

**Avaliação de Políticas Públicas nos Tribunais de Contas brasileiros:
concepções e práticas na perspectiva dos auditores de controle externo.**

Luciana Rocha de Melo Alvim

Brasília, DF

2026

**Avaliação de Políticas Públicas nos Tribunais de Contas brasileiros:
concepções e práticas na perspectiva dos auditores de controle externo.**

Luciana Rocha de Melo Alvim

Dissertação apresentada ao Curso de Mestrado Profissional em Administração Pública da Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas, como requisito parcial para obtenção do título de Mestre em Administração Pública.

Orientador: Adalmir de Oliveira Gomes

Brasília, DF

2026

Luciana Rocha de Melo Alvim

Dissertação apresentada ao Curso de Mestrado Profissional em Administração Pública da Faculdade de Economia, Administração, Contabilidade e Gestão de Políticas Públicas, como requisito parcial para obtenção do título de Mestre em Administração Pública.

Data da defesa: 23/02/2026

Comissão Examinadora:

Prof. Dr. Adalmir de Oliveira Gomes - Orientador
PGAP/UnB

Prof. Dr. João Mendes Rocha Neto - Examinador interno
PGAP/UnB

Prof. Dr. Bruno Batista de Carvalho Luz - Examinador externo
UEG – TCE/GO

Prof. Dr. Tomás de Aquino Guimarães – Examinador Suplente
PGAP/UnB

AGRADECIMENTOS

A **Deus**, autor e sustentação de tudo, em cujas mãos coloco a minha vida e em quem confio plenamente. Agradeço pela graça que me conduziu em cada etapa deste percurso, pelo amparo nos momentos de cansaço e pela fortaleza que me permitiu perseverar até a conclusão deste trabalho.

À **minha família**, especialmente ao meu marido, **Rômulo Alvim**, companheiro de vida e de estudos, inclusive no Mestrado, pelo incentivo constante, pela presença firme e pelo apoio cotidiano; às minhas filhas, **Júlia e Luísa**, pela compreensão diante das horas de convivência a que precisei renunciar ao longo desta jornada e por serem a inspiração diária para meu aperfeiçoamento pessoal; e à minha mãe, **Maria José**, por sempre apoiar meus estudos e meu desenvolvimento profissional e, nesta etapa, oferecer suporte essencial com minhas filhas, permitindo que eu seguisse adiante com serenidade e foco

Ao Professor **Adalmir de Oliveira Gomes**, meu orientador, pela parceria firme e generosa, sem a qual esta dissertação não teria sido possível. Agradeço pelas incontáveis horas de discussão, revisão e apoio (intelectual e humano) que qualificaram as perguntas, aperfeiçoaram os métodos e refinaram os resultados. Seu rigor acadêmico, disponibilidade e confiança foram decisivos em todas as etapas da pesquisa. Registro, com especial gratidão, sua contribuição determinante para a realização deste trabalho.

Aos meus chefes, **Conselheiro Renato Rainha e Dr. José Augusto**, pelo apoio, incentivo e compreensão ao longo de toda a trajetória acadêmica. Minha gratidão pelo exemplo de compromisso com o interesse público e pela confiança depositada, que inspiraram o propósito desta pesquisa e viabilizaram sua conclusão com responsabilidade e dedicação.

Ao Secretário-Geral de Controle Externo do TCDF, **Antônio Alexandre do Nascimento Filho**, pela colaboração decisiva na viabilização dos trabalhos de campo desta pesquisa. Agradeço, em especial, por sua atuação diligente na articulação institucional que possibilitou o contato com os Tribunais de Contas de todo o país e a designação de grande parte dos auditores participantes dos grupos focais. Seu apoio permanente, marcado por disponibilidade, resiliência e incentivo contínuo, foi fundamental para que a etapa de coleta de dados transcorresse com êxito e no prazo estabelecido.

A todos os colegas **Auditores de Controle Externo dos Tribunais de Contas** que gentilmente dedicaram seu tempo e conhecimento para participar dos grupos focais, bem como àqueles que

contribuíram, de diferentes formas, para a realização dos trabalhos de campo.

