não foi atendida, pois como dito por Scriven, quando há conflitos de interesses há consequências para a credibilidade da avaliação: “nem sempre fatal à objetividade, mas sim à credibilidade, e seus efeitos sobre a validade e profissionalismo” (Scriven, 2018, p.217).

Diretriz autonomia

Conforme Ghiano, Wehrle, Fiscarelli, Valencia, Checa Gutiérrez, Amariles Erazo, Guzmán e Soberón Álvarez (2021), é função da avaliação a independência do financiamento e gestão institucional. Por isso a equipe avaliadora deve realizar a avaliação do programa de forma autônoma das entidades responsáveis.

Embora o relatório não ofereça em seu texto de maneira bem explicada o mérito da intervenção, seu juízo de valor e nem ofereceu recomendações que possam melhorar intervenções desse tipo, pareceu que os avaliadores tinham autonomia para realizar a avaliação. Diz o relatório (Itaú Social, 2021, p. 5) que “todos os programas do Itaú Social são periodicamente monitorados e avaliados qualitativamente e quantitativamente”. Há, portanto, que se declarar o mérito desta diretriz como parcialmente atendida.

Diretriz Legalidade

Este relatório não traz nenhum indício claro desses acordos com os municípios dos consórcios COGIVA e CIVAP. Apenas na apresentação do que é o Programa Melhoria da Educação aponta, que este desde 1999, “é comprometido com o desenvolvimento e com a disseminação de tecnologias educacionais e social voltadas ao aprimoramento da educação pública” (Itaú Social, 2018, p. 5). Portanto, considerou-se que não se aplica o instrumento jurídico firmado entre as partes explicitando acordos sobre os alcances da avaliação, os direitos da propriedade intelectual e material conforme as Diretrizes de Avaliação para a América Latina e o Caribe.

DIMENSÃO 3 - COMPREENSÃO CULTURAL

Segundo Cunha (2022, p. 96), “vivemos em um momento de grandes mudanças e a abordagem do *design thinking*, por ser

multidisciplinar, centrada nas pessoas e construída de forma participativa” é importante nos mais diversos processos, inclusive políticas públicas. É uma ferramenta excelente para resolução de problemas, porque inclui as pessoas e a maneira como elas pensam, sentem e gostariam de ver as coisas acontecerem. Essa abordagem durante a etapa de planejamento evitaria frases dos *stakeholders* dizendo que não se sentem contemplados, quando ouvidos no processo de avaliação: “certa avaliação crítica dos municípios em relação ao programa por esse ser considerado distante da realidade local. Além desta, há na avaliação de resultados a informação de dificuldade entre equipe do programa e gestores da Secretaria de Educação devido à alta rotatividade que ocorre neste tipo de função” (Itaú Social, 2018, p. 17). Isto revela a falta de compreensão cultural que deveria estar prevista no planejamento da intervenção.

O problema visto aqui nos resultados desta diretriz é exatamente o que a dimensão compreensão cultural quer que seja evitada quando no planejamento da intervenção. Neste ponto, cabe observar que esta dimensão considera a complexidade cultural e o contexto em que a avaliação será desenvolvida. Portanto, é altamente necessário o respeito com o alto nível de compreensão cultural do ecossistema da intervenção. Depreende-se, portanto, o não atendimento desta diretriz.

Portanto, os processos da avaliação precisam ser sensíveis aos diferentes contextos histórico-culturais (Ghiano; Wehrle; Fiscarelli; Martinic; Gutiérrez; Erazo; Guzmán; Álvarez, 2021). É por essas informações contidas no relatório, objeto desta meta-avaliação, que se considerou essa dimensão e suas três diretrizes não atendidas. Inclusive, isto é observado no relatório do Itaú Social (2018) quando dizem que há resultados contra intuitivos ou visibilidade a resultados não previstos como: “mudanças no atendimento a alunos com deficiência” (Itaú Social, 2018, p. 19).

DIMENSÃO 4 - RELEVÂNCIA E UTILIDADE

Tomando por referência principal as Diretrizes de Avaliação para a América Latina e o Caribe, este critério participação eficiente preconiza que as perspectivas das diferentes partes envolvidas sejam beneficiárias ou não no programa precisam estar cientes do planejamento, dos espaços e tempos explicitamente estabelecidos e da comunicação dos resultados (Ghiano; Wehrle; Fiscarelli; Martinic; Gutiérrez; Erazo; Guzmán; Álvarez, 2021).

No documento objeto desta meta-avaliação isso não se dá de forma eficiente, pois, no capítulo apresentação está dito que: "o presente documento sintetiza as avaliações de curto prazo da edição 2015 a 2017 do Programa Melhoria da Educação em duas Câmaras técnicas de educação de consórcios, um em São Paulo e outro no Paraíba" (Itaú Social, 2018, p. 5).

Enquanto isto na página anterior está apresentado que os gestores educacionais dos municípios brasileiros são o público do programa e o tempo da avaliação é de médio e longo prazo (o que se espera de uma avaliação de impacto). Esses problemas apontados tornam o relatório insuficiente no critério participação eficiente. Considerando-se que a Dimensão se refere à relevância e à utilidade da avaliação e que esses parâmetros não foram contemplados em seu relatório, considera-se que esta diretriz não foi adequadamente atendida

Diretriz propósitos acordados mutuamente

Para as Diretrizes de Avaliação para a América Latina e o Caribe, este critério significa que os propósitos da avaliação geralmente formulados como objetivos, sejam eles gerais ou específicos, devem ser explícitos e ficar documentados nos termos de referência ou no contrato. A equipe ou os profissionais responsáveis pela avaliação participam e contribuem na revisão e precisão, de tal forma

que a formulação final fique estabelecida de comum acordo com a entidade ou agência demandante da avaliação (Ghiano; Wehrle; Fiscarelli; Martinic; Gutiérrez; Erazo; Guzmán; Álvarez, 2021, p. 44-45).

No objeto da meta-avaliação, no capítulo apresentação constam informações sobre o Programa Melhoria da Educação e seus objetivos: “contribuir com as secretarias municipais de educação para o fortalecimento de sua capacidade em garantir acesso, permanência e aprendizagem com equidade para todos e todas” (Itaú Social, 2018, p. 6). E durante a leitura compreende-se como objetivos específicos (não estão expostos de maneira clara): oferecer educação inclusiva e de qualidade nas redes públicas municipais, ofertar formação em serviço e continuada para gestores educacionais e suas equipes. Compreende-se também que o programa procura se adequar à realidade local levando em conta as experiências dos participantes. Portanto, no critério propósitos acordados mutuamente, as informações foram consideradas suficientes.

Diretriz valores explícitos

No objeto meta-avaliado só há evidência de uma pergunta avaliativa: “em que nível os objetivos almejados pelo Programa Melhoria da Educação foram atingidos no curto prazo em 2017”. Não há nenhuma evidência sobre o mérito da chamada avaliação quantitativa que foi preparada para responder a esta pergunta e nem sobre a chegada a resposta (Itaú Social, 2018, p. 8).

Nos subcapítulos que tratam dos indicadores e metodologia as informações sobre o levantamento de dados não estão claras. A metodologia não esclarece quem eram os grupos de controle e o grupo de tratamento e nem os algoritmos de que dispunham em intersecção. Informa, entretanto, que os dados foram analisados por pesquisadores econométricos. Não há nenhum indício de critérios, valores e pontos de vista formulados nas interpretações. Os resultados elencados no documento informam resultados de curto prazo

com evidências numéricas. Quando comparados os indicadores e a metodologia empregada não explicam os números apresentados nos resultados (Itaú Social, 2018, p. 13).

Um exemplo disto está na dimensão gestão-secretarias que diz: “o programa provocou uma melhora substancial no relacionamento entre municípios. Esse efeito foi de mesma ordem de grandeza nos dois consórcios participantes: 18,90 p.p no CIVAP e 13,92 p.p no COGIVA” (Itaú Social, 2018, p. 13) ao passo que retornando à metodologia não há nada que explique a interpretação de como se chegou a esses números e o valor e mérito destes mesmos números).

No processo de avaliação, observações, depoimentos e dados recolhidos são analisados e interpretados para que se consiga chegar às respostas que darão conta das perguntas da avaliação. Isto é importante para que se diminuam os vieses de limitações que possam afetar os resultados da avaliação. É importante ainda que sejam explicitados critérios e pontos de vista (Ghiano; Wehrle; Fiscarelli; Martinic; Gutiérrez; Erazo; Guzmán; Álvarez, 2021),

Quanto a este critério, a meta-avaliação considerou que o relatório não atende a diretriz. Visto que o relatório do Programa Melhoria da Educação diz que:

A metodologia de avaliação de impacto baseia-se sempre na comparação entre os dados do grupo de tratamento e os dados do controle, juntamente com a aplicação de técnicas estatísticas que buscam eliminar fatores que possam distorcer a percepção do pesquisador sobre os reais efeitos do programa avaliado (Itaú Social, 2018, p. 12).

Observou, também, que não há nenhuma referência a autores de avaliação que tenham abordado o assunto. Relatam somente que foram atendidos alguns requisitos sobre os municípios de controle selecionados: localização geográfica, aspectos socioeconômicos e variáveis relativas à estrutura e à qualidade escolar. Mas não há nenhum dado que evidencie e explique detalhadamente essas informações.

O mérito não está presente na análise da lista de verificação; pois de acordo com Leite, Ferreira, Freitas (2023, p. 51) a “avaliação de impacto preocupa-se essencialmente em inferir a causalidade, ou seja, em identificar e registrar os potenciais impactos de um programa, reconhecendo como eles foram produzidos e buscar uma relação causal entre os impactos identificados e o programa avaliado”. No documento meta-avaliado onde podem ser encontrados os resultados obtidos não há nenhum apontamento sobre essa relação causal.

Diretriz informação relevante, pertinente e oportuna.

Neste critério as Diretrizes de Avaliação para a América Latina e o Caribe esclarecem que a avaliação deve dar respostas às partes envolvidas e que a informação produzida deve ser de utilidade prática e deve estar à disposição de forma correta e oportuna (Ghiano; Wehrle; Fiscarelli; Martinic; Gutiérrez; Erazo; Guzmán; Álvarez, 2021).

Quando se procura evidências deste critério no documento meta-avaliado percebe-se que está disponível para qualquer pessoa no *site* do Itaú Social. No entanto, seu texto não produz informações que sejam de fácil entendimento a outros leitores. Informações que delimitem melhor os municípios beneficiados, sua cultura educacional e seus problemas na área de educação não são explicitamente apresentados. E isto pode deixar outros gestores educacionais sem uma menção clara das características de seu município e dos municípios em questão.

Espera-se que os resultados apresentados numa avaliação estejam bem explicitados para que não fiquem somente voltados para a visão somente de atores como os avaliadores ou os contratantes diretos da avaliação. Isto é o que é possível compreender a partir das Diretrizes de Avaliação para a América Latina e o Caribe, sendo ratificado nestas palavras: “A avaliação na educação deve ser entendida, portanto, como um processo amplo com desdobramentos coletivos e institucionais, além de individuais” (Araújo; Rabelo, 2015, p. 444 *apud* Araújo; Oliveira; dos Santos, 2022, p. 874).

Assim sendo, considerando que o texto do relatório da avaliação não produz informações de fácil compreensão para todos os interessados no processo avaliativo, considera-se que esta diretriz não foi adequadamente contemplada em sua documentação divulgada ao público em geral.

Diretriz resultados úteis

À propósito das Diretrizes de Avaliação para a América Latina e o Caribe, diz-se que a avaliação deve ter resultados úteis que contribuam com a reflexão crítica sobre a problemática do objeto de forma que fique conforme as diretrizes de rigorosidade - e contenha informações que auxiliem pessoas a tomar decisões ou a gerenciar projetos com o conhecimento produzido (Ghiano; Wehrle; Fiscarelli; Martinic; Gutiérrez; Erazo; Guzmán; Álvarez, 2021).

Quando se procura estes aspectos no documento meta-avaliado, encontra-se informações de que o Programa Melhoria da Educação se propõe a tratar de problemas educacionais quanto à execução das políticas de educação inclusiva dos municípios brasileiros, a garantia da equidade em educação e da matrícula e da manutenção destes alunos nas salas de aula. Esses problemas são inerentes a diversos municípios do Brasil e não apenas nos municípios envolvidos no Programa Melhoria da Educação no período 2015-2017.

No entanto, seu relatório avaliativo não promove clareza na relação entre os resultados apontados e nos indicadores desses problemas vistos nos municípios envolvidos. Isso compromete a comunicação de aprendizado sobre o tema com gestores de outros municípios. Em alguns resultados mencionados não há dados que esclareçam como se chegou a esse resultado. Por esse, motivo o relatório é considerado insuficiente, a saber: "a avaliação deu ainda visibilidade a resultados não previstos diretamente pela equipe de implementação; é o caso de muitos municípios terem relatado a promoção de mudanças no atendimento ao aluno com deficiência e/ou

na legislação acerca do tema, além do lançamento de um guia que compila documentos e fluxogramas para servir de suporte aos municípios que querem melhorar seus processos de Educação Especial" (Itaú Social, 2018, p. 20).

Diretriz Interesse pelas consequências e incidências

Para as Diretrizes de Avaliação da América Latina e o Caribe o critério interesse pelas consequências e incidência é um modo de contribuir para o uso responsável e sustentável dos resultados ao fazer recomendações práticas e realistas. Assim deve servir à avaliação (Ghiano; Wehrle; Fiscarelli; Martinic; Gutiérrez; Erazo; Guzmán; Álvarez, 2021).

Ao procurar apontamentos dessa diretriz no objeto meta-avaliado não é possível encontrar informações suficientes. Isto se evidencia no capítulo considerações finais onde está escrito: "é preciso que todo e qualquer resultado seja interpretado com cautela mesmo com resultados positivos, como apresentado durante essa avaliação" (Itaú Social, 2018, p. 20). Este é o único parágrafo parecido com uma recomendação e apesar de usar a palavra positivo como um valor atribuído a resultados não especifica claramente o que o relatório considerou como positivo. Deixa este mérito a ser julgado pelo leitor.

Por mais que uma Avaliação Qualitativa possa ter um caráter subjetivo este aspecto está na perspectiva do avaliador e do objeto avaliado e não de quem lê um relatório avaliativo e que poderia utilizá-lo como referencial para novas intervenções em municípios diferentes, por exemplo (Leite; Ferreira; Freitas, 2023, p. 120).

Assim, esta diretriz não foi contemplada no texto do relatório da avaliação.

RESULTADOS FINAIS DA META-AVALIAÇÃO

Os resultados finais deste estudo meta-avaliativo encontram-se expostos no Quadro 4. Pode-se observar que dos 26 pontos verificados no estudo somente um foi “Totalmente Atendido” (TA). Deles, 24 foram classificados como “Não Atendidos” (NA) e um ponto de verificação foi considerado como “Não Aplicável” ao estudo.

Quadro 4 – Atendimento às diretrizes para avaliação para América Latina e o Caribe em relação ao relatório meta-avaliado

Lista de Verificação	TA	PA	NA	NA
Informou as condições de avaliabilidade?			X	
Fez uma avaliação contextualizada?			X	
Mostrou compreensão detalhada da intervenção?			X	
Elaborou perguntas avaliativas relevantes e consensualizadas?			X	
Enunciam abordagens e métodos devidamente explicados e justificados?			X	
Informou as condições de avaliabilidade?			X	
Fez uma avaliação contextualizada?			X	
Mostrou compreensão detalhada da intervenção?			X	
Elaborou perguntas avaliativas relevantes e consensualizadas?			X	
Enunciam abordagens e métodos devidamente explicados e justificados?			X	
Pontuou o respeito dos direitos das pessoas?			X	
Contemplou transparência?			X	
Contemplou a ética e integridade profissional?			X	
Pontuou a legalidade?				X

Lista de Verificação	TA	PA	NA	NA
Apontou autonomia?		X		
Pontuou direitos culturais?			X	
Contemplou igualdade e equidade?			X	
Documentou reciprocidade e interculturalidade?			X	
Pontuou a relevância e utilidade?			X	
Contemplou uma participação efetiva e consciente?			X	
Documentou propósitos acordados mutuamente?			X	
Pontuou valores explícitos?			X	
Permitiu acesso a informação relevante, pertinente e oportuna?			X	
Apontou resultados úteis?			X	
Difundiu comunicação e relatórios pontuais e apropriados?			X	
Apontou interesse pelas coincidências e incidências?			X	

Fonte: Os autores (2023).

CONSIDERAÇÕES E RECOMENDAÇÕES FINAIS

Sabedores de que a meta-avaliação é uma avaliação da avaliação, este estudo meta-avaliativo procurou contribuir positivamente para o aperfeiçoamento de programas semelhantes porque traz consigo fundamentos da disciplina avaliação. Somando ao fato de que há diretrizes críveis que podem nortear as avaliações de programas e que avaliar é produzir mérito e precisa ser realizada por profissionais capacitados para isto.

É por isso que a pergunta avaliativa proposta: “até que ponto a Avaliação Qualitativa e de Impacto do Programa Melhoria da Educação atende às Diretrizes da Avaliação para a América Latina e Caribe em suas quatro dimensões?” traz consigo a resposta e o mérito de que a Avaliação Qualitativa e de Impacto do Programa Melhoria da Educação não atende a essas diretrizes de avaliação.

Dentre as quatro dimensões utilizadas (com vinte diretrizes no total), uma dimensão chama a atenção para qualquer programa avaliativo realizado no Brasil: compreensão cultural. Isso porque o Brasil é um país de dimensões continentais e sua diversidade cultural e social é tão grande quanto seu tamanho. Além disso, a promoção da igualdade e da equidade trata-se de ponto crítico para a convivência social positiva e para a promoção de políticas públicas de relevância e de respeito aos direitos fundamentais e precisa ser ponto importantíssimo de uma avaliação, a fim de que não invalide este tipo de estudo.

Porque procurar compreender a população, os atores sociais que estão envolvidos no ambiente avaliado é atender necessidades e problemas reais destes atores. Estas diferenças sociais, culturais, de grupos, de problemas é o que melhor justifica a escolha dessas Diretrizes de Avaliação para a América Latina e o Caribe como fonte para o quadro de diretrizes e critérios de uma meta-avaliação técnica e justa.

Recomenda-se que os próximos relatórios avaliativos do Programa Melhoria da Educação procurem levar em franca consideração os seguintes aspectos, a saber.

- a. A avaliação tem caráter inter e transdisciplinar. Portanto, uma equipe formada por avaliadores capacitados pode render melhores produtos avaliativos;
- b. Em relação ao período do Programa Melhoria da Educação foi realizado entre 2015 a 2017 sugere-se considerar cada

ano e realizar uma avaliação iluminativa, cujo objetivo é coletar informações detalhadas do programa com uma tendência holística;

- c. Considerar diminuir o número de indicadores a serem mensurados, o período a ser avaliado e documentá-los como fonte de informação. Este procedimento poderá trazer informações mais críveis e tentar resolver o problema da rotatividade na gestão das Secretarias Municipais de Educação, que foi considerado uma grande dificuldade no relatório;
- d. Levar em conta a avaliação formativa durante o período da intervenção e documentá-la como fonte de dados para abordagens futuras de longo prazo focadas no impacto de programas educacionais;
- e. No relatório consta a Avaliação de Impacto, seria valioso utilizar-se também da abordagem Análise de Contribuição, cujo objetivo é avaliar a influência nos resultados dando-lhe a relação de causa e efeito;
- f. Considerar, também, a Colheita de Resultados como uma abordagem avaliativa que auxilie o horizonte temporal de programas como este, pois essa abordagem procura coletar evidências de mudanças e voltar atrás, determinado se a intervenção contribuiu para a mudança.

Pelo exposto, o estudo meta-avaliativo encerra-se procurando servir de apoio ao planejamento, desenvolvimento e fonte de aprendizagem para novos processos avaliativos, em especial de programas - semelhantes ou não à área educacional - como o Programa Melhoria da Educação do Itaú Social, visando ao aprimoramento do estudo da avaliação como ciência. E que novos processos avaliativos de qualidade aconteçam para a promoção da igualdade e equidade em respeito ao direito das pessoas.

MÉRITO DO ESTUDO

Lígia Silva Leite

A prática da avaliação é uma atividade presente no cotidiano de todos nós, seja na esfera pessoal e/ou profissional. Geralmente a avaliação é realizada de maneira informal e assistemática, quando no dia a dia avaliamos a temperatura do ambiente, o nosso estado de ânimo, o sabor da nossa comida, a reação das pessoas etc.

Porém, uma avaliação orientada por uma perspectiva profissional precisa atender a determinados princípios divulgados pela literatura da área da avaliação. Precisa ser sistemática e fundamentada em critérios reconhecidos da área avaliada para que os julgamentos de valor elaborados a partir da avaliação tenham mérito (Whorten; Sanders; Fritzpatrick, 2004).

Dada a importância da avaliação para fundamentar os processos de decisão individuais, pessoais e profissionais, a literatura recomenda a realização de meta avaliações, que são definidas por Elliot (2011) como:

O processo de delinear, obter e aplicar informação descritiva e de julgamento – sobre a utilidade, a viabilidade, adequação e precisão de uma avaliação e sua natureza sistemática, competente conduta, integridade/honestidade, respeitabilidade e responsabilidade social – para orientar a avaliação e divulgar publicamente seus pontos fortes e fracos.

Com esta intenção os alunos do Curso de Mestrado Profissional em Avaliação da Faculdade Cesgranrio desenvolveram um trabalho avaliativo cujo objeto de análise foi o Relatório de Avaliação Qualitativa e de Impacto do Programa Melhoria da Educação, elaborado pelo Itaú Social em relação ao período de 2015 a 2017 (Itaú Social, 2018). Para fundamentar este processo de julgamento o grupo decidiu utilizar como critério as Diretrizes da Avaliação

para a América Latina e Caribe (Ghiano; Wehrle; Fiscarelli; Martinic; Gutiérrez; Erazo; Guzmán; Álvarez, 2021). Este documento é reconhecido na área da avaliação como referência para aferir a qualidade de programas e projetos desenvolvidos na América Latina e no Caribe.

O projeto analisado possui relevância temática por relatar o desenvolvimento de um processo avaliativo realizado por profissionais, cujo objeto de avaliação pertence à área das Políticas Públicas, campo importante para o desenvolvimento social uma vez que:

São consideradas ações, metas e planos adotados pelos governos (federal, estadual e municipal), com a participação da sociedade civil, com o objetivo de garantir os direitos e alcançar o bem-estar da sociedade (Mandu Inovação Social, 2021).

De acordo com esta perspectiva, a busca da qualidade em todas as ações, metas e planos adotados pelo governo justifica a realização de avaliações criteriosas das mesmas e, ainda mais, a realização de meta avaliações, de modo a garantir a qualidade dessas iniciativas governamentais.

O processo avaliativo descrito no Relatório de Avaliação Qualitativa e de Impacto do Programa Melhoria da Educação, elaborado pelo Itaú Social (2015 - 2017) foi desenvolvido por profissionais do Instituto Datafolha, FRONTERES Pesquisa e Formação em Gestão e Políticas Públicas e a Área de Avaliação de Impacto Itaú Unibanco.

O estudo meta-avaliativo identificou que dentre os 26 indicadores propostos pelas Diretrizes da Avaliação para a América Latina e Caribe (Ghiano; Wehrle; Fiscarelli; Martinic; Gutiérrez; Erazo; Guzmán; Álvarez, 2021), apenas um indicador "Apontou autonomia" foi considerado como Plenamente Atendido. O indicador "Pontuou a legalidade?" foi considerado Não Aplicável e os demais 24 foram registrados como Não Atendidos.

Este resultado conduz a algumas reflexões importantes sob o ponto de vista da meta avaliação, são elas: as Diretrizes da Avaliação para a América Latina e Caribe (Ghiano; Wehrle; Fiscarelli; Martinic; Gutiérrez; Erazo; Guzmán; Álvarez, 2021) utilizadas como critério de medida e julgamento do objeto avaliado podem não estar adequadas às políticas públicas que orientaram a elaboração do objeto meta avaliado, daí ter sido identificado apenas um indicador como Plenamente atendido, dentre os 26 considerados; os meta avaliadores podem ter tido dificuldade de identificar no Relatório analisado informações relevantes que fundamentassem o registro de evidência de cada um dos indicadores analisados; o processo avaliativo meta avaliado pode, na verdade, ter sido realizado de modo a não atender aos critérios de qualidade sugeridos pelas Diretrizes, documento reconhecido como parâmetro criterioso e de qualidade para a realidade da América Latina e Caribe.

Qualquer que seja a razão para os resultados encontrados pelos meta-avaliadores este processo de avaliação revela a necessidade e importância de que seja buscada a qualidade dos processos avaliativos realizados, não só na esfera das Políticas Públicas, mas também nas demais áreas de atividade. Revela também a necessidade de que sejam desenvolvidos instrumentos criteriosos que sirvam de base para avaliações sistemáticas que possam gerar resultados meta avaliativos confiáveis e que conduzam à melhoria da qualidade dos projetos, produtos e políticas desenvolvidas, bem como das avaliações e meta avaliações realizadas.

REFERÊNCIAS

BILELLA, Pablo D. Rodríguez; VALENCIA, Sergio Martinic; ALVAREZ, Luis Soberón; KILER, Sarah D.; HERNÁNDEZ, Ana L. Guzmán; TAPELLA, Esteban. **Diretrizes para Avaliação para a América Latina e o Caribe**. 1. ed. Ciudad Autónoma de Buenos Aires, 2016.

CUNHA, Robson Moreira, CAVALLLO, Cristina, Sandra R. H. **Criatividade e atitude empreendedora Mariano**. Niterói, RJ: Universidade Federal Fluminense, 2022. Disponível em: <https://play.google.com/books/reader?id=4qSzEAAAQBAJ&pg=GBS.PA1&hl=pt>. Acesso em: 27 jul. 2023.

DAVIDSON, E. Jane. **Evaluation methodology basics**: the nuts and bolts of sound evaluation. Thousands Oaks, CA: Sage, 2005.

DEMO, Pedro. Avaliação Qualitativa um ensaio introdutório. **Educação e Seleção**, São Paulo, n. 14, p. 5-16, 2013. Disponível em: <https://publicacoes.fcc.org.br/edusel/article/view/2606>. Acesso em: 19 jul. 2023.

ELLIOT, Ligia Gomes (org.) **Instrumentos de avaliação e pesquisa**: caminhos para construção e validação. Rio de Janeiro: Wak Editora, 2012.

ELLIOT, Ligia Gomes; ORLANDO FILHO, Ovídio (org.). **Meta avaliações de políticas públicas do governo**: o que revelam os relatórios de avaliação. São Paulo: Pimenta Cultural, 2019.

ELLIOT, Ligia Gomes. Meta-avaliação: das abordagens às possibilidades de aplicação. **Ensaio**: aval. pol. públ. educ., Rio de Janeiro, v. 19, n. 73, p. 941-964, 2011.

FRANCISCHINI, Paulino G. **Indicadores de desempenho dos objetivos à ação**: métodos para elaborar. Rio de Janeiro. Alta Books, 2017. 448 p.

GHIANO, Maria Celeste; WEHRLE, Andrea María; FISCARELLI, Andrea Peroni; VALENCIA, Sergio Martinic; CHECA GUTIÉRREZ, Marcia Itzel; AMARILES ERAZO, Fabiola; GUZMÁN, Ana Luiza; SOBERÓN ÁLVAREZ, Luis. **Diretrizes de avaliação para a América Latina e o Caribe**. 2. ed. [S. l.]: Relac, 2021. Disponível em: <https://www.deval.org/en/publications/diretrizes-de-avaliacao-para-a-america-latina-e-o-caribe-edicao-2021>. Acesso em: 2 jun. 2023.

ITAU SOCIAL. **Avaliação qualitativa e de impacto**: programa melhoria da educação. Câmaras Técnicas de Consórcio. São Paulo, 2017.

LEITE, Lígia Silva; FERREIRA, Sandra M. Martins Redoválío; FREITAS, Sonia Regina **Natal de abordagens avaliativas**: alternativas para o avaliador. Rio de Janeiro, 2023.

MANDU INOVAÇÃO SOCIAL. **Para que servem as políticas públicas?** São Paulo, 2021. Disponível em: <https://mandusocial.org/blog/sem-categoria/para-que-servem-as-politicas-publicas/>. Acesso em: 17 dez. 2024.

ORLANDO FILHO, Ovidio, OLIVEIRA, Regina Adelaide Carvalhal de. O relatório de avaliação do programa apoio à alimentação escolar na educação básica: um estudo meta-avaliativo. *In: Meta avaliações de políticas públicas do governo: o que revelam os relatórios de avaliação*. ELLIOT, Ligia Gomes, ORLANDO FILHO, Ovidio (Org.). São Paulo: Pimenta Cultural, 2019.

PEREIRA, Daniel. **Mapa de empatia**: o que é e como fazem um em 6 passos. São Paulo, 2021. Disponível em: <https://analistamodelosdenegocios.com.br/mapa-de-empatia-o-que-e>. Acesso em: 17 jul. 2023.

SCRIVEN, Michael. **Avaliação**: um guia de conceitos. Rio de Janeiro: Paz e Terra, 2018.

SCRIVEN, Michael. The methodology of evaluation. *In: TYLER, Ralph Winfred; GAGNÉ, Robert Mills; SCRIVEN, Michael. (ed.). Perspectives of curriculum evaluation*. Chicago: Rand McNally, 1967. p. 39-83.

SOUZA, Celina. Políticas públicas: uma revisão da literatura. **Sociologias**, Porto Alegre, v. 8, n. 16, p. 20-45, 2006. DOI: <https://doi.org/10.1590/S1517-45222006000200003>. Disponível em: <https://www.scielo.br/j/soc/a/6YsWyBWZSdFgSqDVQhc4jm/abstract/?lang=pt>. Acesso em: 29 jun. 2023.

STUFFLEBEAM, Daniel L. The metaevaluation imperative. **American Journal of Evaluation**, Thousand Oaks, v. 22, n. 2, p. 183-209, 2001. DOI: <https://doi.org/10.1177/109821400102200204>. Disponível em: <https://journals.sagepub.com/doi/10.1177/109821400102200204>. Acesso em: 29 jun. 2023.

WORTHEN, Blaine R.; SANDERS, James R.; FITZPATRICK, Jody L. **Avaliação de programas: concepções e práticas**. São Paulo: Editora Gente, 2004.

5

*Hélio Huguenin Tavares Júnior
Virginia Virgínio de Lima
Glauco da Silva Aguiar*

ESTUDO META-AVALIATIVO DA AVALIAÇÃO DE POSSÍVEIS IMPACTOS:

PROFMAT

POLÍTICAS PÚBLICAS E AVALIAÇÃO: CONCEITOS E RELEVÂNCIAS

A partir da primeira metade do século XX, o estudo das políticas públicas emergiu como um campo acadêmico fundamental, consolidando-se como uma área essencial para garantir que as decisões governamentais fossem baseadas em evidências e práticas eficazes. Esse movimento visa maximizar os benefícios sociais e minimizar desperdícios e ineficiências na gestão pública. (Mastrodi; Ifanger, 2019).

De acordo com esses autores, a expressão “políticas públicas” tem sido amplamente utilizada para designar ações estatais eficazes na promoção e garantia de direitos, sendo frequentemente percebida como uma solução para diversas demandas sociais. Esse entendimento está alinhado à Constituição Federal de 1988, que estabelece o dever do Estado de assegurar o acesso igualitário da população aos direitos fundamentais (Brasil, 1988).

No entanto, Souza (2006) destaca que não há uma definição única ou definitiva para o conceito de política pública, refletindo a complexidade e a diversidade de abordagens sobre o tema. Enquanto alguns estudiosos a consideram um campo autônomo, Mead (1995) a classifica como uma vertente do estudo da política, voltada para a análise das decisões governamentais em resposta a desafios sociais, econômicos e políticos. Essa perspectiva enfatiza a necessidade de um exame crítico das ações governamentais e de seu alinhamento com as expectativas da sociedade.

A definição mais amplamente reconhecida do que vem a ser política pública é a de Harold D. Lasswell, que propôs a seguinte questão central para a análise da área: “Quem ganha o quê, por quê e que diferença faz?” (Souza, 2006). Destaca-se que Lasswell (1935-1938), foi o pioneiro na análise de políticas públicas, desenvolvendo

uma abordagem que continua a ser frequentemente citada e debatida. Sua definição destaca a natureza prática e o impacto das políticas públicas, sublinhando a importância de entender quem se beneficia das decisões políticas, quais são os motivos por trás dessas decisões e quais são as suas consequências para a sociedade.

Nas últimas décadas, cresceu a preocupação com a avaliação das ações governamentais, destacando a necessidade de exigir transparência e responsabilidade na aplicação dos recursos públicos voltados para a implementação de políticas públicas (Alfradique; Tavares; Orlando Filho, 2020). Desta forma, em cenários de crescente complexidade e demanda por serviços públicos, a capacidade de avaliar e aprimorar continuamente as políticas torna-se essencial para assegurar que essas realmente cumpram seus objetivos e estejam alinhadas com as necessidades e expectativas da população.

Por ser uma área que abrange múltiplas disciplinas e se inter-relaciona com várias outras, a Avaliação enfrenta dificuldades na definição de seu campo de atuação e na especificação de seu conteúdo. Worthen, Sanders e Fitzpatrick (2004, p. 78) consideram a avaliação como: “Empreitada técnica e política multidimensional que requer tanto novas conceituações, quanto novos olhares de quando e como as metodologias existentes em outros campos podem ser usadas com propriedade”.

Questões relacionadas aos processos de melhorias em uma política pública podem ser obtidas por meio dos resultados de uma avaliação. De fato, as avaliações têm como objetivo emitir juízo de valor sobre um determinado objeto, produzindo dados, que, em forma de recomendações, possuam potencialidade de gerar melhorias naquilo que está sendo avaliado.

Ainda, sobre o propósito de uma avaliação, Scriven (1967) afirmou que seu único objetivo ou propósito é o de “determinar o valor ou mérito de tudo o que é avaliado”. Posteriormente, ele ampliou essa visão, argumentando que a avaliação consiste no “processo

de determinação do mérito, importância e valor das coisas, sendo as avaliações os produtos desse processo” (Scriven, 2018). Outros autores, como Mark, Henry e Julnes (2000), Patton (2008a), Preskill e Torres (1998) e Worthen, Sanders e Fitzpatrick (2004), reforçam que a avaliação deve ter como objetivo central a melhoria organizacional e dos programas avaliados.

Em relação à implementação e condução, Stufflebeam (2002) alerta que toda avaliação pode estar sujeita a fragilidades e variações de qualidade, podendo ser classificada como boa, ruim ou razoável. Ele destaca que avaliações cujos resultados levem a decisões prejudiciais podem comprometer a credibilidade do avaliador. Para evitar esse problema, surge a necessidade de avaliar a própria avaliação – um processo denominado meta-avaliação, definido por Scriven (1991) como a avaliação de uma avaliação.

Stufflebeam (2001) descreve a meta-avaliação como o processo de delinear, obter e aplicar informação descritiva e de julgamento, para orientar a avaliação e tornar público, seus pontos fortes e fracos. Nesse sentido, Schwandt (1989) acrescenta que quando o objeto de estudo de uma avaliação é outra avaliação, este deve ter o propósito de certificar a qualidade do processo avaliativo inicialmente implementado, possibilitando a formulação de juízo de valor, que venha agregar valor e contribuir para o aperfeiçoamento da avaliação original.

Outros estudiosos, como Elliot (2011, p. 960), afirmam que “a necessidade de meta-avaliar uma avaliação, seu relatório ou o próprio processo avaliativo é a forma de atestar-lhes a qualidade ou de evidenciar pontos que ainda precisam ser burilados”. Afirmam, ainda, que a meta-avaliação verifica a qualidade da própria avaliação e para isso, utiliza-se de critérios pré-estabelecidos que trazem luz à forma como o estudo foi conduzido. Berends e Roberts (2003) entendem que existem duas razões para se meta-avaliar: a primeira é possibilitar a verificação do processo de condução da avaliação, ou seja, se

durante todo o estudo, o avaliador esteve pautado na ética e suas boas práticas; e, a segunda, oportuniza a incorporação de novos conhecimentos aos avaliadores e meta-avaliadores, advindos de uma prática reflexiva sobre os procedimentos utilizados na avaliação.

Cabe, portanto, à meta-avaliação a função de atestar se o processo avaliativo foi conduzido adequadamente, utilizando-se de critérios de qualidade, reconhecidamente aceitos, determinados pelos meta-avaliadores.

Diversos autores desenvolveram metodologias para a realização de meta-avaliações. Entre as abordagens mais relevantes, destacam-se: a) Lista Chave de Verificação da Avaliação (Scriven, 2007) – Instrumento que orienta os avaliadores na condução do processo avaliativo, auxiliando na validação de suas conclusões; b) Padrões de Avaliação (*Joint Committee on Standards for Educational Evaluations*, 2011) – Conjunto de diretrizes tanto para a elaboração de avaliações quanto para a realização de meta-avaliações, servindo como referência para garantir a qualidade do estudo; c) Critérios Fundamentais de Davidson (2005) – Utilizados neste estudo, esses critérios abrangem cinco dimensões essenciais para aferir a qualidade de uma avaliação: validade, utilidade, conduta, credibilidade e custos.

Essas abordagens fornecem diretrizes fundamentais para garantir que a meta-avaliação seja conduzida de forma rigorosa e transparente, reforçando sua importância na qualificação das políticas públicas e na melhoria contínua dos processos avaliativos (Elliot, 2011).

CONHECENDO O PROFMAT: MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL

O Mestrado Profissional em Matemática em Rede Nacional (PROFMAT) é um programa de mestrado semipresencial, de abrangência nacional, oferecido por uma rede de Instituições de Ensino Superior no âmbito da Universidade Aberta do Brasil (UAB), sob a coordenação da Sociedade Brasileira de Matemática (SBM) e com o apoio do Instituto Nacional de Matemática Pura e Aplicada (IMPA). O programa conta com financiamento e supervisão da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), garantindo sua qualidade acadêmica e institucional (PROFMAT, 2021).

O PROFMAT foi concebido como uma iniciativa da CAPES, em parceria com a comunidade científica da área de Matemática, e recebeu reconhecimento do Conselho Nacional de Educação (CNE). Seu principal público-alvo são professores de Matemática em exercício na Educação Básica, com prioridade para aqueles que atuam na rede pública de ensino. O programa tem como objetivos: a) Aperfeiçoar a formação profissional dos professores de Matemática da Educação Básica; b) Promover o aprofundamento do conhecimento matemático, com foco em sua aplicação no ensino; c) Elevar a qualidade do ensino de Matemática nas escolas brasileiras, especialmente na rede pública; d) Oferecer capacitação e qualificação continuada, alinhada com as demandas do sistema educacional (PROFMAT, 2021).

O PROFMAT é estruturado para oferecer uma formação robusta em Matemática e Didática da Matemática, combinando teoria e prática. O curso é dividido em disciplinas obrigatórias e optativas, abordando tópicos avançados de Matemática, como: Álgebra, Geometria, Análise Real e Complexa, Probabilidade e Estatística, Matemática Aplicada ao Ensino, Educação Matemática

e Tecnologias Digitais. Além disso, são oferecidos seminários e atividades complementares, que estimulam a troca de experiências entre docentes e pesquisadores. Como trabalho final é exigido a elaboração de uma dissertação, na qual o professor-pesquisador desenvolve um estudo aplicado, voltado para a melhoria do ensino de Matemática na Educação Básica.

Apesar de sua importância, o PROFMAT enfrenta desafios, como a dificuldade de conciliação entre trabalho docente e atividades acadêmicas do mestrado; problemas com a infraestrutura e suporte das instituições parceiras, que variam conforme a localidade e dificuldade de adaptação ao modelo semipresencial, que exige maior autonomia dos alunos.

Apesar disso, o programa continua sendo uma referência nacional na formação de professores de Matemática, contribuindo para a valorização da carreira docente e para a melhoria do ensino no Brasil.

De acordo com a página da *web* do PROFMAT (2021), o programa:

Vem ao encontro do Plano Nacional de Educação – PNE, Lei Nº 13.005, de 25 de junho de 2014, que coloca em sua Meta 16: formar em nível de pós-graduação, 50% (cinquenta por cento) dos professores da Educação Básica, até o último ano de vigência deste PNE, e garantir a todos(as) os(as) profissionais da Educação Básica formação continuada em sua área de atuação, considerando as necessidades, demandas e contextualizações dos sistemas de ensino. Além disso, o PROFMAT também atende as metas 14, 17 e 18, que tratam, respectivamente, de elevar o número de matrículas na pós-graduação *stricto sensu*; valorização do professor; e plano de carreira.

Dessa forma, o programa revela uma relevância significativa para a Educação Básica, uma vez que contribui para o cumprimento das metas estabelecidas no Plano Nacional de Educação (PNE).

Além disso, relatórios apontam para possíveis impactos, que se traduzem em consideráveis acréscimos nos resultados de desempenho dos alunos e na atuação de seus egressos, entre outros efeitos positivos na área da Educação Básica.

METODOLOGIA META-AVALIATIVA

OBJETO, OBJETIVO E QUESTÃO DO ESTUDO

O presente estudo tem como objeto de análise a avaliação apresentada no relatório PROFMAT: Avaliação de possíveis impactos, encomendado pela Sociedade Brasileira de Matemática (SBM). O objetivo é realizar uma meta-avaliação do processo avaliativo, utilizando o referido relatório como principal documento de investigação.

O relatório está estruturado em várias seções. Na Apresentação, define-se o objeto de avaliação: o Programa de Mestrado Profissional em Matemática (PROFMAT). Em seguida, a Introdução contextualiza o programa, apresentando dados estatísticos e seu histórico. Após essa base contextual, a seção de Metodologia detalha os procedimentos adotados na condução do estudo.

A Análise de Dados considera diferentes aspectos do PROFMAT, incluindo indicadores sociais dos egressos, o perfil das escolas que participaram do levantamento, os resultados do programa nas instituições onde os egressos lecionam e seu impacto no desempenho dos alunos em competições como as Olimpíadas de Matemática. Além disso, são discutidas outras contribuições do programa para as escolas, bem como iniciativas e situações relevantes envolvendo professores com mestrado pelo PROFMAT.

Dessa forma, este estudo meta-avaliativo busca responder à seguinte questão: até que ponto a avaliação apresentada no relatório PROFMAT: Avaliação de possíveis impactos está alinhada com os Critérios Fundamentais estabelecidos por Davidson (2005), no sentido de se aferir a qualidade de uma avaliação realizada?

A META-AVALIAÇÃO

Para dar início ao estudo, foi realizada uma revisão de literatura abrangente, investigando publicações diversas relacionadas ao tema. Esse processo seguiu uma abordagem qualitativo-exploratória, caracterizada pela análise de textos acadêmicos, artigos e documentos pertinentes, com o objetivo de compreender diferentes perspectivas sobre a meta-avaliação e sobre o PROFMAT, a fim de identificar referenciais teóricos relevantes. Além disso, conduziu-se uma pesquisa documental, centrada no exame detalhado do Relatório que apresentou os resultados da avaliação original, que será meta-avaliada. Esse estudo documental permitiu uma análise crítica dos dados e da metodologia empregada, contribuindo para a construção da base teórica e analítica do presente estudo.

Para pavimentar este estudo meta-avaliativo do relatório PROFMAT: Avaliação de possíveis impactos, foram adotados os Critérios Fundamentais propostos por Davidson (2005). Esses critérios têm como objetivo aferir o mérito e a qualidade intrínseca do objeto meta-avaliado, garantindo que as conclusões sejam válidas, justificáveis, úteis, éticas, confiáveis e de custo satisfatório.

A análise foi conduzida com base em cinco dimensões essenciais: validade, utilidade, conduta, credibilidade e custo. Esses critérios permitem avaliar não apenas a precisão e relevância das conclusões do relatório, mas também sua aplicabilidade, conformidade ética, confiabilidade metodológica e viabilidade financeira, assegurando uma avaliação rigorosa e fundamentada.

O Quadro 1 apresenta a descrição detalhada dos cinco Critérios Fundamentais propostos por Davidson (2005), que servirão como farol norteador para a condução desta meta-avaliação.

Quadro 1 – Descrição dos critérios fundamentais de Davidson (2005)

Critérios	Descrições
Validade	Diz respeito à relação entre as conclusões apresentadas e sua justificativa. As associações feitas pelo avaliador, entre fatos e valores, devem expressar o quanto as conclusões se justificam.
Utilidade	É determinada pela necessidade que determinado grupo de interessados possui em relação aos resultados obtidos.
Conduta	Consiste na observação do cumprimento dos padrões legais, éticos e profissionais, de adequação cultural e de não impedimento foram respeitados pelo avaliador.
Credibilidade	A obtenção de credibilidade para uma avaliação está diretamente relacionada à aceitação, por parte dos interessados, dos resultados e suas fontes, isto é, em que medida merecem crédito.
Custo	Inclui diferentes tipos de custos, que vão desde a relevância das despesas, pagamento dos avaliadores e de outros itens, além da prestação de contas, que fazem parte de uma equação em que se deseja saber se os resultados valem o que custaram.

Fonte: Os autores (2024) Adaptado de Elliot (2011) e Alfradique, Tavares e Orlando Filho (2020).

Cabe ressaltar que o relatório PROFMAT: Avaliação de possíveis impactos foi solicitado pela Sociedade Brasileira de Matemática (SBM) como um complemento ao estudo realizado em 2017, intitulado PROFMAT: Uma reflexão e alguns resultados (2017). Esse documento faz parte do processo contínuo de acompanhamento do Programa de Mestrado Profissional em Matemática em Rede Nacional (PROFMAT), buscando avaliar e aprimorar seus impactos.

Logo, este estudo foi conduzido após a conclusão da avaliação do programa, caracterizando-se, portanto, como uma meta-avaliação somativa. De acordo com Stufflebeam (2001), esse tipo de meta-avaliação tem como principal objetivo evidenciar pontos fortes

e eventuais fragilidades da avaliação original, contribuindo para seu aprimoramento e agregando mérito e valor ao processo avaliativo.

Para melhor orientar o julgamento dos resultados obtidos na avaliação e validá-los com base nos Critérios Fundamentais estabelecidos por Davidson (2005), adotou-se a escala de conceitos proposta pela autora, que varia de A a E (Davidson, 2005, p. 217). A descrição qualitativa de cada conceito está apresentada no Quadro 2, que será aplicado a cada aspecto investigado nesta meta-avaliação, servindo como referência para a análise e interpretação dos resultados.

Quadro 2 – Escala de conceitos de Davidson

Conceito	Descrição
A	Atinge todos os aspectos principais deste ponto e os expressa de forma clara e concisa.
B	Atinge a maioria dos aspectos relativos a este ponto, mas não cobre um ou dois deles, ou trata dos ingredientes principais, mas não é 100% claro.
C	Trata do ponto, mas perde alguns aspectos cruciais, ou perde vários aspectos importantes, ou é um tanto desorganizado ou pouco claro.
D	Tem um ou dois elementos que parecem tratar implicitamente do ponto, mas o tratamento é deficiente.
E	Perde o ponto totalmente.

Fonte: Os autores (2024) adaptado de Davidson (2005, p. 217).

Como mencionado anteriormente, a meta-avaliação tem o propósito de avaliar a qualidade dos resultados obtidos na avaliação original, utilizando critérios sólidos e respaldados na literatura. Além disso, busca responder às expectativas dos *stakeholders*, garantindo que a avaliação atenda aos padrões esperados.

Nesse contexto, a formulação de uma questão avaliativa central é essencial para manter o foco do avaliador e orientar a condução da meta-avaliação. No caso específico deste estudo, pretende-se responder se a avaliação descrita no Relatório PROFMAT: Avaliação

de possíveis impactos está alinhada com os objetivos do programa, considerando os Critérios Fundamentais de qualidade da avaliação estabelecidos por Davidson (2005). Como já afirmado, essa questão serviu como base para as definições metodológicas que estruturaram todo o processo avaliativo.

Os julgamentos apresentados neste estudo meta-avaliativo são de responsabilidade dos autores, que realizaram a análise de forma detalhada e criteriosa do referido Relatório. As informações contidas no documento foram confrontadas com os critérios estabelecidos para a meta-avaliação, permitindo a formulação de juízos de valor.

Cabe ressaltar que, por se tratar de uma meta-avaliação somativa, baseada exclusivamente no Relatório, não foi possível obter informações adicionais diretamente dos atores envolvidos no contexto da avaliação original ou dos indivíduos impactados por seus resultados. Dessa forma, as conclusões apresentadas refletem uma interpretação fundamentada, mas ainda passível de contrapontos e discussões.

CRITÉRIOS DE VALORAÇÃO DOS RESULTADOS

Os critérios estabelecidos por Davidson (2005) e seus respectivos pontos de verificação estão contemplados no Quadro 3, a seguir.

Quadro 3 – Critérios e pontos de verificação

Critérios	Pontos de verificação
Validade	▪ Cobre todas as fontes relevantes de valor; ▪ Cobre compreensivamente processo e resultado; ▪ Não inclui critérios irrelevantes ou ilícitos; ▪ Inclui análises qualitativas e quantitativas; ▪ Estabelece claramente como os dados são tratados; ▪ Explicita com clareza a procedência das conclusões avaliativas; ▪ Inclui recomendações válidas e pertinentes.

Critérios	Pontos de verificação
Utilidade	▪ Relevante para as decisões dos interessados; ▪ Comunica os resultados de forma apropriada.
Conduta	▪ Cumpre as legislações pertinentes (local, nacional e internacional); ▪ Respeito aos padrões éticos;
Credibilidade	▪ Apresenta familiaridade com o contexto e domínio dos conteúdos apreciados; ▪ Demonstra independência e/ou imparcialidade;
Custos	▪ Especifica os custos financeiros e a relação custo-benefício do processo avaliativo;

Fonte: Os autores (2024), adaptado de Davidson (2005).

Os pontos de verificação foram investigados e classificados, segundo sua conformidade, com base na escala de conceitos proposta por Davidson (2005, p. 217), apresentada no Quadro 2, da seguinte forma:

- a. Completamente Atendidos: pontos que receberam os conceitos A ou B.
- b. Parcialmente Atendidos: pontos que receberam o conceito C.
- c. Não Atendidos: pontos que receberam os conceitos D ou E.

Valendo-se dos conceitos atribuídos a cada ponto de verificação, para a valoração dos resultados de cada um dos critérios e, conseqüentemente, do estudo meta-avaliativo, foram estabelecidos os seguintes parâmetros:

- a. Validade: classificada como Completamente Atendida (CA) quando todos os seus pontos de verificação receberam conceitos A ou B; Parcialmente Atendida (PA) quando pelo menos quatro pontos de verificação foram atendidos; e Não Atendida (NA) quando quatro ou mais pontos não foram cumpridos, ou seja, receberam conceitos D ou E.

- b. Utilidade, Conduta e Credibilidade: considerados Completamente Atendidos (CA) quando todos os seus pontos de verificação receberam conceitos A ou B; Parcialmente Atendidos (PA) quando pelo menos um ponto de verificação foi atendido; e Não Atendidos (NA) quando nenhum dos dois pontos de verificação foi atendido (ambos receberam conceitos D ou E).
- c. Custo: avaliado com base em um único ponto de verificação, sendo classificado como Completamente Atendido (CA), Parcialmente Atendido (PA) ou Não Atendido (NA), conforme o julgamento desse critério específico.

APRESENTAÇÃO DOS RESULTADOS E ANÁLISES

Nesta seção, são apresentados os resultados da meta-avaliação com base nos critérios e nos padrões estabelecidos. A avaliação seguiu a escala de conceitos definida por Davidson (2005, p. 217), que varia de A (conceito mais alto) a E (conceito mais baixo), garantindo uma análise estruturada e comparável.

AVALIAÇÃO DO CRITÉRIO VALIDADE

Ao iniciar a leitura do Relatório PROFMAT: Avaliação de possíveis impactos observou-se que o documento não informa o ano de sua realização na seção de Apresentação. São mencionados apenas os nomes das autoras e o contratante do estudo, a Sociedade Brasileira de Matemática (SBM). No entanto, o texto indica que este estudo complementa o relatório PROFMAT: Uma

reflexão e alguns resultados, produzido e apresentado à SBM em 2017, pelas mesmas autoras.

A ausência da data de realização compromete a precisão temporal dos fatos apresentados, o que pode impactar a interpretação e a confiabilidade do estudo. De acordo com o *Joint Committee* (2011), a informação fidedigna é um dos indicadores do Padrão Precisão, essencial para garantir a transparência e a confiabilidade dos procedimentos de seleção, coleta e análise de dados em uma avaliação. Em geral, a fidedignidade de um estudo aumenta à medida que omissões e erros aleatórios são reduzidos.

A seguir, são apresentadas as análises dos pontos de verificação do critério Validade, buscando responder aos questionamentos levantados.

1) A avaliação cobriu todas as fontes relevantes de valor?

Sim, todas as fontes consultadas e citadas no Relatório meta-avaliado demonstraram ser suficientes para responder ao objetivo da avaliação. As análises qualitativas e quantitativas fundamentaram-se em:

- a. Identificação de evidências do impacto real do programa na qualidade do ensino e aprendizagem dos alunos de Matemática por meio do estudo *PROFMAT: Uma reflexão e alguns resultados*;
- b. Levantamento de informações sobre escolas onde os alunos (PROFMAT) lecionaram e que constam no Sistema de Gestão Acadêmica do PROFMAT (SCA);
- c. Dados e informações gerados por meio do levantamento documental, por meio das fichas de inscrições realizadas para ingresso no PROFMAT;
- d. Informações coletadas em reuniões entre as equipes de consultoria e da SBM;

- e. Dados da Olimpíada Brasileira de Matemática das Escolas Públicas (Obmep), fornecidos pelo Instituto de Matemática Pura e Aplicada (Impa), sobre a participação das escolas e *ranking* na olimpíada;
- f. Dados oriundos de questionários, aplicados aos diretores de escola onde os egressos atuam;
- g. Entrevistas com diretores de escolas onde os egressos do Programa lecionam, a fim de coletar informações *in loco*, segundo a ótica do diretor, sobre o impacto do programa no desempenho do professor e alunos de matemática.

A prospecção dos dados ocorreu de forma a cobrir todas as fontes relevantes de valor, a fim de conseguir respostas e identificação de elementos para compreensão de cada situação proposta, obtendo de acordo com Davidson (2005) conceito A.

II) A avaliação cobre compreensivamente processo e resultado?

Nesse quesito, o estudo gera incertezas quanto à precisão das informações coletadas. Embora aborde o tema proposto, apresenta fragilidades significativas, especialmente em relação à organização e clareza dos dados.

Uma das principais lacunas identificadas foi a ausência de informações sobre a data de realização das entrevistas, observações e recebimento dos questionários, o que comprometeu a análise temporal dos resultados. Além disso, há inconsistências na apresentação dos dados, como a relação entre o tamanho da amostra (2.974) e o número total de egressos do PROFMAT (2.588), gerando dúvidas sobre o número correto de participantes.

Da mesma forma, a quantidade de escolas consideradas no estudo (1.645) foi identificada a partir dos códigos do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

(INEP). No entanto, o próprio relatório (p. 10) destaca um problema metodológico relevante: a duplicação de escolas devido a variações na grafia dos nomes. Conforme mencionado no documento, “a mesma escola possuía diferentes grafias e, portanto, na primeira etapa foi contabilizada mais de uma vez”, dificultando a identificação correta das instituições pelo código do INEP.

A coleta de dados apresentou limitações quanto à abrangência e taxa de resposta dos participantes. Das 1.645 escolas identificadas pelo código do INEP, foram obtidos 1.396 endereços de e-mail de instituições onde os egressos do PROFMAT atuavam. O questionário, elaborado no Google Formulários, foi enviado por e-mail e continha tanto questões objetivas quanto abertas, com o objetivo de avaliar a satisfação dos egressos e diretores de escolas em relação aos resultados do programa, além de coletar sugestões para seu aprimoramento.

Dos 1.396 e-mails enviados, 45 não foram entregues, reduzindo o número efetivo de destinatários para 1.351 escolas. No entanto, apenas 184 escolas responderam ao questionário, correspondendo a uma taxa de resposta de 14% ao longo de 40 dias.

A baixa adesão foi atribuída a diversos fatores, incluindo endereços de e-mail inválidos, caixas postais cheias, egressos que não lecionavam mais na escola, diretores de férias, entre outros. No entanto, o estudo não detalha todas as razões específicas que podem ter influenciado essa baixa taxa de participação.

A baixa adesão ao estudo justifica o questionamento quanto à qualidade dos resultados aferidos pela avaliação. Portanto, de acordo com Davidson (2005), este ponto de verificação obteve conceito C.

III) Houve a preocupação de não incluir critérios irrelevantes ou ilícitos?

Quanto à relevância e licitude dos dados utilizados na avaliação, pode-se considerar que o relatório é uma fonte legítima e

relevante de informações, contribuindo para a prestação de contas à sociedade e aos órgãos públicos interessados no PROFMAT.

O documento foi elaborado com base em dados cadastrais oficiais do próprio PROFMAT, do INEP e do Instituto de Matemática Pura e Aplicada (IMPA), garantindo a procedência e confiabilidade das informações utilizadas. Dessa forma, não foram identificados critérios que pudessem ser considerados irrelevantes ou ilícitos na avaliação.

Conforme os parâmetros estabelecidos por Davidson (2005), o relatório atende plenamente a esse ponto de verificação, sendo classificado com o conceito A.

IV) Foram incluídas análises qualitativas e quantitativas que são apropriadas aos dados coletados?

Os dados analisados incluem informações do estudo *PROFMAT: Uma reflexão e alguns resultados* (SBM, 2017), cruzamento de registros do **Sistema de Gestão Acadêmica do PROFMAT** e do **Instituto de Matemática Pura e Aplicada (IMPA)**, além das respostas obtidas por meio de questionários enviados via e-mail aos **diretores de escolas onde egressos do programa lecionam**.

Entretanto, o relatório **não especifica como foram realizadas as análises qualitativas**, nem detalha de forma objetiva os procedimentos adotados para o **tratamento e categorização dos dados**. As informações coletadas foram apresentadas predominantemente em **gráficos**, organizados conforme os indicadores definidos para a avaliação do estudo *PROFMAT: Avaliação de possíveis impactos*.

Diante da **falta de clareza sobre a metodologia de tratamento e classificação dos dados**, conclui-se que esse ponto de verificação atende à maioria dos requisitos relacionados aos dados coletados, mas **não explicita os critérios de análise adotados**.

Dessa forma, de acordo com **Davidson (2005)**, esse ponto de verificação foi classificado com conceito **B**.

V) A avaliação estabeleceu com clareza como os dados coletados no processo são interpretados?

Esse indicador refere-se à forma de apresentação dos dados coletados no estudo. O Capítulo 4 (p. 11), que trata da análise dos dados do programa, expõe os resultados por meio de gráficos inteligíveis, iniciando com a abrangência das escolas que responderam ao formulário de pesquisa.

Além disso, o estudo apresenta alguns indicadores sociais dos egressos, como gênero, faixa etária e local de residência. Em relação às escolas participantes, o relatório as identifica com base no perfil demográfico das cidades em que estão localizadas.

Em relação aos achados qualitativos do estudo, o relatório apresenta algumas conclusões sobre o ambiente escolar no período pós-PROFMAT, bem como sobre o impacto do programa no comportamento dos professores em sala de aula. Além disso, são discutidos os efeitos do PROFMAT no incentivo à participação dos alunos na Olimpíada Brasileira de Matemática das Escolas Públicas (OBMEP) e os resultados obtidos pelos estudantes que participaram da competição.

No entanto, foi identificada uma inconsistência nos dados apresentados. O gráfico referente à Figura 10, que deveria ilustrar o aumento do número de alunos premiados na OBMEP (p. 20), apresenta valores divergentes em relação aos descritos no corpo do texto. Além disso, o gráfico é idêntico ao da Figura 9 (p. 19), sugerindo um possível erro na apresentação dos resultados.

Por fim, o relatório também destaca as contribuições do PROFMAT para outros indicadores escolares, além de iniciativas

desenvolvidas por professores egressos do programa, que geraram impactos significativos e perceptíveis no ambiente escolar.

Considerando o exposto, este ponto de verificação atingiu a maioria dos aspectos propostos, deixando de atender apenas ao item referente ao aumento do número de alunos premiados nas Olimpíadas de Matemática, que não foi demonstrado na Figura 10 (p. 20), obtendo de acordo com Davidson (2005) conceito B.

VI) A avaliação explicitou com clareza a procedência das conclusões avaliativas?

Acerca desse ponto de verificação, observou-se que o Relatório apresentou as conclusões avaliativas nas considerações finais. A avaliação deixou clara que suas principais fontes de informação foram os dados *PROFMAT: Uma reflexão e alguns resultados* (SBM, 2017) e o questionário enviado por correio eletrônico. Portanto, nesse quesito, a avaliação atendeu plenamente este ponto de verificação, obtendo de acordo com Davidson (2005) conceito A.

VII) A avaliação incluiu recomendações válidas?

“As recomendações vão além de simples conclusões avaliativas” (Scriven, 2018). O relatório meta-avaliado não apresenta recomendações ao final do estudo, e suas conclusões foram sucintas e genéricas, limitando-se à exposição dos resultados. Um dos desafios identificados foi a baixa adesão dos diretores de escolas, o que representa uma fragilidade significativa na avaliação. Para essa fragilidade, por exemplo, poderia ter sido recomendado a adoção de estratégias para aumentar a participação desses gestores em futuras avaliações, garantindo maior representatividade e confiabilidade dos dados.

Dessa forma, as conclusões abordam o tema de forma implícita e superficial, sem explorar adequadamente todos os aspectos

demonstrados no relatório. Como consequência, o estudo não atende plenamente ao ponto de verificação, sendo classificado, de acordo com Davidson (2005), com conceito E.

AVALIAÇÃO DO CRITÉRIO UTILIDADE

De acordo com Davidson (2005), tanto a avaliação quanto os seus resultados, devem ser relevantes para os *stakeholders*, auxiliando no processo decisório, dando publicidade aos resultados e tornando-os disponíveis para o acesso de todos os interessados. Segundo Elliot (2011, p. 951) “é preciso que os resultados sejam úteis a quem encomendou a avaliação. Caso contrário, a avaliação não se justifica, pois não serviu ao cliente ou aos interessados”.

A seguir são apresentadas as análises do critério utilidade, a luz de seus dois pontos de verificação:

1) A avaliação foi relevante para as decisões dos interessados?

O estudo foi encomendado pela Sociedade Brasileira de Matemática (SBM) com o propósito de avaliar os impactos do PROFMAT na atuação de seus egressos e, consequentemente, no desempenho dos alunos em sua prática cotidiana. Como uma ferramenta essencial para a tomada de decisão dos envolvidos no programa, a avaliação permitiu identificar melhorias significativas no desempenho dos professores em sala de aula, bem como um aumento na motivação dos alunos para os estudos de Matemática.

Os resultados também indicaram que a transformação no ambiente escolar pós-PROFMAT se refletiu na adoção de práticas inovadoras de ensino, incluindo o uso de novas metodologias, instrumentos e ferramentas didáticas. Essas mudanças atenderam às expectativas do programa, alinhando-se ao seu objetivo de aprimorar a formação profissional para a docência.

Diante disso, o estudo demonstrou sua relevância e aplicabilidade na orientação de ações futuras, justificando a classificação desse indicador com conceito A, de acordo com Davidson (2005).

II) Os resultados foram comunicados de forma apropriada?

O Relatório não esclarece como os resultados foram comunicados aos principais interessados na avaliação, em especial à SBM, contratante do estudo. Também não há informações sobre a divulgação dos achados à sociedade civil, além da simples disponibilização do documento em formato digital. A ausência de um plano explícito de comunicação dos resultados compromete a transparência e a utilidade da avaliação para os diferentes *stakeholders*.

No entanto, estão presentes os resultados obtidos a partir dos questionários enviados aos diretores de escolas públicas e dos dados extraídos do cadastro do PROFMAT. Além disso, foram considerados os principais indicadores externos, como os resultados das avaliações do Índice de Desenvolvimento da Educação Básica (Ideb), do Exame Nacional do Ensino Médio (Enem) e do Sistema de Avaliação de Rendimento Escolar (Saesp). No âmbito dos indicadores internos, o estudo destacou a redução nas taxas de reprovação em Matemática, o aumento do interesse dos alunos pela disciplina, a realização de mais atividades extracurriculares, a melhoria na relação professor-aluno e o uso de práticas inovadoras de ensino.

Compreende-se, assim, que este ponto de verificação não foi completamente atendido, sendo classificado, de acordo com Davidson (2005), com conceito C.

AVALIAÇÃO DO CRITÉRIO CONDUTA

A seguir, são apresentadas as análises dos dois pontos de verificação do critério conduta, a fim de responder as seguintes questões:

I) A avaliação cumpriu com as legislações pertinentes?

O amparo legal utilizado para o programa é o PNE (Plano Nacional de Educação), instituído pela Lei nº13.005 de 25 de junho de 2014. Esse Plano tem como uma de suas metas o desafio de formar, em nível de pós-graduação, 50% (cinquenta por cento) dos professores da Educação Básica até o último ano de vigência deste PNE e de garantir a todos(as) os(as) profissionais da Educação Básica formação continuada em sua área de atuação, considerando as necessidades, demandas e contextualizações dos sistemas de ensino.

Portanto, pode-se afirmar que o relatório meta-avaliado foi realizado pautado pela legislação em vigor (PNE), cumprindo o propósito de formar profissionais capacitados para atuar no ensino de Matemática na educação básica. Assim, pode-se ajuizar que este ponto de verificação foi atendido plenamente no Relatório meta-avaliado, obtendo de acordo com Davidon (2005), conceito A.

II) A avaliação respeitou os padrões éticos e profissionais?

O critério de Conduta refere-se à avaliação do cumprimento de padrões legais, éticos e profissionais, bem como à adequação cultural e à inexistência de conflitos que possam comprometer a integridade do processo avaliativo, conforme definido por Davidson (2005).

O Relatório teve como objetivo identificar as contribuições do PROFMAT para a melhoria da educação básica, baseando-se em relatos de discentes e coordenadores, além de entrevistas com diretores das escolas onde os egressos do programa atuam (PROFMAT, 2018).

Dessa forma, pode-se afirmar que o relatório foi conduzido em conformidade com os padrões éticos e profissionais, garantindo a idoneidade e a legitimidade do estudo. Além disso, sua execução ficou a cargo de profissionais qualificados, assegurando a credibilidade da avaliação. Diante disso, este ponto de verificação foi

plenamente atendido, sendo classificado, de acordo com Davidson (2005), com conceito A.

AVALIAÇÃO DO CRITÉRIO CREDIBILIDADE

Neste tópico são apresentadas as análises dos dois pontos de verificação do critério Credibilidade, no sentido de responder as seguintes perguntas:

1) A avaliação apresentou familiaridade com o contexto e domínio dos conteúdos apreciados?

De acordo com Davidson (2005, p. 211), os principais fatores que asseguram a credibilidade de uma avaliação incluem o conhecimento do contexto, a independência, a imparcialidade ou a ausência de conflitos de interesse, bem como a competência, tanto na avaliação, quanto na área de estudo em questão. Essa credibilidade relaciona-se com a aceitação dos resultados pelos interessados e do conhecimento profundo do avaliador sobre o projeto ou programa. A imparcialidade e independência do avaliador são essenciais para evitar conflitos de interesse e garantir a confiança externa. Além disso, o domínio técnico do avaliador é crucial para assegurar uma avaliação precisa e confiável (Elliot, 2011).

Em relação à Familiaridade com o contexto foi verificado que os avaliadores conhecem bem o programa que foi avaliado. Verificou-se também que os avaliadores têm pleno conhecimento sobre o seu campo de atuação profissional. Portanto, pode-se afirmar que este ponto de verificação foi atendido no Relatório meta-avaliado, obtendo de acordo com Davidson (2005), conceito A.

II) A avaliação demonstrou ao mesmo tempo independência e/ou imparcialidade?

No que se refere à independência, imparcialidade e possíveis conflitos de interesse, Elliot (2011, p. 949) destaca que “o avaliador não deve ter preferências por qualquer grupo de *stakeholders*, nem apresentar julgamentos tendenciosos”.

Os resultados apresentados neste relatório foram elaborados com o propósito de complementar o estudo “PROFMAT: Uma Reflexão e Alguns Resultados”, que analisou o acompanhamento do programa entre 2013 e 2016. O novo estudo investigou a contribuição do PROFMAT para a melhoria dos indicadores educacionais e do desempenho dos alunos, conduzindo novas análises qualitativas e quantitativas com base nas evidências já existentes, além de dados coletados por meio de questionários, reuniões com equipes e levantamento documental.

Entretanto, a credibilidade dos resultados pode ser questionada, uma vez que as conclusões refletem apenas 14% da amostra total do estudo. Essa baixa taxa de resposta pode comprometer a representatividade dos achados e gerar interpretações sobre possível tendenciosidade, especialmente por parte de alguns *stakeholders*.

Dessa forma, considerando o critério de Credibilidade estabelecido por Davidson (2005), conclui-se que o relatório atende a diversos aspectos desse parâmetro, porém não apresenta um resultado que reflita plenamente o real impacto do programa. Portanto, de acordo com os critérios definidos por Davidson (2005), esse ponto de verificação foi classificado com conceito B.

AVALIAÇÃO DO CRITÉRIO DE CUSTOS

A seguir, são apresentadas as análises do ponto de verificação relacionado ao critério de custos, com o objetivo de responder a seguinte pergunta:

1) A avaliação especificou os custos financeiros e a relação custo-benefício do processo avaliativo?

De acordo com Elliot (2011), todo processo avaliativo envolve diferentes tipos de custos. No caso de avaliações encomendadas, torna-se essencial examinar se os resultados obtidos justificam os gastos despendidos, levando em consideração despesas operacionais, honorários dos avaliadores e demais recursos financeiros envolvidos. Dessa forma, a análise da relação custo-benefício deve ser criteriosa e também é adequada para a meta-avaliação.

Em relação a esse critério, observa-se que o relatório submetido à meta-avaliação carece de informações detalhadas acerca dos custos envolvidos na execução do estudo. Não há referência explícita à relação custo-benefício, tampouco aos aspectos financeiros, que, conforme argumentado por Davidson (2005) constituem parâmetros fundamentais para a avaliação da economicidade de um projeto.

Diante disso, conclui-se que o relatório meta-avaliado não atende ao critério de análise de custos preconizado por Davidson (2005), sendo esse ponto de verificação classificado com conceito E.

QUADRO DE CONCEITOS EM RELAÇÃO AOS CRITÉRIOS E PONTOS ESTABELECIDOS

De acordo com os critérios estabelecidos por Davidson (2005) e a escala de conceitos utilizada na análise dos pontos de

verificação de cada critério observado, o Quadro 4, apresentado a seguir, sintetiza os principais resultados da meta-avaliação.

Quadro 4 – Resultados consolidados dos Critérios

Critérios	Pontos de verificação	Conceito
Validade	I. Cobriu todas as fontes relevantes de valor?	A
	II. Cobriu compreensivamente processo e resultado?	C
	III. Não incluiu critérios irrelevantes ou ilícitos?	A
	IV. Incluiu análises qualitativas e quantitativas?	B
	V. Estabeleceu claramente como os dados são tratados?	B
	VI. Explicitou com clareza a procedência das conclusões avaliativas?	A
	VII. Incluiu recomendações válidas e pertinentes?	E
Utilidade	I. Foi relevante para as decisões dos interessados?	A
	II. Comunicou os resultados de forma apropriada?	C
Conduta	I. Cumpriu as legislações pertinentes (local, nacional e internacional)?	A
	II. Respeitou aos padrões éticos?	A
Credibilidade	I. Apresentou familiaridade com o contexto e domínio dos conteúdos apreciados?	A
	II. Demonstrou independência e/ou imparcialidade?	B
Custos	I. A avaliação especificou os custos financeiros e a relação custo-benefício do processo avaliativo?	E

Fonte: Os autores (2024) adaptado de Davidson (2005).

A classificação da qualidade, obtida através da análise quantitativa, foi definida de acordo com os dados demonstrados no Relatório, que de acordo com o julgamento dos autores desta meta-avaliação, foram classificadas em três níveis: Completamente Atendidos (CA), Parcialmente Atendidos (PA) e Não Atendido (NA). Os critérios que obtiveram conceitos A e B, foram classificados como Completamente Atendidos. Aqueles que obtiveram conceito C, foram classificados como Parcialmente Atendidos. Já, aqueles que obtiveram conceito D e E, foram classificados como Não Atendidos.

Deste modo, o Quadro 5 apresenta o julgamento dos critérios fundamentais de Davidson (2005), de acordo com a análise realizada nesta Meta-avaliação.

Quadro 5 – Resultados consolidados da Meta-avaliação

Critério	Pontos de verificação	Conceito	Julgamento dos Critérios
Validade	Itens: (I), (III), (VI)	A	Parcialmente Atendida (PA) (pelo menos quatro pontos de verificação com conceitos A ou B)
	Item: (IV), (V)	B	
	Item: (II)	C	
	Item: (VII)	E	
Utilidade	Item: (I)	A	Parcialmente Atendidos (PA) (pelo menos um ponto de verificação foi atendido)
	Item: (II)	C	
Conduta	Item: (I)	A	Completamente Atendidos (CA) (todos os pontos de verificação com conceitos A ou B)
	Item: (II)	A	
Credibilidade	Item: (I)	A	Completamente Atendidos (CA) (todos os pontos de verificação com conceitos A ou B)
	Item: (II)	B	
Custos	Item: (I)	E	Não Atendido (NA)

Legenda: Completamente Atendido (CA), Parcialmente Atendido (PA) e Não Atendido (NA).

Legenda: Conceitos A e B (CA); Conceito C (PA); Conceitos D e E (NA).

Fonte: Os autores (2024).

Conforme apresentado no Quadro 5, o critério Validade teve cinco de seus pontos de verificação completamente atendidos, um parcialmente atendido e um não atendido. Dessa forma, seguindo os critérios estabelecidos no estudo, considera-se que esse critério foi Parcialmente Atendido.

O critério Utilidade, composto por dois pontos de verificação, obteve um conceito de completo atendimento e outro de atendimento parcial. Assim, em conformidade com os critérios do estudo, classifica-se esse critério como Parcialmente Atendido.

No caso do critério Conduta, todos os pontos de verificação foram completamente atendidos. Portanto, de acordo com os critérios estabelecidos, esse critério foi classificado como Completamente Atendido.

O critério Credibilidade apresentou conceitos A e B para seus dois pontos de verificação. Embora haja variação nos conceitos, ambos significam completamente atendidos segundo os parâmetros do estudo meta-avaliativo. Assim, o critério foi classificado como Completamente Atendido.

Por fim, o critério Custos, que continha um único ponto de verificação, foi considerado Não Atendido, conforme os critérios estabelecidos neste estudo.

CONSIDERAÇÕES FINAIS

A meta-avaliação do relatório PROFMAT: Avaliação de possíveis impactos confirmou sua relevância para a sociedade e sua importância na capacitação de professores de Matemática da educação básica. Os dados indicam melhorias no desempenho dos docentes em sala de aula, refletindo-se também no aprendizado dos alunos. O aumento da participação estudantil na OBMEP evidencia o impacto positivo do programa, impulsionado pelo maior incentivo por parte dos professores.

No entanto, ao analisar a qualidade do relatório com base nos Critérios Fundamentais de Davidson (2005) – Validade, Utilidade, Conduta, Credibilidade e Custos –, identificaram-se fragilidades em diversos aspectos. Apenas os critérios de Conduta e Credibilidade foram plenamente atendidos. A Validade e a Utilidade apresentaram tanto pontos fortes quanto fracos, sendo classificadas como parcialmente atendidas.

No critério Validade, alguns elementos foram considerados plenamente satisfatórios, como uso de fontes relevantes de valor, a não inclusão de dados ilícitos e/ou irrelevantes, a inclusão de análises qualitativas e quantitativas, tratamento dos dados e a procedência das conclusões apresentadas. Entretanto, o item sobre descrição abrangente do processo e resultado foi apenas parcialmente atendida, e a ausência de recomendações válidas e pertinentes resultou na classificação desse item como não atendido.

O critério Utilidade destacou-se pela relevância da avaliação no processo decisório dos interessados, que foi considerado Completamente Atendido, mas a comunicação dos resultados foi considerada parcialmente atendida, pois não abrangeu todos os achados da avaliação original. Já o critério Custos não foi atendido, pois o relatório não incluiu informações sobre os custos financeiros ou a relação custo-benefício do estudo avaliativo.

Além dessas limitações, a baixa adesão dos participantes (apenas 14% responderam ao convite) afetou especialmente na avaliação do item que trata dos processos e dos resultados, que foi parcialmente atendido. O relatório também apresentou falhas na apresentação dos dados, como a ausência da data de realização do estudo e erros na exibição das informações (por exemplo, na Figura 10, p. 20). A falta de recomendações comprometeu o objetivo da avaliação, que deveria fornecer subsídios para aprimoramentos no PROFMAT.

Conclui-se o objetivo da avaliação não foi completamente atendido. Embora o relatório tenha evidenciado impactos positivos na formação dos professores e no desempenho dos alunos, algumas inconsistências metodológicas e lacunas na análise comprometeram a eficácia da avaliação. A inclusão de recomendações, uma comunicação mais clara dos resultados e a consideração dos custos da avaliação seriam passos essenciais para aprimorar futuros estudos sobre o programa.

RECOMENDAÇÕES

Visando aprimorar a qualidade de avaliações futuras, tornando os resultados mais representativos e úteis para embasar melhorias no PROFMAT, recomenda-se:

- a. Transparência nos custos da avaliação para aumentar a transparência na gestão dos recursos – sejam eles públicos ou privados –, recomenda-se que futuras avaliações explicitem os custos envolvidos no processo avaliativo. Essa prática não apenas reforça a prestação de contas, mas também permite uma análise mais precisa da relação custo-benefício do estudo.
- b. Ampliação da participação dos respondentes. Dado o baixo índice de participação no estudo avaliativo (apenas 14% dos diretores escolares responderam ao convite, de um total de 1.351 e-mails enviados), sugere-se a adoção de estratégias que incentivem uma adesão maior dos interessados. Isso pode incluir múltiplas formas de contato (como ligações, mensagens personalizadas ou encontros virtuais explicativos), além de um prazo maior para respostas.
- c. Inclusão de um público mais amplo, considerando a relevância do estudo para a formação de professores de Matemática da educação básica, recomenda-se que futuras avaliações envolvam não apenas diretores escolares, mas também mestrandos e egressos do PROFMAT. A inclusão desses grupos contribuirá para uma visão mais abrangente e fidedigna dos impactos do programa, fortalecendo a base de dados para análise e aprimoramento dessa política pública.

MÉRITO DO ESTUDO

Glauco da Silva Aguiar

A meta-avaliação desempenha um papel essencial na qualificação e aprimoramento dos processos avaliativos, sendo um instrumento indispensável para assegurar a credibilidade e a eficácia das avaliações realizadas em diversos contextos, incluindo políticas públicas e programas educacionais. O estudo “Meta-avaliação da Avaliação de Possíveis Impactos – PROFMAT” ilustra de maneira clara a relevância desse processo ao analisar criticamente os resultados contidos no relatório de avaliação do programa.

A principal contribuição da meta-avaliação reside na sua capacidade de verificar se os métodos e critérios utilizados em uma avaliação são adequados, confiáveis e eficazes para a tomada de decisão. No caso do relatório PROFMAT, a aplicação dos Critérios Fundamentais de Davidson (2005) permitiu identificar pontos fortes e fragilidades do estudo original, fornecendo informações valiosas para aprimorar futuras avaliações do programa.

Entre os aspectos positivos, destaca-se a relevância do PROFMAT como um programa estruturado para o aprimoramento da formação de professores de Matemática da Educação Básica. Ao oferecer uma qualificação robusta e fomentar o desenvolvimento profissional dos docentes, o programa contribui significativamente para a melhoria do ensino e do aprendizado da Matemática no Brasil. Os impactos positivos do PROFMAT podem ser observados tanto no desempenho dos professores em sala de aula quanto nos avanços dos estudantes, refletidos em indicadores como a participação na Olimpíada Brasileira de Matemática das Escolas Públicas (OBMEP). Além disso, este estudo reforça que a meta-avaliação não apenas valida a qualidade de uma avaliação, mas também contribui para seu aprimoramento contínuo, garantindo que os resultados sejam mais

confiáveis, transparentes e úteis para a gestão de políticas educacionais como o PROFMAT. Esse processo de revisão crítica fortalece a *accountability*, a melhoria contínua dos programas e a tomada de decisões baseadas em evidências sólidas.

Outro ponto de destaque é que tanto a avaliação quanto a meta-avaliação desempenham um papel fundamental no aprimoramento da formação profissional de avaliadores, ampliando sua capacidade crítica e metodológica na condução de estudos avaliativos. Ao proporcionar uma análise rigorosa e fundamentada, a meta-avaliação promove o desenvolvimento de competências essenciais para a realização de avaliações mais precisas e eficazes.

Portanto, a meta-avaliação deve ser incorporada de forma sistemática às avaliações de impacto de políticas públicas e programas educacionais, pois proporciona um olhar refinado sobre a metodologia aplicada e a confiabilidade dos achados. Ao promover maior rigor científico e metodológico, a meta-avaliação contribui para que os estudos avaliativos realmente cumpram seu papel: oferecer subsídios qualificados para a melhoria das políticas educacionais e para o fortalecimento de iniciativas fundamentais como o PROFMAT.

REFERÊNCIAS

ALFRADIQUE, L. M.; TAVARES, T. T.; ORLANDO FILHO, O. **Relatório de Avaliação de Efetividade de Projetos de desenvolvimento Científico e Tecnológico Apoiados pelo Fundo Amazônia/BNDES: um estudo meta-avaliativo**. São Paulo: Pimenta Cultural, 2019.

ASWELL, H. D. **Politics: who gets what, when, how**. Cleveland: Meridian Books, 1958.

BRASIL. Constituição da República Federativa do Brasil, de 5 de outubro de 1988. **Presidência da República**, Brasília, DF, 5 out. 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 17 out. 2024.

BERENDS, L.; ROBERTS, B. Evaluation standards and their application to indigenous programs in victoria, Austrália. **Evaluation Journal of Australasia**, Canberra, Austrália, v. 3, n. 2, p. 54-59, 2003.

DAVIDSON, Jane. **Evaluation methodology basics**. Thousands Oaks, CA: Sage, 2005.

ELLIOT, Ligia G. Meta-Avaliação: das abordagens às possibilidades de aplicação. **Ensaio: avaliação de políticas públicas em educação**. Rio de Janeiro, v. 19, n. 73, p. 941-963, out./dez., 2011.

FITZPATRICK, Jody L.; SANDERS, James R.; WORTHEN, Blaine R. **Avaliação do programa: abordagens alternativas e diretrizes práticas**. 1. ed. Upper Saddle River: Pearson Education, 2011.

LASWELL, H. D. **Politics: who gets what, when, how**. Cleveland, Meridian Books. 1936.

LEITE, Lígia Silva; VILARINHO, Lúcia Regina Goulart; FERREIRA, Sandra Maria Martins Redovalio; FREITAS, Sonia Regina Natal de; SANTOS, Flávia Giffoni de Abreu dos. Avaliação de políticas públicas: uma análise de artigos acadêmicos armazenados na base e-Aval. **Ensaio: aval. públ. Educ.**, Rio de Janeiro, v. 32, n. 123, p. 1-16, abr./jun. 2024.

MASTRODI, Josué; IFANGER, Fernanda Carolina de Araújo. Sobre o conceito de políticas públicas. **Revista de Direito Brasileira**, Belo Horizonte. v. 24, set./dez., 2019. Disponível em: https://go.gale.com/ps/i.do?id=GALE%7CA635666391&sid=googleScholar&v=2.1&it=r&linkaccess=fulltext&issn=2237583X&p=AONE&sw=w&userGroupName=unirio_br&aty=ip. Acesso em: 17 out. 2024.

MEAD, L. M. Public Policy: Vision, Potential, Limits. **Policy Currents**, 1995.

PROFMAT. **Profmat: uma reflexão e alguns resultados**, 2017. Disponível em: http://www.profmatt-sbm.org.br/wp-content/uploads/sites/23/2017/07/PROFMAT-relatorio_DIGITAL.pdf. Acesso em: 10 jan. 2018.

PROFMAT. **Apresentação**. Rio de Janeiro: Sociedade Brasileira de Matemática, 2021. Acesso em: 12 set. 2024. Disponível em: <https://profmatt-sbm.org.br/relatorios-de-avaliacao-profmatt/>. Acesso em: 11 set. 2024.

SOUZA, Celina. **Políticas públicas: questões temáticas e de pesquisa**. Rio de Janeiro: Caderno CRH 39, 2003.

SCHWANDT, T. A. The politics of verifying trustworthiness in evaluation auditing. **American Journal of Evaluation**. 1989. Disponível em: https://www.researchgate.net/publication/249773105_The_Politics_of_Verifying_Trustworthiness_in_Evaluation_Auditing. Acesso em: 11 set. 2024.

SCRIVEN, Michael. The methodology of evaluation. /n: STAKE, R. E. (Ed.), Curriculum evaluation. **American Education Research Association Monograph Series on Evaluation**. Chicago, n. 1, p. 39-83, 1967.

SCRIVEN, Michael. **Evaluation Thesaurus**. 4. Ed. Newbury Park, CA: Sage, 1991.

SCRIVEN, Michael. **Key evaluation checklist**. Kalamazoo, MI: The Evaluation Center, 2007.

SCRIVEN, Michael. **Avaliação**: um guia de conceitos. Editora Paz e Terra, 2018.

STUFFLEBEAM, Daniel L. The metaevaluation imperative. **American Journal Evaluation**, Thousands Oaks, v. 22, n. 2, p. 183-209, 2001.

STUFFLEBEAM, Daniel L. The Methodology of Metaevaluation. /n: STUFFLEBEAM, Daniel L.; MADAUS, George F.; KELLAGHAN, T. (ed.). **Evaluation models**: viewpoints on educational and human services evaluation. 2. ed. New York: Kluwer Academic Publishers, 2002.

WORTHEN, Blaine R.; SANDERS, James R.; FITZPATRICK, Jody L. **Avaliação de programas**: concepções e práticas. São Paulo: Editora Gente, 2004.