RESUMO

A realização de avaliações de políticas públicas tem reposicionado os Tribunais de Contas como atores estratégicos para o aprimoramento da ação governamental, do aprendizado organizacional e da entrega de valor público. No Brasil, esse movimento ocorre em meio a ambiguidades conceituais e metodológicas, sobretudo quanto à distinção entre auditorias operacionais e avaliações de políticas públicas em sentido estrito, bem como aos limites institucionais do Controle Externo na emissão de juízos de valor sobre políticas governamentais. Esta dissertação analisa como auditores de controle externo percebem a avaliação de políticas públicas no âmbito dos Tribunais de Contas brasileiros, identificando metodologias adotadas, critérios para a valoração e estratégias para fomentar o uso das avaliações, com base nos quais foi possível identificar algumas trajetórias emergentes de inovação. A pesquisa adota abordagem qualitativa e baseia-se em dois eixos empíricos complementares: (i) realização de grupos focais de dois a seis auditores de controle externo, totalizando 126 participantes, distribuídos nos 33 Tribunais de Contas do país; e (ii) análise de conteúdo a uma amostra de relatórios de fiscalização indicados pelos participantes dos grupos focais ou identificados nos portais institucionais das Cortes. O desenho da pesquisa ancora-se em referenciais contemporâneos da teoria da avaliação, como a ‘Árvore da Teoria da Avaliação’, revisitada por Lemire et al. (2020) e colaboradores, e o ‘Jardim das Abordagens de Avaliação’ (Montrosse-Moorhead, 2024). Esses referenciais fundamentaram a estrutura analítica do roteiro de discussão do grupo focal, organizado em três ramos centrais – métodos, valores e uso – e orientaram a categorização e a interpretação dos achados empíricos. Os resultados indicam que, embora a maioria dos Tribunais de Contas realize fiscalizações com emissão de juízo de valor sobre políticas públicas, predomina a compreensão da avaliação como extensão da auditoria operacional, com foco nos critérios clássicos de economicidade, eficiência e eficácia. Em relação aos métodos, percebe-se com base nos depoimentos e documentos analisados, que a adoção sistemática de metodologias alinhadas às diretrizes da NBASP 9020 ainda é incipiente, embora diversas experiências inovadoras tenham sido observadas, como o uso de modelos de avaliação baseados na Teoria da Mudança, no Modelo Lógico de Programa, na Síntese de Evidências e em avaliações de impacto. Quanto aos valores, segundo a percepção dos participantes dos grupos focais, a cultura institucional ainda é predominantemente de conformidade. Por último, em relação ao uso, foram relatadas estratégias diferenciadas de envolvimento de *stakeholders* para induzir o efetivo uso de

resultados das avaliações. Ao final, propõe-se uma Tipologia dos Tribunais de Contas de acordo com metodologias, critérios e estratégias para incentivar o uso das avaliações, oferecendo vantagens teóricas, institucionais e práticas. Do ponto de vista teórico, permite construir um diagnóstico dos Tribunais na avaliação de políticas públicas. Sob a perspectiva institucional, fortalece a identidade institucional das Cortes de Contas como avaliadores, evidencia assimetrias e orienta estratégias de aprimoramento. Na dimensão prática, incentiva o aprendizado entre instituições coirmãs, amplia a qualidade das avaliações e facilita a comunicação com os *stakeholders*, favorecendo o uso efetivo das evidências produzidas.

Palavras-chave: Avaliação de Políticas Públicas; Tribunais de Contas; Controle Externo; Auditoria Operacional; Métodos, Valoração e Uso da Avaliação.

ABSTRACT

Public policy evaluation has repositioned Brazil's Courts of Accounts (*Tribunais de Contas*) as strategic actors in strengthening governmental action, organizational learning, and the delivery of public value. In Brazil, this movement has unfolded amid conceptual and methodological ambiguities, particularly regarding the distinction between performance audits and *stricto sensu* public policy evaluation, as well as the institutional boundaries of external oversight bodies when issuing value judgments on government policies. This dissertation examines how external control auditors understand public policy evaluation within Brazilian Courts of Accounts, mapping the methodologies adopted, the criteria used for valuation, and the strategies employed to foster the use of evaluation results, from which emergent trajectories of innovation are identified. The study adopts a qualitative approach and is based on two complementary empirical strands: (i) focus groups with 2 to 6 external control auditors per session, totaling 126 participants across all 33 Courts of Accounts in the country; and (ii) content analysis of a sample of oversight reports either indicated by focus group participants or retrieved from the Courts' institutional portals. The research design is grounded in contemporary evaluation theory frameworks, notably the Evaluation Theory Tree revisited by Lemire et al. (2020) and colleagues, and the Garden of Evaluation Approaches (Montrosse-Moorhead, 2024). These frameworks informed the analytical structure of the focus group protocol (organized around three core branches: methods, values, and use) and guided the categorization and interpretation of empirical findings.