6

*Cristina Borges Alves
Alessandra Baldner Pontes
Ovidio Orlando Filho*

META-AVALIAÇÃO DA AVALIAÇÃO IMPACTO SOCIAL DO PROGRAMA *UNIÃO FAZ A VIDA,* REALIZADA PELA FUNDAÇÃO SICRED

A AVALIAÇÃO DE IMPACTO SOCIAL: CARACTERÍSTICAS E RELEVÂNCIA

Ao longo de sua recente história, os estudos sobre Avaliação de Impacto se tornam cada vez mais relevantes para investidores sociais de organizações filantrópicas de diferentes campos de atuação. Os financiadores ou doadores estão procurando avaliar se seus recursos estão sendo investidos em projetos que efetivamente trazem benefícios à sociedade, assim como as empresas que recebem esses recursos estão preocupadas em calcular e demonstrar o resultado do impacto de suas ações na sociedade, analisando se o programa mudou a vida das pessoas envolvidas.

As organizações que desenvolvem projetos sociais, sejam públicos, privados e do terceiro setor, são cada vez mais solicitadas a fornecer evidências sobre a relevância das suas iniciativas, não apenas aos apoiantes e investidores, mas também pelos diversos *stakeholders* envolvidos nas ações. Não se trata mais apenas de discutir o número de beneficiários, mas de compreender a transformação efetiva que esses programas podem trazer para a vida desses beneficiários e para a sociedade como um todo (Barki; Torres; Barros, 2023).

Dentro desse cenário, a Avaliação de Impacto vem ajudando a manter um relacionamento transparente, favorecendo assim a ampliação de programas sociais.

Segundo Amaral (2013, p. 7):

Avaliar o impacto de uma determinada intervenção de desenvolvimento não se limita a verificar o que aconteceu, mas também verificar porque é que isso aconteceu. Ao determinar os efeitos de uma intervenção, a avaliação do impacto está a contribuir diretamente para a responsabilização e prestação de contas, tanto na dimensão ascendente como na sua dimensão descendente, ou seja,

tanto em relação aos financiadores, doadores e cidadãos, assegurando que as suas contribuições são aplicadas em intervenções eficazes, como em relação aos beneficiários e comunidades locais, ao informá-los sobre se a intervenção está a beneficiá-los ou não.

Cabe se destacar que a avaliação é considerada uma ferramenta essencial para entender a viabilidade de programas e projetos. Ela permite redirecionar objetivos quando necessário e reformular propostas e atividades.

Conforme observa Cavalcanti (2006):

A avaliação se revela como um importante mecanismo de gestão, uma vez que fornece informações e subsídios para tomada de decisão dos gestores, formuladores e implementadores de programas, pois possibilita conhecer o que está acontecendo e atuar sobre os fatos de forma a realizar ajustes necessários, economizando-se dessa forma tempo e recursos, o que eleva a credibilidade das ações públicas. Entretanto, apesar de ser indiscutível a necessidade de avaliação, no âmbito do setor público brasileiro essa prática ainda é muito incipiente, especialmente quando focaliza a avaliação de desempenho e de resultados de uma instituição pública ou de um programa governamental, no qual a política pública se materializa.

O Instituto para o Desenvolvimento de Investimento Social (2018) define a Avaliação de Impacto como uma ferramenta estratégica de grande valor. É entendido que essa avaliação proporciona às organizações várias informações e evidências que possibilitam a reflexão sobre as metodologias utilizadas, além de oferecer apoio durante o processo decisório. Inclui-se que ela também investiga a relação de causa e efeito entre as intervenções realizadas e os impactos que foram observados, permitindo a identificação de elementos essenciais para promover as mudanças, outros que não contribuem de forma tão direta e, ainda, limitadores e fatores que criam obstáculos.

De acordo com um estudo feito em 2020 para avaliar o impacto na Educação Básica, publicado no Jornal de Políticas Educacionais, Parente (2020) observou que embora das pesquisas analisadas em seu estudo, a maioria tenha sido formulada pelo Estado (89%), ele também pôde constatar que o Terceiro Setor tem sido alvo de pesquisas de avaliação de impacto na educação básica, chamando a atenção de pesquisadores, a fim de avaliar seus efeitos na educação básica pública.

Em uma pesquisa feita pelo Instituto para Desenvolvimento do Investimento Social, (2024), com a contribuição de representantes da Fundação Avina, Plano CDE e Yunus Negócios Sociais constatou-se que mesmo sendo considerado um tema importante, a Avaliação de Impacto de projetos e negócios sociais ainda é incipiente no Brasil. O principal desafio encontrado para medição do impacto foi o custo elevado, o longo prazo de execução e a insegurança em relação à adequação e à eficácia dos métodos de medição de impacto no país. Identificaram também que as metodologias ainda são desconhecidas, e ainda é vista como “custo” e não como aliada à captação de recursos.

Segundo Carvalho e Fabiani (2023, p. 2):

A avaliação de impacto é uma prática já disseminada e adotada amplamente como ferramenta de tomada de decisão em países como Estados Unidos, Canadá, Alemanha e Inglaterra. No Brasil, ainda apresenta uma aplicação menos destacada, especialmente em decorrência da complexidade e do tempo e custos envolvidos. Entretanto, vemos avanços nos últimos anos, sendo cada vez mais valorizada e aplicada, em especial no contexto do Terceiro Setor. Essa prática permite que sejam mensurados os efeitos gerados por projetos e ações dessas organizações, além de ser um importante instrumento para tomada de decisões e aprimoramento de estratégias. Mensurar o impacto é, portanto, o caminho para se desenhar intervenções cada vez mais eficazes e ajudar a edificar sociedades mais justas e sustentáveis.

POLÍTICAS PÚBLICAS E SUAS AVALIAÇÕES

A avaliação de políticas públicas é concebida como um meio para aprimorar a eficiência, eficácia e efetividade na utilização de recursos e no controle social da atuação do Estado. Constitui-se como atividade estratégica quando integrada ao planejamento governamental, guiando as decisões em processos que envolvem a definição, ajustes, continuidade ou interrupção das ações do governo (Oliveira, 2016).

De acordo com o Tribunal de Contas da União (TCU) (Brasil, 2021), as avaliações realizadas possibilitam aferir o mérito a respeito da importância, utilidade, resultados, impactos e sustentabilidade, além de proporcionar transparência sobre a qualidade e o valor público do gasto. Permitem, também, orientar decisões acerca da viabilidade e da necessidade de manter, otimizar ou descontinuar a política pública. Por isso, é fundamental avaliar a relevância e a utilidade da política pública, ou seja, verificar se os objetivos e as ações governamentais atendem às necessidades dos beneficiários, sejam diretos e indiretos, e aos contextos político, econômico, social e ambiental.

O TCU (Brasil, 2021, p. 10) ainda destacam:

Políticas públicas podem ser definidas como o conjunto de diretrizes e intervenções emanadas do estado, feitas por pessoas físicas e jurídicas, públicas e/ou privadas, com o objetivo de tratar problemas públicos e que requerem, utilizam ou afetam recursos públicos. São crescentes as demandas sociais pela transparência, racionalização dos gastos públicos, bens e serviços de qualidade e investimentos em infraestrutura, saúde, educação, mobilidade urbana, habitação e segurança. Elas requerem uma atuação governamental pautada em boas práticas para formulação, implementação e avaliação das políticas públicas.

A Casa Civil da Presidência da República (Brasil, 2018, p. 14) traduz os conceitos de políticas públicas e de avaliação de políticas públicas da seguinte forma:

um conjunto de programas ou ações governamentais necessárias e suficientes, integradas e articuladas para a provisão de bens ou serviços à sociedade, financiadas por recursos orçamentários ou por benefícios de natureza tributária, creditícia e financeira. A avaliação de políticas públicas, executada como um processo sistemático, integrado e institucionalizado, tem como premissa básica verificar a eficiência dos recursos públicos e, quando necessário, identificar possibilidades de aperfeiçoamento da ação estatal, com vistas à melhoria dos processos, dos resultados e da gestão.

Segundo Cavalcanti (2006), a avaliação de políticas públicas vai além de ser apenas um meio para aprimorar ou redirecionar os programas do governo. Trata-se, principalmente, de um recurso capaz de prestar contas à sociedade das ações governamentais. Sob essa ótica, o diálogo público e o controle social se tornam essenciais na sua implementação, pois a qualidade dos programas só aumenta quando a participação dos cidadãos é intensificada e efetivamente acontece. De acordo com publicação sobre Avaliação de Políticas Públicas realizada pela Casa Civil da Presidência da República (Brasil, 2018), a abordagem de avaliação escolhida precisa estar alinhada ao ciclo implementação da política pública. A avaliação de impacto só deverá ser realizada quando a política já tiver alcançado tempo de execução suficiente para que os resultados e impactos esperados possam ser observados. Observa, ainda que, seguindo as melhores referências internacionais, as avaliações de políticas públicas devem ser executadas em caráter permanente e vinculadas ao ciclo de políticas públicas, deixando claro que apenas com a efetiva utilização dos resultados das avaliações na adoção de medidas de melhoria na execução das políticas e no processo orçamentário e financeiro, os princípios de eficiência, efetividade e eficácia serão aplicados.

Complementado, o órgão cita que:

A experiência internacional mostra que utilizar metodologias bem-aceitas e roteiros padronizados gera resultados confiáveis e permite maior comparabilidade entre as diferentes políticas públicas, por homogeneizar e sistematizar a linguagem e as informações. A credibilidade desses resultados é que contribui para a expansão, o reinvestimento ou a economia de recursos nas políticas públicas avaliadas (Brasil, 2018, p. 15).

META-AVALIAÇÃO: FUNDAMENTANDO SUA RELEVÂNCIA

Scriven (1991) definiu a meta-avaliação de forma pragmática, definindo-a como a avaliação de uma avaliação. Já Stufflebeam (2001) amplia o conceito, apresentando a meta-avaliação como:

O processo de delinear, obter e aplicar informação descritiva e de julgamento – sobre a utilidade, a viabilidade, adequação e precisão de uma avaliação e sua natureza sistemática, competente conduta, integridade/honestidade, respeitabilidade e responsabilidade social – para orientar a avaliação e divulgar publicamente seus pontos fortes e fracos (Stufflebeam, 2001, p. 185).

Elliot (2011) destaca que uma das razões para a realização de uma meta-avaliação é a “reflexão sobre todos os procedimentos utilizados na avaliação, dando oportunidade à incorporação de novos conhecimentos por parte não só dos meta-avaliadores, como também dos avaliadores, quando tomam ciência da análise realizada” (Elliot, 2011, p. 943).

Já Oliveira e Rothen (2024, p. 6), justificam o uso das meta-avaliações resumindo-as em três funções principais:

As que dizem respeito à verificação da qualidade de avaliações; as que se referem à melhoria das avaliações por meio da verificação de seus pontos fortes e vulnerabilidades; e, as que apontam para o atestado da utilidade das avaliações para o público interessado.

Woodside e Sakai (2001, p. 22), por sua vez, citam que as meta-avaliações têm três características principais

1. São sínteses de achados e inferências de pesquisas avaliativas sobre o desempenho do programa. Eles relatam a eficácia do gerenciamento das metas alcançadas pelos programas e fornecem informações sobre duas características: programas bem gerenciados e programas mal gerenciados.
2. Eles informam sobre a validade e utilidade dos métodos de avaliação, oferecendo orientação sobre métodos de avaliação úteis.
3. Eles fornecem fortes evidências sobre o impacto do programa, subsidiando o processo de tomada de decisão a respeito. Assim, os resultados da meta-avaliação auxiliam e justificam o aumento da confiança das partes interessadas e dos gestores dos programas nos resultados da avaliação.

OBJETO DO ESTUDO E OBJETIVO DO ESTUDO META-AVALIATIVO

O objeto desse estudo meta-avaliativo é a Avaliação do Retorno Social do Investimento para o Programa A União Faz a Vida. O processo avaliativo está exposto no relatório realizado em 2020 pelo Instituto para o Desenvolvimento do Investimento Social, junto a Fundação Sicredi (Fabiani; Altemani; Faleiros; Otani, 2020). Desse modo, o citado relatório foi a principal fonte de consulta deste artigo.

O documento analisou o impacto do Programa A União Faz a Vida, de modo a estimar o retorno social dos recursos investidos e analisar sua contribuição para a sociedade.

Seu objetivo é meta-avaliar a avaliação do Programa. Nesse aspecto, cabe se ressaltar que o Instituto para o Desenvolvimento do Investimento Social realizou uma avaliação de impacto social, seguindo os preceitos do protocolo Retorno Social sobre Investimento (SROI), aplicado em projetos sociais de inclusão, na procura de poder ajudar a demonstrar a importância e o valor social gerado.

O PROGRAMA A UNIÃO FAZ A VIDA E SUA AVALIAÇÃO

O Programa foi lançado em 1995, por meio de um programa piloto implantado na cidade de Santo Cristo, no Rio Grande de Sul, desenvolvido pela Fundação de Desenvolvimento Educacional e Cultural do Sistema de Crédito Cooperativo (Fundação Sicredi). Trata-se de uma Organização da Sociedade Civil de Interesse Público, responsável pelo desenvolvimento de iniciativas sociais e culturais vinculadas aos valores do Sistema de Crédito Cooperativo (Sicredi), que trabalha com ações diretamente ligadas à educação e cultura (Fundação Sicredi, [2021]).

Segundo a Fundação Sicredi, sua metodologia propõe que sua abordagem de ensino seja baseada no desenvolvimento de projetos, com ênfase em interações, nos interesses e curiosidades de crianças e adolescentes e a integração dos campos de conhecimento com os aspectos da vida real local. O Programa A União Faz a Vida (2024) é a principal iniciativa de responsabilidade social da Fundação, que o define como um programa educacional que foi criado para promover valores e atitudes de cooperação e cidadania junto aos adolescentes e crianças no Brasil.

Segundo consta no site do Programa, sua intenção é:

Transformar para melhor a realidade dos alunos, de suas famílias e comunidades. Ao promover a educação para um mundo mais cooperativo, inspira professores, estudantes e comunidades, a partir de metodologia educacional própria, fundamentada no desenvolvimento de projetos (A união faz a vida, 2024).

O Programa apresenta como resultado desejado, a logo prazo, a transformação de crianças e jovens em figuras centrais de seu processo de aprendizagem e sua atuação cidadã, para que promovam em suas comunidades impactos positivos. Para atingir esse objetivo, o Programa busca em seus alunos: aumentar do interesse, envolvimento e protagonismo no processo de aprendizagem; aumentar o senso de cidadania e coletividade; fortalecer os relacionamentos interpessoais; e, desenvolver habilidades socioemocionais. Já quanto aos professores, busca desenvolvimento profissional pela vivência de novas formas de ensinar e aprender (A união faz a vida, 2024).

Em 2019, a Fundação Sicredi, em parceria com o Instituto para o Desenvolvimento do Investimento Social, realizou um estudo que procurou mensurar o impacto social do Programa A União Faz a Vida. Ela desejava comprovar o impacto que o Programa causava em alunos e professores e o quanto isso refletia em impactos para a sociedade, como também:

- a. Auxiliar o planejamento estratégico e o processo de tomada de decisão do Sicredi quanto ao Programa A União Faz a Vida.
- b. Entender o impacto do Programa A União Faz a Vida e demonstrar a efetividade do investimento do Sicredi, apoiadores e parceiros.
- c. Identificar os pontos de melhoria e otimização dos impactos gerados pelo Programa.

No sentido de se aferir a qualidade desse Programa, foi empregado em sua avaliação a ferramenta denominada Retorno Social do Investimento (SROI). Esse método de análise vem sendo adotado pelo Instituto para o Desenvolvimento do Investimento Social para avaliar o impacto social de programas sociais. Destaca-se que existem diversos métodos avaliativos que podem ser aplicados para achar os elementos que estabeleçam uma relação de causa e efeito, podendo ser por exemplo análise de retorno sobre o investimento, análise custo-benefício ou análise custo-efetividade.

A partir da coleta de dados qualitativa e quantitativa com os alunos, seus familiares e professores, foi elaborado a Teoria de Mudança do Programa A União Faz a Vida. “A Teoria de Mudança é uma síntese sobre o impacto que se objetiva atingir e a estratégia proposta para que as mudanças efetivamente aconteçam” (A união faz a vida, 2024, p. 6).

Conforme sintetiza Barki, Torres e Barros (2023), a Teoria da Mudança tem a finalidade de identificar de forma clara três dimensões principais: os impactos que uma organização almeja para a sociedade e para os beneficiários do projeto; os resultados de longo e curto prazos que permitem que estes objetivos sejam atingidos; e as principais atividades a serem realizadas pela organização para alcançar o impacto.

A Teoria de Mudança é uma abordagem mais abrangente e narrativa que busca descrever como e por que se espera que uma intervenção específica cause uma mudança desejada em um contexto particular. Ela explora as suposições subjacentes, os mecanismos de mudança e as relações de causa e efeito que conectam as atividades do programa aos resultados e ao impacto desejado. A teoria de mudança envolve a identificação de uma cadeia lógica de eventos e resultados que se baseia em evidências e conhecimentos teóricos sobre como a mudança ocorre (Joppert; Braz; Queiroz; Lamas; Gripp, 2024, p. 13).

O objetivo da Teoria da Mudança, segundo Barki, Torres e Barros (2023, p. 5), é ser uma “ferramenta básica de planejamento e de preparação para avaliação de impacto”. Quanto ao método utilizado nesse estudo avaliativo, ele tem sido muito utilizado como uma ferramenta de avaliação de impacto social e financeiro de iniciativas de negócios sociais e de organizações sem fins lucrativos. Assim, o *Social Return On Investment* (SROI) ou Retorno Social do Investimento é um tipo de análise de custo-benefício que ajuda as organizações a avaliarem os aspectos intangíveis dos seus programas, ou seja, os aspectos que criam valor social real que geram nas vidas das pessoas, nos bens e serviços públicos e no meio ambiente, mas que muitas vezes são difíceis de medir devido aos seus aspectos subjetivos. Ou seja, ela tem como premissa que projetos ou programas devem ser avaliados não apenas pelo retorno financeiro, mas também pelo valor social que geram na comunidade e no meio ambiente (Fabiani; Carvalho, 2023).

Segundo a Avaliação de Impacto Social do Programa A União Faz a Vida (Fabiani; Altemani; Faleiros; Otani, 2020), as avaliações tradicionais nem sempre refletem as mudanças mais importantes e mais valorizadas pelos participantes. E defende que a riqueza do SROI está justamente em medir o impacto vivenciado pelos *stakeholders* e que é considerado relevante por eles.

No relatório apresentado da Avaliação do Retorno Social do Investimento para o Programa A União Faz a Vida, ele esclarece o motivo da escolha de seu método avaliativo, afirmando que o protocolo SROI permite, por conta das características listadas abaixo, que os objetivos determinados no projeto sejam atingidos:

Os resultados da avaliação SROI retratam em que medida as intervenções são eficientes e de que forma os resultados são percebidos pelos grupos de interesse (também chamados de *stakeholders* ou públicos-alvo); As informações geradas pela avaliação SROI podem auxiliar a Fundação Sicredi a maximizar o impacto de um determinado recurso (financeiro ou não-financeiro), uma vez

que combinam em seu processo avaliativo métodos de pesquisa qualitativa e quantitativa, possibilitando analisar quais fatores, intrínsecos e extrínsecos ao Programa, estão dificultando ou potencializando o seu sucesso (Fabiani; Altemani; Faleiros; Otani, 2020, p. 13).

Conforme orientação do *A guide to Social Return on Investment*, realizar uma análise de impacto social utilizando a metodologia SROI envolve seis etapas:

Estabelecer escopo e identificar as principais partes interessadas. É importante ter limites claros sobre o que sua análise SROI cobrirá, quem estará envolvido no processo e como.

Mapeando resultados. Ao se envolver com suas partes interessadas, você desenvolverá um mapa de impacto, ou teoria da mudança, que mostra a relação entre entradas, saídas e resultados.

Evidenciando resultados e dando-lhes um valor. Este estágio envolve encontrar dados para mostrar se os resultados aconteceram e, em seguida, valorizá-los.

Estabelecendo impacto. Tendo coletado evidências sobre os resultados e monetizando-os, os aspectos da mudança que teriam acontecido de qualquer maneira ou são resultado de outros fatores são eliminados da consideração.

Calculando o SROI. Esta etapa envolve somar todos os benefícios, subtrair quaisquer negativos e comparar o resultado com o investimento. É também aqui que a sensibilidade dos resultados pode ser testada.

Relatórios, uso e incorporação. Facilmente esquecida, esta última etapa vital envolve compartilhar descobertas com as partes interessadas e responder a elas, incorporando processos de bons resultados e verificação do relatório (The SROI Network, 2012, p. 9).

METODOLOGIA APLICADA NO ESTUDO META-AVALIATIVO

Este estudo possui caráter qualitativo com componentes descritivo e exploratório, pois procura conhecer mais sobre uma avaliação de impacto de um programa de política pública, apresentando e discutindo padrões de avaliação, possibilitando assim identificar oportunidades de melhoria em seus processos avaliativos. Para tal objetivo, inicialmente, realizou-se pesquisa bibliográfica para identificar os principais aspectos conceituais sobre a temática de Avaliação de Impacto, Teoria da Mudança, Políticas Públicas, sobre o método do Retorno Social sobre Investimento (SROI), e sua efetiva utilização, bem como para identificar os padrões a serem aplicados para análise das avaliações selecionadas neste estudo.

A meta-avaliação realizada possui caráter somativo e sua aplicação foi realizada a partir de uma lista de verificação considerando os padrões do *Joint Committee on Standards for Educational Evaluation* com questionamento se a avaliação do Programa atende ou não aos critérios elencados. É importante ressaltar que o *Joint Committee* é uma instituição de referência internacional em padrões para a realização de avaliações e de meta-avaliações (Yarbrough: Shulha: Hopson e Caruthers, 2011).

QUESTÕES AVALIATIVAS

A questão meta-avaliativa se apresenta como um importante guia que orientará a avaliação, no intuito de nortear este estudo. Considerando a importância de se entender mais sobre avaliação de impacto em programas sociais, e a qualidade da avaliação de um programa, essa avaliação se propôs a responder a seguinte questão avaliativa:

Em que ponto a avaliação descrita no Relatório da Avaliação do Retorno Social do Investimento para o Programa A União Faz a Vida atendeu aos padrões de avaliação definido pelo *The Joint Committee on Standards for Educational Evaluation (JCSEE)*, considerando como referência a terceira edição da obra *Program Evaluation Standards* publicada em 2011?

Como toda avaliação dessa natureza, este estudo trilha pelos caminhos da análise qualitativa, baseado na interpretação que os autores utilizaram para a formulação de seu juízo de valor. Assim, ele está sujeito a pontos e contrapontos realizados por outros atores que, porventura, possuam uma visão diferente da realidade estudada. Também deve ser considerado o fato de o estudo ter se baseado somente na leitura, mesmo que minuciosa, de seu relatório final, sobre o qual, considerando seu conteúdo, foram formuladas as análises que conduziram à resposta da questão avaliativa.

OS PADRÕES DE *JOINT COMMITTEE*

O *Joint Committee on Standards for Educational Evaluation (JCSEE)* é considerado a fonte oficial de padrões de avaliação nos Estados Unidos e no Canadá, sendo composto por representantes de organizações profissionais e especialistas em avaliação. Os padrões definidos pelo *Joint Committee* são utilizados em diversos tipos de avaliação de programas, embora tenham origem na avaliação educacional (*Joint Committee on Standards for Educational Evaluation*, 2020), tratando-se de um comitê conjunto de padrões para avaliação educacional que:

Foi formado em 1975 como uma coalizão de associações profissionais nos Estados Unidos e Canadá preocupadas com a qualidade da prática de avaliação. A participação no comitê inclui representação da comunidade educacional, avaliadores profissionais e especialistas no assunto,

cada um com interesse no uso e na prática da avaliação. O primeiro conjunto de padrões, *The Program Evaluation Standards*, foi publicado em 1981. Desde então, o comitê continuou a revisar e atualizar os padrões originais, bem como a desenvolver novos padrões para áreas especializadas de avaliação educacional.

Os atributos de avaliação, usadas internacionalmente pelos avaliadores, definidos pelo *Joint Committee*, estão divididas em cinco. O Quadro 1 descreve de modo simplificado as características de cada um deles:

Quadro 1 – Atributos do *Joint Committee*

Atributos	Características
Utilidade	Os padrões de utilidade visam aumentar a extensão em que as partes interessadas do programa consideram os processos e produtos de avaliação valiosos para atender às suas necessidades.
Viabilidade	Os padrões de viabilidade visam aumentar a eficácia e a eficiência da avaliação.
Propriedade	Os padrões de propriedade apoiam o que é apropriado, justo, legal, certo e equitativo nas avaliações.
Precisão	Os padrões de precisão visam aumentar a confiabilidade e a veracidade das representações, proposições e descobertas de avaliação, especialmente aquelas que apoiam interpretações e julgamentos sobre qualidade.
Responsabilidade	Os padrões de responsabilização da avaliação incentivam a documentação adequada das avaliações e uma perspectiva meta avaliativa focada na melhoria e responsabilização pelos processos e produtos de avaliação.

Fonte: As autoras (2024) adaptado de Yarbrough, Shulha, Hopson e Caruthers (2011).

Observa-se que as categorias elencadas pelo *Joint Committee* buscam avaliar de forma abrangente os diferentes tipos de avaliação, podendo ser aplicados a diversas metodologias e tipos de avaliação.

Para dar valor aos atributos foram utilizados os seguintes critérios:

- a. Os atributos de Utilidade e de Precisão foram considerados como atendidos se todos os seus padrões foram atendidos; foram considerados como parcialmente atendidos se, pelo menos, cinco de seus padrões foram atendidos; e, foram considerados como não atendidos se menos de cinco padrões não foram atendidos;
- b. O atributo da Adequação foi considerado como atendido se todos os seus padrões foram atendidos; foram considerados como parcialmente atendidos se, pelo menos, quatro de seus padrões foram atendidos; e, foram considerados como não atendidos se menos de quatro de seus padrões não foram atendidos.
- c. O atributo da Exequibilidade foi considerado como atendido se todos os seus padrões foram atendidos; foram considerados como parcialmente atendidos se, pelo menos, dois de seus padrões foram atendidos; e, foram considerados como não atendidos se menos de dois de seus padrões não foram atendidos.
- d. O atributo da Responsabilização foi considerado como atendido se todos os seus padrões foram atendidos; foram considerados como parcialmente atendidos se, pelo menos, dois de seus padrões foram atendidos; e, foram considerados como não atendidos se menos de dois de seus padrões não foram atendidos.

RESULTADOS E ANÁLISES

A avaliação do Programa A União Faz a Vida, realizado pelo Instituto para o Desenvolvimento do Investimento Social (Fabiani;

Altemani; Faleiros; Otani, 2020), junto à Fundação Sicredi, foi meta-avaliado com base nos padrões estabelecidos pelo *Joint Committee*.

Para isso, foram utilizados os critérios: atende totalmente, atende parcialmente ou não atende. Apresentamos abaixo os padrões propostos pelo *Joint Committee* e seus conceitos, conforme o Quadro 2:

Quadro 2 – Padrões do atributo de utilidade da avaliação

Padrões	Características
Credibilidade do Avaliador	As avaliações devem ser conduzidas por processos qualificadas que estabeleçam e mantenham a credibilidade no contexto da avaliação.
Atenção aos Interessados (Stakeholders)	As avaliações devem dedicar atenção a toda a gama de indivíduos e grupos investidos no programa e afetados por sua avaliação.
Propósitos Negociados	Os propósitos da avaliação devem ser identificados e continuamente negociados com base nas necessidades das partes interessadas.
Explicitações de valores	As avaliações devem esclarecer e especificar os valores individuais e culturais que sustentam propósitos, processos e julgamentos.
Informação Relevante	As informações de avaliação devem atender às necessidades identificadas e emergentes das partes interessadas.
Processos e Produtos Significativos	As avaliações devem construir atividades, descrições e julgamentos de maneiras que incentivem os participantes a redescobrir, reinterpretar ou revisar seus entendimentos e comportamentos.
Comunicação e Relatórios apropriados e no prazo	As avaliações devem atender às necessidades contínuas de informações de seus diversos públicos.
Preocupação com consequências e influência	As avaliações devem promover o uso responsável e adaptável, ao mesmo tempo em que protegem contra consequências negativas não intencionais e uso indevido.

Fonte: As autoras (2024) adaptado de Yarbrough, Shulha, Hopson e Caruthers (2011).

O relatório (Fabiani; Altemani; Faleiros; Otani, 2020) aponta que estudo avaliativo foi realizado pelo Instituto para o Desenvolvimento do Investimento Social, junto a Fundação Sicredi. Esse instituto é uma organização social independente, voltada para projetos de impacto social, reconhecida no Brasil, premiada como uma das “Melhores Ongs de 2022”. Logo, entende-se que o padrão de **credibilidade do avaliador** (U1) foi atendido plenamente, pois a organização avaliadora apresentava conhecimento técnico e experiência para conduzir a avaliação.

Segundo Yarbrough, Shulha, Hopson e Caruthers (2011, p. 44) as “discussões bem fundamentadas com os *stakeholders* sobre intenções, processos e valores integrados ao programa representam uma importante fonte de dados para a compreensão do programa no contexto”.

Fabiani, Altemani, Faleiros e Otani (2020) esclarecem que foi possível identificar os *stakeholders* por meio da análise da documentação disponível acerca do Programa e das entrevistas iniciais com a equipe da Fundação Sicredi. O relatório listou a ligação que os entrevistados tinham dentro do programa:

As entrevistas iniciais foram muito importantes para identificar os *stakeholders*, entender em detalhe as atividades pertencentes ao Programa A União Faz a Vida e mapear as percepções e expectativas dos entrevistados em relação ao Programa (...). Assim, para esta avaliação, foi priorizada a análise dos *stakeholders* significativamente afetados pelas atividades do Programa A União Faz a Vida: os alunos e professores (Instituto para o Desenvolvimento do Investimento Social, 2018, p. 17 e 19).

Depois que foram identificados dez grupos de interesse, como assessoria pedagógica, equipe de gestão das escolas e equipe da cooperativa local, a avaliação priorizou apenas a análise de dois grupos de *stakeholders*: alunos e professores. Antes de chegar a esses dois grupos, o estudo avaliativo seguiu a recomendação do

SROI. Primeiramente elaborou um quadro com esses grupos para analisar quais foram o envolvimento deles no projeto, ou seja, quais deles foram significativamente afetados pelas atividades do programa. Depois decidiram quais desses grupos de *stakeholders* envolvidos deveriam ser excluídos, e quais as razões para tal decisão.

Porém, não levaram em conta os agentes viabilizadores da execução, incluíram apenas os beneficiários do programa. Justificaram que estavam seguindo a recomendação do protocolo SROI. Segundo o *The Guide to Social Return on Investment* (The SROI Network, 2015), “você precisará incluir todas as partes interessadas que você considerar poder sofrer mudanças materiais com o resultado da atividade, ou seja, resultados relevantes e significativos”.

Avaliando todas essas informações relacionadas ao padrão **atenção às partes interessadas** (*stakeholders*) (U2), ficou entendido que esse padrão foi atendido parcialmente, pois a decisão tomada na escolha de apenas dois grupos de *stakeholders* talvez possa ter comprometido um pouco a qualidade da avaliação, pois compreender e interpretar as vivências de um grupo maior de *stakeholders* que participaram do programa, coletando suas opiniões e experiências, seria importância para a avaliação. Elas acreditam que os agentes que viabilizaram a execução do programa, como por exemplo a equipe de gestão das escolas e assessoria pedagógica, também teriam muito a agregar.

O relatório (Fabiani; Altemani; Faleiros; Otani, 2020) informa que os impactos dos alunos e professores foram desenvolvidos e validados junto aos principais *stakeholders* do Programa. O relatório descreveu três eixos de mudança e explicou o que cada impacto significaria.

As hipóteses de impacto social geradas pelo Programa A União Faz a Vida sinalizadas na Teoria de Mudança foram validadas, detalhadas e mensuradas juntos aos beneficiários nas fases de coleta de dados qualitativos.

Os relatos e informações coletadas durante a realização dos grupos focais demonstraram evidências de que todas as mudanças representadas na Teoria de Mudança foram, de fato, percebidas pelos alunos, seus familiares e professores do Programa. No entanto, além das mudanças previstas na primeira versão da Teoria de Mudança, também foram identificados outros impactos que passaram a compor as variáveis avaliativas dos eixos de mudança de cada público (Instituto para o Desenvolvimento do Investimento Social, 2018, p. 27).

Logo, com relação ao padrão **Propósitos Negociados** (U3), concluiu-se que essa avaliação atendeu a esse padrão, as necessidades foram identificadas pelos usuários.

O documento explica que o programa A União Faz a Vida se dedica à construção de cidadania, atitudes e valores de cooperação entre crianças e adolescentes, a partir de uma metodologia de ensino própria:

Com base na interlocução entre o currículo escolar, a valorização dos interesses e conhecimentos dos alunos e a conscientização acerca do contexto social de comunidades locais, a metodologia propõe o desenvolvimento de cidadãos críticos e cooperativos (Instituto para o Desenvolvimento do Investimento Social, 2018, p. 27).

Fabiani, Altemani, Faleiros e Otani (2020) informam que o programa contribui diretamente para o alcance de dois objetivos para o Desenvolvimento Sustentável (ODS) promovidos pela Organização das Nações Unidas, como parte da Agenda 2030:

Educação de Qualidade: Contribuindo para o desenvolvimento de uma educação mais inclusiva e equitativa e de qualidade, e para a promoção do acesso a oportunidades de aprendizagem para todos; Parcerias e Meios de Implementação: referente ao es forço para fortalecer os meios de implementação e revitalizar a parceria global para o desenvolvimento sustentável, isto é, estabelecer compromissos de cooperação entre diferentes atores da

sociedade de modo a impulsionar o desenvolvimento sustentado das nações (Instituto para o Desenvolvimento do Investimento Social, 2018, p. 11).

Pode-se, então, concluir que o padrão **Explicitação de valores** (U4) também foi plenamente atendido.

Conforme demonstrado no relatório, as informações coletadas através dos grupos focais e de seus relatos “demonstraram evidências de que todas as mudanças representadas na Teoria de Mudança foram, de fato, percebidas pelos alunos, seus familiares e professores do Programa” (Fabiani; Altemani; Faleiros; Otani, 2020. p. 27).

Segundo Yarbrough, Shulha, Hopson e Caruthers (2011), “a aceitabilidade de produtos da avaliação, tais como representações, julgamentos e recomendações, está intimamente ligada ao que os *stakeholders* pensam da adequação e relevância das informações utilizadas”.

O relatório (Fabiani; Altemani; Faleiros; Otani, 2020) relata detalhadamente todo o processo de desenvolvimento da avaliação, estabelecendo o escopo, identificando os *stakeholders*, a coleta de dados, descreve o modelo SROI, além de apontar os principais resultados obtidos com detalhes. Logo, o padrão **Informação Relevante** (U5) foi atendido plenamente.

A avaliação de impacto elaborou a Teoria de Mudança do Programa A União Faz a Vida, que precedeu a avaliação de impacto propriamente dita, gerando mais segurança para os proponentes dessa avaliação em termos de indicadores e possíveis metodologias a serem utilizadas. O documento ressalta ainda que essa análise foi validada junto às principais partes interessadas do Programa.

Para aferir a qualidade desse Programa, foi empregada em sua avaliação a ferramenta denominada Retorno Social do Investimento (SROI), desenvolvido pela organização *Social Value UK*, e teve como intuito mensurar o impacto social gerado pelo Programa.

O protocolo (SROI) se tornou cada vez mais popular nos anos seguintes, especialmente pelo papel fundamental desempenhado pela Social Value UK na sua evolução e difusão, sendo adotado por outras organizações ao redor do mundo como uma forma de medir e comunicar o valor social gerado pelas suas atividades. Hoje em dia, o SROI é amplamente utilizado como uma ferramenta de avaliação de impacto social e financeiro de iniciativas de negócios sociais e de organizações sem fins lucrativos (Fabiani; Carvalho, 2023, p. 3).

Atualmente, segundo a Revista Brasileira de Avaliação (Fabiani; Carvalho, 2023, p. 3), “o SROI é amplamente utilizado como uma ferramenta de avaliação de impacto social e financeiro de iniciativas de negócios sociais e de organizações sem fins lucrativos”. Ele é uma análise de custo-benefício reconhecida pelo *Cabinet Office do Reino Unido* (The SROI Network, 2024).

Logo, foi concluído que o padrão **Processos e Produtos Significativos** (U6) foi atendido plenamente, com métodos reconhecidos internacionalmente.

O relatório (Fabiani; Altemani; Faleiros; Otani, 2020) publicado descreve detalhadamente todo o processo avaliativo, ajudando os interessados a requererem profunda compreensão do Programa. Trata-se de um relatório extenso, que contempla o roteiro e sistematização dos grupos focais, questionário e sistematização da coleta de dados quantitativos, cálculo das proxies financeiras, resultados da avaliação SROI do programa e as Conclusões e Recomendações. Yarbrough, Shulha, Hopson e Caruthers (2011), cita as contribuições da continuidade de abordagens, atividades e produtos da avaliação para o entendimento das partes interessadas:

Quando abordagens, atividades e produtos da avaliação são informados continuamente, os *stakeholders* conseguem entender de maneira adequada os procedimentos e os achados, o que possibilita utilizações não antecipadas e contribui para que a avaliação responda a mudanças

no cronograma e nas necessidades (Yarbrough; Shula; Hopson; Caruthers, 2011).

No entanto, mesmo que o relatório tenha informado que o SROI considera vital a etapa de compartilhamento de resultados com as partes interessadas, não ficou claro se apresentaram relatórios aos interessados ao longo da avaliação. Também não ficou muito claro se todos os prazos foram informados e cumpridos. Por esses motivos, houve o entendimento que o relatório atendeu parcialmente ao padrão de **Comunicação e Relatórios Adequados e no Prazo** (U7).

De acordo com Yarbrough, Shulha, Hopson e Caruthers (2011), “a avaliação sistemática colhe e analisa dados em resposta a perguntas consideradas importantes pelos *stakeholders*, além de orientá-los na utilização, de maneira lógica, significativa, ética e transparente, das informações resultantes”.

O relatório (Fabiani; Altemani; Faleiros; Otani, 2020) aponta que, durante a coleta de dados qualitativos, foi checada com as partes interessadas a existência de resultados negativos e/ou inesperados que tenham ocorrido a partir do programa analisado. O documento menciona que:

Na avaliação do Programa A União Faz a Vida, não houve nenhum impacto negativo que possa ter se deslocado, nem tampouco foi identificado algum impacto positivo proveniente de outros locais/regiões, eliminando, portanto, a necessidade da variável de Deslocamento do modelo de avaliação SROI (Fabiani, Altemani, Faleiros; Otani, 2020. p. 56).

Porém, não foi esclarecido se algumas condutas éticas ocorreram ou não durante a avaliação. As Diretrizes para Conduta Ética das Avaliações da *Australasian Evaluation Society* (2013, p. 8) definem a necessidade de antecipação ou discussão da avaliação de riscos ou danos potenciais às partes interessadas, conforme extrato a seguir:

A decisão de realizar uma avaliação ou procedimentos específicos dentro de uma avaliação deve ser cuidadosamente considerada à luz dos riscos ou danos potenciais para os clientes, grupos-alvo ou funcionários do programa. Na medida do possível, estas questões devem ser antecipadas e discutido durante a negociação inicial da avaliação.

Essas diretrizes também orientam que os respondentes devem ser informados sobre quais informações serão solicitadas, como as informações serão registradas e utilizadas, e que devem, também, informar os riscos e benefícios prováveis decorrentes da sua participação na avaliação (Australasian Evaluation Society, 2013).

Quanto à relação ética do avaliador com os participantes do estudo avaliativo, também fazem parte dos princípios a garantia da confidencialidade das informações prestadas pelos participantes, protegendo seus dados pessoais e mantendo sua privacidade e assegurar o anonimato dos participantes.

Os Princípios orientadores da *American Evaluation Association* também reforçam essas orientações aos avaliadores “cumprir a ética profissional, os padrões e os regulamentos atuais (incluindo consentimento informado, confidencialidade e prevenção de danos) relativos aos participantes da avaliação”.

Embora as orientações éticas da *Social Research Association* (Social Value International, 2021, p. 6), defendem que o “consentimento pode ser contado como informado quando os participantes em potencial recebem informações suficientes para você presumir que eles entendem o que a participação implica”, eles também esclarecem que é necessário informar os riscos, quem terá acesso aos dados pessoais e no que implicará sua participação, o que nessa avaliação não aconteceu.

Não foram encontrados pontos que pudessem causar um impacto negativo grande. Porém, não foram apresentados os riscos

da avaliação, e se houve algum termo de consentimento assinado pelos respondentes quando aplicados os instrumentos avaliativos. Tampouco acharam qualquer documento que comunicasse ao participante os procedimentos que seriam usados para garantir a confidencialidade e proteção dos seus dados durante uma avaliação, como orienta a Rede Brasileira de Monitoramento e Avaliação (Silva, Joppert; Gasparini, 2020).

Por esses motivos, ficou apreendido que o relatório atendeu parcialmente esse padrão de **Preocupação com Consequências e Influência** (U8).

Em síntese, o atributo **Utilidade** foi considerado como parcialmente atendido, tendo em vista que cinco padrões foram atendidos, de acordo com os critérios estabelecidos neste estudo avaliativo.

Elliot (2011, p. 951-952) aponta que os quatro padrões de exequibilidade do *Joint Committee* "são voltados para procedimentos que assegurem que a avaliação será gerenciada de modo efetivo, será prática e realística, será cuidadosa ao tratar do contexto e seus interesses políticos e culturais, será tanto efetiva quanto eficiente ao utilizar os recursos" (Quadro 3).

Quadro 3 – Padrões do atributo de exequibilidade da avaliação

Padrões	Características
E1 - Gerenciamento do Projeto	As avaliações devem usar estratégias eficazes de gerenciamento de projetos.
E2 - Procedimentos Práticos	Os procedimentos de avaliação devem ser práticos e adequados à maneira como o programa opera.
E3 - Viabilidade do Contexto	As avaliações devem reconhecer, monitorar e equilibrar os interesses e necessidades culturais e políticas de indivíduos e grupos.
E4 - Uso dos Recursos	As avaliações devem usar os recursos de forma eficaz e eficiente.

Fonte: As autoras (2024) adaptado de Yarbrough, Shulha, Hopson e Caruthers (2011).

Antes de iniciar a avaliação utilizando os princípios do protocolo SROI, o Instituto para o Desenvolvimento do Investimento Social elaborou a Teoria da Mudança do Programa A União Faz a Vida, ajudando a traçar metas e objetivos:

A Teoria de Mudança facilita o entendimento sobre o que o Programa A União Faz a Vida busca alcançar e levanta as hipóteses de mudança a serem validadas ou ajustadas junto aos *stakeholders*, na fase qualitativa de coleta de dados. A partir da elaboração da Teoria de Mudança do Programa A União Faz a Vida, foram criadas hipóteses quanto às mudanças geradas na vida dos beneficiários por meio do Programa, que levaram à formulação de um modelo teórico que explica de que forma esse processo ocorre. Neste sentido, são estabelecidas as ligações de causa-efeito entre cada iniciativa e seus respectivos impactos (Fabiani; Altemani; Faleiros; Otani, 2020, p. 22).

Segundo o relatório (Fabiani; Altemani; Faleiros; Otani, 2020), a Teoria de Mudança é uma síntese sobre o impacto que se objetiva atingir e a estratégia que foi proposta para que as mudanças efetivamente aconteçam. Ela tem como objetivos apoiar o planejamento estratégico e a tomada de decisão da Fundação com relação ao Programa, e ajudá-la a identificar possíveis áreas de melhoria e otimização dos efeitos gerados pelo Programa.

Já as seis etapas recomendadas pelo SROI, que foram seguidas pelo Instituto, ajudaram muito no planejamento estratégico: “1. Estabelecer o escopo e identificar os *stakeholders*-chave, 2. Mapear resultados, 3. Evidenciar resultados e atribuir-lhes um valor, 4. Estabelecer o impacto, 5. Calcular o SROI, 6. Relatar, usar e incorporar as conclusões” (Fabiani; Altemani; Faleiros; Otani, 2020, p. 16).

Foi reconhecido a importância dessas etapas, porém não foi demonstrado o cronograma das fases e suas estimativas, para evitar futuros riscos, assim como os custos envolvidos na avaliação não foram mencionados e como os respondentes foram convidados a participar da avaliação. Por essas razões, elas entenderam que esse padrão **Gerenciamento de Projetos** (E1) foi parcialmente atendido.

Segundo Yarbrough, Shulha, Hopson e Caruthers (2011), “procedimentos de coletas de dados responsivos à diversidade contextual melhoram a exequibilidade da avaliação e colaboram para interpretações válidas”.

O relatório (Fabiani; Altemani; Faleiros; Otani, 2020) esclarece que o objetivo principal da coleta de dados foi envolver os principais beneficiários do Programa para entender o que mudou em suas vidas deles através dele.

Os grupos focais permitiram verificar a materialidade das hipóteses de impacto levantadas na Teoria de Mudança, para a definição das variáveis a serem consideradas na avaliação. Os objetivos específicos dos grupos focais foram:

- Entender as mudanças que ocorreram na vida dos beneficiários geradas a partir da experiência com o Programa A União Faz a Vida;
- Verificar a materialidade das hipóteses de impacto estabelecidas na primeira versão da Teoria de Mudança;
- Definir os impactos que serão considerados na avaliação e, quando necessário, realizar ajustes na hipótese da Teoria de Mudança previamente construída;
- Dar insumos para a definição de indicadores que serão quantificados.

(Fabiani; Altemani; Faleiros; Otani, 2020. p. 31).

Na fase de coleta de dados qualitativos, as hipóteses foram validadas, detalhadas e mensuradas junto aos beneficiários do Programa (Fabiani; Altemani; Faleiros; Otani, 2020). Logo, como esse procedimento junto ao grupo focal conseguiu identificar as necessidades dos *stakeholders* avaliados, que no caso foram os alunos, seus familiares e professores do Programa, ficou entendido que esse padrão **Procedimentos Práticos** (E2) foi plenamente atendido.

Yarbrough, Shulha, Hopson e Caruthers (2011) citam duas categorias importantes como influentes no padrão **Viabilidade do contexto** (E3): a primeira são “os interesses políticos – aqueles relacionados à situação econômica e autoridade, por exemplo – e sistemas preexistentes de valores, como os religiosos ou culturais”, e que “normalmente, essas influências controlam os recursos, as regulamentações e outros fatores que afetam diretamente o programa e sua avaliação”; a segunda consiste nos valores que os grupos ou os indivíduos levam para o contexto da avaliação.

Os autores apontam como um dos riscos da **viabilidade do contexto** o atendimento às necessidades e apenas um dos grupos principais das partes interessadas, dando assim, a impressão de ser aquele mais importante que os outros. Observa-se que o relatório do Instituto para o Desenvolvimento do Investimento Social (Fabiani; Altemani; Faleiros; Otani, 2020) lista dez partes interessadas na avaliação do programa. No entanto, são priorizadas apenas duas partes interessadas: alunos e professores.

Além disso, segundo Yarbrough, Shulha, Hopson e Caruthers (2011), indivíduos e grupos mais tímidos podem permanecer à parte, não tendo seus interesses incluídos. Dessa forma, uma das recomendações do *Joint Committee* é identificar os *stakeholders* que não estão envolvidos no processo e incentivá-los a fornecer informações necessárias à avaliação.

No Brasil, a RBMA cita os níveis de participação adequados aos interessados como sua primeira diretriz, devendo ser a mais plural possível, conforme texto abaixo:

Os processos de concepção e realização das avaliações devem incluir as perspectivas dos interessados e acolher suas demandas, perguntas e preocupações. Tal inclusão deve ser a mais plural possível, com atenção aos grupos historicamente mais excluídos e marginalizados da sociedade, e fruto de pactuação que defina tipos e níveis de participação que assegurem espaços e mecanismos de

efetivo compartilhamento de poder e responsabilidades (Silva; Joppert; Gasparini, 2020, p. 14).

Como o relatório de Fabiani, Altemani, Faleiros e Otani (2020) prioriza somente duas partes interessadas num universo de dez partes identificadas, chegou-se à conclusão de que a avaliação não atende ao padrão apontado.

Com relação ao padrão **Uso de recursos** (E4), Yarbrough, Shulha, Hopson e Caruthers (2011) mencionam que “uma avaliação efetiva e eficiente não consome recursos além dos necessários à concretização de seus objetivos”, e que “o equilíbrio entre efetividade e eficiência na utilização dos recursos contribui para garantir que a avaliação compense os gastos e para que os patrocinadores percebam que o investimento valeu a pena”.

Esses custos se relacionam ao valor total dos recursos que são empregados na avaliação: tempo e recursos humanos (avaliadores, participantes, equipes, voluntários); quantias cobradas; equipamentos, instalações e comunicação. Esses custos da avaliação, segundo o *Joint Committee*, devem ser levados tanto em conta quanto os benefícios, e todos precisam estar alinhados aos propósitos ou às necessidades das partes interessadas.

Constatou-se que a avaliação feita pelo Instituto para o Desenvolvimento do Investimento Social (Fabiani; Altemani; Faleiros; Otani, 2020) utiliza a análise de custo-benefício chamada de Retorno Social do Investimento (SROI) e descreve o cálculo utilizado, como pode ser visto:

As *proxies* financeiras, dessa forma, são utilizadas para estimar o valor social de bens, que embora não sejam exatos, são suficientemente adequados para a avaliação da mudança social gerada por determinado Programa. Esses valores são estimados para diferentes grupos de pessoas (*stakeholders*), uma vez que, por se tratar de um valor subjetivo, diferentes *stakeholders* poderão ter percepções diferentes a respeito do valor atribuído a

alguns benefícios. Ao estimar esse valor por meio de *proxies* financeiras e combinar essas valorações, é possível estimar o valor social total criado por uma intervenção, a partir do cálculo da combinação de incidência dos resultados (o número de pessoas impactadas e a intensidade das mudanças) com os valores monetários atribuídos a estas mudanças, definidos através das *proxies* financeiras (Fabiani; Altemani; Faleiros; Otani, 2020. p. 38).

Contudo, o relatório apresentado pelo Instituto para o Desenvolvimento do Investimento Social (Fabiani; Altemani; Faleiros; Otani, 2020) não menciona os custos da avaliação realizada, somente os benefícios gerados pelo Programa, conforme cálculo baseado no SROI, conforme descrito acima. Dessa forma, concluiu-se que o padrão foi atendido parcialmente.

Por fim, o atributo **Exequibilidade** foi considerado pelas autoras como não atendido, porque somente um de seus padrões foi atendido, de acordo com os critérios estabelecidos neste estudo avaliativo.

Os padrões de adequação, segundo Elliot (2011), devem assegurar que a avaliação respeitará aspectos morais, éticos e legais relativos à qualidade da avaliação: “esses padrões são relativos ao bem-estar das pessoas envolvidas na avaliação ou que possam ser afetadas por seus resultados” (Elliot, 2011. p. 952), conforme Quadro 4:

Quadro 4 – Padrões do atributo de adequação

Padrões	Características
A1 - Orientação Responsiva e Inclusiva	As avaliações devem ser responsivas às partes interessadas e suas comunidades.
A2 - Acordos Formais	Os acordos de avaliação devem ser negociados para tornar as obrigações explícitas e levar em consideração as necessidades, expectativas e contextos culturais dos clientes e outras partes interessadas.

Padrões	Características
A3 - Direitos e Respeito Humanos	As avaliações devem ser elaboradas e conduzidas para proteger os direitos humanos e legais e manter a dignidade dos participantes e outras partes interessadas.
A4 - Clareza e Equidade (Justiça)	As avaliações devem ser compreensíveis e justas ao abordar as necessidades e propósitos das partes interessadas.
A5 - Transparência e Abertura	As avaliações devem fornecer descrições completas das descobertas, limitações e conclusões a todas as partes interessadas, a menos que isso viole obrigações legais e de propriedade.
A6 - Conflitos de Interesses	As avaliações devem identificar e abordar de forma aberta e honesta conflitos de interesses reais ou percebidos que possam comprometer a avaliação.
A7 - Responsabilidade Fiscal	As avaliações devem contabilizar todos os recursos gastos e cumprir procedimentos e processos fiscais sólidos.

Fonte: As autoras (2024) adaptado de Yarbrough, Shulha, Hopson e Caruthers (2011).

Yarbrough, Shulha, Hopson e Caruthers (2011) descrevem que as avaliações responsivas dedicam atenção às necessidades e às características das partes interessadas e aos seus contextos, conforme explicitado abaixo:

Uma avaliação inclusiva reconhece todos os *stakeholders* e seus contextos, necessidades e experiências – e não somente um grupo deles. Assim, uma avaliação que seja responsiva apenas para alguns *stakeholders* e contextos pode não ser adequadamente inclusiva (Yarbrough; Shulha; Hopson; Caruthers, 2011, p. 35).

De acordo com os autores, uma **Orientação responsiva e inclusiva** (A1) produz consequências éticas e morais de longo alcance, e servem para melhorar os outros padrões do *Joint Committee*: “este padrão reconhece que, em uma sociedade democrática, os *stakeholders* devem ser incluídos e atendidos, independentemente de opção política, status ou poder”. Yarbrough, Shulha, Hopson e Caruthers

(2011) afirmam reconhecer limitações impostas pelas circunstâncias, quando nem todas as partes interessadas consigam participar com o mesmo grau de inclusão ou receber igual atenção.

A RBMA descreve a diversidade e a equidade como a sua nona diretriz, segundo a qual:

As avaliações devem assegurar tratamento equânime a todos os atores envolvidos em seus processos, valorizando suas necessidades, posições e demandas, respeitando e adequando-se a sua condição étnico-racial, de gênero, social, física, geracional, econômica ou cultural (Silva; Joppert; Gasparini; 2020, p. 17).

No que concerne ao padrão **Orientação responsiva e inclusiva**, o relatório de Fabiani, Altemani, Faleiros e Otani (2020) aponta que a avaliação privilegiou a análise de apenas duas partes interessadas em detrimento de outras oito: alunos e professores. Concluiu-se, então, que a avaliação atendeu parcialmente ao padrão A1, já que não houve a inclusão de todos os interessados.

Em referência ao padrão **Acordos formais** (A2), Yarbrough, Shulha, Hopson e Caruthers (2011) esclarecem que os acordos formais na avaliação são os contratos que especificam obrigações e vinculam os avaliadores às partes responsáveis pelo processo:

De modo geral, acordos formais explicitam obrigações e condições éticas, legais e profissionais, e em muitos casos servem como documentos. (...) Acordos formais exigem processos de comunicação e resultam em produtos palpáveis. O processo de elaboração de acordos formais pode favorecer a credibilidade e destacar o profissionalismo do avaliador (U1 Credibilidade do Avaliador), além de contribuir para revelar diferenças de pensamento e influências culturais que, de outro modo, permaneceriam ocultas (Yarbrough; Shula; Hopson; R.K.; Caruthers, 2011).

Os autores mencionam, ainda, que os avaliadores devem renegociar os acordos sempre que for necessário, particularmente

se houver alteração na abrangência do trabalho. Além disso, acordos verbais também podem ser considerados válidos.

O relatório do Instituto para o Desenvolvimento do Investimento Social aponta que durante a pandemia de COVID-19 foi acordado o desenvolvimento de uma etapa de coleta piloto de dados quantitativos junto aos alunos de 10 e 11 anos, mais quatro centrais participantes do programa (Fabiani; Altemani; Faleiros; Otani, 2020). Não há menção ao acordo ser verbal ou expresso. Contudo, não há qualquer outra menção a acordos no relatório. Concluiu-se que o padrão **Acordos Formais** não foi considerado atendido.

No que tange a **Direitos Humanos e Respeito** (A3), Yarbrough, Shulha, Hopson e Caruthers (2011) mencionam que este padrão faz com que os avaliadores sejam elevados a um grau de compromisso com o respeito e os direitos humanos superior ao que é exigido pelos legisladores, por agências externas ou por conselhos de ética, durante o processo avaliativo. Segundo os autores:

A proteção dos direitos humanos e da legalidade, bem como o respeito pela dignidade dos participantes, tem de ser parte integrante dos processos e produtos da avaliação. Os procedimentos, por exemplo, devem incentivar interações justas com os *stakeholders* e entre eles. Os participantes devem ser informados de seus direitos, inclusive de participar, de desistir de participar ou de questionar decisões tomadas em qualquer etapa do processo de avaliação. Além disso, por delicadeza, o avaliador precisa familiarizar-se com os usos e costumes na instituição ou comunidade, em especial com aqueles inteiramente novos para ele, de modo que suas práticas reflitam atenção e valorização. Direitos individuais, costumes, convenções sociais e maneiras de expressar respeito às vezes são determinados pelos contextos culturais dos participantes, da organização ou da comunidade (Yarbrough; Shulha; Hopson; Caruthers, 2011).

A respeito da ética no comportamento do avaliador, o Guia de princípios da *American Evaluation Association* cita que:

É, em primeiro lugar, da responsabilidade do avaliador iniciar a discussão e o esclarecimento de questões éticas com as partes relevantes para a avaliação. Os princípios podem ser usados para comunicar aos clientes e outras partes interessadas o que podem esperar em termos do comportamento ético profissional de um avaliador (Australasian Evaluation Association, 2013, p. 3).

Davidson (2005, p. 205) define como um dos critérios fundamentais para orientar a meta-avaliação a condução das avaliações de “maneira ética, legal, profissional e apropriada”. O relatório do Instituto para o Desenvolvimento do Investimento Social (Fabiani; Altemani; Faleiros; Otani, 2020), por sua vez, não cita proteção de direitos humanos ou de direitos legais, tampouco dignidade dos participantes ou das partes interessadas. Portanto, conclui-se que o padrão não foi atendido.

Yarbrough, Shulha, Hopson e Caruthers (2011) apontam que **clareza e equidade** (A4) são componentes que interagem, devendo as avaliações serem claras e justas quando tratarem dos propósitos e das necessidades das partes interessadas. Os autores ressaltam que, “para que haja equidade, é preciso atenção adaptável às necessidades de grupos e indivíduos comparadas aos custos, garantindo ao mesmo tempo o direito de todos”.

Embora o relatório do Instituto para o Desenvolvimento do Investimento Social (Fabiani; Altemani; Faleiro; Otani, 2020) mencione que o SROI tem a característica de medir o impacto que foi vivenciado de fato pelas partes interessadas e o que é interessante para esses grupos, só foram levados em consideração dois dos dez grupos de partes interessadas do Programa. Ainda assim, sobre o SROI, o Instituto para o Desenvolvimento do Investimento Social explica que:

Um aspecto chave desse protocolo é seu foco na percepção do beneficiário – o envolvimento dos *stakeholders* é um dos princípios da SROI, o que significa que o impacto social deve ser avaliado a partir do ponto de vista daqueles que estão diretamente envolvidos no projeto

social (Instituto para o Desenvolvimento do Investimento Social, 2018, p. 2).

Da mesma forma, Fabiani, Altemani, Faleiro e Otani (2020) não mencionam se as avaliações foram inteligíveis para as partes. Portanto, concluiu-se que o item não foi atendido com relação ao estabelecido pelo *Joint Committee*.

De acordo com Yarbrough, Shulha, Hopson e Caruthers (2011), **Transparência e abertura** (A5) é um padrão que objetiva o fornecimento de informações confiáveis, completas e honestas às partes interessadas, inclusive a revelação dos principais patrocinadores da avaliação. Mas, essas informações necessitam da sensibilidade do avaliador na divulgação de dados ou da não divulgação, conforme as previsões legais.

A abertura se refere à comunicação franca com as partes interessadas do que elas têm direito a saber em todos os ciclos de avaliação, desde que não haja violação de legislação ou de códigos éticos ou profissionais. Já a transparência expõe as limitações da metodologia, da lógica, do projeto e da análise das avaliações. “O clima de transparência favorece a aceitação, a credibilidade, a utilização e a precisão” (Yarbrough; Shulha; Hopson; Caruthers, 2011, p. 35).

O relatório do Instituto para o Desenvolvimento do Investimento Social (Fabiani; Altemani; Faleiro; Otani, 2020, p. 15) explicita que “ser transparente” é um dos princípios do Protocolo SROI, mas não esclarece como é feita a transparência e a abertura dos dados às partes interessadas. Portanto, concluiu-se que o padrão **Transparência e abertura** (A5) não foi atendido quando analisado sob a ótica dos padrões do *Joint Committee*.

No que tange a **Conflitos de interesses** (A6), Yarbrough, Shulha, Hopson e Caruthers (2011, p. 29) afirmam que eles são inevitáveis entre os grupos e os indivíduos, mas que devem ser administrados, e não impedidos durante a avaliação: “Frequentemente o

modo como são administrados os conflitos de interesses torna-se mais importante do que a existência deles". Os autores ainda destacam que "nas avaliações, os conflitos de interesses afetam todos os grupos de *stakeholders*, inclusive os avaliadores".

Com relação aos conflitos de interesses entre os grupos de partes interessadas, o relatório do Instituto para o Desenvolvimento do Investimento Social (Fabiani; Altemani; Faleiro; Otani, 2020. p. 29, 40 e 43) cita diversas vezes a "priorização do diálogo como ferramenta para solução de conflitos". No entanto, o documento prioriza a avaliação de apenas dois grupos em detrimento de outros oito, seguindo o Protocolo SROI. Assim, entendeu-se que haveria mais conflitos caso fosse mantida a avaliação de todas as partes interessadas, o que poderia deixar a avaliação do Programa mais complexa.

De toda forma, ao analisar o padrão, percebeu-se que o relatório avaliativo do Programa não atendeu aos padrões, mesmo seguindo o Protocolo SROI, porque todos os conflitos deixaram de ser reconhecidos, retirando das partes interessadas a oportunidade de reconhecê-los e de lidar com eles.

Segundo Yarbrough, Shulha, Hopson e Caruthers (2011), o padrão **Responsabilidade fiscal** (A7) na avaliação do *Joint Committee*:

se refere aos deveres e obrigações para um bom gerenciamento de todos os recursos da avaliação, de modo que sejam efetiva e adequadamente utilizados, sem desperdício. Além disso, responsabilidade fiscal inclui o monitoramento da verdadeira utilização dada aos recursos, de acordo com as melhores práticas de contabilidade e auditoria, e a disponibilização dos relatórios para fins de supervisão (Yarbrough; Shula; Hopson; Caruthers, 2011, p. 35).

Vaitsman e Paes-Sousa (2011, p. 22) explicam que o termo *accountability* (prestação de contas, responsabilidade fiscal) "conota a própria obrigação de se prestar contas, assumida voluntariamente pelo sujeito, sem necessidade da intervenção de um terceiro para

exigi-la. Essa distinção é crucial, pois a obrigação faz parte dos valores e, portanto, da cultura (...). Os autores apontam, ainda, que “a institucionalização de diferentes mecanismos de responsabilidade fiscal, entre os quais a avaliação, tornou-se parte da governança nos Estados democráticos e onde existem diferentes mecanismos de prestação de contas”.

O relatório do Instituto para o Desenvolvimento do Investimento Social (Fabiani; Altemani; Faleiro; Otani, 2020) não menciona prestação de contas ou responsabilidade fiscal. Ele menciona apenas o termo “responsabilidade social”. Portanto, em relação aos padrões do *Joint Committee*, foi verificado o não atendimento pelo relatório de avaliação.

O atributo **Adequação** foi considerado como não atendido, porque seis de seus sete padrões não foram atendidos, de acordo com os critérios estabelecidos neste estudo avaliativo.

Com relação à precisão (Quadro 5), Yarbrough, Shulha, Hopson e Caruthers (2011) explicam que:

A precisão resulta da exatidão das representações, proposições e resultados, em especial aqueles que fundamentam julgamentos acerca da qualidade dos programas ou de seus componentes. Em geral, alcança-se a precisão por meio de teorias, justificativas, métodos e delineamentos solidamente fundamentados. A avaliação deve buscar o máximo de precisão que for exequível, adequada e útil ao embasamento de conclusões e decisões sólidas em uma situação específica (Yarbrough; Shula; Hopson; Caruthers, 2011).

Quadro 5 - Padrões do atributo de precisão

Padrões	Conceito
P1 - Conclusões e Decisões Justificadas	As conclusões e decisões da avaliação devem ser explicitamente justificadas nas culturas e contextos onde têm consequências.
P2- Informação Válida	As informações de avaliação devem atender aos propósitos pretendidos e dar suporte a interpretações válidas.

Padrões	Conceito
P3 - Informação Fidedigna	Os procedimentos de avaliação devem produzir informações suficientemente confiáveis e consistentes para os usos pretendidos.
P4 - Explicitação das Descrições do Propósito e Contexto	As avaliações devem documentar programas e seus contextos com detalhes e escopo apropriados para os propósitos da avaliação.
P5 - Gerenciamento da Informação	As avaliações devem empregar métodos sistemáticos de coleta, revisão, verificação e armazenamento de informações.
P6 - Análises e Planos Sólidos	As avaliações devem empregar projetos e análises tecnicamente adequados e apropriados para os propósitos da avaliação.
P7 - Explicitação do Raciocínio da Avaliação	O raciocínio de avaliação que leva de informações e análises a descobertas, interpretações, conclusões e julgamentos deve ser documentado de forma clara e completa.
P8 - Comunicação e Relatório	As comunicações de avaliação devem ter escopo adequado e proteger contra equívocos, vieses, distorções e erros.

Fonte: As autoras (2024) adaptado de Yarbrough, Shulha, Hopson e Caruthers (2011).

Os autores também ressaltam a importância do atendimento dos padrões de precisão (Quadro 5) pelas avaliações: “nenhuma avaliação consegue eliminar por completo todas as fontes ocasionais ou sistemáticas de imprecisão, mas o objetivo principal dos padrões de precisão é minimizar, na medida do possível, a ocorrência e o impacto de tais fatores” (Yarbrough; Shula; Hopson; Caruthers, 2011).

Yarbrough, Shulha, Hopson e Caruthers (2011) explicam que, com relação ao padrão **conclusões e decisões justificadas** (P1), as conclusões e as decisões da avaliação devem ser explicitamente justificadas nas culturas e nos contextos em que têm consequências, porque:

Ideias e significados expressos em linguagens específicas podem mudar conforme a época, o local, as pessoas e os grupos. Para que as conclusões se mostrem precisas em um sentido prático, com suas consequências justificadas, devem ser entendidas dentro dos contextos nos quais

possam ser aplicados os conceitos e linguagens que as sustentam (Yarbrough; Shula; Hopson; Caruthers, 2011).

Esse entendimento decorre do fato de que um Programa pode passar por variações de verba, de participantes, de necessidade, ou de importância política, não sendo possível a sua reprodução para os períodos seguintes (Yarbrough; Shula; Hopson; Caruthers, 2011). Compreende-se, dessa forma, que nenhum programa é igual, e que as projeções baseadas em um programa não podem ser aplicadas em outro. Por isso, as conclusões e as suas consequências precisam ser justificadas nos contextos e na cultura.

O relatório de Fabiani, Altemani, Faleiros e Otani (2020) cita o contexto de vida dos alunos e o contexto social das comunidades locais, para os quais foi proposta uma metodologia de avaliação conforme o SROI, e encaixada na Teoria da Mudança, citada no relatório. Entendeu-se que o padrão **conclusões e decisões justificadas** (P1) foi atendido pelo relatório do Instituto para o Desenvolvimento do Investimento Social.

De acordo com Yarbrough, Shulha, Hopson e Caruthers (2011), a **Informação Válida** (P2) é definida como a informação da avaliação que deve servir aos propósitos pretendidos e sustentar interpretações válidas. Os autores explicam a importância da interpretação pessoal de experiências:

No cenário da avaliação, diferenças na interpretação de linguagens, experiências e outras informações devem ser explicitadas e examinadas. Qualquer conceito incompleto ou mal explicado pode criar mal-entendidos capazes de prejudicar a precisão.

Felizmente, na maior parte das situações ligadas à avaliação, muitas ideias expostas são básicas e amplamente aceitas, tornando-se o mais próximo do que se pode chamar de verdades universais (Yarbrough; Shula; Hopson; Caruthers, 2011).

Yarbrough, Shulha, Hopson e Caruthers (2011, p. 35) explicam, ainda, que “o consenso sobre o significado exato dos termos utilizados na avaliação geralmente exige comunicação e explicação adicionais” e que “além dos significados conceituais, o significado exato inclui a operacionalidade que conferimos aos termos na prática”.

O relatório de Fabiani, Altemani, Faleiros e Otani (2020, p. 21 e 32) aponta que: de forma alinhada aos princípios do protocolo avaliativo SROI, a Teoria de Mudança do Programa A União Faz a Vida, e os impactos nela descritos, foram desenvolvidos e validados junto aos principais *stakeholders* do Programa, e que “os relatos dos grupos focais corroboraram para a validação da lista dos *stakeholders* materialmente impactados pelo Programa (...)”. O documento destaca ainda, que somente o material referente ao retorno financeiro foi validado junto à Fundação Sicredi (Fabiani; Altemani; Faleiros; Otani, 2020, p. 56).

Dessa forma, entendeu-se que o padrão **Informação Válida** (P2) foi atendido parcialmente, pelo fato de a Teoria de Mudança desenvolvida não ter sido validada por especialistas no assunto.

Com relação a **Informação Fidedigna** (P3), Yarbrough; Shulha; Hopson; Caruthers (2011) definem a fidedignidade como a coerência nas informações coletadas por meios de procedimentos confiáveis. Sendo, portanto, um alicerce à validade da informação. Os autores destacam que o oposto da fidedignidade é o erro aleatório, e que:

A eliminação do erro aleatório tem igual importância para os resumos de informações qualitativas ou quantitativas. Medidas padronizadas, observações, descrições detalhadas, classificações, transcrições de entrevistas e outros tipos de informações qualitativas e quantitativas exigem atenção quanto à consistência e à confiabilidade (Yarbrough; Shula; Hopson; Caruthers, 2011, p. 35).

Worthen, Sanders e Fitzpatrick (2011) mencionam que documentos públicos e bases de dados são importantes fontes de

informações existentes, bem como as partes interessadas. Fabiani, Altemani, Faleiros e Otani (2020) por sua vez, citam os grupos focais, nos quais os beneficiários do *A União Faz a Vida* fazem parte, como a principal fonte de informação para a construção da avaliação deste Programa. E que, com relação ao índice de confiança da amostra resultante da aplicação de questionários com um grupo de professores, foi calculada, também, a margem de erro. Dessa forma, pode-se concluir que o padrão **Informação Fidedigna** (P3) foi atendido.

Segundo Yarbrough, Shulha, Hopson e Caruthers (2011), o padrão **Explicitação das Descrições do Programa e do Contexto** (P4) "recomenda que se esclareça exatamente o que está sendo avaliado e como é feita a contextualização". Os autores destacam que um componente importante da avaliação é a teoria do Programa, que descreve o seu funcionamento. Além disso, Yarbrough, Shulha, Hopson e Caruthers (2011) enfatizam que "a opinião dos *stakeholders*-chave é a principal fonte de informação sobre o programa. A documentação de suas experiências se faz necessária, para descrever o que realmente provocou os resultados e produtos observados".

Fabiani, Altemani, Faleiros e Otani (2020) descrevem o processo de modelagem da Teoria da Mudança de acordo com o Modelo *SROI* de forma detalhada. Pode-se concluir, então, que o padrão **Explicitação das Descrições do Programa e do Contexto** (P4) foi atendido.

Yarbrough; Shulha; Hopson; Caruthers (2011) explicam que o padrão **Gerenciamento da informação** (P5) é referente aos métodos utilizados para garantir seleção, equilíbrio, coleta e armazenamento corretos da informação, e que:

Enfatiza a necessidade de um gerenciamento da informação que mantenha a precisão; discute semelhanças e diferenças entre tipos e métodos de informação; trata de questões ligadas à quantidade de informações a ser coletada, ao armazenamento das informações e de como pensar em métodos sob a perspectiva da precisão (Yarbrough; Shula; Hopson; Caruthers, 2011, p. 35).

O relatório de Fabiani, Altemani, Faleiros e Otani (2020) cataloga e considera as fontes de informação da avaliação, traduzidas nas partes interessadas, embora sejam utilizadas apenas duas na avaliação, em detrimento de outras oito, devido à metodologia *SROI*. Pelo fato de o relatório não abranger todas as partes interessadas como fontes de informação, as autoras entendem que o padrão foi parcialmente atendido.

Ao abordar o padrão **Análises e delineamentos sólidos** (P6), Yarbrough, Shulha, Hopson e Caruthers (2011, p. 35) enfatizam a necessidade da precisão nas análises:

Análises precisas são necessárias também para um raciocínio eficiente, levando às conclusões. Análises apuradas (precisas) permitem que as informações sejam reunidas e reduzidas, descritas, sumarizadas (resumidas) e compreendidas, explicitando suas limitações no contexto das questões e usos específicos da avaliação. Sem análises tecnicamente adequadas, o raciocínio eficiente fica prejudicado.

Fabiani, Altemani, Faleiros e Otani (2020) apontam no relatório que o Programa foi baseado em critérios definidos pelo modelo *SROI*, com adoção de variáveis adotadas na mensuração do valor social gerado. Este modelo é validado pelo *Cabinet Office* do Reino Unido. Ficou entendido, dessa forma, que foram utilizadas análises e delineamentos sólidos, conforme as bases do modelo *SROI*, embora não tenham sido citados no relatório. O padrão **Análises e delineamentos sólidos** (P6), portanto, pode ser considerado como atendido em relação ao proposto pelo *Joint Committee*.

Yarbrough, Shulha, Hopson e Caruthers (2011) mencionam sobre o padrão **Explicitação do raciocínio da avaliação** (P7) que “para chegar a conclusões, avaliações sistemáticas devem ir além de modelos informais, implícitos ou intuitivos. O padrão P7 chama atenção para a necessidade de um raciocínio sólido e explícito, como base de resultados, conclusões e julgamentos”. Os autores também

citam o mapeamento inversos para a identificação dos componentes do raciocínio: “Uma técnica às vezes bastante útil à identificação dos componentes do raciocínio e das evidências que levaram a determinados tipos de conclusão é o mapeamento inverso”.

O relatório de Fabiani, Altemani, Faleiros e Otani (2020) aponta a utilização do modelo *SROI* como base da linha de raciocínio para a avaliação e para a conclusão e as recomendações. Ficou entendido, portanto, que o padrão **Análises e delineamentos sólidos** (P6) foi atendido.

Com relação ao padrão **Comunicação e relatório** (P8), Yarbrough, Shulha, Hopson e Caruthers (2011, p. 35) enfatizam que as avaliações precisas são construídas sobre a comunicação precisa:

Todas as fases da avaliação necessitam de uma comunicação precisa. (...) Na falta de instruções e explicações precisas durante a coleta de dados, por exemplo, os entrevistados podem não entender o que as tarefas querem deles. A comunicação precisa é necessária para sustentar o raciocínio da avaliação, levando a resultados confiáveis, e a interpretações e conclusões que atendam aos propósitos da avaliação.

O relatório de Fabiani, Altemani, Faleiros e Otani (2020) não menciona a redação ou o uso de relatórios formais ou informais nos estágios de avaliação, embora enfatize o diálogo entre as partes interessadas na solução de problemas. Por tanto, concluiu-se que não foi atendido o padrão Comunicação e Relatório.

O atributo de Precisão foi considerado como não atendido, porque apenas três de seus sete padrões não foram atendidos, de acordo com os critérios estabelecidos neste estudo avaliativo.

Yarbrough, Shulha, Hopson e Caruthers (2011) definem o atributo **responsabilização** como o uso responsável de recursos nas avaliações de programas, a fim de produzir valor: “a documentação e a melhoria da responsabilização da avaliação exigem esforços

semelhantes àqueles exigidos pela responsabilização do programa. É preciso uma avaliação da avaliação – a “meta-avaliação”. Os três padrões de responsabilização orientam o processo de meta-avaliação (Quadro 6):

Quadro 6 - Padrões de responsabilização

Padrões	Conceito
R1 - Documentação da Avaliação	As avaliações devem documentar completamente seus propósitos negociados e projetos, procedimentos, dados e resultados implementados.
R2 - Meta-avaliação Interna	Os avaliadores de meta avaliação interna devem usar esses e outros padrões aplicáveis para examinar a responsabilização do projeto de avaliação, os procedimentos empregados, as informações coletadas e os resultados.
R3 - Meta-avaliação Externa	Os patrocinadores, clientes, avaliadores e outras partes interessadas do Programa de Meta avaliação Externa devem incentivar a realização de meta avaliações externas usando esses e outros padrões aplicáveis.

Fonte: As autoras (2024) adaptado de Yarbrough, Shulha, Hopson e Caruthers (2011).

Yarbrough, Shulha, Hopson e Caruthers (2011, p. 35) definem o atributo **responsabilização** como o uso responsável de recursos nas avaliações de programas, a fim de produzir valor: “a documentação e a melhoria da responsabilização da avaliação exigem esforços semelhantes àqueles exigidos pela responsabilização do programa. É preciso uma avaliação da avaliação – a “meta-avaliação”. Os três padrões de responsabilização orientam o processo de meta-avaliação.

Com relação ao padrão **Documentação da avaliação** (R1), Yarbrough, Shulha, Hopson e Caruthers (2011) mencionam que a documentação da avaliação e seus contextos em detalhe e abrangência adequados são fundamentais para formar uma base para o julgamento da qualidade da avaliação. Essa documentação favorece a responsabilização.

O relatório de Fabiani, Altemani, Faleiros e Otani (2020) afirma que a documentação do Programa União Faz a Vida foi responsável pela identificação das partes interessadas da avaliação externa realizada pelo Instituto para o Desenvolvimento do Investimento Social, juntamente com as entrevistas iniciais feitas com a equipe da Fundação Sicredi. Observa-se que essa é a única menção à documentação neste relatório, sem detalhes ou aprofundamento. Foi definido que o padrão é considerado como atendido parcialmente.

Yarbrough, Shulha, Hopson e Caruthers (2011, p. 33) definem o padrão **Meta-avaliação interna** (R2) como um complemento da **Meta-avaliação externa** (R3), mas ressaltam as suas diferenças:

A distinção entre interno e externo existe tanto para avaliadores quanto para meta-avaliadores. Em geral, define-se “avaliador interno” como alguém que é empregado da organização onde é aplicado o programa. O “avaliador externo”, ao contrário, é empregado de outra organização e não tem outras funções no programa avaliado. Avaliadores externos podem trabalhar em colaboração com avaliadores internos e com a equipe do programa, sem que sua situação ou independência sejam afetadas. A colaboração entre avaliadores internos e externos produz avaliações combinadas, com uma perspectiva ao mesmo tempo interna e externa.

O relatório de Fabiani, Altemani, Faleiros e Otani (2020) se trata de uma meta-avaliação externa, tendo sido realizada pelo Instituto para o Desenvolvimento do Investimento Social, com base no modelo *SROI*. Observa-se que não há qualquer menção à realização de uma avaliação interna ou eexterna. Dessa forma, julgou-se que os padrão **Meta-avaliação interno** (R2) não foi atendido, enquanto o padrão **Meta-avaliação externa** (R3) é considerado como Atendido.

As autoras concluíram que, de acordo com os critérios estabelecidos neste estudo, o atributo **responsabilização** foi considerado como não atendido.

CONCLUSÕES E RECOMENDAÇÕES

Com referência aos resultados da meta-avaliação obtidos do estudo, respondendo à questão avaliativa em que ponto a avaliação descrita no Relatório da Avaliação do Retorno Social do Investimento para o Programa A União Faz a Vida obedeceu aos padrões de avaliação definido pelo *The Joint Committee on Standards for Educational Evaluation (JCSEE)*, dos cinco atributos avaliados do *Joint Committee* (utilidade, exequibilidade, adequação, precisão e responsabilização), segundo os critérios elaborados pelas autoras, apenas o atributo utilidade foi considerado como parcialmente atendido.

Os demais foram considerados como não atendidos conforme os critérios estabelecidos neste estudo. É importante observar que diversos padrões foram atendidos parcialmente, com alguns ajustes a se fazer. Conclui-se, portanto, que o relatório atendeu parcialmente aos atributos e padrões estabelecidos para avaliações pelo *Joint Committee*.

Como pontos fortes do relatório de avaliação destacam-se os esclarecimentos dos procedimentos e conceitos sobre métodos utilizados na Avaliação de Impacto como Teoria da Mudança e o protocolo *Social Return On Investment (SROI)*, informações estas que foram fundamentais para o entendimento da metodologia aplicada. As etapas também foram muito detalhadas.

Atualmente existem muitas discussões entre associações profissionais e organismos internacionais sobre dilemas éticos e suas soluções, o que coloca a ética como essencial nas avaliações:

A ética é uma questão fundamental para a avaliação. Sem uma cultura ética sólida, a credibilidade e confiança nos resultados da avaliação sairão necessariamente prejudicadas. Neste sentido, é fundamental definir os princípios éticos e respectivas regras de conduta que devem reger a atuação de todos os atores envolvidos na função avaliação (Amaral, 2013, p. 1).

Com base nas orientações dos códigos de ética direcionados para os avaliadores, pode-se apontar no estudo alguns procedimentos éticos que não foram seguidos e que prejudicaram o padrão avaliado, como orienta os códigos de ética. Como exemplo, se pode citar a informação dos riscos e benefícios prováveis decorrentes de sua participação, a elaboração de um Termo de Consentimento Livre e Esclarecido (TCLE) por escrito e o compartilhamento de resultados com as partes interessadas, inclusive com relatórios ao longo da avaliação. Sugere-se que os custos decorrentes do processo avaliativo sejam explicitados, para demonstrar transparência. É muito importante o compartilhamento de resultados com as partes interessadas, inclusive com relatórios ao longo da avaliação.

Foi relatado que as pessoas que participaram das entrevistas para a identificação dos *stakeholders* foram selecionadas pelo próprio Sicredi. Além disso, o Instituto para o Desenvolvimento do Investimento Social deveria ter escolhido os participantes com base em critérios que incluíssem diversos grupos de interesses diferentes, para oferecer maior credibilidade à avaliação. Da mesma forma, a instituição deveria ter optado por analisar mais grupos de *stakeholders*, como, por exemplo, as pessoas que participaram do programa. Outro ponto de melhoria seria a formulação de um cronograma das etapas e de seus prazos.

Mostra-se necessário que seja aplicada periodicamente outras avaliações para que se tenha retorno das opiniões dos maiores interessados na avaliação, possibilitando uma melhoria continuada do Programa a União faz a Vida e que seus padrões avaliativos sejam revistos pelos avaliadores responsáveis pelo planejamento e implementação do processo avaliativo.

MÉRITO DO ESTUDO

Ovidio Orlando Filho

A meta-avaliação é um processo estruturado que tem como objetivo analisar e medir a qualidade de uma avaliação já concluída. Essa abordagem permite verificar se a avaliação seguiu critérios adequados, garantindo que os seus resultados sejam precisos e relevantes.

Para que a meta-avaliação seja eficaz, é necessário considerar aspectos como a clareza dos objetivos da avaliação original, a adequação dos métodos utilizados, a qualidade dos dados coletados e a coerência das conclusões. Além disso, a transparência do processo e a forma como os resultados foram comunicados também são analisadas.

Dentre as metodologias mais utilizadas na meta-avaliação, destaca-se o uso de padrões estabelecidos, como os propostos pelo *Joint Committee on Standards for Educational Evaluation*, que definem critérios para a qualidade das avaliações em diversas áreas. Esses critérios incluem atributos como: a utilidade (se a avaliação atende às necessidades dos usuários), a viabilidade (se foi conduzida de forma eficiente), a adequação (se respeitou princípios éticos) e a exatidão (se os resultados refletem a realidade avaliada).

Os benefícios da meta-avaliação são diversos. Ela permite identificar falhas e pontos de melhoria no processo avaliativo, contribuindo para a realização de avaliações futuras mais robustas e confiáveis. Além disso, fortalece a credibilidade das avaliações realizadas e assegura que suas conclusões sejam utilizadas de forma adequada para a tomada de decisões.

Desse modo, a meta-avaliação se constitui em uma importante ferramenta estratégica que garante a qualidade e a efetividade

das avaliações, promovendo um ciclo contínuo de aperfeiçoamento nos processos avaliativos.

O artigo foi adequadamente elaborado, considerando-se, principalmente se tratar de um estudo meta-avaliativo somativo, que possui a principal característica de trilhar seu caminho metodológico pautado em critérios que visam aferir boas práticas avaliativas postas em práticas em avaliações já encerradas, por meio do estudo de sua documentação.

O seu título reflete bem o estudo, a avaliação realizada originalmente é apresentada de forma sucinta, porém com detalhes que facilitam seu entendimento, sua metodologia é clara e criteriosamente bem elaborada, é bem fundamentado, citando autores de referência sobre o assunto para estabelecer argumentos, é coerente na organização das ideias e no desenvolvimento do raciocínio e a questão do estudo é coerentes com o objetivo.

REFERÊNCIAS

AMARAL, Pedro. **Avaliação de impacto**: breve introdução. Portugal: Camões, Instituto de Cooperação e da Língua, 2013. Disponível em: https://www.instituto-camoes.pt/images/cooperacao/doc_trabalho8_2013.pdf. Acesso em: 2 set. 2024.

AUSTRALASIAN EVALUATION SOCIETY. **Guidelines for the ethical conduct of evaluations**. Austrália: AES, 2013. Disponível em: https://www.aes.asn.au/images/AES_Guidelines_web_v2.pdf. Acesso em: 17 set. 2024.

BARKI, Edgard; TORRES, Haroldo da Gama; BARROS, Octavio Augusto de. **Teoria da mudança**: o que é e para que serve? Bibliotecas Sebrae, [S. l.], 2023. Disponível em: [https://bibliotecas.sebrae.com.br/chronus/ARQUIVOS_CHRONUS/bds/bds.nsf/a41351890935edd6be3e90387966f7ee/\\$File/31806.pdf](https://bibliotecas.sebrae.com.br/chronus/ARQUIVOS_CHRONUS/bds/bds.nsf/a41351890935edd6be3e90387966f7ee/$File/31806.pdf). Acesso em: 4 set. 2024.

BRASIL. Tribunal de Contas da União. **Política pública em dez passos**. Brasília: TCU, 2021. Disponível em: https://portal.tcu.gov.br/data/files/1E/D0/D4/DF/12F99710D5C6CE87F18818A8/Politica%20Publica%20em%20Dez%20Passos_web.pdf. Acesso em: 2 out. 2024.

CAVALCANTI, Mônica Maria de Arruda. Avaliação de políticas públicas e programas governamentais: uma abordagem conceitual. **Revista Interfaces de Saberes**, Caruaru, 2006. Disponível em: <https://sinapse.gife.org.br/download/avaliacao-de-politicas-publicas-e-programas-governamentais-uma-abordagem-conceitual>. Acesso em: 3 set. 2024.