Results indicate that, although most Courts of Accounts conduct oversight activities that involve evaluative judgments about public policies, evaluation is predominantly understood as an extension of performance auditing, centered on the classical criteria of economy, efficiency, and effectiveness. With respect to methods, evidence from both testimonies and documents suggests that systematic adoption of methodologies aligned with NBASP 9020 guidelines remains incipient, even though several innovative experiences were identified, including the use of T Theory of Change-based models, program logic models, evidence synthesis, and impact evaluations. Regarding values, focus group participants reported that the prevailing institutional culture remains largely compliance oriented. Finally, in terms of use, participants described differentiated stakeholder engagement strategies aimed at inducing the effective uptake of evaluation results. The dissertation concludes by proposing a typology of Courts of Accounts based on their methodologies, valuation criteria, and strategies to encourage the use of evaluations, offering theoretical, institutional, and practical benefits. Theoretically, the typology enables a diagnostic assessment of evaluative practices within the Courts. Institutionally, it strengthens the Courts' identity as evaluators, highlights asymmetries, and informs improvement strategies. Practically, it promotes peer learning among counterpart institutions, enhances evaluation quality, and improves communication with stakeholders, thereby fostering the effective use of the evidence produced.

Keywords: Public Policy Evaluation; Courts of Accounts; External Control; Performance Audit; Methods; Valuation; Use.

SUMÁRIO

ABSTRACT	7
SUMÁRIO	9
1. INTRODUÇÃO	11
1.1. Objetivos Geral e Específicos	13
2. QUADRO TEÓRICO-CONCEITUAL	15
2.1 Fundamentos e institucionalização da avaliação de políticas públicas.....	15
2.1.1. Conceitos e delimitações.....	15
2.1.2 Evolução internacional e institucionalização no Brasil	16
2.2 Dimensões analíticas da avaliação.....	19
2.2.1 Classificações das avaliações.....	19
2.2.2 Diversos caminhos possíveis em avaliação de políticas públicas.....	20
2.2.3 Marcos teóricos	21
2.2.3.1 A Árvore da Teoria da Avaliação.....	21
2.2.3.2 O Jardim das Abordagens de Avaliação	25
2.3 Avaliação no âmbito dos Tribunais de Contas.....	27
2.3.1 Legitimidade e inovação no Controle Externo	27
2.3.2 Auditorias Operacionais	29
2.3.3 Sistematização de categorias de acordo com o quadro teórico-conceitual da pesquisa	30
3. MÉTODOS E TÉCNICAS	33
4. RESULTADOS E DISCUSSÃO.....	40
4.1 Concepções e práticas avaliativas dos auditores: caracterização do panorama nacional	40
4.2 Métodos, critérios e usos das avaliações: experiências avançadas e trajetórias emergentes nos TCs	41
4.2.1 Métodos: instrumentos predominantes, adesão à NBASP 9020 e inovação metodológica	41
4.2.2 Valoração: critérios usados, análise de contexto e valores emergentes	45
4.2.3 Uso: estratégias de indução, consensualismo e gargalos do monitoramento	47
4.3 Padrões e variações institucionais: proposição de tipologia das trajetórias avaliativas..	50
4.3.1 Tribunais incipientes	51
4.3.2 Tribunais incrementais	52
4.3.3 Tribunais normativos-estruturantes.....	55
4.3.4 Tribunais disruptivos	58
4.4 Implicações metodológicas para o Controle Externo: discussão integradora dos achados	63
5. CONCLUSÕES E RECOMENDAÇÕES.....	66
6. PRODUTO TÉCNICO-TECNOLÓGICO (PTT): Guia de Arquiteturas e Trilhas para Avaliação de Políticas Públicas (APP) em Tribunais de Contas. Modelo modular, não hierárquico,	

com autodiagnóstico, entregáveis mínimos e indicadores por trilha.....	70
REFERÊNCIAS	84
Apêndice A.....	89
Apêndice B.....	91
Anexo A.....	92

1. INTRODUÇÃO

A Constituição Federal de 1988 (Brasil, 1988) instituiu os Tribunais de Contas como estruturas fundamentais de Controle Externo no Brasil, dotando-os de autonomia e competências expressas para a fiscalização contábil, financeira, orçamentária, operacional e patrimonial da União e das entidades da administração direta e indireta. Historicamente focada na legalidade e na conformidade, a literatura e as práticas recentes sugerem que a atuação dessas Cortes vem passando por uma transformação paradigmática nas últimas décadas, impulsionada pela necessidade de verificar não apenas a regularidade dos gastos, mas a efetividade das entregas à sociedade (Brasil, 1988; Leoni, 2023; Luz et al., 2019, 2021; Medeiros, 2023).

No contexto examinado, a avaliação tem sido tratada como instrumento para o aprimoramento da ação governamental e da tomada de decisão baseada em evidências. Organismos internacionais, como a Organização para a Cooperação e Desenvolvimento Econômico (OCDE), reconhecem a relevância das Instituições Superiores de Controle (ISC) na condução de avaliações que buscam alcançar impactos positivos no debate público e na boa governança. Contudo, a transição do controle estritamente burocrático para uma atuação focada no desempenho e na avaliação de políticas públicas *stricto sensu* impõe desafios conceituais, metodológicos e de legitimidade que ainda carecem de plena sistematização no cenário nacional (OCDE, 2022).

Para fins desta pesquisa, e alinhando-se às Normas Brasileiras de Auditoria do Setor Público (NBASP 9020), a Avaliação de Políticas Públicas (APP) é compreendida como um exame sistemático, baseado em evidências, que se concentra no exame da relevância e da utilidade da política, enquanto a Auditoria Operacional (ANOP) avalia principalmente a economicidade, a eficácia e a eficiência. A avaliação de política pública também examina essas três dimensões, todavia, visa ‘resultados amplos’ e ‘impactos globais e socioeconômicos’ (NBASP 9020, 2021).

Atualmente, observa-se que a ANOP permanece como a principal alternativa metodológica utilizada pelos Tribunais de Contas para examinar políticas públicas. A pesquisa buscou avaliar a suficiência da ANOP para capturar a complexidade da avaliação de políticas públicas, e discorrer sobre as alternativas inovadoras que têm surgido, como a incorporação da Teoria da Mudança, o uso da ferramenta Modelo Lógico de Programa para análise de desenho de políticas, a realização de levantamentos econométricos para avaliação de impacto, e o uso

da Síntese de Evidências para identificar estratégias baseadas em evidências para o atendimento a uma demanda social específica.

A expansão do escopo de atuação dos Tribunais de Contas para a avaliação de políticas públicas suscita debates acerca de sua legitimidade. Por um lado, há o questionamento sobre o ‘déficit democrático’ de um órgão técnico avaliar decisões de gestores eleitos. Parte da doutrina sustenta que a legitimidade para avaliar políticas públicas decorre da competência constitucional e da capacidade dos Tribunais de induzirem o aperfeiçoamento da gestão, entregando valor público e assegurando a efetividade de direitos fundamentais. A avaliação, nesse sentido, não visa substituir a discricionariedade do gestor, mas qualificar o processo decisório e fomentar o controle social (Gomes, 2020; Leoni, 2023; Macedo, 2023).

Como Entidades Fiscalizadoras Superiores (EFS), incumbe também aos Tribunais de Contas a realização de avaliações de políticas públicas, dada sua capacidade de desenvolver ou ter acesso ao conhecimento metodológico para a avaliação, bem como o conhecimento adquirido de políticas públicas a partir de outros trabalhos realizados, além de disporem da independência funcional necessária: objetividade e independência do governo e de interesses privados (NBASP 9020, 2021). Todavia, assim como os governos, também os Tribunais de Contas ainda enfrentam desafios na sistematização de práticas de avaliação de políticas públicas a fim de assegurar um padrão mínimo de qualidade e, por conseguinte, a realização de avaliações que resultem em deliberações que efetivamente contribuam para a melhoria das políticas públicas fiscalizadas (Oliveira & Mendonça, 2020).