BRASIL. [Constituição (1988)]. Constituição da República Federativa do Brasil. [Diário Oficial da República Federativa do Brasil], Brasília, DF, 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 3 set. 2024.

BRASIL. Ministério da Educação. **Organização para a Cooperação e Desenvolvimento Econômico – OCDE**. Brasília, DF, 2018. Disponível em: http://portal.mec.gov.br/index.php?option=com_contentview=articleid=20746&Itemid=1319. Acesso em: 30 ago. 2024.

DAVIDSON, E. Jane. **Evaluation methodology basics**: the nuts and bolts of sound evaluation. Thousands Oaks, CA: Sage, 2005.

ELLIOT, Ligia Gomes. Meta-avaliação: das abordagens às possibilidades de aplicação. **Ensaio**: aval. pol. públ. educ., Rio de Janeiro, v. 19, n. 73, p. 941-964, 2011.

FABIANI, Paula; ALTEMANI, Raquel; FALEIROS, Laís; OTANI, Ana Paula Lie. **Avaliação do retorno social do investimento do Programa A União Faz a Vida da Fundação Sicredi**. [S. l.]: IDIS, 2020. Disponível em: https://idis.org.br/files/RelatorioSROI_SICREDI.pdf. Acesso em: 2 set. 2024.

FABIANI, Paula Jancso; CARVALHO, Denise de Freitas. Avaliação de impacto para inclusão produtiva: metodologias, desafios e limitações. **Revista Brasileira de Avaliação**, [S. l.], v. 12, n. 2 especial, 2023. DOI: <http://dx.doi.org/10.4322/rbaval202312011>. Disponível em: <https://rbaval.org.br/article/10.4322/rbaval202312011/pdf/rbaval-12-2+spe-e121123.pdf>. Acesso em: 2 set. 2024.

INSTITUTO PARA O DESENVOLVIMENTO DO INVESTIMENTO SOCIAL. **Avaliando o impacto social dos cursos técnicos CEAP**. São Paulo: IDIS, 2018. Disponível em: <https://www.idis.org.br/publicacoesidis/avaliando-o-impacto-social-dos-cursos-tecnicos-ceap/>. Acesso em: 12 set 2024.

JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION. Sobre o JCSEE. **Evaluationstandards.org**, [S. l.], 2020. Disponível em: <https://evaluationstandards.org/>. Acesso em: 30 ago. 2014.

JOPPERT, Márcia Paterno; BRAZ, Márcia Malaquias; QUEIROZ, Adriana; LAMAS, Bárbara; GRIPP, Demétrio. **Competências para a prática de avaliação no Brasil**: documento síntese. Rio de Janeiro: Rede Brasileira de Monitoramento e Avaliação, 2024. Disponível em: https://drive.google.com/file/d/1rT-KY_iLuxDpRD5f5F_L2qPJVV3UWhf/view. Acesso em: 13 set. 2024.

OLIVEIRA, Flávia Cristina Gomes. **Avaliação dos resultados da política de compensação ambiental nas unidades de conservação federais**. 2016. 32 f. Trabalho de Conclusão de Curso (Especialista em gestão Pública) – Escola nacional de Administração Pública, Brasília, 2016. Disponível em: <https://repositorio.enap.gov.br/jspui/bitstream/1/2468/1/FI%c3%a1via%20Cristina.pdf>. Acesso em: 13 set. 2024.

OLIVEIRA, Ivan dos Santos; ROTHEN, José Carlos. Meta-avaliação de aspectos emancipatórios: uma proposta teórico-metodológica. **Estudos em Avaliação Educacional**, São Paulo, v. 35, 2024. DOI: <https://doi.org/10.18222/eae.v35.10631>. Disponível em: <https://publicacoes.fcc.org.br/eae/article/view/10631>. Acesso em: 1 out. 2024.

PARENTE, Cláudia da Mota Darós; PARENTE, Juliano Mota; HERNANDES, Elianeth Dias Kanthack. Avaliação de impacto na educação Básica. **Jornal de Políticas Educacionais**, Brasília, DF, v. 14, 2020. DOI: <https://doi.org/10.5380/jpe.v14i0.71054>. Disponível em: <http://educa.fcc.org.br/pdf/jpe/v14/1981-1969-jpe-14-e71054.pdf>. Acesso em: 28 ago. 2024.

FUNDAÇÃO SICRED. **Impacto social**: cooperação que promove impacto e transformação. [S. l.]: Sicred, [2021]. Disponível em: https://www.auniaoofazavida.com.br/impacto/estudo/Revista_Impactos%20PUFV_pagina.pdf. Acesso em: 3 set. 2024.

SCRIVEN, Michael. **Evaluation thesaurus**. Newbury Park: Sage Publications, 1991.

A UNIÃO faz a vida. **Fundação Sicredi**, [S. l.], [2024]. Disponível em: <https://www.sicredi.com.br/site/fundacao/uniao-faz-vida/>. Acesso em: 28 ago. 2024.

THE SROI NETWORK. **A guide to social return on investment**. [S. l.]: SROI, 2012. Disponível em: <https://static1.squarespace.com/static/60dc51e3c58ae4f43ae5c975/t/60f7fa286b9c6a47815bc3b2/1626864196998/The-SROI-Guide-2012.pdf>. Acesso em: 15 set. 2024.

SOCIAL VALUE INTERNATIONAL. **The principles of social value**. [S. l.]: Social Value International, 2021. Disponível em: https://socialvalueuk.org/wp-content/uploads/2022/12/Principles_of_Social_Value.pdf. Acesso em: 25 set. 2024.

STUFFLEBEAM, Daniel L. The metaevaluation imperative. **American Journal of Evaluation**, Thousand Oaks, v. 22, n. 2, p. 183-209, 2001. DOI: <https://doi.org/10.1177/109821400102200204>. Disponível em: <https://journals.sagepub.com/doi/10.1177/109821400102200204>. Acesso em: 10 jun. 2022.

YARBROUGH, Donald B.; SHULLA, Lyn M.; HOPSON, Rodney K.; CARUTHERS, Flora A. **The program evaluation standards**: a guide for evaluators and evaluation users. 3. ed. Thousands Oaks, Califórnia: Sage, 2011.

SILVA, Rogério R.; JOPPERT, Márcia P.; GASPARINI, Max F. V. (org.). **Diretrizes para a prática de avaliação no Brasil**. Rio de Janeiro: Rede Brasileira de Monitoramento e Avaliação, 2020. Disponível em: https://rbma.site/wp-content/uploads/2021/02/Diretrizes-para-a-pratica-de-avaliacao-no-Brasil_20ago20.pdf. Acesso em: 8 jan. 2025.

WORTHEN, Blaine R.; SANDERS, James R.; FITZPATRICK, Jody L. **Avaliação de programas: concepções e práticas**. São Paulo: Editora Gente, 2004.

WOODSIDE; Arch G.; SAKAI; Marcia Y. Meta-evaluations of performance audits of government tourism-marketing programs. **Journal of Travel Research**, [S. l.], v. 39, n. 4, 2001. DOI: <https://doi.org/10.1177/004728750103900403>. Disponível em: <http://jtr.sagepub.com/cgi/content/abstract/39/4/369>. Acesso em: 14 out. 2024.

7

*Sheila Regina da Rocha de Magalhães
Luciana Vieira Santos
Lucí Hildenbrand*

META-AVALIAÇÃO DA AVALIAÇÃO DE IMPACTO SOCIAL DO PROJETO PALCOS PARA A INCLUSÃO FINANCIADO PELO PROGRAMA OPERACIONAL PARA A INCLUSÃO SOCIAL E EMPREGO DE PORTUGAL

INTRODUÇÃO

O contexto da crise econômica e social que Portugal vem enfrentando nas últimas décadas tem apresentado desafios relacionados às intervenções públicas nos domínios do combate à pobreza, da promoção da inclusão social e do emprego. Nesta perspectiva, segundo o Programa Operacional para a Inclusão Social e Emprego (POISE), de Portugal, tem-se assistido a um aumento da exposição de grande contingente de pessoas a situações de risco de pobreza e/ou exclusão social, decorrentes, em primeira instância, da menor participação dos sujeitos no mercado de trabalho, resultante do aumento do desemprego e da sua duração.

Em verdade, o aumento do desemprego, tanto em sua componente conjuntural como estrutural, aliado aos baixos níveis de educação escolar e de qualificação da população para o trabalho, tem levado a uma maior fragilização de vários segmentos da sociedade portuguesa. Portugal continua, assim, a apresentar elevados níveis de pobreza monetária e de exclusão social, conforme atestam dados do POISE.

Para o alcance das metas traçadas no sentido de se atingir seus objetivos sociais com a devida qualidade, os governos, apoiados em colaborações de especialistas, traçam políticas públicas que são direcionadas para o atendimento das necessidades cotidianas e emergentes nas sociedades. Em geral, tais políticas são implementadas por meio de programas e projetos que visam, de forma sistemática, atender às necessidades de grupos particulares - pessoas, organizações, instituições - que formam a base que sustentam o desenvolvimento das nações.

No POISE, as Iniciativas de Inovação e Empreendedorismo Social são definidas como projetos que preconizam respostas inovadoras que se distinguem das respostas tradicionais na resolução de

problemas sociais pelo seu potencial de impacto e sustentabilidade. Nele, os investidores sociais são identificados como entidades privadas, públicas ou da economia social, com objetivos filantrópicos ou comerciais, que contribuem com recursos financeiros para o desenvolvimento de uma iniciativa de inovação e empreendedorismo social, objetivando favorecer o impacto social. Ou seja, criar valor para a sociedade a partir de um projeto ou de uma intervenção, deduzindo os custos de oportunidade dos recursos utilizados POISE.

Neste contexto, a tipologia de Parcerias para o Impacto apresenta duplo objetivo.

Estimular a criação, desenvolvimento e/ou crescimento de iniciativas de inovação e empreendedorismo social de elevado potencial de impacto, promovendo a sua robustez operacional e financeira e dinamizar a prática de investimento social ao alavancar financiamento privado ou público de investidores sociais e aprofundar a sua vocação de filantropia de impacto, a qual implica o apoio financeiro plurianual a iniciativas, disponibilizando acompanhamento e requerendo medição de resultados e promoção da sua sustentabilidade financeira (Faria 2003, p. 5).

Autores especializados em políticas públicas, como Faria (2003) e Souza (2006), afirmam que as políticas públicas eram consideradas quase exclusivamente destinadas a responder às necessidades sociais definidas pelos sistemas políticos. A atenção nos processos de implementação dos programas e projetos dessas políticas era, pois, centrada nas demandas e articulações de interesses e luta pelo poder político. E, no caso, a avaliação se resumia naquilo que os governos faziam ou deixavam de fazer (Dye, 1984), não sendo à época relevante se avaliar o modo como faziam.

Entretanto, a partir das últimas décadas do século XX, emergiu, inicialmente no campo educacional, uma preocupação que rápida e fortemente se expandiu para diferentes contextos sócio-organizacionais. Por meio dessa tendência, ficou patente a necessidade de se avaliar aquilo que os governos realiza(va)m, favorecendo

a prestação de contas sobre o modo pelo qual se utiliza(va)m dos recursos públicos na implementação das políticas públicas. Passou a existir, então, não só uma preocupação com o uso da avaliação dos programas e projetos, implementados em função dessas políticas, mas também com o posicionamento das partes interessadas (*stakeholders*) em relação aos resultados alcançados.

Desse modo, considera-se que tanto os programas de políticas públicas quanto as suas avaliações devem possuir a necessária qualidade, em termos da eficiência e eficácia de suas conduções (mérito), assim como haver utilidade e efetividade (valor) de seus resultados.

Orlando Filho e Oliveira (2019, p. 154) corroboram esse pensamento, afirmando que:

A avaliação de uma política pública necessita se tornar uma prática consagrada, incorporando uma cultura institucional possuidora de conceitos adequados a cada situação, metodologias específicas e de sistemas de informações desenvolvidos, visando sustentar a implantação de um sistema contínuo de avaliações de qualidade. Deve-se avaliar não somente os programas, mas, também, os processos avaliativos postos em práticas, objetivando aferir-lhes a qualidade. Ou seja, as avaliações também necessitam ser avaliadas.

Como forma de atestar a qualidade do Relatório de Avaliação de Impacto Social do Projeto Palcos para a Inclusão e/ou de evidenciar pontos ainda passíveis de aprimoramento, na sequência, o presente estudo apresenta a etapas metodológicas que cumpriu para atingir a esses propósitos: conceituação de meta-avaliação e sua importância para aferir a qualidade das avaliações; definição do objeto e do objetivo do estudo em pauta; descrição da metodologia meta-avaliativa baseada nos critérios de Davidson (2005); apresentação dos resultados alcançados, das conclusões e recomendações.

A META-AVALIAÇÃO E A SUA IMPORTÂNCIA

A avaliação sistematizada e formal é essencial ao desenvolvimento de qualquer programa ou projeto, tendo em vista sua finalidade: determinar o valor, o mérito ou a qualidade do objeto avaliado, mediante a “identificação, esclarecimento e aplicação de critérios defensáveis” (Worthen; Sanders; Fitzpatrick, 2004, p. 35). Por meio das práticas avaliativas, é possível verificar pontos fortes e fracos do avaliando, tendo em vista a tomada de decisões no que diz respeito às recomendações relativas à continuidade do programa, ao seu encerramento ou à procedência de alterações. Esse método que tem por objetivo avaliar uma avaliação é definido por Scriven (2018) como sendo uma meta-avaliação.

O processo meta- avaliativo julga se a avaliação atende aos seus propósitos. Assim, são notórias as suas vantagens, pois além de verificar se o processo realizado possui mérito e valor, ainda permite constatar, dentre outros pontos, se houve sua condução envolveu conduta ética adequada, por parte dos avaliadores, contribuindo para conferir ainda mais valor à avaliação ou sugerindo melhorias. Segundo Elliot (2011, p. 960), “em diversas instâncias, a necessidade de meta-avaliar uma avaliação, seu relatório ou o próprio processo avaliativo é a forma de atestar-lhe a qualidade ou de evidenciar pontos que ainda precisam ser aperfeiçoados”. Em paralelo, Stufflebeam (2001) salienta que a meta-avaliação somativa possibilita aos *stakeholders*, isto é, aos interessados no processo, a percepção de pontos fortes e aqueles que necessitam de melhoria na avaliação.

Por se tratar de um material acadêmico, elaborado em função de estudos realizados no Programa de Pós-Graduação da Faculdade Cesgranrio, pode-se assegurar o distanciamento dos meta-avaliadores com os principais atores do contexto da avaliação, pois foi mantida “em relação ao programa, uma distância estrutural cuja

finalidade é reduzir conflitos de interesses e limitar as tendências” (Yarbrough; Shulha; Hopson; Caruthers, 2011, p. 13).

No conceito de Penna Firme e Letichevsky (2010, p. 185, tradução das autoras), o propósito da meta-avaliação é o de:

Assegurar a qualidade da avaliação que, por sua vez, tem na sua essência: julgar o mérito e a relevância de seu foco de atenção, por meio de um processo de indagação sistemática; o mérito no que se refere às qualidades intrínsecas do objeto; e a relevância no que se refere ao impacto e à repercussão do objeto-foco em questão.

A partir da segunda metade do século XX, os estudos meta-avaliativos têm se convertido em importante foco de interesse dos estudiosos dos processos avaliativos. Para Schwandt (1989) meta-avaliar é a aplicação de um método que certifica a qualidade de uma avaliação, exigindo o emprego de um conjunto de procedimentos, padrões e critérios, estabelecidos por meio de juízos críticos, utilizados em seu julgamento. Ou seja, “emitir uma nova avaliação sobre o estudo avaliativo” (Machado; Chaise; Elliot, 2016, p. 3). Já Stufflebeam (2001, p. 185) define a meta-avaliação como:

O processo de delinear, obter e aplicar informação descritiva e de julgamento - sobre a utilidade, a viabilidade, adequação e precisão de uma avaliação e sua natureza sistemática, competente conduta, integridade/honestidade, respeitabilidade e responsabilidade social - para orientar a avaliação e divulgar publicamente seus pontos fortes e fracos.

Para Balzan e Dias Sobrinho (2000), a necessidade de se meta-avaliar está pautada na necessidade de se melhorar e inovar a própria avaliação. Por sua vez, Barros e Elliot (2018) determinam que, para se realizar uma meta-avaliação, é preciso seguir parâmetros e determinar critérios definidos e estruturados que permitam aferir o valor, a ética, a relevância e a utilidade da avaliação executada. Nesse sentido, para conduzir a meta-avaliação da forma mais

objetiva e imparcial possível, foi necessário ao presente estudo adotar critérios pautados em padrões, ou diretrizes norteadoras, que são tecnicamente aceitos e adotados nesse campo de atuação acadêmico-profissional.

OBJETIVO DO ESTUDO

Além da realização de processos avaliativos dos projetos e programas de políticas públicas, espera-se também que seja aferida a qualidade de suas avaliações. Em outras palavras, é preciso que a avaliação realizada sobre o programa original seja avaliada, considerando que este procedimento “representa uma obrigação ética, bem como científica, quando envolve o bem-estar de outras pessoas” (Scriven, 2018, p. 361).

Apoiando-se no pensamento de Scriven (2018), pode-se afirmar que este estudo meta-avaliativo somativo teve por objetivo fornecer evidências sobre a qualidade técnica do Relatório de Avaliação de Impacto Social - Palcos para a Inclusão, representando um serviço de avaliação que pretende gerar recomendações úteis que possam auxiliar na elaboração de futuros relatórios por parte da empresa avaliadora. Ou seja, o estudo procura mostrar se, de modo geral, a avaliação realizada cumpriu sua missão de julgar quão adequadamente: a) foram realizadas as atividades previstas no Projeto; e b) expressaram os achados avaliativos, considerando as necessidades dos interessados no processo.

O OBJETO DO ESTUDO

O objeto deste estudo meta-avaliativo é o Relatório de Avaliação de Impacto Social - Palcos para a Inclusão. Esse Projeto

iniciou suas atividades em junho de 2017, nas cidades de Porto, Trofa e Maia em Portugal para o público residente em habitações sociais e territórios vulneráveis, consistindo na apresentação do trabalho artístico desenvolvido por alunos do Espaço t, onde “t” significa todos, e se caracteriza como espaço aberto que não se restringe a um grupo delimitado, mas que se dedica inteiramente ao ser humano. O Espaço t surge em 1994, como um instrumento de combate à exclusão, promovendo a apresentação de trabalhos artísticos com crianças, jovens e adultos em diversas situações de vulnerabilidade social, reforçando a coesão social e o respeito a diversidade, com vistas à mudança social. O Projeto Palcos para a Inclusão alcançou 10.569 espectadores, por meio de 304 ações, em parceria com 78 entidades durante os 41 meses de execução. Parte das ações se desenvolveram no contexto de pandemia mundial da Covid-19 e por isso foram adaptadas para o ambiente *online*.

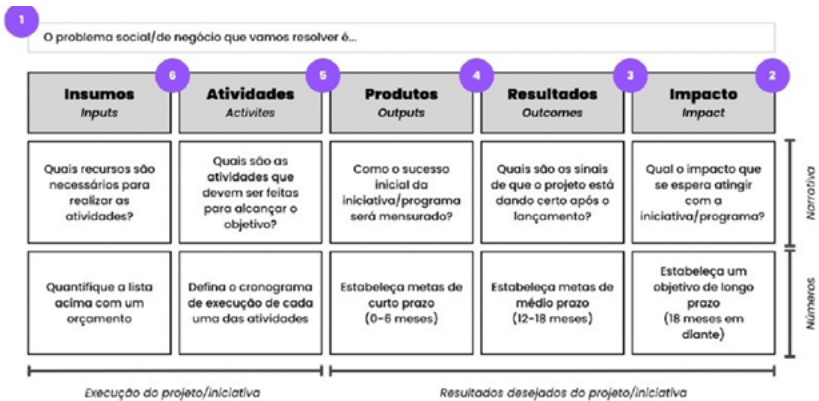
A avaliação de impacto do Projeto Palcos para a Inclusão foi realizada pela empresa *Stone Soup Consulting* em 2021. O sistema de avaliação utilizado foi baseado na Teoria de Mudança e nos parâmetros internacionais adotados pelo *Social Return On Investment* (SROI), que afere retorno social de um projeto, comparando o valor dos recursos investidos, com o valor do impacto social gerado.

Em resumo, essa teoria é um mapa; uma representação da forma pela qual a realidade pode ser mudada e inclui as etapas ou pré-condições que devem ser atingidas no curto e médio prazos para viabilizar o alcance do objetivo final de longo prazo. Neste sentido, cada passo ou pré-condição é um objetivo em si, sendo importante considerar que, ao longo do processo de mudança, fatores externos, atuando de modo paralelo e independente, podem influenciar os resultados do projeto, facilitando, dificultando ou impedindo a mudança objetivada.

Segundo Okleina (2023) a “Teoria da Mudança é um modelo de planejamento usado por empreendedores sociais e organizações

para descrever o impacto desejado causado por uma intervenção ou programa em uma determinada comunidade.” A Figura 1 ilustra o entendimento.

Figura 1 – Modelo de teoria da mudança



Fonte: Okleina (2023).

Na maioria dos projetos sociais, a mudança não ocorre de modo linear, significando que os resultados de curto e médio prazos não deixam de existir ao longo do tempo, mas continuam a se autoalimentar e reforçar o objetivo de longo prazo. Assim, a Teoria de Mudança, adotada para avaliar o projeto Palcos para a Inclusão, constituiu importante ferramenta que permitiu não apenas uma compreensão robusta das mudanças que o projeto pretendeu gerar na sociedade, mas o direcionamento do processo avaliativo.

Conforme descrito no Relatório, os objetivos do processo avaliativo do projeto foram:

- a. medir os resultados da iniciativa Palcos para a Inclusão de forma participativa, em colaboração com a equipe do Espaço t e os parceiros do projeto, promovendo oportunidades de aprendizagem para a melhoria contínua;

- b.** construir e implementar um sistema de avaliação e medição dos impactos sociais de maneira imparcial, transparente e clara, por uma equipe externa, experiente e independente;
- c.** desenvolver um sistema que possibilitasse a verificação e o monitoramento do impacto social do projeto no futuro, identificando os principais indicadores e referenciando recomendações para a geração de maior impacto;
- d.** consolidar os resultados da avaliação e medição dos impactos sociais para ampliar o reconhecimento e interesse público da iniciativa e do impacto gerado na sociedade.

O relatório foi estruturado da seguinte forma:

- 1.** Introdução
 - a. Sumário
 - b. Glossário
- 2.** Enquadramento do Programa de Avaliação
- 3.** Estratégia de Impacto Planejada
- 4.** Atividade e Caracterização dos Participantes
 - a. Atividades
 - b. Participantes
- 5.** Sistemas de Avaliação Implementado
 - a. Metodologia
 - b. Instrumentos de Avaliação
 - c. Definição da Amostra
 - d. Caracterização da Amostra

6. Análise de Dados

- a. Análise da linha de base dos principais beneficiários
- b. Análise *Social Return on Investment* (SROI)
- c. Análise da iniciativa Corpo Evento
- d. Análise de dados qualitativos

7. Conclusões e Recomendações

- a. Conclusões
- b. Recomendações

A implementação do sistema de avaliação considerou as ações propostas pelo projeto tais como: Pintura/Expressão Plástica, Dança/Experimentação Musical, Tai Chi, Grupo de Teatro Aceno de Acenar, Fotografia, Teatro Terapêutico, Canto e Drama, Teatro e Dança Inclusiva.

Segundo os consultores da empresa Stone Soup Consulting (2021, p. 15), pretendeu-se com essa avaliação:

- a. aferir as mudanças causadas pelos alunos do Espaço t nas crianças, jovens e adultos em diversas situações de vulnerabilidade social na região onde ocorreram as ações artísticas do projeto;
- b. avaliar a eficiência do projeto no desenvolvimento de impacto social através de um cálculo específico que compreende o valor social e valor investido;
- c. identificar e propor estratégias que permitam maximizar o impacto social gerado.

Com base na Teoria de Mudança, foram desenvolvidos indicadores de impacto social para medir o impacto ocorrido nos beneficiários do projeto; esses indicadores foram utilizados para fins do

desenvolvimento dos quatro instrumentos avaliativos (Stone Soup Consulting, 2021, p. 16), sendo um para cada beneficiário do projeto, alunos do Espaço t, moradores, crianças e dinamizadores do projeto.

Os questionários semiestruturados foram a principal tipo de instrumento utilizado para colher os dados quantitativos e qualitativos. Além disso, foi desenvolvido um instrumento para avaliar o impacto de iniciativa relacionada à ação considerada como atividade geradora de mudança. Inicialmente, todos esses instrumentos foram aplicados presencialmente pela equipe do Espaço t e, posteriormente, de forma virtual. Durante o processo, foi considerada a necessidade de complementar a avaliação quantitativa com componentes qualitativos e, assim, foram inseridas mais duas técnicas avaliativas: entrevistas telefônicas e grupo focal.

PROCEDIMENTOS METODOLÓGICOS DA META-AVALIAÇÃO

Este trabalho teve seu início por meio de uma análise bibliográfica e documental realizada em materiais acadêmicos e documentais cujo teor se relacionava ao assunto abordado. As consultas tiveram caráter qualitativo-exploratório de modo a propiciar maior aprofundamento na temática pertinente aos estudos avaliativos.

Conforme afirmou Schwandt (1989), caso o objeto de estudo de uma avaliação seja uma outra avaliação, a proposta deve verificar a qualidade do processo avaliativo inicialmente implementado. Para tal, é estabelecido um conjunto de procedimentos, padrões e critérios que possibilitem a formulação de juízo de valor que possa agregar valor à avaliação original, contribuindo para o seu aperfeiçoamento.

No caso do presente estudo os padrões críticos, eleitos para nortear essa avaliação que incidirá sobre outra avaliação, foram estabelecidos por Davidson (2005) e centram-se no julgamento do mérito ou da qualidade intrínseca da avaliação realizada, considerando as suas validades, utilidade, conduta, credibilidade e os seus custos.

A autora afirma que:

Avaliações devem produzir conclusões válidas e justificáveis; devem ser úteis ao cliente e a outros públicos relevantes; ser conduzidas de uma maneira ética, legal, profissional e apropriada; ser digna de crédito para os públicos relevantes; ser tão econômica, rápida e não obstrutiva quanto possível (Davidson, 2005, p. 205).

O Quadro 1 apresenta as definições de cada critério fundamental que uma avaliação de qualidade deve possuir.

Quadro 1 - Descrição dos critérios fundamentais de Davidson

Critérios	Descrição
Validade	Recomendações válidas, se existirem, ou seja, há base suficiente para propô-las e permite reconhecer que elas vão funcionar se forem implementadas.
Utilidade	Demonstra que uma avaliação precisa ser útil para aqueles que a encomendaram e/ou que nela têm interesse.
Conduta	Consiste na observação, pelo meta-avaliador, se padrões legais, éticos e profissionais, de adequação cultural e de não impedimento foram respeitados pelo avaliador.
Credibilidade	Dividido em (a) familiaridade com o contexto; (b) independência, imparcialidade, e/ou falta de conflito de interesse; e (c) expertise (especialidade) em avaliação e na matéria sob investigação.
Custos	Uma avaliação sempre inclui diferentes tipos de custo. No caso de avaliações encomendadas, a prestação de contas, a relevância das despesas, o pagamento dos avaliadores e de outros itens fazem parte de uma equação em que se deseja saber se os resultados valem o que custaram. É uma relação custo-benefício que se procura e deve ser equilibrada.

Fonte: Elliot (2011).

Para a apresentação dos resultados finais da meta-avaliação, serão considerados os critérios de classificação definidos por Davidson (2005): *Conceito A* (atinge todos os aspectos principais deste ponto (leia-se do teor avaliado pelo item) e os expressa de forma clara e concisa; *Conceito B* (atinge a maioria dos aspectos relativos a este ponto, mas não cobre um ou dois deles, ou trata dos ingredientes principais, mas não é 100% claro); *Conceito C* (trata do ponto, mas perde alguns aspectos cruciais, ou perde vários aspectos importantes, ou é um tanto desorganizado ou pouco claro); *Conceito D* (tem um ou dois elementos que parecem tratar implicitamente do ponto, mas o tratamento é deficiente); *Conceito E* (perde o ponto totalmente).

Conforme destacado anteriormente, a meta-avaliação possui o propósito de, por meio do estabelecimento de critérios sólidos, oferecer *feedback* em relação à qualidade da avaliação original. Um importante critério para se atingir o objetivo traçado pela meta-avaliação é a formulação de uma ou mais questões avaliativas.

Para Simões (2015, p. 24), "essas questões avaliativas são a *alma* da avaliação. Não existe boa avaliação sem uma boa pergunta [...]. As perguntas avaliativas são a base para as definições metodológicas que orientarão todo o projeto de avaliação". Nesse sentido, esse estudo meta-avaliativo pretendeu responder à seguinte questão: Em que medida o conteúdo do Relatório de Avaliação de Impacto Social Palcos para a Inclusão está em conformidade com os critérios fundamentais de qualidade da avaliação, estipulados por Davidson (2005)?

Observa-se que os conceitos (A, B, C, D e E), aferidos para expressar o atendimento de cada critério, foram definidos pelas autoras desse estudo com base em juízos de valor subjetivos, a partir do confronto entre o conteúdo manifesto pelo Relatório de Avaliação de Impacto Social do Projeto Palcos para Inclusão e os conceitos inerentes aos Critérios Fundamentais de Davidson (2005). O teor das análises teve, por consequente, caráter qualitativo, representando a

forma com que as autoras interpretaram, criticamente, a realidade expressa no Relatório da Avaliação do projeto.

RESULTADOS DO ESTUDO

Nesta seção, apresentam-se os resultados da meta-avaliação realizada sobre o Relatório de Avaliação de Impacto Social Palcos para a Inclusão. Ao final, explicita-se o desempenho de cada critério considerando a classificação estabelecida por Davidson.

CRITÉRIO DE VALIDADE

Para Davidson (2005, p. 208-209), alguns indicadores podem ser usados como norteadores ao atendimento do critério de validade da avaliação:

Cobre todas as fontes relevantes de valor (ex.: necessidades, padrões relevantes, definições de senso comum sobre valor). Cobre compreensivamente processo, resultado e custo. Não inclui critérios irrelevantes ou ilícitos. Inclui análises qualitativas e quantitativas que são apropriadas aos dados coletados. Estabelece claramente como os dados são interpretados. Explicita com clareza a procedência das conclusões avaliativas (nada de saltos lógicos ou *fumaça no espelho*). Inclui recomendações válidas (se existirem, ou seja, há base suficiente para propô-las e permite reconhecer que elas vão funcionar se forem implementadas).

Nessa perspectiva, Elliot (2011, p. 947) afirma que a “validade da avaliação diz respeito à relação entre as conclusões apresentadas pelo avaliador e sua justificativa”. Por buscar estabelecer a validade da avaliação, a meta-avaliação deve questionar: “O avaliador fez inicialmente as questões corretas? O avaliador dimensionou a

avaliação apropriadamente? O que foi deixado de fora que deveria ter sido incluído e vice-versa?” (Davidson, 2005, p. 206)

A estratégia de avaliação de impacto social da iniciativa Palcos para a Inclusão, adotada pelos consultores da empresa *Stone Soup Consulting*, se baseou em um processo participativo, cujo eixo de trabalho se apoiou na Teoria da Mudança e nos parâmetros internacionalmente acordados no âmbito do *Social Return On Investment* (SROI), que aferem o retorno social de um projeto, comparando os valores dos recursos investidos com os do impacto social gerado.

A Teoria da Mudança, proposta pelos consultores da *Stone Soup Consulting*, possibilitou uma compreensão aprofundada das mudanças, que o projeto pretendeu gerar na sociedade, sendo construída em três etapas: *workshop* com a equipe de gestão do projeto no Espaço t para fins da elaboração do mapa de atores e da construção da Teoria da Mudança; realização de grupo focal, que reuniu os principais *stakeholders* para fins da validação do pré-alinhamento dessa teoria e da elaboração do questionário a ser preenchido pelos dinamizadores das atividades; realização de sessão dinâmica com os alunos do Espaço t, para validar as mudanças identificadas para as crianças que participaram no projeto (Stone Soup Consulting, 2021).

O desenho da Teoria da Mudança do projeto foi apresentado em um quadro, contendo os seguintes elementos: recursos, atividades, resultados diretos, beneficiários, resultados indiretos a curto, médio e longo prazos e os impactos.

O sistema de avaliação do projeto seguiu os sete princípios integrantes da Metodologia do SROI: envolver os *stakeholders*; compreender as mudanças geradas; valorizar apenas o que importa; incluir apenas aquilo que é relevante; reivindicar apenas o valor criado pelas atividades que efetuou; efetuar uma análise transparente; precisa e honesta; verificar os resultados por meio de certificação independente, a fim de trazer maior segurança ao processo de cálculo do rácio (Roux, 2010).

O rácio SROI é obtido dividindo-se o valor do impacto económico e social atualizado pelo valor dos investimentos efetuados. O valor do impacto social gerado é obtido por meio da valorização de mudanças intangíveis e de difícil contabilização ou transação em resultados monetários com o recurso a aproximações financeiras (*proxys*). Por exemplo, um rácio de 2:1 indica que para cada 1 euro investido no projeto, o retorno social é de dois euros (Roux, 2010).

Percebe-se que os consultores estabeleceram objetivos geral e específicos para a avaliação do projeto, porém as questões avaliativas, mesmo não aparecendo de maneira explícita, foram apresentadas na metodologia do relatório a título de pretensões a serem atingidas pela avaliação, plenamente compatíveis com os seus propósitos.

Pode-se afirmar que o relatório apresenta informações detalhadas e compreensíveis sobre os processos, as análises, os resultados e os custos do projeto Palcos para a Inclusão e cobre todas as fontes relevantes de valor. Observa-se que não fora identificado nenhum indicativo de critérios irrelevantes ou ilícitos. Os resultados foram apresentados em tabelas, gráficos e figuras, o que corroborou para a compreensão das interpretações.

Observa-se, por exemplo, que as respostas dos alunos abrangidos pelo projeto revelaram que as mudanças mais importantes foram as relacionadas às mudanças intrínsecas, como o aumento do *sentimento de utilidade* (à sociedade), manifestado por 79% dos respondentes; a diminuição do sentimento de isolamento, expresso por 73% de mudanças positivas; ou o aumento da autoestima, por 69% dos respondentes. Por outro lado, os temas como aceitação das diferenças das pessoas com deficiências e a empatia em relação às que têm deficiências físicas, psíquicas e sociais apresentaram o menor valor de mudanças, ambas com 49% (Stone Soup Consulting, 2021).

Quanto às recomendações, pode-se inferir que foram objetivas, coerentes e válidas, correspondendo ao objetivo, ao contexto e às necessidades da avaliação. O relatório apresenta seção específica

para as recomendações que poderão contribuir para futuros projetos de similar natureza. Dessa forma, após análise dos indicadores que nortearam o julgamento da validade da avaliação, considera-se que o critério foi plenamente atendido.

CRITÉRIO UTILIDADE

A utilidade de uma avaliação é determinada pela necessidade que seus *stakeholders* possuem em relação aos seus resultados. Nesse sentido, tanto a avaliação quanto os seus resultados devem ser: relevantes para as questões ou decisões dos interessados; pontuais, de modo que os resultados estejam disponíveis quando as pessoas interessadas precisarem tomar decisões, e comunicados claramente em linguagem, mídia e canais de comunicação apropriados (Davidson, 2005, p. 209). Nesta perspectiva, Elliot (2011, p. 948) assevera que quanto “mais os resultados forem usados pelos participantes ou pelos interessados na avaliação, mais forte se torna o atendimento ao critério de utilidade”.

Os avaliadores da *Stone Soup Consulting* (2021) apresentaram as Conclusões e Recomendações no item 7 do relatório. É importante salientar que o projeto foi exitoso, tendo atingido 3,16 de SROI, quando extrapolado para todos os beneficiários (Stone Soup Consulting, 2021), evidenciando que significa para cada 1,6 euros investidos, o retorno social foi de 3 euros. No documento, destaca-se uma das recomendações traçadas pelos consultores da empresa avaliadora:

Sendo as artes entendidas como instrumento de transformação social, de forma a maximizar e promover a sustentabilidade das mudanças, seria pertinente, no futuro, o Espaço conceber projetos que possam capacitar os agentes dos territórios de atuação, de forma a conseguirem, após a conclusão do projeto, produzir outras atividades e dinâmicas semelhantes no território (Stone Soup Consulting, 2021, p. 50).

Assim, os avaliadores da *Stone Soup Consulting* conseguiram organizar e pontuar claramente os resultados da avaliação, oferecendo conteúdo robusto e útil para servir de subsídios aos *stakeholders* nas tomadas de decisão. Com base nessa análise, pode-se inferir que o critério de utilidade foi atendido pela avaliação expressa no relatório meta-avaliado.

CRITÉRIO DE CONDUTA

O critério de conduta consiste na observação pelo meta-avaliador se os padrões legais éticos e profissionais, de adequação cultural e de não impedimento do próprio processo da avaliação foram respeitados pelo avaliador (Davidson, 2005). Em complementação, Elliot (2011) diz que os padrões de conduta se relacionam, diretamente, com a validade e a confiabilidade no tocante aos resultados da avaliação. No entanto, lembra que conduta é um critério mais aplicável ao processo da avaliação do que ao relatório completo e acabado.

No relatório, não foi verificado nenhum tipo de desvio de conduta ou de impedimento do próprio processo de avaliação. O conteúdo do Relatório, escrito por consultores da empresa *Stone Soup Consulting*, apresentou todos os dados e resultados com clareza, concisão, transparência, precisão e responsabilização. Além disso, os avaliadores seguiram padrões éticos e profissionais, observando a legislação e a adequação cultural. Por esses motivos, considera-se que o critério de conduta foi atendido em sua íntegra.

CRITÉRIO DE CREDIBILIDADE

Elliot (2011, p. 949) observa que “a obtenção de credibilidade para uma avaliação está diretamente relacionada à aceitação, por parte dos interessados, dos resultados e suas fontes, isto é, em que

medida merecem crédito". Davidson (2005, p. 211), por sua vez, afirma que os elementos fundamentais de credibilidade de uma avaliação são: "a) familiaridade com o contexto; b) independência, imparcialidade, e/ ou falta de conflito de interesse; e c) *expertise* em avaliação e na matéria sob investigação".

Um avaliador familiarizado com o contexto conhece bem o projeto ou programa que está sendo avaliado, as suas estrutura, organização e normas, e, em geral, ainda conhece o grupo onde o objeto avaliado se insere. Essa familiaridade faz com que a apresentação de resultados seja mais realista e próxima do entendimento dos interessados (Elliot, 2011). De acordo com a autora:

Independência e imparcialidade são fatores que se tornam mais 2011, p. presentes quando a avaliação é divulgada para públicos externos ou para grupos de interesse (*stakeholders*). São importantes no trato de conflito de interesses que poderiam prejudicar a credibilidade da avaliação. O avaliador não deve ter preferências por qualquer grupo de *stakeholders*, nem apresentar julgamentos tendenciosos. Tais posturas fazem parte de sua atuação como profissional (Elliot, 2011, p. 949).

A avaliação foi realizada por uma equipe formada por consultores pertencentes à empresa *Stone Soup Consulting*. Essa escolha garantiu a imparcialidade, não sendo notado nenhum conflito de interesses. Ficou bastante evidente a *expertise* em avaliação demonstrada pelos consultores da *Stone* em relação à intervenção e ao seu contexto, evidenciado que A partir da estrutura de organização do relatório deixou nítida a familiaridade dos avaliadores com o contexto da avaliação. Segundo Elliot (2019), avaliadores capacitados e experientes na área de avaliação e conhecedores sobre o tema que está sendo analisado, são essenciais para garantir a credibilidade da avaliação. Nesse sentido, os avaliadores demonstraram conhecimento sobre o conteúdo analisado, a metodologia aplicada e com a prática avaliativa.

Diante do exposto, o critério credibilidade foi considerado plenamente atendido, perante os indicadores familiaridade, imparcialidade.

CRITÉRIO DE CUSTOS

“No caso de avaliações encomendadas, a prestação de contas, a relevância das despesas, o pagamento dos avaliadores e de outros itens fazem parte de uma equação em que se deseja saber se os resultados valem o que custaram” (Elliot, 2011, p. 949-950). Na meta-avaliação deve-se procurar a relação equilibrada de custo-benefício. Segundo Elliot (2019), a prestação de contas permite identificar os custos referentes à realização da avaliação são viáveis e se justificaram os seus resultados. Além disso, este padrão complementa a prestação de contas, esclarecendo sobre a função dos itens que compõem as despesas, o que auxilia os *stakeholders* na elaboração de um planejamento financeiro adequado e eficaz para as futuras avaliações que pretendam contratar.

O relatório não menciona os custos da avaliação, o pagamento dos avaliadores e a relevância das despesas. É importante registrar que a metodologia aplicada (SROI) calcula os custos tangíveis e intangíveis em relação aos benefícios sociais gerados pelo projeto avaliado. Portanto, como não há no relatório qualquer alusão aos custos do processo avaliativo realizado, o critério de custos foi considerado não atendido.

O Quadro 2 apresenta os respectivos julgamentos efetivados por esta meta-avaliação, de acordo com os critérios de classificação de Davidson (2005). Ou seja, exhibe um demonstrativo geral dos resultados obtidos por cada critério fundamental por meio desta meta-avaliação.

Quadro 2 – Resultado geral da meta-avaliação

Critérios	Julgamento (A-E)
1 – Validade	A
2 – Utilidade	A
3 – Conduta	A
4 – Credibilidade (Familiaridade, imparcialidade e especialidade)	A
5 – Custos	E

Fonte: As autoras (2022).

CONCLUSÕES

As conclusões do estudo respondem à questão avaliativa formulada: Em que medida o conteúdo do Relatório de Avaliação de Impacto Social “Palcos para a Inclusão” está em conformidade com os critérios fundamentais de qualidade da avaliação, estipulados por Davidson (2005)?

Face aos resultados apresentados, pode-se concluir que o conteúdo do relatório atende plenamente aos critérios validade, utilidade, conduta e credibilidade, estabelecidos por Davidson (2005). No entanto, no que se refere ao critério referente aos custos, o relatório não faz alusão aos gastos efetuados com os avaliadores e com a avaliação em si. Nesse aspecto, é importante destacar que o detalhamento dos custos de uma avaliação se faz necessário para verificar a eficiência do processo avaliativo e definir se a relação custo-benefício foi profícua para a organização contratante.

Toda avaliação representa um custo a mais para seus contratantes e sua definição permite um melhor conhecimento das relações entre custo de sua implementação

e qualidade dos resultados alcançados, ou seja, do que foi obtido como insumo da avaliação. O custo serve, também, como uma importante ferramenta da avaliação, possibilitando se obter indicadores de desempenho e impacto, que, por sua vez, mensuram se os resultados obtidos foram adequados às necessidades do programa (Alfradique; Tavares; Orlando Filho, 2020).

Por fim, o Relatório de avaliação contemplou plenamente quatro dos cinco critérios estabelecidos por Davidson (2005) e utilizados na metodologia desta meta-avaliação. A *expertise* em avaliação dos consultores da *Stone Soup Consulting* ficou evidente na estruturação do relatório, aplicação da metodologia, análise dos resultados, e na descrição das conclusões e recomendações.

O Relatório demonstrou atender com excelência aos objetivos traçados no processo avaliativo, tendo sido estruturado de forma clara e objetiva, sugerindo recomendações válidas que poderão servir de subsídios úteis aos *stakeholders* para a tomada de decisão.