De forma geral, as avaliações de políticas públicas nos Tribunais de Contas são realizadas por meio de ANOPs, as quais têm apresentado resultados positivos significativos, especialmente nas áreas sociais mais relevantes para a sociedade, como saúde, educação, segurança públicas e infraestrutura (Carvalho et al., 2022; Sachsida, 2018). Contudo, há discussões acerca da adequabilidade do instrumento para a avaliação, especialmente devido às dimensões não alcançadas pela ANOP: relevância e utilidade (NBASP 9020, 2021; Nunes et al., 2022).

No plano normativo é possível identificar a diferenciação a partir da comparação entre a NBASP 300, voltada à auditoria operacional, e a NBASP 9020, específica para avaliação de políticas públicas. A diversidade de práticas no âmbito dos Tribunais de Contas exige a sistematização de orientações metodológicas para assegurar qualidade e robustez às fiscalizações (NBASP 300, 2013; NBASP 9020, 2021; Nelson, 2020).

Nessa linha, não foram identificados estudos que sistematizem, em âmbito nacional, a percepção de auditores de controle externo dos 33 Tribunais de Contas do Brasil quanto (i) à suficiência e adequação das metodologias utilizadas, (ii) aos critérios empregados para a emissão de juízos de valor e (iii) às estratégias adotadas para fomentar o uso das avaliações por gestores públicos.

Diante desse cenário, o problema central que orienta este trabalho é: **Como os Tribunais de Contas brasileiros concebem e operacionalizam a avaliação de políticas públicas, na perspectiva dos auditores de controle externo?**

1.1. Objetivos Geral e Específicos

O objetivo geral da pesquisa é **analisar como os Tribunais de Contas brasileiros concebem e operacionalizam a avaliação de políticas públicas, na perspectiva dos auditores de controle externo, identificando padrões e implicações para o Controle Externo.**

Os objetivos específicos são os seguintes:

- a) Caracterizar o panorama nacional das percepções e práticas avaliativas dos auditores de controle externo nos Tribunais de Contas brasileiros;
- b) Analisar experiências consolidadas e trajetórias emergentes, com foco nos métodos, nos critérios de valoração e nos usos das avaliações;
- c) Construir uma tipologia das trajetórias avaliativas a partir dos padrões e variações institucionais identificados;
- d) Discutir as implicações para o aprimoramento da avaliação de políticas públicas no âmbito do Controle Externo; e
- e) Propor, como Produto Técnico-Tecnológico, um Guia de Maturidade em Avaliação de Políticas Pública para os Tribunais de Contas.

A relevância deste estudo reside na sua contribuição para a formulação e o aprimoramento de estratégias voltadas ao fortalecimento da qualidade das avaliações realizadas no âmbito do Controle Externo. Ao analisar como os Tribunais de Contas brasileiros concebem e operacionalizam a avaliação de políticas públicas, sob a perspectiva dos auditores de controle externo, a pesquisa busca identificar padrões, práticas e implicações dessas avaliações para o exercício do Controle Externo.

Nesse sentido, ao sistematizar as práticas e percepções de auditores de todo o país, o estudo tem por objetivo oferecer subsídios para que os Tribunais de Contas atuem de forma mais assertiva no atendimento das demandas sociais prioritárias, fomentando o aprimoramento das políticas públicas ou, quando necessário, sua descontinuidade, com base em evidências robustas.

Ao contribuir para o aprimoramento das avaliações de políticas públicas conduzidas pelos Tribunais de Contas, a pesquisa fortalece um relevante mecanismo institucional de aprendizado organizacional no setor público e na própria Administração Pública. Isso ocorre na medida em que tais avaliações detêm potencial para influenciar diretamente o processo decisório, o planejamento governamental, a alocação de recursos públicos e a gestão orientada a resultados.

A sistematização das avaliações de políticas públicas vai ao encontro das transformações no papel dos Tribunais de Contas, diante da superação de um controle predominantemente focado na legalidade, de natureza reativa e orientado por uma lógica sancionatória, para a incorporação do controle de desempenho e resultados, de uma atuação preventiva e orientadora, bem como da lógica centrada na geração de valor público.

O Guia de Maturidade em Avaliação de Políticas Pública apresentado ao final da pesquisa, enquanto Produto Técnico-Tecnológico, não deve ser compreendido apenas como um instrumento técnico, mas como um mecanismo de indução de mudança organizacional.

Por fim, cumpre destacar que a dissertação está estruturada em seis capítulos. Após esta introdução, o capítulo 2 apresenta o quadro teórico-conceitual sobre avaliação de políticas públicas e a atuação dos Tribunais de Contas. O capítulo 3 detalha os métodos e técnicas de pesquisa. O capítulo 4 dedica-se à apresentação e discussão dos resultados obtidos nos grupos focais e na análise documental de relatórios de avaliação. O capítulo 5 traz as conclusões e recomendações, seguido pelo capítulo 6, que descreve o Produto Técnico-Tecnológico desenvolvido.

2. QUADRO TEÓRICO-CONCEITUAL

Este capítulo delinea o quadro teórico-conceitual do estudo, abrangendo as definições, a evolução histórica, as metodologias e usos mais frequentes da avaliação de políticas públicas. Subsequentemente, examina-se a literatura sobre a avaliação de políticas públicas no Brasil, com ênfase na legitimidade e na relevância da atuação dos Tribunais de Contas nesse campo.

2.1 Fundamentos e institucionalização da avaliação de políticas públicas

2.1.1. *Conceitos e delimitações*

Embora não exista consenso quanto à definição da terminologia ‘avaliação de políticas públicas’, observa-se uma certa convergência quanto ao objetivo da avaliação (Queiroz & Capelari, 2020; Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008). Sob uma ótica normativa, a avaliação de políticas públicas é vista como um conjunto de regras ou princípios incorporados na legislação ou outros documentos que são usados para orientar a prática da avaliação (Bourgeois & Maltais, 2023). Em uma perspectiva mais abrangente, compreende-se a avaliação de políticas públicas como a emissão de juízo de valor de uma iniciativa organizacional, fundamentada em parâmetros comparativos previamente definidos (Costa & Castanhar, 2003; Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008).

Utilizando-se conceitos mais recentes da literatura, obtém-se a definição de avaliação de políticas públicas como uma tarefa de investigação sistemática, baseada em evidências e devidamente documentada, com emissão de julgamentos em relação aos resultados obtidos com sua implementação, a fim de gerar conhecimentos para informar futuras decisões e processos de elaboração de políticas públicas (Boullosa et al., 2021; Costa & Castanhar, 2003; Jannuzzi, 2018; Patton, 2023; Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008).

No âmbito dos Tribunais de Contas, imprescindível destacar a definição constante das Normas Brasileiras de Auditoria do Setor Público, NBASP 9020 – Avaliação de Políticas Públicas (NBASP 9020, 2021), de que a avaliação de políticas públicas consiste em um exame que objetiva avaliar a utilidade de uma política por meio da análise de seus objetivos, implementação, produtos, resultados e impactos o mais sistematicamente possível, com o objetivo de aferir seu desempenho (Lima, 2023; Medeiros, 2023; NBASP 9020, 2021).

O referido normativo apresenta uma distinção entre os objetos da Auditoria Operacional (ANOP) e da Avaliação de Políticas Públicas (APP). Enquanto a ANOP trata da economicidade, da eficiência e da efetividade até o nível do resultado imediato, a APP, além

dessas mesmas dimensões, expande a análise para a análise da relevância e dos impactos socioeconômicos de longo prazo (NBASP 9020, 2021; OCDE, 2022). Conforme a NBASP 9020, pode-se interpretar que a avaliação de políticas públicas por Tribunais de Contas deve ir além do realizado por meio de auditorias operacionais, conforme será discutido oportunamente.

A Organização para a Cooperação e Desenvolvimento Econômico (OCDE) preconiza que a avaliação de políticas públicas deve focar na análise da relevância, no alcance dos objetivos, na eficiência, na efetividade, no impacto e na sustentabilidade do desenvolvimento proposto pela política (OCDE, 2022). Também destaca que a avaliação de políticas públicas deve proporcionar informação verossímil e útil para permitir a incorporação da experiência adquirida no processo de tomada de decisão governamental (Salama & Picalarga, 2024; Trevisan & Van Bellen, 2008).