RECOMENDAÇÕES

Face aos resultados e às conclusões estabelecidas, sugere-se aos avaliadores da empresa *Stone Soup Consulting*, a título de contribuição, as seguintes recomendações:

Que, nas próximas avaliações, sejam inseridos em seus relatórios os custos que compõem o processo avaliativo, levando-se em conta que esse atributo, em especial, permite evidenciar o grau de responsabilização presente no contexto da avaliação, aumentando, assim, a credibilidade tanto dos avaliadores como de seus contratantes.

MÉRITO DO ESTUDO

Lucí Hildenbrand

Como produto da meta-avaliação afirma-se que o título atribuído ao relatório meta-avaliado reflete o teor do conteúdo abordado pela avaliação, bem como instiga o leitor a conhecê-la dada a sua relevância social. Em sua parte introdutória, contempla aspectos essenciais à iniciação no tema. Assim, de um lado, destaca o objetivo, a questão avaliativa e a justificativa do estudo meta-avaliado; e, de outro, discorre acerca da importância crescente da meta-avaliação como metodologia que oportuniza a qualificação das avaliações realizadas e, por conseguinte, consolidando-se como necessária à validação de seus achados.

No caso, o estudo meta-avaliado fez uso e menção literal e não literal a referencial teórico procedente, incluindo obras recentes e clássicas em quantidade suficiente para bem fundamentar toda a sua argumentação. A metodologia nele adotada – também considerada clássica – foi cumprida com fidelidade, respeitando não só o aporte teórico proposto por Jane Davidson, mas também as orientações por ela fornecidas para nortear o julgamento da avaliação relatada em seu próprio relatório. Com isso, a avaliação mostrou-se consistente e robusta, quando: a) foi adequada para promover a apreciação do objeto avaliativo, cujo percurso foi facilitado pelo fato de a meta-avaliação explicitar, de forma clara e precisa, o seu objetivo e a sua questão avaliativa; b) relacionou e citou falas literais e não literais de autores referência; c) fez uso de análise qualitativa apropriada ao tipo de achados do estudo. Quanto à apreciação dos resultados e às análises realizadas, observa-se que a meta-avaliação percorreu ponto a ponto as orientações metodológicas assumidas para favorecer o juízo de valor do objeto meta-avaliado. Em relação à comunicação dos resultados: a) os dados obtidos foram apresentados de forma adequada e completa; b) os resultados foram

consoantes à teorização escolhida; o recurso das representações gráficas foi incorporado no texto, facilitando a apreensão dos conteúdos emergentes da meta-avaliação. Dentre as limitações, duas possíveis fragilidades poderiam ser apontadas. A primeira devido ao fato de a meta-avaliação não comparar os achados com aqueles encontrados em estudos atuais e/ou afins. E, a segunda, por não contribuir com a geração de conhecimento novo. Neste sentido, cabe um só argumento: como a meta-avaliação é a primeira produção acadêmica das mestrandas, há de se convir que seus conhecimentos em Avaliação e Meta-avaliação se encontram em fase de construção inicial, o que impede estabelecer relações entre o seu estudo e outros estudos, bem como contribuir para a formação de novos conhecimentos saberes em Meta-Avaliação, porque a aprendizagem humana sustenta-se em conhecimentos prévios que se estabelecem e se consolidam com o tempo. No que tange às conclusões, mostram-se suficientes enquanto apoiadas nos resultados e respectivas análises e, também, remetem ao objetivo do estudo. Por fim, em relação às referências, evidenciam-se consistentes para aportar o estudo e acham-se em conformidade com a normalização da ABNT.

REFERÊNCIAS

ALFRADIQUE, L. M.; TAVARES, T. T.; ORLANDO FILHO, O. **Relatório de Avaliação de Efetividade de Projetos de desenvolvimento Científico e Tecnológico Apoiados pelo Fundo Amazônia/BNDES**: um estudo meta-avaliativo. São Paulo: Pimenta Cultural, 2019.

BALZAN, N. C.; DIAS SOBRINHO, J. **Avaliação institucional**: teoria e experiência. 2. ed. São Paulo: Cortez, 2000.

BARROS, L. F. R. B; ELLIOT, L. G. Meta-avaliação do relatório de avaliação da execução de programas de governo nº 8: programa saúde da família. **Meta: Avaliação**, Rio de Janeiro, ed. esp., n. 2, p. 197-224, 2018. DOI: <http://dx.doi.org/10.22347/2175-2753v10i0.1933>. Disponível em: <https://revistas.cesgranrio.org.br/index.php/metaavaliacao/article/view/1933>. Acesso em: 30 out. 2023.

DAVIDSON, E. J. **Evaluation methodology basics**: the nuts and bolts of sound evaluation. Thousands Oaks, CA: Sage, 2005.

DYE, T. D. **Understanding public policy**. Englewood Cliffs, N.J.: Prentice- Hall, 1984.

ELLIOT, L. G. Meta-avaliação: das abordagens às possibilidades de aplicação. **Ensaio: aval. pol. públ. educ.**, Rio de Janeiro, v. 19, n. 73, p. 941-963, 2011.

ELLIOT, L. G.; ORLANDO FILHO, O. (org.). **Meta avaliações de políticas públicas do governo**: o que revelam os relatórios de avaliação. São Paulo: Pimenta Cultural, 2019.

FARIA, C. A. P. Ideias, conhecimento e políticas públicas: um inventário sucinto das principais vertentes analíticas recentes. **Revista Brasileira de Ciências Sociais**, São Paulo, v. 18, n. 51, p. 21-29, 2003. DOI: <https://doi.org/10.1590/S0102-69092003000100004>. Disponível em: <https://www.scielo.br/j/rbcsoc/a/KwfGqDcrtDXgxRjDGgZPYjc/abstract/?lang=pt>. Acesso em: 30 out. 2023.

MACHADO, T. R.; CHAISE, R. M.; ELLIOT, L. G. A meta-avaliação como instrumento de qualidade nas políticas públicas: o programa segundo tempo. **Meta: Avaliação**, Rio de Janeiro, v. 8, ed. esp., p. 1-20, 2016. DOI: <http://dx.doi.org/10.22347/2175-2753v8i0.1034>. Disponível em: <https://revistas.cesgranrio.org.br/index.php/metaavaliacao/article/view/1034>. Acesso em: 30 out. 2023.

OKLEINA. **Conheça a teoria da mudança, ferramenta para avaliar impacto social**. *Posdigitak.pucpr*, Curitiba, 15 fev. 2023. Disponível em: <https://posdigital.pucpr.br/blog/teoria-da-mudanca#:~:text=A%20aplica%C3%A7%C3%A3o%20da%20Teoria%20da,acontecer%20em%20um%20dado%20contexto>. Acesso em: 30 out. 2023.

ORLANDO FILHO, O.; OLIVEIRA, R. A. C. O relatório de avaliação do programa apoio à alimentação escolar na educação básica: um estudo meta-avaliativo. *In*: ELLIOT, L. G.; ORLANDO FILHO, O. (org.). **Meta avaliações de políticas públicas do governo**: o que revelam os relatórios de avaliação. São Paulo: Pimenta Cultural, 2019. p. 151-170.

PENNA FIRME, T.; LETICHEVSKY, A. C. O desenvolvimento da capacidade de avaliação no século XXI: enfrentando o desafio através da meta-avaliação. **Ensaio: aval. pol. públ. Educ.**, Rio de Janeiro, v. 10, n. 36, p. 289-300, 2002.

ROUX, H. N. El SROI (social return on investment) : un método para medir el impacto social de las inversiones. **Análisis Financiero**, [S. l.], n. 113, p. 34-43, 2010.

SCHWANDT, T. A. The politics of verifying trustworthiness in evaluation auditing. **American Journal of Evaluation**, Thousand Oaks, v. 10, n. 4, 1989. DOI: <https://doi.org/10.1177/109821408901000405>. Disponível em: <https://journals.sagepub.com/doi/10.1177/109821408901000405>. Acesso em: 10 jun. 2022.

SCRIVEN, M. **Avaliação**: um guia de conceitos. Rio de Janeiro: Paz e Terra, 2018.

SIMÕES, A. A. **Avaliação de programas e políticas públicas**. Brasília, DF: ENAP, 2015.

SOUZA, C. Políticas públicas: uma revisão da literatura. **Sociologias**, Porto Alegre, v. 8, n. 16, p. 20-45, 2006. DOI: <https://doi.org/10.1590/S1517-45222006000200003>. Disponível em: <https://www.scielo.br/j/soc/a/6YsWyBWZSdFgfSqDVQhc4jm/abstract/?lang=pt>. Acesso em: 10 jun. 2022.

STONE SOUP CONSULTING. **Relatório de avaliação de impacto social**: palcos para a inclusão. Lisboa: Stone Soup Consulting, 2021. Disponível em: https://issuu.com/leonelmoraes/docs/stone_soup_-_relatorio_final_palcosinclus_o. Acesso em: 10 jun. 2022.

STUFFLEBEAM, D. L. The metaevaluation imperative. **American Journal of Evaluation**, Thousand Oaks, v. 22, n. 2, p. 183-209, 2001. DOI: <https://doi.org/10.1177/109821400102200204>. Disponível em: <https://journals.sagepub.com/doi/10.1177/109821400102200204>. Acesso em: 10 jun. 2022.

WORTHEN, B.; SANDERS, J.; FITZPATRICK, J. **Avaliação de programas**: concepções e práticas. São Paulo: Editora Gente, 2004.

YARBROUGH, D. B.; SHULA, L. M.; HOPSON, R. K.; CARUTHERS, F. A. **The program evaluation standards**: a guide for evaluators and evaluations users. 3. ed. Califórnia: Sage, 2011.

8

Flávia de Andrade Viveiros Quadrelli
Paula Magalhães da Silva Santos
Elaine Maria de Andrade Senra

META-AVALIAÇÃO DA AVALIAÇÃO DE IMPACTO SOCIAL DOS CURSOS TÉCNICOS CEAP

A AVALIAÇÃO DE IMPACTO SOCIAL E A METODOLOGIA *SOCIAL RETURN ON INVESTMENT – SROI*

A finalidade específica de um programa social é contribuir para o alcance de um objetivo coletivo. Uma vez que esse objetivo seja definido e formalizado em lei ou em algum tipo de contrato social, o programa é criado para influenciar os resultados de interesse correspondentes. Para isso, bens e serviços são fornecidos aos beneficiários, seguindo a teoria da mudança que embasa o desenho do programa, a fim de modificar efetivamente esses resultados.

Entretanto, apenas acompanhar esses indicadores não basta: é essencial mensurar a contribuição de cada programa para os resultados que se pretende alcançar. Nesse sentido, as avaliações de impacto são fundamentais, pois permitem isolar e identificar o quanto cada iniciativa efetivamente contribui para o sucesso coletivo. Na última análise, o que de fato importa numa avaliação de impacto não é uma evolução global, mas sim o quanto cada programa contribui para o cumprimento dos objetivos.

Segundo o Itaú Social:

Para uma avaliação de impacto, o contexto definitivamente importa, assim como a presença de outros programas complementares e substitutos. Mas o contexto e a presença de outros programas importam apenas conforme influenciam a contribuição do programa que está sendo avaliado para o progresso no resultado de interesse coletivo que foi desenhado para modificar (Menezes; Pinto, 2017, p. 17).

A magnitude do impacto pode ser determinada como a diferença entre o resultado que o grupo de referência alcançou ou teria alcançado na presença do programa e o resultado que esse mesmo

grupo alcançou ou teria alcançado na inexistência deste programa (Menezes; Pinto, 2017). Ponto importante refere-se à definição do grupo de referência, cuja situação o programa pretende alterar. Essa escolha deve decorrer dos objetivos da avaliação e não de conveniências metodológicas. Vale salientar que embora o objetivo de uma avaliação de impacto seja obter o impacto de uma intervenção sobre a população de referência, essa intervenção também tem impacto sobre populações que não foram, e muitas vezes que nunca serão beneficiadas pelo programa. A magnitude do impacto não é fixa, ela é diferente para diferentes beneficiários, varia ao longo do tempo e de acordo com o ambiente em que está inserida, como também depende da disponibilidade de outros programas.

Avaliações de impacto podem ocorrer: (i) antes do início do programa, denominadas avaliações *ex ante*; (ii) durante o período de execução dele, denominadas *ex post* de percurso; e (iii) após a conclusão dele, denominadas *ex post* de encerramento. O objetivo das avaliações *ex ante* de impacto é prever a magnitude dos impactos antes que o programa seja implementado. As avaliações *ex post* de recurso, devem ser utilizadas para verificar a validade das previsões realizadas pelas avaliações *ex ante*. Já a avaliação de impacto após o encerramento de um programa social é útil para determinar quão adequada foi a decisão de alocação de recursos àquele programa e para identificar as melhores práticas e programas com a melhor relação custo-efetividade para outros programas que busquem modificar o mesmo resultado.

Algumas das motivações para avaliar o impacto de um programa social são: validar as teorias que serviram de base ao desenho do programa, confirmar se as percepções dos usuários sobre a eficácia do programa coincidem com a realidade, conhecer a magnitude do impacto, assim como saber quais dimensões o programa tem impacto e quais canais possibilitam a manifestação desse impacto, avaliar quão elevado é o valor dos benefícios do programa para o grupo mais pobre, determinar em que medida a eficácia de um

programa depende das características da população beneficiada e da natureza do ambiente em que é implementado, investigar o impacto efetivo do programa, verificar qual parcela do impacto potencial foi desperdiçada devido a falhas no processo de implementação e avaliar a relação custo-efetividade de um programa.

Os usos que avaliações de impacto têm podem ser divididos em dois grandes grupos: (i) uso interno e (ii) uso externo. Os usos internos da avaliação de impacto são: (i) justificar a existência de um programa, (ii) decidir sobre desativação, continuidade ou expansão do programa e (iii) fornecer insumos para promoção de melhorias no desenho do programa. Já o uso externo da avaliação de impacto é auxiliar no desenho, na implementação ou na gestão de outros programas similares em contextos diferentes.

Cabe destacar que o sucesso de uma avaliação de impacto depende principalmente de dois fatores: o controle sobre o processo de seleção dos beneficiários e um bom sistema de monitoramento. Para sua realização existem algumas metodologias disponíveis para identificar e estimar o impacto das intervenções implementadas. Entre elas a metodologia *Social Return on Investment (SROI)* ou Retorno Social sobre Investimento.

Segundo o Instituto para o Desenvolvimento do Investimento Social (2012), o Retorno Social sobre Investimento é um protocolo de avaliação que propõe uma análise comparativa entre o valor dos recursos investidos em um projeto ou programa e o valor social gerado para a sociedade com essa iniciativa. Ele é utilizado para mensurar e contabilizar o valor do impacto das ações sociais que não podem ser mensuradas, aquelas que são intangíveis. De acordo com o Instituto para o Desenvolvimento do Investimento Social, organizações sem fins lucrativos e empreendimento social usam o SROI como uma ferramenta de gestão para melhorar o desempenho, informar despesas e ressaltar o valor agregado (Instituto para o Desenvolvimento do Investimento Social, 2012). Sua utilização busca a redução da desigualdade e da degradação ambiental e a melhoria do bem-estar,

ao incorporar custos e benefícios sociais, ambientais e econômicos (Instituto para o Desenvolvimento do Investimento Social, 2012).

O foco desta metodologia está na percepção dos *stakeholders*. Ela realiza uma integração entre informações qualitativas e quantitativas, e, através de critérios claros e abrangentes, é realizada a monetização do impacto social para a comparação com o valor investido no Projeto ou Programa. O SROI mede a mudança que é relevante para as pessoas ou organizações que experimentaram ou contribuíram para tal mudança (Instituto para o Desenvolvimento do Investimento Social, 2018).

No estudo em referência foi utilizado o SROI de avaliação, que é aquele conduzido retrospectivamente e baseado em resultados reais (Instituto para o Desenvolvimento do Investimento Social, 2012).

O estudo realizado para o CEAP caracteriza-se como um estudo de avaliação, já que sua estrutura de coleta de dados envolveu alunos dos cursos técnicos desde 1994 e, portanto, é capaz de capturar impactos concretos e de longo prazo que já aconteceram, relatados pelas pessoas que passaram pelos cursos técnicos.

O protocolo SROI prevê a realização de grupos focais e entrevistas, além da aplicação de questionário. É importante ressaltar que não são utilizadas informações estatísticas. E o ponto de saturação é baseado na quantidade de pessoas que são necessárias para se obter as mesmas respostas.

Sete princípios norteiam o SROI representando a base de como deve ser aplicado. São eles: (i) envolver os *stakeholders*; (ii) entender o que muda; (iii) valorizar as coisas que importam; (iv) incluir somente o que for material; (v) não reivindicar em excesso; (vi) ser transparente; e, (vi) verificar o resultado. Além dos princípios, são definidas na metodologia SROI seis etapas, conforme a seguir: (i) estabelecer o escopo e identificar os *stakeholders* chave; (ii) mapear resultados; (iii) evidenciar resultados e atribuir-lhes um

valor; (iv) estabelecer o impacto; (v) calcular o SROI; e (vi) relatar, usar e incorporar as conclusões.

OBJETIVO E OBJETO DO ESTUDO

Este estudo possuiu o objetivo de realizar uma avaliação de uma avaliação em um processo definido como meta-avaliação, tendo como fonte de investigação o seu relatório final, o Relatório de Avaliação do Impacto Social dos Cursos Técnicos CEAP, cujas ações são executadas no âmbito do Centro Educacional Assistencial Profissionalizante. Para Elliot (2011, p. 943) a realização de uma avaliação que tem como objeto avaliado uma outra avaliação representa a “verificação da qualidade da própria avaliação à luz de diversos critérios, quando o meta-avaliador contratado ao término da avaliação focaliza o Relatório produzido”.

Pretende-se, assim, por meio deste estudo meta-avaliativo somativo, produzir verificações e realizar recomendações “que visem à melhoria da qualidade da avaliação e, como consequência, o melhor aproveitamento dos resultados alcançados” (Souza; Souza; Elliot; Hildenbrand, 2018, p. 26).

O objeto avaliado é a Avaliação transcrita no Relatório de Retorno Social do Investimento para os cursos técnicos do Centro Educacional Assistencial Profissionalizante (CEAP), organização social sem fins lucrativos, sediada em São Paulo, focada na educação técnica e profissionalizante de jovens entre 10 e 18 anos, no período do contraturno escolar. A instituição oferece três cursos distintos: Técnico em Administração e Empreendedorismo; Técnico em Redes de Computadores; e Técnico em Informática, todos com duração de 2 anos (4 semestres letivos).

A sua elaboração se deu a partir da decisão do CEAP de avaliar e monetizar o impacto de seus cursos técnicos para estimar o retorno social dos recursos investidos e analisar sua contribuição para a sociedade. A instituição responsável pela avaliação foi o Instituto para o Desenvolvimento do Investimento Social, organização social independente pioneira no apoio técnico ao investidor social privado e seu impacto. Sua atuação baseia-se no tripé geração de conhecimento, consultoria e realização de projetos de impacto, que contribuem para o fortalecimento da filantropia estratégica e da cultura de doação.

Os principais objetivos da avaliação foram: (i) compreender o impacto dos cursos técnicos por meio de um estudo avaliativo que demonstre a efetividade do investimento que o CEAP realiza nos cursos técnicos; (ii) oferecer análises relevantes para o processo de planejamento e tomada de decisões do CEAP quanto aos desdobramentos e continuidade dos cursos técnicos; e, (iii) auxiliar o CEAP a identificar potenciais pontos de aprimoramento e focos prioritários de ação nos cursos técnicos oferecidos pela escola.

Os achados da avaliação buscaram comprovar que a iniciativa do CEAP traz benefícios sociais relevantes para seus participantes em todos os eixos de mudança pretendidos: ética e responsabilidade na relação com famílias, amigos e sociedade; perspectivas de futuro e disposição para perseguir objetivos; formação técnica de excelência; e habilidades sociais desenvolvidas.

Sobre o conceito de meta-avaliação somativa, segundo (Elliot, 2011, p. 943):

Nesse tipo de abordagem, o meta-avaliador lida com Relatórios prontos e pode também utilizar bancos de dados coletados, documentação analisada, registros feitos, depoimentos, enfim todo o material disponível que diz respeito e retrate a avaliação desenvolvida.

Penna Firme e Letichevsky (2002, p. 292) argumentam que esse tipo especial de avaliação "irá julgar, por excelência, o mérito e a relevância da avaliação na sua conclusão, sempre à luz dos standards de avaliação" e, também, "o que se pretende é indagar e julgar se a avaliação em foco satisfaz os critérios de uma verdadeira avaliação". Complementando, Yarbrough, Shulha, Hopson e Caruthers (2011) afiançam que esse tipo de meta-avaliação constitui um julgamento de qualidade realizado em um programa já concluído. Já para Worthen, Sanders e Fitzpatrick (2004), a meta-avaliação somativa pode aumentar a credibilidade dos resultados.

META-AVALIAÇÃO: EXPANDINDO O CONCEITO

A meta-avaliação é um campo emergente que busca analisar e avaliar processos de avaliação, oferecendo uma reflexão crítica sobre a eficácia e a qualidade dos métodos utilizados. Com o crescimento das práticas avaliativas em diversas áreas, como educação, saúde e políticas públicas em geral, a necessidade de garantir que essas avaliações sejam realizadas de maneira rigorosa e sistemática tornou-se evidente. Como primeiro passo para a realização deste estudo, esta seção explora os conceitos centrais da meta-avaliação, citando alguns de seus principais autores que contribuíram para o desenvolvimento desta área.

Stufflebeam (2011, p. 117, tradução nossa) afirma que:

Uma boa avaliação requer que seu próprio esforço seja avaliado. Muitas coisas podem sair, e frequentemente saem errado em sua realização. Portanto, é necessário checar problemas como vieses, erros técnicos, dificuldades administrativas e uso inadequado.

Berends e Roberts (2003 *apud* Elliot, 2011, p. 943) destacam duas razões para se proceder a uma meta-avaliação:

Uma é possibilitar a verificação da aplicação de um código de comportamento adequado durante todo o estudo, ou seja, a ética adotada pelos avaliadores, se ela existe e como foi praticada. A outra concebe a meta-avaliação como uma prática de reflexão sobre todos os procedimentos utilizados na avaliação, dando oportunidade à incorporação de novos conhecimentos por parte não só dos meta-avaliadores, como também dos avaliadores, quando tomam ciência da análise realizada.

Explicando sobre meta-avaliar avaliações de programas de políticas públicas, Ribeiro, Carabajal e Orlando Filho (2019, p. 14) destacam que:

A meta-avaliação coopera para a melhoria e o aperfeiçoamento das políticas e ações institucionais por meio de um acompanhamento e revisão permanente e sistemática do objeto que está sendo avaliado. Sendo assim, uma meta-avaliação tem como objetivo assegurar o mérito e o valor do processo avaliativo.

Segundo pensamento formulado por Furtado e Laperrière (2012, p. 696), a meta-avaliação:

É de natureza qualitativa e pretende estabelecer um julgamento de valor, com critérios pré-definidos, com base em informações existentes ou construídas sobre uma determinada avaliação (em curso ou já realizada) para subsidiar eventuais tomadas de decisão. Veja que temos aí os componentes essenciais da avaliação – informação, estabelecimento de critérios, julgamento e tomada de decisão – aplicados sobre a própria avaliação.

Logo, por meio da meta-avaliação vislumbra-se a possibilidade de revisões, reflexões, melhorias e aperfeiçoamento, o que justifica a sua importância no ciclo avaliativo. Trata-se de uma ferramenta poderosa que permite uma reflexão crítica sobre os processos avaliativos, contribuindo para a melhoria contínua e a eficácia das

avaliações. Os principais autores desse campo, como Michael Scriven, J. Bradley Cousins e Thomas A. Schwandt, oferecem fundamentos teóricos e práticos essenciais para a compreensão e implementação de práticas meta-avaliativas. Ao promover um olhar crítico sobre a avaliação, essa ferramenta não apenas fortalece a credibilidade dos resultados, mas também enriquece o diálogo sobre melhores práticas em diversas áreas, assegurando que as avaliações cumpram seu papel de forma justa e eficaz.

Por fim, seguindo o pensamento de Scriven (2018), cabe observar um aspecto importante da meta-avaliação: sua intencionalidade encontra-se alinhada à expectativa de se construir juízo de valor sobre uma avaliação, no sentido de contribuir para a sua melhoria por meio de recomendações que possam cooperar para o aperfeiçoamento de seus processos avaliativos, não se destinando, assim, a simplicidade em se apontar falhas ou incongruências.

METODOLOGIA DO ESTUDO META-AVALIATIVO

Para elaboração deste trabalho foram realizadas consultas à bibliografia especializada, em um processo de pesquisa qualitativa-exploratória em livros, periódicos acadêmicos, aos sites relacionados à avaliação estudada e à documentação pertinente ao objeto de estudo. Não houve a realização de outros tipos de investigação. De acordo com Araújo, Santos e Aguiar (2020), esse aspecto apesar de comum em avaliações somativas, pode ser classificado como uma limitação do estudo.

Como critérios adotados para o estudo foram utilizados os padrões do *Joint Committee on Standards for Educational Evaluation*, instituição norte-americana responsável pela revisão, atualização,

desenvolvimento e disseminação de padrões para atendimento às necessidades do campo da avaliação, que desde 1975 é a fonte oficial de padrões de avaliação na América do Norte e fornece padrões que são aplicáveis para avaliações conduzidas em ambientes educacionais e não educacionais (Joint Committee on Standards for Educational Evaluation, 2020).

As normas de avaliação do programa, em sua terceira edição publicada em 2011, estabelecem cinco Categorias, que se subdividem em um total de 30 padrões. Os padrões são divididos em cinco categorias: Utilidade, Exequibilidade, Adequação, Precisão e Responsabilização (Yarbrough; Shulha; Hopson; Caruthers, 2011). No entanto, neste estudo foram contempladas as três primeiras categorias: Utilidade, Exequibilidade, Adequação.

Yarbrough, Shulha, Hopson e Caruthers (2011) afirmam que a utilidade de uma avaliação se mede pelo nível de satisfação das necessidades dos interessados no programa - *stakeholders* - em relação a processos e produtos

O Quadro 1 apresenta algumas das principais características dos oito padrões da Categoria da Utilidade da Avaliação.

Quadro 1 – Características da categoria da utilidade da avaliação

Padrões	Características
U1 - Credibilidade do Avaliador	As avaliações devem ser conduzidas por pessoal qualificado que estabeleça e mantenha credibilidade no contexto da avaliação.
U2 - Atenção aos Interessados (<i>stakeholders</i>)	As avaliações devem dedicar atenção a todos os indivíduos e grupos envolvidos no programa e por elas afetados.
U3 - Propósitos Negociados	Os propósitos da avaliação devem ser identificados e continuamente negociados com base nas necessidades dos <i>stakeholders</i> .

Padrões	Características
U4 - Explicitação de Valores	As avaliações devem esclarecer e especificar os valores culturais e individuais que fundamentam objetivos, processos e julgamentos.
U5 - Informação Relevante	A informação obtida por meio da avaliação deve servir às necessidades identificadas e emergentes dos <i>stakeholders</i> .
U6 - Produtos e Processos - Significativos entendimentos e comportamentos.	As avaliações devem construir atividades, descrições e julgamentos, de modo que os participantes se sintam estimulados a redescobrir, reinterpretar ou rever seus entendimentos e comportamentos.
U7 - Comunicação e Relatórios Apropriados e no Prazo	As avaliações devem atender às constantes necessidades de informação de seus vários públicos.
U8 - Preocupação com Consequências e Influências	As avaliações devem promover o uso responsável e adequado de seus resultados, além de prevenir consequências negativas indesejáveis e má utilização.

Fonte: Os autores (2024).

As avaliações são exequíveis quando têm a possibilidade de alcançar um nível adequado de efetividade e eficiência. Existem três conceitos-chaves pertinentes à discussão da exequibilidade da avaliação: avaliabilidade, contexto, valores e responsabilização (Yarbrough; Shulha; Hopson; Caruthers, 2011).

O Quadro 2 apresenta algumas das principais características dos quatro padrões da Categoria da Exequibilidade da Avaliação.

Quadro 2 – Características da categoria da exequibilidade da avaliação

Padrões	Características
E1 - Gerenciamento do Projeto	As avaliações devem adotar estratégias eficazes para o gerenciamento do projeto.
E2 - Procedimentos Práticos	Os procedimentos da avaliação devem ser práticos e responsivos, de acordo com o modo de operação do programa.

Padrões	Características
E3 - Viabilidade do Contexto	As avaliações devem identificar, monitorar e equilibrar interesses políticos e culturais, e necessidades individuais e de grupos.
E4 - Uso dos Recursos	As avaliações devem utilizar os recursos com efetividade e eficiência.

Fonte: Os autores (2024).

Na visão geral da categoria adequação, Yarbrough, Shulha, Hopson e Caruthers (2011) a definem se referindo ao que é apropriado, válido, legal, correto, aceitável e justo na avaliação. São compreendidos três campos que se sobrepõem: (i) direitos, deveres e responsabilidades dos avaliadores; (ii) sistemas de leis, regulamentos e regras; (iii) e, funções e obrigações inerentes à prática profissional da avaliação.

O Quadro 3 apresenta algumas das principais características dos sete padrões da Categoria da Adequação da Avaliação.

Quadro 3 – Características da categoria da adequação da avaliação

Padrões	Características
A1 - Orientação Responsiva e Inclusiva	As avaliações devem ser responsivas aos <i>stakeholders</i> e as suas comunidades.
A2 - Acordos Formais	Os acordos da avaliação devem ser negociados de modo que as obrigações fiquem explícitas, e sejam considerados os contextos culturais, as necessidades e as expectativas dos clientes e de outros <i>stakeholders</i> .
A3 - Respeito e Direitos Humanos	As avaliações devem ser planejadas e conduzidas com a preocupação de proteger os direitos humanos e a legalidade, e manter a dignidade de participantes e outros <i>stakeholders</i> .
A4 - Clareza e Equidade	As avaliações devem ser inteligíveis e justas, ao tratar dos propósitos e necessidades dos <i>stakeholders</i> .
A5 - Transparência e Abertura	As avaliações devem fornecer a todos os <i>stakeholders</i> descrições completas das constatações, limitações e conclusões, a não ser que isso represente violação das leis ou das convenções sociais.

Padrões	Características
A6 - Conflitos de Interesses	As avaliações devem identificar franca e honestamente e trabalhar conflitos de interesses, sejam reais ou disfarçados, que possam comprometer a avaliação.
A7 - Responsabilidade Fiscal	As avaliações devem prestar contas de todos os recursos despendidos e cumprir processos e procedimentos fiscais determinados pela lei.

Fonte: Os autores (2024).

A classificação do nível de atendimento de cada padrão de cada categoria foi estabelecida conforme os critérios a seguir:

- Atendimento à categoria de Utilidade: foram consideradas como atendidas plenamente se todos os seus padrões foram atendidos; foram consideradas como atendidas parcialmente se, pelo menos, quatro de seus padrões foram atendidos; e, foram consideradas como não atendidas se menos de quatro de seus padrões foram atendidos.
- Atendimento da Categoria de Exequibilidade: esta categoria foi atendida plenamente quando todos os seus padrões foram atendidos; foi considerada como atendida parcialmente quando, pelo menos, dois de seus padrões foram atendidos; e, foi considerada como não atendida, quando menos de dois de seus padrões foram atendidos.
- Atendimento da Categoria de Adequação: esta categoria foi atendida plenamente quando todos os seus padrões foram atendidos; foi considerada como atendida parcialmente quando, pelo menos, quatro de seus padrões foram atendidos; e, foi considerada como não atendida, quando três ou menos de seus padrões foram atendidos.
- Quando o padrão não foi identificado no estudo, mas o fato não configurar seu não atendimento, não será considerada sua classificação como Não Atendido.

O juízo de valor sobre o atendimento, ou não, dos padrões de todas as categorias será construído em função da leitura rigorosa do relatório de avaliação por parte dos meta-avaliadores. Nesse aspecto, cabe ressaltar que essa construção teve caráter qualitativo, estando, portanto, sujeita a contrapontos elaborados a partir de outras interpretações.

Para nortear seu processo de construção o estudo procurou responder a seguinte questão meta-avaliativa:

Em que medida o Relatório de Retorno Social do Investimento para os cursos técnicos do Centro Educacional Assistencial Profissionalizante (CEAP), está de acordo com os padrões de qualidade da avaliação do *Joint Committee* (Yarbrough; Shulha; Hopson; Caruthers, 2011), considerando as Categorias de Utilidade, Exequibilidade, Adequação?

RESULTADO E ANÁLISE

Os resultados da meta-avaliação, com suas concernentes análises, seguem relacionados e estruturados por categoria com seus respectivos padrões.

CATEGORIA UTILIDADE

A seguir, estão expressos os resultados alcançados por cada padrão desta categoria:

- a. Credibilidade do Avaliador: o primeiro Padrão de Utilidade é a Credibilidade do Avaliador e por estar ligado à qualificação do avaliador, considera-se que esse padrão foi **Atendido**, visto que o Relatório foi desenvolvido pelo Instituto para o

Desenvolvimento Social que é uma organização social independente fundada em 1999 e pioneira no apoio técnico ao investidor social no Brasil, possuidora de grande *expertise* profissional. Este órgão já analisou o impacto social, através da metodologia *SROI*, de instituições como Amigos do Bem, Doutores da Alegria e Gerando Falcões, o que comprova que ele possui relevante capacidade técnica e experiências em processos avaliativos de programas sociais da iniciativa privada

- b. Atenção aos Interessados (*stakeholders*): apesar de a metodologia *SROI* incluir na avaliação somente os alunos dos cursos técnicos e seus familiares, que experimentam mudanças relevantes como resultado das atividades do projeto em análise, o Relatório deixa claro que o escopo do estudo foi delimitado por meio de entrevistas com pessoas indicadas pela própria organização, com diferentes visões a respeito dos cursos, como: ex-diretor de formação atualmente trabalhando como professor, preceptor e orientador na formação de pais; professor/preceptor; coordenador pedagógico; ex-aluno e atual supervisor em uma empresa parceira do CEAP; a diretora de operação desta empresa, que é a maior contratante de ex-alunos do CEAP; e a mãe de um aluno que frequenta o CEAP há 4 anos e que estava prestes a iniciar o Curso de Técnico em Redes de Computadores, o que demonstra a participação dos principais *stakeholders* na avaliação. Considera-se, assim, o padrão **Atendido**.
- c. Propósitos Negociados: de acordo com o padrão Propósitos Negociados, são os propósitos da avaliação que conduzem as decisões seguintes acerca do programa/projeto. Segundo ele, a finalidade da avaliação deve ser clara. E o relatório deixa bem claro quais são seus objetivos, conforme trecho a seguir:

Os principais objetivos desta avaliação são:

Compreender o impacto dos cursos técnicos através de um estudo avaliativo que demonstre a efetividade do investimento do CEAP;

Apoiar o planejamento estratégico e o processo de tomada de decisões do CEAP quanto aos desdobramentos e continuidade dos cursos técnicos;

Auxiliar o CEAP a identificar potenciais pontos de aprimoramento e focos prioritários de ação nos cursos técnicos oferecidos pela escola” (Brasil, 2018, p. 10). Sendo assim, considera-se que o padrão Propósitos Negociados foi atendido.

Logo, considera-se o padrão Propósitos Negociados classificado como **Atendido**.

- d. Explicitação de Valores: o relatório explicita os valores culturais do CEAP, como: inserção dos jovens no mercado de trabalho através da formação profissionalizante; preocupação com a formação humana dos jovens atendidos, seja através da preceptoria ou das disciplinas de Ética e Religião; e, além disso, por acreditar que as famílias dos alunos cumprem importante papel em sua formação humana, o CEAP proporciona também o curso de Formação de Pais, visando ajudá-los a desempenhar melhor a missão de orientar seus filhos. Por todo o exposto, entende-se que o padrão Explicitação de Valores foi **Atendido**.
- e. Informação Relevante: segundo o padrão Informação Relevante, a informação deve atender aos objetivos estabelecidos. O estudo avaliativo tinha como objetivo demonstrar a efetividade do investimento feito pelo CEAP, o que consequentemente apoiaria sua tomada de decisões. Além disso, deveria auxiliar o CEAP a identificar potenciais pontos de aprimoramento e focos prioritários de ação nos cursos técnicos oferecidos pela escola. E o Relatório apresentou como

resultados a comprovação que a iniciativa traz benefícios sociais relevantes para seus participantes que, inclusive, excedem em três vezes o investimento realizado. O relatório aponta ainda, pontos de aprimoramento relativos à preceptoria, às normas da escola, ao ensino religioso e às aulas de ética. Desta forma, considera-se que o padrão Informação Relevante foi **Atendido**.

- f. Produtos e Processos Significativos entendimentos e comportamentos: o Relatório apresenta recomendações úteis e plausíveis, que têm o objetivo de rever entendimentos e comportamentos da Instituição. Quanto à preceptoria, foi identificado o desejo por parte dos alunos, de que a atividade ocorresse mais do que uma vez por mês, durante 50 minutos. Logo, uma forma de tornar a atuação do CEAP ainda mais efetiva poderia ser o aumento da frequência ou da duração dessas sessões. Os beneficiários apontaram alguns pontos de rigidez das normas da escola, como a proibição de se utilizar adereços como bonés e brincos dentro da escola. Os alunos também manifestaram seu interesse por momentos de reflexão sobre outras religiões além do catolicismo. Eles também manifestaram interesse por mais espaço para diálogo durante as aulas de ética. Logo, pode-se considerar que o padrão foi **Atendido**.
- g. Comunicação e Relatórios Apropriados e no Prazo: o documento tem linguagem simples e clara e seus resultados foram expostos de forma bastante compreensível para todos os interessados. Entretanto, o relatório não contemplou informações sobre prazos e período de realização. Desta forma, considera-se o padrão **Atendido Parcialmente**.
- h. Preocupação com Consequências e Influências: apesar de entendermos a abrangência e a importância dos resultados desta avaliação, não localizamos no Relatório nenhuma

informação sobre a utilização deles. Sabe-se que o *SROI* é menos útil quando estiver sendo feito somente para provar o valor de uma iniciativa e não houver oportunidade de mudar a maneira como as coisas são feitas de acordo com o resultado da análise (Instituto para o Desenvolvimento do Investimento Social, 2012). Sendo assim, considera-se o padrão **Atendido Parcialmente**.

Considerando os critérios estabelecidos para a valoração deste atributo de utilidade da avaliação, tendo em vista que dos oito padrões, seis foram atendidos e dois foram atendidos parcialmente, conclui-se que sua avaliação é a de "Atendida Parcialmente".

CATEGORIA EXEQUIBILIDADE

A seguir, estão expressos os resultados alcançados por cada padrão desta categoria:

- a. Gerenciamento do Projeto: A estratégia apresentada no Relatório assegurou o atingimento dos objetivos da avaliação. Além disso, apresentou boa comunicação dos resultados e das recomendações da avaliação, o que corrobora seu bom gerenciamento. No entanto, o relatório não contemplou informações sobre prazos, período de realização, custos e orçamento. Logo, este padrão, que está intimamente ligado ao padrão U7- Comunicação e Relatórios Apropriados e no Prazo, também foi **Atendido Parcialmente**.
- b. Procedimentos Práticos: o Relatório apresentou caráter prático e responsivo em relação a seus princípios, valores, objetivos e eixos de mudança, como se pode ver no trecho a seguir:

A avaliação de Retorno Social do Investimento dos cursos técnicos do CEAP comprova que a iniciativa traz benefícios sociais relevantes para seus participantes em

todos os eixos de mudança pretendidos, que excedem em 3 vezes o investimento realizado (Instituto para o Desenvolvimento do Investimento Social, 2018).

Sendo assim, considera-se que este padrão foi **Atendido**.

- c. Viabilidade do Contexto: pode-se considerar o padrão como **Atendido Parcialmente**, tendo em vista que, em função da metodologia *SROI*, a avaliação não considerou todos os *stakeholders* influenciados pelos cursos técnicos, ou os que contribuem para a sua realização, conforme pode ser constatado no trecho a seguir:

Para a metodologia *SROI*, deve-se incluir na avaliação somente os *stakeholders* que experimentam mudanças relevantes como resultado das atividades do projeto em análise. Assim, para esta avaliação, foram considerados somente os *stakeholders* significativamente afetados pelas atividades dos cursos técnicos do CEAP, definidos através da quantidade de mudança material experimentada e sua permanência ao longo do tempo (Instituto para o Desenvolvimento do Investimento Social, 2018).

- d. Uso dos Recursos: não constam no Relatório quaisquer informações sobre a utilização de recursos. Sendo assim, considera-se o padrão como **Não Atendido**. Portanto, em relação à exequibilidade, um padrão foi avaliado como atendido, dois foram considerados como atendidos parcialmente e um como não atendido. Desse modo, considerando os critérios adotados, esta categoria foi avaliada como **Não Atendida**.

CATEGORIA ADEQUAÇÃO

A seguir são expressos os resultados da Categoria da Adequação.

- a. Orientação Responsiva e Inclusiva: nesse padrão a avaliação em referência evidencia que os *stakeholders* foram

claramente identificados, suas informações foram coletadas e consideradas, e foram incluídos nas decisões acerca da avaliação. Os avaliadores demonstraram procurar conhecer os *stakeholders* e a cultura, bem como buscaram o equilíbrio adequado entre responsividade e inclusão. Nesse aspecto, se deve considerar que os,

Processos de avaliação responsivos e inclusivos produzem consequências éticas e morais de longo alcance, além de servirem para melhorar a utilidade, exequibilidade, precisão e responsabilização (Yarbrough; Shulha; Hopson; Caruthers, 2011, p. 5).

Desse modo, pode-se inferir que o padrão foi contemplado no relatório da avaliação e, assim, avaliado como **Atendido**.

b. Acordos Formais: Em relação a esse padrão, considera-se:

Em geral, acordos formais são escritos, mas acordos e compromissos verbais também podem ser legítimos. [...] em circunstâncias especiais, nas quais somente é possível o compromisso verbal, testemunhas confiáveis ou registros feitos em colaboração servem como comprovação das condições negociadas (Yarbrough; Shulha; Hopson; Caruthers, p. 9).

Na análise do relatório não foram identificados documentos que comprovem as obrigações de forma explícita. No entanto, considera-se que a documentação de avaliações dessa natureza, comumente gera outros documentos em cujo teor pode ter contemplado esse aspecto (Yarbrough; Shulha; Hopson; Caruthers, 2011). Assim, o padrão foi julgado como **Atendido Parcialmente**, pois considera-se que o Relatório final poderia citar esse aspecto, informando a existência de um acervo de documentação gerado pela avaliação.

c. Respeito e Direitos Humanos: o Padrão Respeito e Direitos Humanos está presente com o reconhecimento dos valores e bagagens sociais e culturais dos *stakeholders* e de suas

comunidades. Contudo, uma ressalva deve ser feita sobre a inexistência de comprovação da informação aos participantes de que eles devem ser informados de seus direitos, inclusive de participar, desistir de participar, ou de questionar decisões tomadas em qualquer etapa do processo de avaliação (Yarbrough; Shulha; Hopson; Caruthers, p. 12). Tendo em vista que no relatório não consta qualquer observação sobre esse aspecto e considerando a relevância de sua exposição no documento, considera-se que o padrão **Não foi Atendido**.