As recomendações da OCDE se apoiam em três pilares considerados fundamentais para a melhoria da utilização das avaliações de políticas públicas: a) institucionalização da avaliação, numa perspectiva de envolvimento de todo o governo; b) promoção da qualidade das avaliações; e c) realização de avaliações que impactem efetivamente a tomada de decisões (OCDE, 2022; Salama & Picalarga, 2024).

Para fins do presente trabalho, adota-se a definição de avaliação de políticas públicas baseada na NBASP 9020 e na literatura mencionada, de que se trata de um exame sistemático, baseado em evidências, com o objetivo de avaliar a relevância e a utilidade de uma política por meio da análise de seus objetivos, desenho, implementação, produtos, resultados (imediatos) e impactos (longo prazo), com a finalidade de aferir o desempenho de determinada política pública (Medeiros, 2023; NBASP 9020, 2021).

2.1.2 Evolução internacional e institucionalização no Brasil

Inicialmente, impende distinguir análise de políticas públicas de avaliação de políticas públicas. Enquanto a primeira abrange todo o ciclo de formulação (*policy cycle*), a segunda constitui um tipo específico de análise sistemática voltada a determinar o mérito e o valor de uma intervenção, podendo assumir caráter formativo (foco nos processos) ou sumativo (foco nos resultados e impactos) (Lemire et al., 2020; Lima & Mendes, 2021; Oliveira & Passador, 2018; Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008).

A temática de avaliação de políticas públicas foi pioneira nos EUA, na década de 1950, sob a denominação de *policy science*. Na Europa, notadamente na Alemanha, a prática ganhou robustez apenas na década de 1970 (Trevisan & Van Bellen, 2008). É plausível que esse

intervalo temporal tenha contribuído para a predominância da literatura norte-americana na consolidação do campo.

Em uma perspectiva macro-histórica da evolução da avaliação de políticas públicas, consideradas as transformações ocorridas em períodos específicos pelo mundo de forma geral, a literatura classifica a evolução da avaliação de políticas públicas em quatro ondas ou gerações de avaliações no cenário mundial: a onda orientada pela ciência; a onda orientada pelo diálogo; a onda neoliberal; e a onda de evidências (Lima & Mendes, 2021; Oliveira & Passador, 2019; Vedung, 2010).

A primeira onda (1965-1975), orientada pela ciência e pelo racionalismo, predominou em países como EUA, França e Alemanha. Focada na mensuração técnica de variáveis e resultados, buscava determinar os meios mais eficientes para o alcance dos fins (custo-benefício). O objetivo era promover uma governança racional e científica, fundamentada em avaliações conduzidas por pesquisadores acadêmicos (Lima & Mendes, 2021; Oliveira & Passador, 2019; Vedung, 2010).

A segunda onda (meados dos anos 70), orientada pelo diálogo, emergiu do declínio da confiança na avaliação puramente experimental. Passou-se a defender um modelo pluralista e construtivista, pautado na racionalidade comunicativa e no envolvimento dos *stakeholders*. O escopo da avaliação deslocou-se da produção de ‘verdades’ absolutas para a construção de consensos, aceitabilidade política e legitimidade democrática (Oliveira & Passador, 2019; Vedung, 2010).

A terceira onda (década de 80), de cunho neoliberal, refletiu o movimento de redução do Estado e a promoção de parcerias público-privadas. Caracterizou-se pela introdução de lógicas de mercado na administração pública, incluindo descentralização, desregulamentação, privatização, sociedade civil e foco na satisfação do cidadão-cliente (Oliveira & Passador, 2019; Vedung, 2010).

Sob a influência da Nova Gestão Pública (NGP), priorizou-se a gestão por resultados e a autonomia gerencial. A avaliação fortaleceu-se como instrumento de *accountability* e medição de desempenho, enfatizando a economicidade, a eficácia e a eficiência na aplicação dos recursos públicos (Oliveira & Passador, 2019; Vedung, 2010). A terceira onda trouxe um fortalecimento da avaliação, com maior ênfase na responsabilidade dos agentes no uso de recursos públicos (economia, eficácia e eficiência de custos) e na medição de desempenho (Oliveira & Passador, 2019; Vedung, 2010).