- d. Clareza e Equidade: para Yarbrough, Shulha, Hopson, Caruthers (2011), para atendimento ao padrão Clareza e Equidade, os avaliadores precisam estar cientes dos papéis que exercem, usando-os positivamente, não devendo exagerar na autoridade, quando em contato com os *stakeholders*, não podem esconder detalhes do processo, promover agenda própria nem agir com parcialidade, dentro ou fora de sua área de atuação. Logo, em função do relatório não expressar qualquer fato contrário ao conceito do padrão, entende-se que ele foi **Atendido**.
- e. Transparência e Abertura: o atendimento ao quinto padrão da Adequação, Transparência e Abertura, possibilita que injustiças sejam evitadas, proporcionando aos envolvidos a possibilidade de agirem com o acesso às informações disponíveis. O relatório não expõe qualquer fato que não justifique o atendimento deste padrão, informando somente que os relatos dos desenvolvimentos normais dos processos implementados. Assim, esse padrão foi considerado como **Atendido**.
- f. Conflitos de Interesses: para o padrão de Adequação, Conflitos de Interesse, não foram relatados no relatório informações a serem aplicadas, podendo-se considerar que esse tipo de ocorrência não foi identificada ou não houve casos de conflitos no contexto da avaliação. Logo o padrão não foi julgado como **Atendido**.

- g. Responsabilidade Fiscal: o Padrão Responsabilidade Fiscal não foi identificado no processo de meta-avaliação, não permitindo acesso às informações e registros sobre a viabilidade financeira da avaliação (Yarbrough; Shulha; Hopson; Caruthers, 2011, p. 25). Devido à importância de a avaliação transcorrer registrando-se o controle fiscal durante todo desenvolvimento do processo, inclusive com registro em seu relatório final, considera-se que esse padrão como **Não Atendido**.

O Quadro 8 apresenta os resultados da Categoria da Adequação.

Quadro 8 - Resultado do estudo meta-avaliativo categoria adequação

Padrão	Nível de Atendimento
A1 - Orientação Responsiva e Inclusiva	A
A2 - Acordos Formais	AP
A3 - Respeito e Direitos Humanos	NA
A4 - Clareza e Equidade	A
A5 - Transparência e Abertura	A
A6 - Conflitos de Interesses	A
A7 - Responsabilidade Fiscal	NA

Legenda: A - Atende; AP - Atende Parcialmente; NA - Não atende.

Fonte: Os autores (2024).

Em relação à Adequação, dos sete padrões, quatro padrões foram **Atendidos**, um foi **Atendido Parcialmente** e dois **Não Foram Atendidos**, logo, de acordo com os critérios estabelecidos, a categoria Adequação foi **Parcialmente Atendida**.

CONCLUSÕES E RECOMENDAÇÕES

Após a realização de uma análise detalhada da Avaliação transcrita no Relatório de Retorno Social do Investimento para os cursos técnicos do Centro Educacional Assistencial Profissionalizante (CEAP), a partir de uma visão dos Categorias do *Joint Committee on Standards for Educational Evaluation*, os meta-avaliadores realizaram a meta-avaliação conforme resultados acima detalhados e chegaram às seguintes conclusões e recomendações.

No que se refere à **Categoria Utilidade** da avaliação, foi observado que o Relatório avaliado poderia ter contemplado mais informações sobre prazos e período de realização da avaliação. Além disso, não foram localizadas no Relatório informações sobre a utilização dos resultados obtidos. Vale destacar que uma avaliação deve ser totalmente transparente e, sempre que possível, deve servir de modelo para outras avaliações.

Para a **Categoria Exequibilidade**, foi observado que o Relatório avaliado poderia ter contemplado informações sobre prazos, período de realização, custos, orçamento e utilização de recursos. Além disso, em função da metodologia *SROI*, a avaliação não considerou todos os *stakeholders* envolvidos.

Os padrões da **Categoria Adequação** devem priorizar aquilo que é apropriado, válido, legal, correto, aceitável e justo na avaliação (Yarborough; Shulha; Hopson; Caruthers, 2011). A seguir, alguns comentários que impactaram essa categoria na meta-avaliação realizada:

O período utilizado para o cálculo do investimento foi do ano de 1994 ao ano de 2017 (Instituto para o Desenvolvimento do Investimento Social, 2018). Devemos observar que este período contempla um total de 23 (vinte e três) anos. Em um país com as

características de instabilidade socioeconômica como o Brasil, torna-se questionável e imprecisa a comparação de valores, mesmo que atualizados pelo IPCA por tão longo período (Instituto para o Desenvolvimento do Investimento Social, 2018). Assinalamos que o IPCA (Índice Nacional de Preços ao Consumidor Amplo) tem por objetivo medir a renda da população brasileira com renda entre 1 e 40 salários-mínimos (IBGE, [2024]). Para a avaliação em referência, considerando o perfil socioeconômico dos envolvidos, o entendimento é que esse não foi o melhor índice a ser utilizado.

A avaliação do CEAP mede o impacto social gerado pelos cursos de educação profissional técnica: Técnico em Administração e Empreendedorismo, Técnico em Redes e Técnico em Informática (Instituto para o Desenvolvimento do Investimento Social, 2018). Contudo, no relatório em referência, não foram discriminados de forma clara e precisa, quantos alunos ingressaram por ano/curso, quantos alunos não concluíram os respectivos cursos, quantos alunos foram formados por curso/ano. Não tem informações sobre a partir de que ano os cursos foram oferecidos, e se foram realizadas mudanças de currículo nos referidos cursos. E, considerando que não foram mencionadas outras avaliações realizadas na, ou pela instituição, entendemos que a ausência dessas informações impacta na composição final de uma avaliação que tem por um dos objetivos auxiliar o CEAP a identificar potenciais pontos de aprimoramento e focos prioritários de ação nos cursos técnicos oferecidos pela instituição (Instituto para o Desenvolvimento do Investimento Social, 2018).

Nas orientações do Instituto para o Desenvolvimento do Investimento Social sobre o *SROI*, é mencionado que a falta de bons dados acerca de resultados é um dos principais desafios quando se faz um *SROI* pela primeira vez (Instituto para o Desenvolvimento do Investimento Social, 2012). E essa informação também se repete em outra questão: no questionário que foi aplicado, na questão 3, os respondentes tinham que assinalar em qual curso se formou, onde, além dos cursos mencionados acima, também constam os cursos Técnico

em Telecomunicações, Técnico em Eletrônica ou outro curso técnico. Considerando que o foco da avaliação eram os cursos Técnico em Administração e Empreendedorismo, Técnico em Redes e Técnico em Informática (Instituto para o Desenvolvimento do Investimento Social, 2018), ficamos em dúvida se os respondentes dos demais cursos foram excluídos da análise.

Uma informação que não ficou claramente definida foi a relação da instituição CEAP com a comunidade do entorno. A seleção dos alunos é realizada com base em critérios que incluem a priorização de estudantes com perfil socioeconômico vulnerável e que vivam na região de Pedreira, bairro onde a escola está situada, e em outros bairros próximos (Instituto para o Desenvolvimento do Investimento Social, 2018). Contudo, no quadro sumário dos *stakeholders* envolvidos e não envolvidos, na análise das razões para exclusão da “Comunidade do entorno” foi a de que não foram encontrados elementos que evidenciem que esse impacto é material (Instituto para o Desenvolvimento do Investimento Social, 2018). Para uma avaliação que tem por objetivo analisar a sua contribuição para a sociedade, devemos entender que a comunidade do entorno é uma das principais vertentes impactadas, precedida somente pela família dos respondentes.

Quanto aos participantes dos grupos focais, observamos que são ex-alunos, formados entre 2014 e 2016, alunos atuais e pais de ex-alunos formados em 2017. Os respondentes do questionário são em maioria ex-alunos de outros anos e permanece a dúvida de como foram analisadas as referências dos cursos, conforme mencionado anteriormente.

No que se refere à formatação do relatório, cabe ressaltar que as normas da Associação Brasileira de Normas Técnicas (ABNT) não foram atendidas para a construção dos Apêndices. Para a ABNT, o apêndice é o texto ou documento elaborado pelo autor, a fim de complementar sua argumentação, sem prejuízo da unidade nuclear

do trabalho, (Associação Brasileira de Normas Técnicas, 2011) contudo, com exceção do Apêndice 3, os demais Apêndices apresentaram características a serem comentadas no quadro a seguir:

Quadro 11 – Comentários Apêndices

Apêndice	Comentário
Apêndice 1 - Referências para saber mais sobre a metodologia <i>SROI</i>	Um dos sites mencionados não está válido.
Apêndice 2 - Sumário das justificativas para a não inclusão dos demais <i>stakeholders</i> nesta avaliação <i>SROI</i>	Informações importantes para a unidade nuclear do trabalho.
Apêndice 3 - Roteiros das Entrevistas de Diagnóstico	Ok.
Apêndice 4 - Roteiro e Sistematização dos Grupos Focais (1ª Fase de trabalho de campo-qualitativa)	Informações importantes para a unidade nuclear do trabalho. Considerando que os grupos focais tiveram um mínimo de 3 participantes e um máximo de 12 participantes, recomendamos atenção para a utilização de vocabulário não adequado para avaliações nos destaques dos comentários dos participantes dos grupos focais, conforme a seguir: vários participantes, consenso, alguns participantes, muitos participantes.
Apêndice 5 - Roteiro e Sistematização de Resultados do Questionário Quantitativo	Informações importantes para a unidade nuclear do trabalho. As informações da sistematização de resultados do questionário quantitativo são fundamentais para a avaliação.
Apêndice 6 - Cálculo das proxies financeiras	Informações importantes para a unidade nuclear do trabalho.
Apêndice 7 - Sugestões e expectativas dos <i>stakeholders</i>	Ausência de justificativa para a não utilização das sugestões de melhoria e expectativas para o futuro.
Apêndice 8 - Bibliografia	Padrão ABNT de Referência e necessidade de incluir os sites mencionados nas <i>proxies</i> utilizadas.
Apêndice 9 - Glossário	Além de não atender ao padrão da ABNT, nem todas as palavras/ conceitos utilizados no relatório constam do Glossário.

Fonte: Os autores (2024).

Concluindo a meta-avaliação, em resposta à questão avaliativa: “Em que medida o Relatório de Retorno Social do Investimento para os cursos técnicos do Centro Educacional Assistencial Profissionalizante (CEAP), está de acordo com os padrões de qualidade da avaliação do *Joint Committee* (Yarbrough; Shulha; Hopson; Caruthers, 2011), considerando as Categorias de Utilidade, Exequibilidade, Adequação?”, os resultados apontaram para uma avaliação negativa já que nenhuma das três categorias foi considerada atendida. As categorias Utilidade e Adequação foram **Atendidas Parcialmente**, já a categoria Exequibilidade, **Não foi Atendida**.

MÉRITO DO ESTUDO

Elaine Maria de Andrade Senra

O artigo intitulado “Meta-avaliação da Avaliação do Impacto Social dos Cursos Técnicos CEAP” apresenta uma contribuição bastante significativa para o campo da meta-avaliação, principalmente pela escolha de um referencial teórico sólido, sustentado nos padrões do *Joint Committee on Standards for Educational Evaluation*.

Ao investigar como se dá a aderência do Relatório de Avaliação do Retorno Social sobre o Investimento (SROI) às categorias de Utilidade, Exequibilidade e Adequação, o texto revela atenção aos aspectos essenciais da boa prática avaliativa e oferece um panorama claro dos pontos fortes e das limitações da avaliação em questão.

A relevância temática é evidenciada pela oportunidade de ampliar o debate sobre a qualidade das avaliações de impacto social, cada vez mais exigidas em cenários educacionais e de políticas públicas. O foco na meta-avaliação — isto é, na avaliação da própria avaliação — se alinha a tendências contemporâneas que exigem transparência, rigor metodológico e prestação de contas em diferentes

tipos de programas e projetos. Assim, o estudo não apenas contribui para a teoria avaliativa, mas também repercute positivamente na prática, ao propor recomendações que podem ser aplicadas por gestores e organizações responsáveis para implementar e monitorar iniciativas sociais.

Outro ponto de mérito é a clareza do referencial metodológico: os padrões do *Joint Committee* são apresentados de forma clara, incluindo exemplos de como cada padrão se materializa no Relatório em análise. Ao adotar um sistema de classificação (Atendido, Atendido Parcialmente e Não Atendido), o artigo gera um resultado objetivo e didático que, por sua vez, contextualiza a análise e permite aos leitores compreenderem facilmente as lacunas encontradas, como a ausência de dados sobre prazos, custos e abrangência completa dos interessados. Esse recurso também fornece subsídios para que avaliações futuras evitem as mesmas limitações, contribuindo para uma cultura de melhoria contínua.

Além disso, o texto revela originalidade ao escolher como objeto de meta-avaliação um estudo de impacto social concreto, desenvolvido de forma cronológica por uma organização independente, com expertise reconhecida na metodologia SROI. Esse panorama empírico permite estabelecer um diálogo construtivo entre teoria e prática, aspecto fundamental para evitar que a meta-avaliação se restrinja ao plano meramente conceitual. Ao evidenciar inconsistências tanto na coleta de dados quanto na apresentação dos resultados, o artigo ressalta a importância de registrar, de maneira sistemática e transparente, não apenas os resultados obtidos, mas também os processos e recursos mobilizados ao longo da investigação.

O trabalho ancora-se ainda em referências teóricas contemporâneas, como Scriven, Stufflebeam e Yarbrough, entre outros autores que se dedicam ao aprofundamento da meta-avaliação. Essa base teórica sólida, aliada à organização coerente do texto, assegura uma comunicação clara e de fácil compreensão para diferentes

perfis de leitores, reforçando o valor da proposta apresentada. Embora a pesquisa seja essencialmente qualitativa e se concentre na análise de um relatório final — sem realizar coletas primárias de dados —, a forma sistematizada como o estudo se desenvolve e as orientações que propõem oferecem uma contribuição significativa para a literatura sobre avaliação de impacto social.

Em síntese, a força do artigo reside na articulação dos referenciais teóricos com a aplicação prática de um caso real, oferecendo subsídios tanto para pesquisadores quanto para avaliadores e gestores de programas sociais. O rigor analítico, a coerência com a linha editorial que valoriza estudos avaliativos e o potencial de gerar impacto positivo no refinamento de práticas de avaliação são fatores que tornam o trabalho meritório e relevante para a área.

REFERÊNCIAS

ARAÚJO, Nayanna Moreno Miranda; SANTOS, Olívia Meireles Ribeiro dos; AGUIAR, Glauco da Silva. Relatório de avaliação de impactos do programa de desenvolvimento e crescimento integrado com sustentabilidade: um estudo meta avaliativo. *In*: ELLIOT, Lígia Gomes; ORLANDO FILHO, Ovídio (org.). **Estudos meta avaliativos: aplicações metodológicas**. São Paulo: Pimenta Cultural, 2020. p. 145-176.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 15287**: 2011: Projeto de pesquisa: apresentação. 2. ed. Rio de Janeiro: ABNT, 2011.

ELLIOT, Lígia Gomes. Meta-avaliação: das abordagens às possibilidades de aplicação. **Ensaio: aval. pol. públ. educ.**, Rio de Janeiro, v. 19, n. 73, p. 941-964, 2011.

FURTADO, Juarez Pereira; LAPERRIÈRE, Hálène. Parâmetros e paradigmas em meta-avaliação: uma revisão exploratória e reflexiva. **Ciênc. Saúde Coletiva**, Rio de Janeiro, v. 17, n. 3, p. 695-705, 2012. DOI: <https://doi.org/10.1590/S1413-81232012000300016>. Disponível em: <https://www.scielo.br/j/csc/a/jpWHQCWR6RtDzkZDcFrL6qv/#>. Acesso em: 5 out. 2024.

IBGE. IPCA: índice nacional de preços ao consumidor amplo. **ibge.gov**, Brasília, DF, [2024]. Disponível em: <https://www.ibge.gov.br/estatisticas/economicas/precos-e-custos/9256-indice-nacional-de-precos-ao-consumidor-amplo.html?=&t=o-que-e>. Acesso em: 21 set. 2024.

INSTITUTO PARA O DESENVOLVIMENTO DO INVESTIMENTO SOCIAL. **Avaliando o impacto social dos Cursos Técnicos CEAP**. São Paulo: IDIS, 2018. Disponível em: <https://www.idis.org.br/publicacoesidis/avaliando-o-impacto-social-dos-cursos-tecnicos-ceap/>. Acesso em: 8 out. 2024.

INSTITUTO PARA O DESENVOLVIMENTO DO INVESTIMENTO SOCIAL. **Um guia para o retorno social do investimento**. São Paulo: IDIS, 2012. Disponível em: https://www.idis.org.br/wp-content/uploads/2016/09/GUIA_SROI_PT_2.pdf. Acesso em: 8 out. 2024.

JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION. Sobre o JCSEE. **Evaluationstandards.org**, [S. l.], 2020. Disponível em: <https://evaluationstandards.org/>. Acesso em: 5 out. 2024.

MENEZES FILHO, Naercio; PINTO, Cristiane Campos de Xaxier (org.). **Avaliação econômica de projetos sociais**. 3. ed. São Paulo: Fundação Itaú Social, 2017.

PENNA FIRME, Thereza; LETICHEVSKY, Ana Carolina. O desenvolvimento da capacidade de avaliação no século XXI: enfrentando o desafio através da meta-avaliação. **Ensaio: aval. pol. públ. educ**, Rio de Janeiro, v. 10, n. 36, p. 289-300, 2002.

RIBEIRO, Tatiana de Sousa; CARABAJAL, Yves de Carvalho; ORLANDO FILHO, Ovidio. Relatório de avaliação do projeto futebol: uma nova visão do jogo: meta-avaliação. *In*: ELLIOT, Ligia Gomes; ORLANDO FILHO, Ovidio. (org.) **Meta avaliações de políticas públicas do governo**: o que revelam os relatórios de avaliação. São Paulo: Pimenta Cultural, 2019. p. 12-44.

SCRIVEN, Michael. **Avaliação**: um guia de conceitos. Rio de Janeiro: Paz e Terra, 2018.

SOUZA, Francisco José de; SOUZA, Jurema Santos; ELLIOT, Ligia Gomes; HILDENBRAND, Luci. Avaliação da educação superior no Brasil: do Provão ao ENADE: um estudo meta-avaliativo. **Meta: Avaliação**, Rio de Janeiro, Edição Especial, n. 2, p. 22-42, 2018. DOI: <http://dx.doi.org/10.22347/2175-2753v10i0.1956>. Disponível em: <https://revistas.cesgranrio.org.br/index.php/metaavaliacao/article/view/1956>. Acesso em: 9 set. 2024.

STUFFLEBEAM, Daniel L. Meta-evaluation. **Journal of MultiDisciplinary Evaluation**, [S. l.], v. 7, n. 15, p. 99-158, 2011. Disponível em: https://journals.sfu.ca/jmde/index.php/jmde_1/article/view/300/295. Acesso em: 9 jan. 2025.

WORTHEN, Blaine R.; SANDERS, James R.; FITZPATRICK, Jody L. **Avaliação de programas: concepções e práticas**. São Paulo: Editora Gente, 2004.

YARBROUGH, Donald B.; SHULLA, Lyn M.; HOPSON, Rodney K.; CARUTHERS, Flora A. **The program evaluation standards: a guide for evaluators and evaluation users**. 3. ed. Thousands Oaks, Califórnia: Sage, 2011.

9

*Andréa Cristina Tavares Moreira
José Rafael Ayres da Motta
Mariana Sá Alcântara Gomes*

META-AVALIAÇÃO DA AVALIAÇÃO DO PROJETO DE PROMOÇÃO DO ECOSSISTEMA DE INVESTIMENTO E NEGÓCIOS DE IMPACTO SOCIOAMBIENTAL NO BRASIL

A AVALIAÇÃO COMO FERRAMENTA PARA A MELHORIAS DOS PROGRAMAS DE POLÍTICAS PÚBLICAS

A avaliação se dá por meio de um processo de coleta de informações que visa mensurar o valor e a eficiência de um determinado programa, projeto, serviço. Este processo é essencial em diversas áreas, na medida em que fornece dados para tomada de decisões, implementações e definição de ações voltadas para o desenvolvimento e a melhoria daquilo que se avalia ou que se quer avaliar. Segundo autores como Patton (2010) e Scriven (2018), a avaliação é um processo cuja análise de dados permite ou possibilita determinar o valor ou o mérito de um programa, projeto ou intervenção.

Quanto à avaliação de políticas públicas, que embasam o Relatório Executivo Avaliação de Impacto Promoção do Ecossistema de Investimento e Negócios de Impacto Socioambiental no Brasil, objeto deste estudo, importa destacar sua importância uma vez que expressa a geração de valor para os serviços prestados pela administração pública e por seus agentes à sociedade civil. Nele, a partir da análise dos resultados alcançados e da mensuração dos impactos promovidos por essas políticas, é possível identificar se seus objetivos foram atingidos junto a determinada população; se os recursos destinados e investidos foram empregados de forma eficiente; se as ações governamentais contribuem para a melhoria das condições sociais, econômicas e/ou institucionais. Tais análises podem, assim, permitir ajustes e melhorias na política avaliada.

De acordo com Secchi (2010), a avaliação de políticas públicas é essencial não só para verificar se seus objetivos iniciais foram cumpridos, mas também para possibilitar ajustes e reformulações necessários. O autor destaca que, ao avaliar os resultados e os impactos, é possível medir a eficácia, a eficiência e a equidade da

política implementada, favorecendo a *accountability* e a transparência pública. A avaliação contínua facilita a identificação de boas práticas, que podem ser replicadas, quanto falhas ou ineficiências, que podem ser corrigidas. Desta forma, as políticas públicas tornam-se mais eficientes e eficazes, alinhando-se às reais necessidades da população.

Os processos avaliativos permitem uma melhor manutenção de projetos e campanhas públicas, além de viabilizarem a criação de novas iniciativas, pautadas em dados fundamentados na realidade objetiva, de modo a garantir que os recursos públicos disponíveis sejam direcionados e bem aplicados nas áreas que realmente precisam ou se fazem necessários, promovendo, assim, maior impacto no atendimento das necessidades populacionais.

Bardach (2006) afirma que a avaliação de políticas públicas oferece dados cruciais para o processo de tomada de decisão, permitindo que gestores e administradores possam identificar seus funcionamentos, otimizando a alocação de recursos e aprimorando as melhores práticas administrativas.

Contudo, não basta apenas avaliar as políticas públicas. É importante avaliar também a qualidade das próprias avaliações realizadas. Esta meta processo, conhecido como avaliação da avaliação ou meta-avaliação, verifica se os critérios avaliativos, os métodos e as técnicas metodológicas aplicados pelos avaliadores foram adequados e imparciais, ajudando e corroborando para a precisão, fidedignidade e confiabilidade dos resultados, de modo a orientar e aprimorar as políticas públicas.

Fitzpatrick, Sanders e Worthen (2011) destacam a importância da avaliação das políticas públicas para o fortalecimento das instituições sociais, afirmando que, por meio dela, desde que criteriosamente conduzida e concebida, é possível garantir se, de fato, as políticas públicas estão atendendo as reais necessidades da população-alvo. Desse modo, a avaliação deve ser vista como uma

ferramenta de aprendizado organizacional que permite ajustes em tempo real, contribuindo para uma administração pública eficaz.

Chelimsky (2009) argumenta que a meta-avaliação, ou seja, a avaliação da avaliação, é fundamental para assegurar a credibilidade do processo avaliativo. O autor também destaca que, ao revisar os métodos utilizados, os avaliadores garantem que as avaliações sejam conduzidas de forma rigorosa e imparcial, o que reforça a confiabilidade dos resultados obtidos e sua aplicação na melhoria das políticas públicas.

CONCEITUANDO META-AVALIAÇÃO

Existem duas formas de avaliar uma avaliação: durante o seu desenvolvimento ou após a sua conclusão. Scriven (2018) destaca que a avaliação formativa, que ocorre durante o andamento do programa, visa melhorar sua aplicabilidade e, a avaliação somativa, que ocorre no final do programa, tem como objetivo julgar os seus impactos e a sua eficiência.

Independentemente da avaliação ser formativa ou somativa, Michel Patton (2010) se concentra em estudar como os resultados avaliativos podem ser utilizados para a tomada de decisões, ou seja, como os achados da avaliação podem contribuir para a melhoria do objeto avaliado.

O processo avaliativo denominado de meta-avaliação, que consiste na avaliação de uma avaliação, foi descrito por volta da década de 1970, por Michael Scriven. O conceito se pautou na necessidade de se garantir que as avaliações realizadas fossem válidas, confiáveis e úteis, promovendo a melhoria contínua dos métodos e das práticas de avaliação.

Segundo Elliot (2011, p. 943) a meta-avaliação ocupa-se da verificação da qualidade da avaliação,

À luz de diversos critérios, quando o meta-avaliador contratado ao término da avaliação focaliza o relatório produzido, realizando no caso, uma meta-avaliação somativa. Além dessa, Berends e Roberts (2003) destacaram duas razões para se proceder a uma meta-avaliação. Uma é possibilitar a verificação da aplicação de um código de comportamento adequado durante todo o estudo, ou seja, a ética adotada pelos avaliadores, se ela existe e como foi praticada. Outra concebe a meta-avaliação como uma prática de reflexão sobre todos os procedimentos utilizados na avaliação, dando oportunidade à incorporação de novos conhecimentos por parte não só dos meta-avaliadores, como também dos avaliadores, quando tomam ciência da análise realizada.

Essa ideia é também compartilhada por Schwandt (1989), que afirma que a meta-avaliação é um método para a certificação da qualidade da avaliação realizada, exigindo exame dos procedimentos que nela foram utilizados, o que envolve fundamentações, critérios avaliativos, análises diversas e cálculos para o julgamento da qualidade do estudo avaliativo.

O termo meta-avaliação foi cunhado pelo pesquisador Michael Scriven, da Claremont Graduate University, no artigo *An introduction to meta evaluation* (Scriven, 1969). Embora existam registros da expressão avaliação de avaliações, ou *evaluation of evaluations*, descrito pelo filipino Pedro Orata, nos anos 1940 (Cook; Gruder, 1978 *apud* Furtado; Laperriere, 2012), o termo meta-avaliação é o mais utilizado na literatura especializada. Em parte, começou a ser empregado como resposta às críticas que incidiam sobre a eficácia e a qualidade das avaliações, dos métodos empregado para realizá-las... assim, na década de 1960, a meta-avaliação surgiu no campo da avaliação de programas, tornando evidente a necessidade de se avaliar não apenas os programas em si, mas também os processos e as práticas utilizadas nas avaliações de programas.

Durante a década de 1970, pesquisadores começaram a formalizar o conceito de meta-avaliação. Estudos e contribuições diversas, de nomes reconhecidos em avaliação, como o de Michael Scriven, por exemplo, foram extremamente significativos para o desenvolvimento do campo de conhecimento. Já nas décadas de 1980 e 1990, a meta-avaliação começou a se expandir para além da avaliação de programas, abrangendo áreas como avaliação educacional e avaliação de políticas públicas. Nesta fase, foram desenvolvidas metodologias e critérios específicos para conduzir meta-avaliações, objetivando identificar pontos fortes e fracos nas práticas avaliativas.

A partir dos anos 2000, a meta-avaliação se consolidou como prática essencial nos mais diversos campos de saber: saúde, educação, administração pública, entre outros. O avanço das tecnologias da informação e comunicação e a globalização, por sua vez, aumentaram a complexidade e a quantidade de dados disponíveis para avaliação, ampliando, em consequência, o campo da meta-avaliação. Com isso, foram desenvolvidas novas ferramentas capazes de favorecer que o julgamento das avaliações se assentasse em bases de elevada qualidade. No elenco das possibilidades constava os padrões do *Joint Committee on Standards for Educational Evaluation* (Yarbrough; Schulha; Hopson; Caruthers, 2011), os Critérios fundamentais de Davidson (2005), as Diretrizes para avaliação na América Latina e no Caribe (Ghiano; Wehrle Martínez; Peroni Fiscarelli; Martinic Valencia; Checa Gutiérrez; Amariles Erazo; Guzmán; Soberón Álvarez, 2021) e as Diretrizes para a prática de avaliação do Brasil (Silva; Joppert; Gasparini, 2020).

Cabe destacar que todas essas diretrizes, padrões e critérios se constituem em conjuntos de saberes que podem propiciar boas práticas avaliativas, promovendo avaliações e meta-avaliações de qualidade.

O OBJETO DO ESTUDO E AS ENTIDADES ENVOLVIDAS

O objeto deste estudo meta-avaliativo é a avaliação do Projeto de Promoção do Ecossistema de Investimento e Negócios de Impacto Socioambiental no Brasil, realizada pelo Instituto de Cidadania Empresarial (2021) e registrado no documento intitulado “Relatório Executivo Avaliação de Impacto: Promoção do Ecossistema de Investimento e Negócios de Impacto Socioambiental no Brasil”, que apresenta o resumo da avaliação final do projeto.

O Projeto supracitado tem o objetivo de contribuir com a criação e o crescimento sustentável de negócios de impacto, além de estabelecer inovações focadas na resolução de problemas sociais e ambientais por meio da aceleração do desenvolvimento do ecossistema emergente de investimento e negócios, visando o desenvolvimento e a escalabilidade de soluções inovadoras para os problemas socioambientais do Brasil (Instituto de Cidadania Empresarial, 2021).

O Relatório, apoiado em modelo executivo de avaliação de impacto, foi desenvolvido pelo Instituto de Cidadania Empresarial, em convênio com o Banco Interamericano de Desenvolvimento (BID). Os estudos avaliativos realizados foram conduzidos pelo Instituto para o Desenvolvimento do Investimento Social. O Instituto de Cidadania Empresarial é uma organização da sociedade civil que atua na articulação de líderes e no fomento de inovações sociais capazes de gerar impactos positivos na população de baixa renda, tendo em vista não só a promoção da inclusão social, mas também a redução da pobreza no país.

As organizações sociais envolvidas no projeto são entidades sem fins lucrativos que atuam em áreas de interesse público como saúde, assistência social, meio ambiente, entre outras. Dentre os pontos abordados no documento, o Instituto de Cidadania Empresarial

analisou aspectos relacionados à questões econômicas, sociais e ambientais, favorecendo análises que ajudam na formulação de políticas públicas e no fomento à inovação, incentivando o uso de tecnologias inovadoras e práticas de gestão eficazes para enfrentar os desafios de desenvolvimento na região.

Com relação ao Banco Interamericano de Desenvolvimento (BID), trata-se de uma instituição financeira internacional que visa promover o desenvolvimento econômico e social na América Latina e no Caribe e sua missão é reduzir a pobreza e melhorar a qualidade de vida na região, contribuindo para o desenvolvimento econômico mais inclusivo e sustentável. Possui como prioridades a redução da desigualdade e a melhoria dos serviços públicos, incluindo a eficiência nos gastos de suas administrações.

Quanto ao Instituto para o Desenvolvimento do Investimento Social, trata-se de uma organização social independente, sem fins lucrativos, pioneira no apoio técnico ao investidor social no Brasil e que tem em sua missão a preocupação de contribuir para a redução das desigualdades socioambientais no país. A instituição atua em favor da promoção do investimento social e da mobilização de recursos para promover iniciativas comprometidas com o desenvolvimento da sociedade.

METODOLOGIA UTILIZADA NA META-AVALIAÇÃO

Este estudo teve seu início a partir de pesquisa de caráter qualitativo-exploratório realizada em livros, artigos e documentos relacionados ao tema abordado. Como critérios meta-avaliativos, foram utilizados aqueles estabelecidos por Scriven (2007), em sua Lista Chave de Verificação, conforme citado em estudo de Elliot (2011).

O estudo possui abordagem somativa, pois foi realizado depois de a avaliação original ter sido finalizada, tendo como principal fonte de consulta o relatório avaliativo final. Cabe ressaltar que, segundo Scriven (2018) e de acordo a Stufflebeam (2011), o tipo de abordagem adotado auxilia os interessados a perceber tanto os pontos fortes como as fragilidades da avaliação realizada.

Ainda sobre a metodologia utilizada neste trabalho, a Lista Chave de Verificação de Scriven, tem por finalidade orientar a avaliação de programas, projetos e políticas, subsidiando a tomada de decisões capazes de promover melhorias nesses mesmos objetos avaliativos.

Desse modo, a Lista Chave de Verificação da Avaliação de Scriven configura-se como instrumento fundamental para aprimorar a qualidade e a eficácia das intervenções e/ou das políticas públicas.

Segundo Scriven (2007), a Lista Chave de Verificação de Avaliação é uma ferramenta que ajuda a garantir que as avaliações sejam realizadas de maneira sistemática e abrangente, clarificando os objetivos da avaliação e o que se espera alcançar com ela. Aborda quatro componentes principais assim denominados: Preliminares, Fundamentos, Subavaliações e Conclusões. Tais componentes se fracionam em 18 pontos de verificação, que orientam julgamentos e demandam justificativas para os julgamentos realizados, tal como se fez ao longo desse estudo meta-avaliativo. Os julgamentos são classificados por meio das letras A a E (Quadro 1), devidamente definidas por Davidson (2005), e acessadas pelo estudo a partir da Lista Chave de Verificação de Scriven (2007), adaptada por Elliot (2011).

Quadro 1 - Modelo da lista chave de verificação de Scriven

	Ponto de Verificação	Julgamento (A-E)	Justificativa do Julgamento
Preliminares	I. Sumário Executivo		
	II. Prefácio		
	III. Metodologia		
Fundamentos	1. Antecedentes e Contexto		
	2. Descrições e Definições		
	3. Consumidores		
	4. Recursos		
	5. Valores		
SubAvaliações	6. Processo		
	7. Resultados		
	8. Custos		
	9. Comparações		
	10. Exportabilidade		
Conclusões	11. Síntese		
	12. Recomendações e Explicações		
	13. Responsabilidade e Justificativa		
	14. Relatório de Apoio		
	15. Meta-avaliação		

Legenda: A = atinge todos os aspectos principais deste ponto e os expressa de forma clara e concisa; B = atinge a maioria dos aspectos relativos a este ponto, mas não cobre um ou dois deles, ou trata dos ingredientes principais, mas não é 100% claro; C = trata do ponto, mas perde alguns aspectos cruciais, ou perde vários aspectos importantes, ou é um tanto desorganizado ou pouco claro; D = tem um ou dois elementos que parecem tratar implicitamente do ponto, mas o tratamento é deficiente; E = perde o ponto totalmente (Davidson, 2005, p. 217).

Fonte: Elliot (2011) adaptado de Scriven (2007).

QUESTÃO META-AVALIATIVA

Para atingir seus objetivos, este estudo procurou responder à seguinte questão meta-avaliativa: Até que ponto a avaliação do Projeto de Promoção do Ecossistema de Investimento e Negócios de Impacto Socioambiental no Brasil atendeu aos componentes da Lista Chave de Verificação de Scriven?

CRITÉRIOS DE VALORAÇÃO ADOTADOS PELA META-AVALIAÇÃO

No sentido de valorar cada componente da Lista Chave de Verificação e, assim responder à questão meta-avaliativa do estudo, foram estipulados os seguintes critérios avaliativos:

- a. O componente Preliminares, composto por três pontos de verificação, foi considerado Atendido Totalmente quando todos os seus pontos de verificação obtiveram conceito A; foi Atendido em Maior Parte quando todos os seus pontos de verificação obtiveram, no mínimo, conceito B; foi considerado Atendido em Menor Parte quando um de seus pontos de verificação foi A ou B e os outros dois C; nos demais casos, o componente foi considerado Não Atendido.
- b. Os componentes Fundamentos, SubAvaliações e Conclusões, compostos por cinco pontos de verificação cada um, foram considerados Atendido Totalmente quando todos os pontos de verificação obtiveram conceito A; foram considerados Atendidos em Maior Parte quando todos os seus pontos de verificação obtiveram, no mínimo, conceito B ou quando três dos seus cinco pontos de verificação contemplaram o valor

de A ou B; foram considerados Atendidos em Menor Parte quando, pelo menos, três dos seus cinco pontos de verificação obtiveram conceito igual ou inferior a C; e, foram considerados Não Atendido nos demais casos.

RESULTADOS DA META-AVALIAÇÃO

Tendo como premissa que o primeiro ponto de verificação do componente Preliminares, constante da Lista Chave de Verificação da Avaliação de Scriven, é o Sumário Executivo - que "deve trazer uma visão geral dos resultados de forma sintética e em poucas páginas" (Elliot, 2011, p. 944) -, os meta-avaliadores, amparados na análise dos dados, classificaram-no como A na medida em que atinge todos os aspectos principais deste ponto e os expressa de forma clara e concisa, conforme recomenda Davidson (2005).

O Sumário Executivo trouxe informações claras, concisas e relevantes a este aspecto inicial da Lista Chave de Verificação, consistindo em resumo da avaliação final do Projeto de Promoção do Ecosistema de investimento e Negócios de Impacto Socioambiental no Brasil.

Sobre o prefácio, que integra o componente Preliminares, foram mencionados, a partir de quadros resumo, citados nas páginas 09, 13, 17 e 21 do relatório, os objetivos, os públicos-alvo, atividades realizadas e os impactos do projeto, pontos de verificação deste item. Segundo Elliot (2011, p. 945), ao prefácio cabe

Informar quem encomendou a avaliação e [quais são/ foram] as razões principais [dessa solicitação], quais as questões avaliativas [que se busca ou buscou responder] e quais [foram ou são] os interessados ou quais públicos que aguardam os resultados da avaliação e/ ou deles necessitam.

Assim, mediante a análise realizada por este estudo, os meta-avaliadores atribuíram ao item a classificação B, pois ainda que atinja a maioria dos aspectos relativos a este ponto, não explicita o ponto que demanda informar sobre quem ou qual instituição encomendou o relatório.

Outro Ponto de Verificação é a metodologia avaliativa que, referida a página 7 do relatório, deve mencionar o desenho e a abordagem adotados pela avaliação, dentre outros procedimentos. Assim, devido ao fato de essas informações estarem disponíveis em capítulo especificamente voltado para este ponto, os meta-avaliadores o classificaram como A.

Conforme ilustra o Quadro 2, que remete ao componente Preliminares do relatório, os meta-avaliadores, utilizando-se dos critérios de valoração da meta-avaliação discriminados no item 6 deste estudo, classificaram o componente como Atendido em Maior Parte.

Quadro 2 – Avaliação do componente preliminares: julgamento dos níveis

Ponto de Verificação	Julgamento (A-E)
I. Sumário Executivo	A
II. Prefácio	B
III. Metodologia	A

Fonte: Os autores (2024).

A proposta de trabalho do Instituto para o Desenvolvimento do Investimento Social para avaliar o Projeto de Promoção do Ecossistema de Investimento e Negócios de Impacto Socioambiental buscou, a partir dos objetivos e estratégias planejadas para a iniciativa, definir um conjunto de procedimentos metodológicos capaz de verificar os resultados e impactos gerados pelo Projeto, a saber: análise de documentação; levantamento de indicadores; reuniões participativas; entrevistas exploratórias; entrevistas; questionário; análise de dados.

O primeiro item referido no Quadro 3 e relativo ao componente Fundamentos, apresenta “a razão da existência do objeto, seja ele um projeto, um programa, um material, um serviço prestado dentre outros” (Elliot, 2011, p. 946) e, de acordo com a Lista Chave de Verificação de Scriven é denominado Antecedentes e Contexto. Ao ser meta-avaliado o referido item obteve a classificação A, visto que atinge todos os aspectos principais deste ponto e os expressa com clareza concisão.

O item Descrições e Definições, deve contemplar os componentes que relatam detalhadamente o contexto e o próprio cenário do programa em relação a sua criação e seu desenvolvimento. Também deve identificar seus principais *stakeholders* e respectivos interesses na avaliação. Por terem sido avaliados os aspectos considerados relevantes ao julgamento, o item foi classificado como A.

Seguindo a Lista Chave de Verificação da Avaliação de Scriven, o próximo item, denominado Consumidores, também obteve julgamento A, pois o público-alvo objetivado pela avaliação foi referido de forma clara, ao longo do relatório.

Os itens 4 e 5 dos Fundamentos que tratam dos Recursos e Valores compreendem, dentre os primeiros, os recursos financeiros, humanos, intelectuais e sociais alinhados ao programa avaliado. E, aos segundos, os “valores relativos a requisitos legais, éticos, profissionais, valores culturais, históricos, tradicionais, científicos, tecnológicos, políticos e econômicos” (Elliot, 2011, p. 946), que sejam condizentes com as demandas e os objetivos dos programas avaliados. Ao serem julgados pelos meta-avaliadores obtiveram classificação E, pois, os pontos requeridos não foram devidamente contemplados.

AVALIAÇÃO DA COMPONENTE FUNDAMENTOS

Conforme mostra o Quadro 3, que trata do componente Fundamentos do relatório, os meta-avaliadores, utilizando os critérios

de valoração da meta-avaliação considerados no estudo, classificam o componente como Atendido em Maior Parte.

Quadro 3 – Avaliação do componente fundamentos: julgamento dos níveis

Ponto de Verificação	Julgamento (A-E)
1. Antecedentes e Contexto	A
2. Descrições e Definições	A
3. Consumidores	A
4. Recursos	E
5. Valores	E

Fonte: Os autores (2024).

No componente SubAvaliações, o item Processo foi classificado como B pois, ainda que atinja a maioria dos aspectos relativos a este ponto, não cobre um ou dois deles, ou trata dos ingredientes principais, sem plena clareza. Neste aspecto, também não foram explicitados os custos e valores do projeto.

Quanto aos Resultados ou Impactos da avaliação, no relatório, que correspondem ao ponto de verificação do item 7 da Lista Chave de Verificação de Avaliação de Scriven (2007), a classificação do julgamento feito pelos meta-avaliadores foi A, pois, não só foram atingidos todos os aspectos destes pontos, como também os mesmos foram apresentados de maneira suficientemente clara e concisa.

Custos, que é o oitavo item da Lista de Verificação da Avaliação, foi, do ponto de vista dos meta-avaliadores, classificado como D, uma vez que o relatório meta-avaliado apresenta de forma deficitária os critérios relevantes para o seu julgamento. Na abordagem do item, o texto trata implicitamente do ponto em um ou dois de seus elementos, caracterizando o tratamento deficiente.

No Capítulo 5 são apresentados os impactos percebidos pelos beneficiários do programa, porém, os meta-avaliadores consideraram que em Custos, os critérios de avaliação desse item podem abarcar outros que sejam monetários, a exemplo dos que envolvem espaço, tempo, vidas, estresse, capital político, social ou pessoal, e custos ambientais a longo prazo, essa amplitude de entendimento reforçou a classificação D ao item.

No relatório não foram estabelecidas comparações entre a avaliação do programa e a de outros programas mais efetivos, o que se tornou empecilho para que os meta-avaliadores pudessem ter acesso, por exemplo, a possibilidades de aprimoramento do projeto e de sua própria avaliação. Portanto, ao intervir no julgamento dos meta-avaliadores, este impeditivo fez com que ao item fosse atribuída a classificação E, que representa a perda total da pontuação.

No que tange à Exportabilidade, último item do componente SubAvaliações, o relatório apresenta, à página 13, as necessidades e as possibilidades de expansão do projeto para outras regiões do Brasil, enquanto seu tema retrata realidades que integram o País. No entanto, ainda que apresente, em suas considerações finais, algumas recomendações para ampliação do projeto, os meta-avaliadores consideraram que o relatório não apresentou aspectos relevantes nesse sentido, repercutindo sobre o julgamento e a classificação C da Exportabilidade.

AVALIAÇÃO DA COMPONENTE SUBAVALIAÇÕES

Conforme exhibe o Quadro 4 que se refere ao componente SubAvaliações do relatório, os metas-avaliadores, utilizando os critérios de valoração da meta-avaliação, anteriormente apresentados a partir do item 6 deste estudo, classificaram o componente como Atendido em Menor Parte.

Quadro 4 – Avaliação do componente subavaliação: julgamento dos níveis

Ponto de Verificação	Julgamento (A-E)
1. Processo	B
2. Resultados	A
3. Custos	D
4. Comparações	E
5. Exportabilidade	C

Fonte: Os autores (2024).

Para o último componente meta avaliado Conclusões, os meta-avaliadores classificaram com a letra B o nível referente à Síntese, pois, ainda que o relatório tenha apresentado aspectos de forma sintética, demonstrando, por exemplo, os resultados referentes a cada instrumento avaliativo, por meio de quadro resumo e apontado pontos fortes e fracos da avaliação, as conclusões, as recomendações e os *stakeholders*, ao final de cada instrumento, não foram mencionadas particularidades dos recursos, custos e valores, o que justifica a sua classificação.

As Recomendações e Explicações, relativas ao componente Conclusões da Lista Chave de Verificação da Avaliação acham-se voltadas para o aprimoramento do programa e, igualmente, fundamentadas no relatório, nas etapas 5 de todos os capítulos e nas considerações finais, do capítulo 6, que trazem significativas contribuições para o projeto, a classificação atribuída ao julgamento deste item foi A.

No item treze da Lista de Verificação, que diz respeito à Responsabilidade e Justificativa dos resultados, os mesmos não foram identificados no relatório meta-avaliado. Além deles, também não constaram no relatório os seus destinatários. Por estas razões, os meta-avaliadores atribuíram a classificação E ao seu julgamento.

De acordo com a Lista Chave de Verificação de Avaliação de Scriven, o décimo quarto item - Relatório e Apoio, integrantes do componente conclusões, deve explicitar os seus formatos e destinatários, além de especificar o tipo de apoio dado à divulgação dos achados da avaliação. Ao apreciar o relatório, os meta-avaliadores consideram que as informações demandadas foram tratadas de forma implícita e/ou deficiente, ao longo de todo texto. Exemplo ilustrativo é dado na seguinte situação: em nenhum momento, o relatório explicita que o Instituto de Cidadania Empresarial foi o organismo responsável pelo projeto. Em decorrência dessas omissões coube ao item a classificação D.

Meta-avaliação é o último item do componente Conclusões, da Lista Chave de Verificação da Avaliação. Este item remeteu às avaliações conduzidas pelo Instituto para o Desenvolvimento do Investimento Social, de forma independente e em duas etapas distintas, com intervalo de aproximadamente um ano entre elas, tendo ocorrido a sua fase intermediária entre os meses de abril e outubro de 2019 e a final entre junho e setembro de 2020 (Instituto de Cidadania Empresarial, 2021). No julgamento dos meta-avaliadores a justificativa para a classificação máxima (A) decorreu do fato de o item ter zelado pela cobertura de todos os aspectos da avaliação, que incluem, entre outros, a resposta à questão avaliativa, a coerência entre a metodologia e os seus fundamentos, os instrumentos avaliativos utilizados e respectivos padrões de avaliação, os pontos fortes e fracos da avaliação.

AVALIAÇÃO DA COMPONENTE CONCLUSÕES

Conforme consta no Quadro 5, que mostra a classificação do julgamento do componente Conclusão do relatório meta-avaliado - Atendido em Maior Parte, à luz dos critérios de valoração adotados pela presente meta-avaliação.

Quadro 5 – Avaliação do componente conclusões: julgamento dos níveis

Ponto de Verificação	Julgamento (A-E)
1. Síntese	B
2. Recomendações e Explicações	A
3. Responsabilidade e Justificativa	E
4. Relatório de Apoio	D
5. Meta-avaliação	A

Fonte: Os autores (2024).

RECOMENDAÇÕES E CONCLUSÕES

Pautando-se nas análises e julgamentos realizados em torno do teor do Relatório Executivo Avaliação de Impacto: promoção do Ecossistema de Investimento e Negócios de Impacto Socioambiental no Brasil, o presente estudo meta-avaliativo apresenta recomendações que podem ser utilizadas com o intuito de propor melhorias a serem incorporadas nas próximas avaliações do projeto. Em conformidade com essa intenção recomenda-se que:

- a. O Prefácio do relatório seja revisado, incluindo claramente, não apenas quem ou qual organismo encomendou o relatório e as motivações para fazê-lo, além da menção ao público-alvo, privilegiando de um lado a transparência do projeto e, de outro, o entendimento acerca dos beneficiários da avaliação;
- b. A apresentação dos recursos, valores e custos da avaliação seja declarada e documentada, considerando todos os desdobramentos de cada um desses níveis para, assim, cobri-los em suas complexidades;

- c. Estabelecer comparações com os resultados de outros projetos, com o objetivo de refinar o desenho da prática avaliativa e, por conseguinte, a sua implementação e avaliação;
- d. Desenvolver a portabilidade do projeto e o detalhamento dos desafios e oportunidades nas suas adaptações para cada região do país na qual será implantado;
- e. Detalhar todos os recursos financeiros, humanos e materiais envolvidos e mencionar a quem ou a qual instituição cabe a responsabilidade sobre o projeto e sua avaliação;
- f. Especificar quais são os destinatários finais do relatório de avaliação de maneira a favorecer, não somente o *accountability*, mas também as decisões e iniciativas a serem tomadas pelos *stakeholders* e demais interessados.

Concluindo, o relatório analisado Atende em Maior Parte os seguintes componentes da Lista Chave de Verificação da Avaliação de Scriven: Preliminares, Fundamentos e Conclusões e Atende em Menor Parte apenas o componente SubAvaliações. Dentre os 18 pontos de verificação considerados, 8 corresponderam à classificação A, indicando que atingem todos os aspectos principais do ponto e os expressam de forma clara e concisa, a saber: em Preliminares, o sumário executivo e a metodologia; em Fundamentos, os antecedentes e contexto, as descrições e definições e os consumidores; em SubAvaliações, os resultados; em Conclusões, as recomendações e explicações e a meta-avaliação.

A classificação B corresponderam três pontos, a saber: em Preliminares, o prefácio; em SubAvaliações, o processo; em Conclusões, a síntese, indicando que atingem a maioria dos aspectos relativos a estes pontos, mas não cobrem dois ou mais deles, ou trata dos ingredientes principais, mas não é 100% clara.

Assim, a meta-avaliação realizada evidenciou que o desempenho global da avaliação, a partir da perspectiva de seus quatro componentes, obteve classificação Atende em Maior Parte em 3 deles: Preliminares, Fundamentos e Conclusões. No que tange ao componente SubAvaliações, a classificação Atende em Menor Parte revelou as maiores fragilidades.

MÉRITO DO ESTUDO

Mariana Sá Alcântara Gomes

O Projeto de Promoção do Ecossistema de Investimento e Negócios de Impacto Socioambiental no Brasil, desenvolvido pelo Instituto de Cidadania Empresarial, com convênio do Banco Interamericano de Desenvolvimento, objetiva estimular negócios de impacto, impulsionando ações inovadoras no sentido do enfrentamento de questões no âmbito socioambiental em todo o Brasil. Trata-se, portanto, de um projeto de políticas públicas em área de extrema relevância para o desenvolvimento do país.

A avaliação desse projeto representa um importante esforço de monitoramento da qualidade e da eficácia de políticas públicas voltadas para o desenvolvimento sustentável e que visam a busca de soluções inovadoras, adequadas às necessidades do contexto socioambiental do Brasil do século XXI, uma vez que é, por meio dela, que é possível verificar se os objetivos foram alcançados, bem como propor reformulações a partir dos resultados aferidos.

No entanto, para garantir a eficácia e a segurança na replicação de modelos de políticas, projetos e programas é fundamental avaliar também a qualidade das próprias avaliações realizadas sobre eles, assegurando, dessa forma, que os critérios, os métodos e as técnicas aplicados nessas avaliações sejam adequados, garantindo resultados precisos e confiáveis.

O artigo intitulado “Meta-avaliação da avaliação do Projeto de Promoção do Ecossistema de Investimento e Negócios de Impacto Socioambiental no Brasil”, apresenta o estudo meta-avaliativo que evidência com clareza, desde o título, o seu objeto do estudo, revelando o assunto do qual o texto trata.

Além disso, o artigo exibe uma descrição suficientemente detalhada do objeto da avaliação, fornecendo uma visão geral bastante completa, informando com clareza o objeto da meta-avaliação e destacando que a fonte principal do estudo meta-avaliativo consiste no “Relatório Executivo Avaliação de Impacto: Promoção do Ecossistema de Investimento e Negócios de Impacto Socioambiental no Brasil”, publicado pelo Instituto de Cidadania Empresarial, em convênio com o Banco Interamericano de Desenvolvimento (BID), sendo tal estudo avaliativo conduzido pelo Instituto para o Desenvolvimento do Investimento Social. O artigo ainda explicita que o estudo teve início com uma pesquisa de caráter qualitativo-exploratório, tendo como fontes livros, artigos e documentos selecionados, relacionados ao tema do estudo.

A justificativa também é apresentada de forma clara e concisa, apontando o caráter de contribuição, próprio da meta-avaliação, aos processos de avaliação, uma vez que, por meio dela, viabiliza-se a verificação dos procedimentos e das técnicas utilizadas e, eventualmente, pode-se sugerir possibilidades de aprimoramento.

O artigo deixa claro que a referida meta-avaliação possui caráter somativo e apresenta fundamentação teórica sólida, referenciando autores e estudos consistentes, corroborando a metodologia escolhida para a realização do estudo.

Quanto à metodologia adotada, o artigo se baseia na literatura, destacando que o modo de condução de uma meta-avaliação deve ser definido de acordo com o tipo de avaliação sobre o qual se pretende debruçar e aponta a terceira edição do Joint Committee

on Standards for Educational Evaluations que define uma série de padrões aplicáveis inclusive para meta-avaliações. Dentre as abordagens que se aplicam a meta-avaliações somativas, o estudo elege a Lista Chave de Verificação da Avaliação de Scriven, a partir do trabalho de Elliot (2001).

O estudo meta-avaliativo em questão que teve, portanto, sua metodologia pautada na Lista Chave de Verificação da Avaliação de Scriven com o propósito de responder a seguinte questão meta-avaliativa:

Até que ponto a avaliação do Projeto de Promoção do Ecossistema de Investimento e Negócios de Impacto Socioambiental no Brasil atendeu aos componentes da Lista Chave de Verificação de Scriven?

Os autores recorreram à literatura para fundamentar a classificação de cada um dos critérios de verificação da Lista Chave de Verificação da Avaliação de Scriven, produzindo, assim, um estudo pautado na conformidade e isenção inerentes ao caráter da atividade de meta-avaliação.

A apresentação dos resultados é detalhada e bem fundamentada. Os itens da Lista Chave de Verificação da Avaliação de Scriven que não alcançaram classificação satisfatória foram apontados e compuseram as recomendações do estudo.

O estudo meta-avaliativo em questão contribui, portanto, para o aperfeiçoamento de processos avaliativos e meta-avaliativos na medida em que apresenta uma metodologia clara e uma fundamentação sólida destaca-se também a originalidade e atualidade do estudo, elaborado sobre uma avaliação recente e relevante no âmbito das políticas públicas.

REFERÊNCIAS

BARDACH, E. **A practical guide for policy analysis**: the eightfold path to more effective problem solving. Washington: CQ Press College, 2006.

CHELIMSKY, E. Integrating evaluation units into the political environment of government: the role of evaluation policy. **New Directions for Evaluation**, [S. l.], v. 2009, n. 123, p. 51-66. DOI: <https://doi.org/10.1002/ev.305>. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/ev.305>. Acesso em: 20 fev. 2025.

ELLIOT, L. G.; Meta-avaliação: das abordagens às possibilidades de aplicação. **Ensaio: aval. pol. públ. educ.**, Rio de Janeiro, v. 19, n. 73, p. 941-964, 2011. Disponível em: <https://www.scielo.br/j/ensaio/a/djGLfPzM5PkV5JkL9BdcZ5k/?format=pdf&lang=pt>. Acesso em: 8 jan. 2025.

FITZPATRICK, J. L.; SANDERS, J. R.; WORTHEN, B. R. **Program evaluation**: alternative approaches and practical guidelines. Upper Saddle River. NJ: Pearson, 2011.

FURTADO, J. P.; LAPERRIÈRE, H. Parâmetros e paradigmas em meta-avaliação: uma revisão exploratória e reflexiva. **Ciênc. Saúde Coletiva**, Rio de Janeiro, v. 17, n. 3, p. 695-705, 2012. DOI: <https://doi.org/10.1590/S1413-81232012000300016>. Disponível em: <https://www.scielo.br/j/csc/a/jpWHQCWR6RtDzkZDcFrl6qv/#>. Acesso em: 5 out. 2024.

GHIANO, M. C.; WEHRLE MARTÍNEZ, A. M.; PERONI FISCARELLI, A.; MARTINIC VALENCIA, S.; CHECA GUTIÉRREZ, M. I.; AMARILES ERAZO, F.; GUZMÁN, A. L.; SOBERÓN ÁLVAREZ, L. **Diretrizes de avaliação para a América Latina e o Caribe**. 2. ed. [S. l.]: Relac, 2021. Disponível em: <https://www.deval.org/en/publications/diretrizes-de-avaliacao-para-a-america-latina-e-o-caribe-edicao-2021>. Acesso em: 5 set. 2024.

INSTITUTO DE CIDADANIA EMPRESARIAL. **Relatório executivo avaliação de impacto**: promoção do ecossistema de investimento e negócios de impacto socioambiental no Brasil. [S. l.]: IDIS, 2021. Disponível em: <https://www.idis.org.br/publicacoesidis/relatorio-executivo-de-avaliacao-de-impacto-promocao-do-ecossistema-de-investimento-e-negocios-de-impacto-socioambiental-no-brasil/>. Acesso em: 8 out. 2024.

SILVA, R. R.; JOPPERT, M. P.; GASPARINI, M. F. V. (org.). **Diretrizes para a prática de avaliação no Brasil**. Rio de Janeiro: Rede Brasileira de Monitoramento e Avaliação, 2020. Disponível em: https://rbma.site/wp-content/uploads/2021/02/Diretrizes-para-a-pratica-de-avaliacao-no-Brasil_20ago20.pdf. Acesso em: 8 jan. 2025.

PATTON, M. Q. **Developmental evaluation**: applying complexity concepts to enhance innovation and use. Nova York: Guilford Press, 2010.

SCHWANDT, T. A. The politics of verifying trustworthiness in evaluation auditing. **American Journal of Evaluation**, Thousand Oaks, v. 10, n. 4, 1989. DOI: <https://doi.org/10.1177/109821408901000405>. Disponível em: <https://journals.sagepub.com/doi/10.1177/109821408901000405>. Acesso em: 8 jan. 2025.

SCRIVEN, M. **Key evaluation checklist**. Kalamazoo, MI: Western Michigan University, 2007.

SCRIVEN, M. **Avaliação**: um guia de conceitos. Tradução Marília Sette Câmara. Rio de Janeiro: Paz e Terra, 2018.

SCRIVEN, M. An introduction to meta-evaluation. **Educational Product Report**, Nova York, v. 2, n. 5, p. 36-38, 1969.

SECCHI, L. **Políticas públicas**: conceitos, esquemas de análise, casos práticos. São Paulo: Cengage Learning, 2010.

STUFFLEBEAM, D. L. Meta-evaluation. **Journal of MultiDisciplinary Evaluation**, [S. /], v. 7, n. 15, p. 99-158, 2011. Disponível em: https://journals.sfu.ca/jmde/index.php/jmde_1/article/view/300/295. Acesso em: 9 jan. 2025.

YARBROUGH, D. B.; SHULA, L. M.; HOPSON, R. K.; CARUTHERS, F. A. **The program evaluation standards**: a guide for evaluators and evaluations users. 3. ed. Califórnia: Sage, 2011.

SOBRE OS AUTORES E AS AUTORAS

Albertina Thereza Pacheco Rodrigues

Mestranda em Avaliação, Faculdade Cesgranrio, com previsão de defesa do projeto de dissertação em abril de 2025. Especialização em Gestão Educacional Integrada, Instituto Superior de Educação Afonso Claudio, FAAC, 2010. Especialização em Administração Escolar, S B I, Universidade Cândido Mendes, 2006. Graduada em Letras, Pontifícia Universidade Católica do Rio de Janeiro, 2010. Graduada em Pedagogia, Universidade Federal do Rio de Janeiro, 2004, com habilitação em Magistério das Séries Iniciais do Ensino Fundamental e Magistério da Educação Infantil. Curso Normal Superior em Magistério das Séries Iniciais do Ensino Fundamental com ênfase em Educação de Jovens e Adultos, 2002. Atualmente é professora da Rede Municipal de Educação do Rio de Janeiro. Tem experiência em coordenação pedagógica, onde contribuiu para a orientação pedagógica e estratégias de aprendizado dos alunos da Instituição e avaliação da equipe. Tem experiência na área de Educação, com ênfase em Coordenação e Administração Escolar, Tecnologia Educacional, Séries Iniciais do Ensino Fundamental, Educação Infantil, Educação de Jovens e Adultos e Avaliação da aprendizagem.

E-mail: therezarodrigues@yahoo.com.br

Alcineidy Nascimento Falcão

Mestranda em Avaliação pela Faculdade Cesgranrio, com previsão de defesa do projeto dissertação em 2025. Pós-Graduada – Curso de Especialização em Psicopedagogia Clínica e Institucional, Universidade Estácio de Sá, 2022. Pós-Graduada em Gestão Empreendedora – Educação pela Universidade Federal Fluminense, 2016. Pós-graduada *Lato-Sensu* em Docência Superior, UNISUAM, 2006. Graduada em Pedagogia com habilitação em Administração Escolar, UNIABEU, 1999. Experiência como docente da Educação Básica ao Ensino Superior. Vivência como Gestora Escolar, Coordenadora Pedagógica, Orientadora Educacional e Supervisora Educacional. Mestranda em Avaliação pela Faculdade Cesgranrio, com previsão de defesa em 2025. Servidora pública há 30 anos. Atuou como Técnica na Coordenadoria de Gestão Pedagógica (COOGEP) - da Superintendência de Gestão Pedagógica (SUPGE), na Secretaria de Educação do Estado do Rio de Janeiro. Atualmente, é Assessora do Diretor Regional Administrativo da Metropolitana I.

E-mail: alcineidyfalcao@gmail.com

Alessandra Baldner Pontes

Mestranda em Avaliação, Faculdade Cesgranrio, com previsão de defesa do projeto de dissertação em agosto de 2025. Pós-Graduada em Relações Internacionais, Universidade Cândido Mendes, UCAM, 2007. Graduada em Comunicação Social, Centro Universitário da Cidade, 2005. Atualmente é Especialista em Previdência Social do Fundo Único de Previdência Social do Estado do Rio de Janeiro, e atua como Coordenadora de Desenvolvimento de Pessoas e Educação Corporativa. Tem experiência nas áreas de Administração, Recursos Humanos e Previdência Social.

E-mail: alebaldner@gmail.com

Andréa Cristina Tavares Moreira

Mestranda em Avaliação pela Faculdade Cesgranrio, com previsão de defesa da dissertação do projeto de dissertação em agosto de 2025. Especialização em Língua Espanhola: Práticas Docentes, Universidade Veiga de Almeida, UVA/RJ, 2010. Graduada em Letras – Espanhol, Centro Universitário da Cidade, UniverCidade, 2005. Professora pós-graduada em “Língua Espanhola - Práticas Docentes” pela Universidade Veiga de Almeida (2011). Apresentou como título de conclusão de curso “A Compreensão Textual de Propaganda Publicitária no Ensino de Língua Estrangeira” tendo a teoria da Análise do Discurso e da Leitura Interativa como marcos em sua construção monográfica e em sua prática docente. Atualmente é professora regente de espanhol na Prefeitura da Cidade do Rio de Janeiro e professora regente de espanhol na Secretaria de Educação do Estado do Rio de Janeiro. Tem experiência com o ensino do idioma espanhol (idioma adicional) em classes de Ensino Fundamental I, Ensino Fundamental II e Ensino Médio também em escolas de administração privada. Acredita na Educação como meio principal para o desenvolvimento do ser humano, bem como, para o desenvolvimento de uma sociedade capaz de agir e interagir em prol da evolução.

E-mail: andreatavaresm@gmail.com

Cristina Borges Alves

Mestranda em Avaliação pela Faculdade Cesgranrio com previsão de defesa em agosto de 2025. Graduada em Administração pela Universidade do Estado do Rio de Janeiro, 1997. Atualmente é Especialista em Previdência Social do Fundo Único de Previdência Social do Estado do Rio de Janeiro. Tem experiência na área de Administração.

E-mail: borgescriss@hotmail.com

Dalva Carneiro

Mestra em Avaliação pela Faculdade Cesgranrio, 2024. Especialização em Psicologia do Trabalho, Fundação Getúlio Vargas/Instituto de Seleção e Orientação Profissional, FGV/ISOP, 1980. Graduada em Psicologia pela Universidade Federal do Rio de Janeiro, 1977. Atualmente é Especialista em Previdência Social - Fundo Único de Previdência Social do Estado do Rio de Janeiro. Tem experiência na área de Psicologia, com ênfase em Psicologia do Trabalho e Gestão de Recursos Humanos.

E-mail: dalvacarneiro.dc@gmail.com

Elaine Maria de Andrade Senra

Doutora em História das Ciências e das Técnicas e Epistemologia (HCTE), Universidade Federal do Rio de Janeiro, UFRJ, 2009; Pós-doutorado no Programa de Pós-Graduação em História das Ciências e das Técnicas e Epistemologia (HCTE/UFRJ), 2013. Especialização em Ensino de Física pela Universidade Federal Fluminense - UFF, 2004. Licenciatura em Matemática pela FEAP, 2000. Atuou como Professora Colaboradora no Programa de Pós-Graduação em História das Ciências e das Técnicas e Epistemologia - HCTE, da Universidade Federal do Rio de Janeiro - UFRJ (2013-2016). Foi Professora Titular no Centro Universitário Serra dos Órgãos, UNIFESO, Teresópolis, (1995-2021), atuando na Educação Básica e nos cursos de graduação em Pedagogia, Farmácia, Licenciatura em Matemática, Engenharias de Produção, Ambiental e Civil. Foi coordenadora do Curso de Licenciatura em Matemática, UNIFESO, Teresópolis, RJ (2008-2013), Diretora do Centro de Ciências e Tecnologia CCT, UNIFESO, Teresópolis, RJ (2010-2018), Diretora de Pós-Graduação, Pesquisa e Extensão, UNIFESO, Teresópolis, RJ (2018-2021). Professora Adjunta na Faculdade Cesgranrio (2021). Atuou nos cursos de Pedagogia e Sistemas da Informação (2021 - 2022). Atua no Mestrado Profissional em Avaliação, Faculdade Cesgranrio, desde 2023, nas disciplinas Avaliação e Epistemologia e Experiências Multidisciplinares em Avaliação. É Membro do Conselho Editorial da Revista "Meta: Avaliação" conceito A4 Qualis/Capes. Atua principalmente nas Áreas de Educação, Epistemologia e Avaliação.

E-mail: elainemariapandrade@gmail.com

Elaine Rodrigues Perdigão

Doutora em Ciências Sociais pela PUC-Rio, Mestre em Antropologia Social e graduada em Ciências Sociais pela UFF, com ênfase em Teoria Social, Metodologia das Ciências Sociais e Pesquisa Qualitativa. Especialista em Projeto, Implementação e Gestão da Educação a Distância (PIGEAD/UFF). Professora colaboradora na área de Tecnologia Educacional da Diretoria de Extensão da Fundação Cecierj. Atualmente, é Professora Adjunta no Mestrado Profissional em Avaliação da Faculdade Cesgranrio, onde se dedica ao estudo e à avaliação de programas sociais, contribuindo para projetos de pesquisa na área e orientando dissertações de mestrado. Leciona as disciplinas Avaliação Qualitativa e Ética na Avaliação, com foco em metodologias e modelos avaliativos. Integra o Conselho Editorial da Revista Meta: Avaliação.

E-mail: elaineperdigao@cesgranrio.org.br

Flávia de Andrade Viveiros Quadrelli

Mestranda em Avaliação pela Faculdade Cesgranrio, com previsão de defesa do projeto de dissertação até agosto de 2025. Especialização em Gestão da Qualidade Total pela Universidade Católica de Petrópolis, 1996. Graduada em Administração pela Universidade Católica de Petrópolis, 1993. Possui ampla experiência em ambientes de gestão empresarial (Siemens/Chemtech) e universitária (UCP, Estácio e Pró-Saber), tendo atuado como gerente administrativa, coordenadora de compras, coordenadora acadêmica e coordenadora de extensão. Atualmente é professora das disciplinas de Tecnologias de Informação e Comunicação (TICs) no Curso Normal Superior do Pró-Saber e Gestora do Sistema Administrativo-acadêmico do Pró-Saber. O seu principal interesse de estudo é voltado para as TICs no contexto da formação docente e do cidadão, bem como as perspectivas de utilização dessas tecnologias de forma consciente e responsável.

E-mail: flaviavq@gmail.com

Glauco da Silva Aguiar

Doutor em Educação (2008) e Mestrado em Educação pela Pontifícia Universidade Católica do Rio de Janeiro (2001). Graduado em Licenciatura Plena em Matemática pela Universidade do Estado do Rio de Janeiro (1991). Professor Adjunto do Mestrado Profissional em Avaliação da Faculdade Cesgranrio(2014 - ...). Integrou o Quadro Complementar de Oficiais do Magistério do Exército Brasileiro e participa como membro da Sociedade Brasileira de Educação Matemática SBEM-RJ, GT Avaliação e Educação Matemática. Desenvolve trabalhos e pesquisas na área de Educação, Educação Matemática e Estatística, com ênfase em Análise e Interpretação de Dados, Avaliação de Sistemas e Programas Educacionais, Currículo de Matemática e Desigualdades Educacionais.

E-mail: glauco@cesgranrio.org.br

Hélio Huguenin Tavares Júnior

Mestrando em Avaliação pela Faculdade Cesgranrio, com previsão de defesa do projeto dissertação para agosto de 2025. Pós-Graduado em Ambiente Organizacional, Saúde e Ergonomia pela Escola Superior Aberta do Brasil, 2012. Graduado em Odontologia pela Faculdades São José, 2016. Graduado em Fisioterapia pelo Centro Universitário Augusto Motta, 2007. Técnico de Anatomia e Necropsia da Universidade Federal do Estado do Rio de Janeiro - UNIRIO desde junho de 2010.

E-mail: heliohuguenin@gmail.com

José Rafael Ayres da Motta

Mestrando em Avaliação pela Faculdade Cesgranrio, com previsão de defesa do projeto de dissertação em agosto de 2025. Aluno Especial em Desenvolvimento Regional do Mestrado da Universidade do Estado da Paraíba – UEPB, 2019; Especialização MBA em Gestão Estratégica em Recursos Humanos, Universidade Federal Rural do Rio de Janeiro - UFRJ, 2007; Bacharel em Administração pela Universidade Católica de Petrópolis – UCP, 1998. Servidor público federal a mais de vinte anos, atualmente Chefe do Serviço de Administração e Contrato da Comissão Nacional de Energia Nuclear – CNEN. Foi membro da equipe de projeto e implantação do planejamento estratégico institucional no Laboratório Nacional de Computação Científica – LNCC; participou, na Universidade Federal do Rio de Janeiro - UFRJ, como membro da equipe de projeto e proposta legislativa para institucionalizar os Serviços de Saúde Ocupacional; Foi designado representante do Ministério da Ciência Tecnologia e Inovação – MCTI em atividades de análise, avaliação, acompanhamento e fiscalização das transferências de recursos por meio de convênios, contratos de repasse e termos de parcerias celebrados com entidades públicas e privadas sem fins lucrativos. Possui experiência na área de Gestão Pública, principalmente nos subtemas: Logística na Administração Pública (convênios, licitações, contratos, planilhas de formação custo e preço, suprimentos, patrimônio); Desenvolvimento de Pessoas (legislação e treinamentos); Transparência e Sistemas Estruturantes da Administração Pública.

E-mail: rafael.motta@cnen.gov.br

Karla Verônica Mendes da Silva

Graduada em Direito pela Sociedade Unificada de Ensino Superior e Cultura Ltda., 1990; Especialização em Gestão Pública pela AVM Educacional Ltda., 2014; Especialização em Direito Administrativo pelo Centro Universitário União das Américas, 2022; Mestra em Avaliação pela Faculdade Cesgranrio, 2023. Atualmente é Assistente em Administração da Universidade Federal do Estado do Rio de Janeiro. Tem experiência na área de Administração.

E-mail: karlaeisen@hotmail.com

Ligia Gomes Elliot

Ph D em Educação/Avaliação (1980) e Mestrado em Estudos Latino Americanos (1979), Universidade da Califórnia, Los Angeles (UCLA). Mestrado em Educação, UFRJ (1976), com pesquisas de Pós-Doutorado na área de avaliação, apoio do CNPq (1981-1983; 1983-1985). Graduada em Ciências Sociais, UERJ (1969). Profª Visitante (1981-1984) e Profª Adjunta da Faculdade de Educação, UFRJ (1995-1984), Dept. de Metodologia da Avaliação e Pesquisa e de Fundamentos da Educação; Sub-Diretora Adjunta de Ensino para Graduados (1990-1993) e Coordenadora do Programa de Pós-Graduação em Educação (até 1995). Pesquisadora e avaliadora da Fundação Cesgranrio (1997-...). Docente e Coordenadora do Mestrado Profissional em Avaliação da Faculdade Cesgranrio (2007- ...). Temas: avaliação de sistemas, programas e instituições; metodologias e abordagens avaliativas; meta avaliação.

E-mail: ligiaelliot@yahoo.com.br

Lígia Silva Leite

Doutorado em Educação, Temple University (1990). Mestrado em Educação, Universidade Federal do Rio de Janeiro (1982). Pesquisa de Pós-Doutorado em Tecnologia Educacional, Universidade de Pittsburgh (1994). Graduada em Pedagogia, UFRJ (1974). Profª Adjunta do Programa de Mestrado Profissional em Avaliação da Faculdade Cesgranrio (2009 - ...). Foi orientadora do Doutorado em Educação, Nova Southeastern University, EUA. Foi vice-presidente da Associação Brasileira de Tecnologia Educacional, Profª Adjunta do Dep. de Estudos Aplicados ao Ensino, da Faculdade de Educação da UERJ e da Nova Southeastern University. Experiência na área de Educação, com ênfase em Educação a Distância e Tecnologia Educacional. Tem atuado na área da Avaliação como docente, pesquisadora e autora de artigos e livros. Temas: avaliação, ensino-aprendizagem, formação de professores, construção de conhecimento.

E-mail: ligialeite@terra.com.br

Lucí Mary Araújo Hildenbrand

Doutora em Ciências da Comunicação, Universidade de São Paulo (1995). Mestrado em Educação, Universidade Federal do Rio de Janeiro (1984). Especialização em Planejamento, Implementação e Gestão de Educação a distância, Universidade Federal Fluminense (2012), em Design Instrucional para EaD Virtual, Universidade Federal de Itajubá (2010). Licenciada em Ciências/Biologia, Universidade Federal Rural do Rio de Janeiro (1975) e em Pedagogia, Universidade Estácio de Sá (2007 e 2008). Professora Adjunta do Mestrado Profissional em Avaliação da Faculdade Cesgranrio (2010-...). Temas: avaliação; meta-avaliação; avaliação de programas e de tecnologias; metodologias e abordagens avaliativas. Experiência em ensino e avaliação na área de Avaliação de Tecnologias de Informação e Comunicação, Avaliação de Programas de Comunicação e Educação e em Adaptação de Instrumentos Avaliativos.

E-mail: lucihildenbrand@yahoo.com.br

Luciana Vieira Santos

Mestre em Avaliação pela Faculdade Cesgranrio, 2024; Especialização em Administração e Supervisão Escolar, Universidade Candido Mendes, 2010; Graduada em Pedagogia pela Universidade do Estado do Rio de Janeiro, 2003. Atualmente é Coordenadora da DIPEP Divisão Pedagógica da Educação Profissional, SENAI RJ. Tem experiência de 18 anos na área de Educação, com ênfase em Educação Profissional. Interesse em atuar com Educação Profissional e Corporativa com experiência em projetos estratégicos, avaliação, liderança técnica e assessoria de processos pedagógicos, legislação educacional, tendências educacionais e inovação.

E-mail: lvieira05@yahoo.com.br

Maria das Graças Madeira

Mestra em Avaliação pela Faculdade Cesgranrio, 2024; Pós-Graduada Lato Sensu em Gestão Pública pela Universidade Cândido Mendes, 2012; Graduada em Letras (Bacharelado Português/Espanhol) pela Universidade do Estado do Rio de Janeiro – UERJ, 1999. Tem experiência na área de Letras, com ênfase em Língua Portuguesa. Já ministrou aulas de Espanhol. Ingressou no Serviço Público Federal em 1994 e, desde então, trabalha na Universidade Federal do Estado do Rio de Janeiro (UNIRIO), exercendo cargos de gestão a partir do ano 2000.

E-mail: magramadeira@gmail.com

Mariana Sá Alcântara Gomes

Doutora em Psicossociologia de Comunidades e Ecologia Social pela Universidade Federal do Rio de Janeiro (EICOS -2007); Mestra em Educação em Ciências e Saúde pela Universidade Federal do Rio de Janeiro (NUTES - 2001), Bacharela em Comunicação Social, com habilitação em Radialismo, pela Universidade Federal do Rio de Janeiro (ECO - 1997), formação em Arteterapia (Atelier Eveline Carrano - AARJ - 2015) e Licenciada em Filosofia (UNIP - 2024). Atua como arteterapeuta e professora colaboradora no Curso de Formação em Arteterapia no Atelier de Artes e Terapias Eveline Carrano e como professora adjunta do Curso de Licenciatura em Pedagogia da Faculdade Cesgranrio, membro do NDE do Curso de Pedagogia, membro do NEXT (Núcleo de Extensão e Treinamento) e coordenadora dos projetos de extensão Fala Escola! Intercâmbio de Saberes com a Educação Básica e Pedagogia nas Redes e professora no Mestrado Profissional em Avaliação da Faculdade Cesgranrio.

E-mail: mariana.cesgranrio@gmail.com

Michelle Terra Duarte

Doutoranda em Educação pela UNESA. Mestra em Avaliação pela Faculdade Cesgranrio, 2023. Pós-graduada em Metodologia do Ensino de História e Geografia, Faculdade Educamais, 2021. Pós-graduada em Educação Especial Inclusiva com Ênfase em Tecnologia Assistiva e Comunicação Alternativa, Faculdade Educamais, 2021. Pedagoga pela Universidade do Estado do Rio de Janeiro, 2014. Comprometida com a transformação da educação pública, é professora alfabetizadora voluntária no programa Tempo de Aprender, do Governo Federal, atuando em escolas no município do Rio de Janeiro. Essa experiência tem permitido a aplicação de abordagens personalizadas e o uso de tecnologias para melhorar a qualidade do ensino da alfabetização.

E-mail: terrapedagoga@gmail.com

Ovidio Orlando Filho

Doutor em Ciências da Educação pela Universidade do Minho, em Portugal (2014). Mestre em Administração e Desenvolvimento Empresarial pela Universidade Estácio de Sá (2005). Especialização em Análise de Sistemas e Processamento de Dados (1992) e em Organização e Métodos (1993) pelo Centro de Instrução Almirante Newton Braga. Aperfeiçoamento em Eletrônica pelo Centro de Instrução Almirante Wandenkolk (1993). Graduado em Engenharia Op. Eletrônica pelo Centro Federal de Educação Tecnológica Celso Suckow da Fonseca (1976). Editor associado da revista Meta: Avaliação. Professor Adjunto do Mestrado Profissional em Avaliação da Faculdade Cesgranrio (2015), Docente das disciplinas Políticas Públicas em Avaliação, Meta-Avaliação e Avaliação do Desempenho Profissional. Desenvolve trabalhos e pesquisas nas áreas de Políticas Públicas em Avaliação e Estudos Meta-avaliativos.

E-mail: ovidio@cesgranrio.org.br

Paula Magalhães da Silva Santos

Mestranda em Avaliação pela Faculdade Cesgranrio, com previsão de defesa do projeto dissertação em agosto de 2025. Graduada em Serviço Social pela Universidade Federal do Estado do Rio de Janeiro, 2023. Graduada em Estatística pela Universidade do Estado do Rio de Janeiro, 2011. Atualmente é Especialista em Previdência Social do Fundo Único de Previdência Social do Estado do Rio de Janeiro.

E-mail: paulamssf@hotmail.com

Patrícia Sousa de Souza

Mestranda em Avaliação pela Faculdade Cesgranrio, com previsão de defesa projeto de dissertação em outubro de 2025; Especialista em Empreendedorismo para o Ensino Médio pela Universidade Federal Fluminense, 2023; Especialização em Gestão do Trabalho Pedagógico (Supervisão, Gestão, Inspeção, Coordenação) pela Faculdade Venda Nova do Imigrante, 2019; Graduação em andamento em Tecnólogo em Ciência da Felicidade, Centro de Ensino Superior de Maringá.; Graduada em Letras - Análise Semiótica, Universidade Gama Filho, 2004. Atua na SEEDUCRJ como professora formadora e docente do ensino fundamental 2 e médio. Tem experiência em coordenação de cursos de extensão, no ensino de Língua Portuguesa e Redação (presencial e EAD); coordenação de equipes, articulador pedagógico propondo projetos interdisciplinares, bem como na gestão de processos administrativos.

E-mail: profpatriciasouza@gmail.com

Sheila Regina da Rocha de Magalhães

Mestra em Avaliação pela Faculdade Cesgranrio, 2024. Graduada em Letras - Inglês e respectivas literaturas pela Universidade do Estado do Rio de Janeiro (UERJ), 1983. Servidora pública há 42 anos, com atuação nas áreas de gestão e pedagógica de projetos, cursos e eventos na UERJ e na UEZO (Centro Universitário Estadual da Zona Oeste), no Rio de Janeiro. Na UERJ: Laboratório de Idiomas do Instituto de Letras - professora e coordenadora de cursos livres de idiomas (1982 a 1989); Centro de Produção da UERJ - CEPUERJ - Coordenadora de cursos e eventos e Diretora do departamento de treinamento profissional (1990 a 1999); Sub-reitoria de Extensão e Cultura - chefe de seção e coordenadora do evento "UERJ sem Muros" (2000 a 2003); Laboratório de Ciências Radiológicas do Instituto de Biologia Roberto Alcântara Gomes (IBRAG) - coordenadora da área de cursos, eventos e gestora de projetos especiais (2004 a 2007); Policlínica Piquet Carneiro - diretora de administração e planejamento (2008 a jun/2010). Na UEZO: Reitoria - chefe de gabinete (jul/2010 a fev/2013) e coordenadora de relações institucionais (mar/2013 a 2014). No Departamento de Ciências Radiológicas (DCR) do IBRAG/ UERJ - coordenadora de administração acadêmica do Programa Nacional de Formação em Radioterapia, contemplando o curso de Mestrado Profissional em Física Médica e o Curso de Qualificação Profissional para Técnicos em Radioterapia (2015 a 2018), assessora do chefe do Departamento (2019) e analista da qualidade no Laboratório de Metrologia (2020 a 2024). A partir de janeiro de 2025, passou a exercer a função de Gerente da Qualidade do Laboratório de Metrologia do DCR.

E-mail: sheila.regina.magalhaes@gmail.com

Virginia Virgínio de Lima

Mestranda em Avaliação pela Faculdade Cesgranrio, com previsão de defesa do projeto de dissertação em agosto de 2025. Graduada em Administração pela Universidade Estácio de Sá, 2021. Pós-graduada em Gestão empresarial e em Nutrição e suplementação nutricional e esportiva, 2020. Graduada em Nutrição pela Universidade Federal do Estado do Rio de Janeiro, 2014. Desde 2020 trabalha como Assistente Administrativo no Instituto Nacional de Educação de Surdos, INES, Brasil.

E-mail: virginialimanutri@gmail.com

ÍNDICE REMISSIVO

A	
Acordos Formais	51, 61, 65, 225, 228, 287, 295, 297
Administração Pública	25, 246, 335
A União faz a Vida	8, 78
B	
Bolsa Família	8, 42, 44, 45, 46, 53, 57, 61, 64, 66, 67, 68, 69, 71, 73, 76, 77
C	
Clareza e Equidade	51, 62, 65, 226, 287, 296, 297
Conflitos de Interesses	51, 63, 65, 226, 288, 296, 297
E	
Ensino Superior	15, 21, 25, 31, 38, 165, 332, 336, 338
Exequibilidade	50, 53, 59, 60, 72, 74, 211, 225, 285, 286, 288, 293, 298, 302
G	
Gerenciamento	50, 51, 58, 59, 60, 67, 70, 220, 221, 233, 236, 286, 293
I	
impacto social	79, 83, 85, 92, 96, 106, 115, 203, 204, 205, 206, 207, 213, 214, 216, 217, 229, 245, 250, 255, 257, 258, 263, 264, 273, 274, 279, 290, 299, 302, 303, 304, 305
M	
Melhoria da Educação	9, 121, 122, 123, 124, 125, 126, 132, 136, 137, 139, 143, 145, 146, 147, 149, 153, 154, 155, 156
meta-avaliação	9, 11, 12, 16, 17, 18, 19, 20, 21, 23, 31, 32, 33, 34, 35, 37, 39, 40, 46, 48, 58, 59, 65, 70, 72, 73, 74, 75, 80, 81, 84, 87, 88, 89, 111, 113, 114, 115, 118, 120, 121, 126, 127, 132, 137, 140, 141, 144, 145, 146, 147, 152, 153, 163, 164, 167, 168, 169, 170, 171, 173, 185, 186, 188, 191, 192, 201, 202, 208, 229, 239, 240, 241, 243, 251, 252, 253, 261, 262, 268, 269, 270, 271, 272, 273, 280, 281, 282, 283, 284, 289, 297, 298, 302, 303, 304, 305, 309, 310, 311, 312, 319, 321, 322, 324, 326, 327, 328, 329, 330, 336
O	
Orientação Responsiva e Inclusiva	51, 60, 65, 225, 287, 294, 297
P	
Política pública	244
processo avaliativo	16, 18, 19, 20, 26, 36, 43, 49, 57, 59, 74, 80, 90, 95, 102, 111, 114, 119, 121, 126, 141, 149, 156, 157, 163, 164, 167, 170, 171, 172, 182, 185, 186, 190, 202, 207, 217, 228, 242, 243, 252, 256, 259, 268, 269, 270, 283, 310
PROFMAT	9, 160, 165, 166, 167, 168, 169, 170, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 184, 188, 189, 190, 191, 192, 193
R	
Responsabilidade Fiscal	51, 63, 64, 65, 226, 288, 297
S	
sistema de avaliação	255, 257, 258, 263
T	
Teoria de Mudança	111, 205, 214, 215, 216, 221, 222, 235, 255, 256, 258
Transparência e Abertura	51, 62, 65, 226, 287, 296, 297
V	
Viabilidade do Contexto	50, 59, 60, 220, 287, 294

WWW.PIMENTACULTURAL.com

ESTUDOS META-AVALIATIVOS

diretrizes,
categorias,
padrões
e critérios