A quarta onda (anos 2000), baseada em evidências (*evidence-based policy*), consolidou-se pela exigência de comprovação da eficácia das intervenções estatais (o que funciona). Estabeleceu-se uma hierarquia de evidências, privilegiando experimentos randomizados e quase-experimentais para inferência causal. Esse ‘renascimento científico’ instrumentalizou a avaliação como mecanismo de legitimação para a continuidade, reformulação ou descontinuidade de programas governamentais (Oliveira & Passador, 2019; Vedung, 2010). As quatro ondas informam o caso brasileiro, no qual coexistem desempenho (NGP), diálogo (práticas consensuais) e evidências (inferência causal), como será demonstrado mais adiante.

No Brasil, a avaliação de políticas públicas ganhou relevância na transição democrática (final dos anos 80 e início dos 90). O processo foi impulsionado pela reforma do Estado e pela pressão de organismos internacionais de fomento, que condicionavam o financiamento à demonstração de resultados (Lima & Mendes, 2021; Trevisan & Van Bellen, 2008).

A redemocratização deslocou a agenda pública para temas como descentralização, participação e transparência. A persistência de desafios na efetividade das políticas sociais e a difusão internacional das teses de reforma administrativa reforçaram a necessidade da avaliação como ferramenta de gestão, orientando a produção acadêmica e governamental do período (Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008).

Nota-se que a institucionalização da avaliação no Brasil é recente, alinhada à terceira onda global. Contudo, a despeito do crescimento quantitativo de estudos, a avaliação ainda não se consolidou plenamente como etapa indissociável da gestão pública nacional, haja vista o baixo índice de utilização efetiva dos resultados pelos gestores públicos (Oliveira & Passador, 2019; Queiroz & Capelari, 2020; Salama & Picalarga, 2024; Trevisan & Van Bellen, 2008).

Observa-se o fortalecimento do campo no país, com predominância de avaliações qualitativas, de escopo local e focadas na prestação de contas (*accountability*) e na qualidade dos serviços. Recentemente, emergem experiências baseadas no processamento, integração e análise de grandes volumes de dados (*big data*) (Oliveira & Passador, 2019).

A matéria adquiriu *status* constitucional com as Emendas nº 108/2020 e nº 109/2021. A primeira inseriu o parágrafo único ao artigo 193 da CF/88, estabelecendo o planejamento e a avaliação de políticas públicas como dever estatal. A segunda acrescentou o § 16 ao art. 37, tornando obrigatória a avaliação pelos órgãos da administração pública, com a devida publicidade do objeto e dos resultados alcançados (Bitencourt & Tavares, 2022).

Não obstante o avanço normativo, persistem obstáculos práticos: o desenho deficiente de programas (*evaluability*), a prevalência de metodologias frágeis no controle de variáveis, a

resistência cultural de gestores (Ramos & Schabbach, 2012) e a baixa confiabilidade ou descontinuidade das bases de dados e indicadores (Oliveira & Passador, 2019; Ramos & Schabbach, 2012).

2.2 Dimensões analíticas da avaliação

2.2.1 Classificações das avaliações

As avaliações de políticas públicas são tradicionalmente classificadas conforme diferentes critérios: a) momento de implementação, b) posição do avaliador, e c) natureza do objeto (Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008); d) tipo de problema respondido (Ramos & Schabbach, 2012); e) uso destinado (Oliveira & Passador, 2019; Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008); e f) abordagem metodológica adotada (Oliveira & Passador, 2019).

Quanto ao momento, classifica-se em *ex ante* ou *ex post*. A avaliação *ex ante* precede a implementação, subsidiando a formulação e o exame de custo-benefício e custo-efetividade, visando adequar recursos aos objetivos e amparar a decisão sobre a implementação, ou não, da política (Lima & Mendes, 2021; Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008). A avaliação *ex post* ocorre durante ou após a execução, baseando-se nos resultados alcançados. Seus achados orientam decisões sobre a manutenção, reformulação ou descontinuidade da política pública (Lima & Mendes, 2021; Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008).

Quanto ao agente que avalia, a avaliação pode ser: externa, quando conduzida por atores alheios à instituição gestora (e.g. Tribunais de Contas); interna, executada pela própria instituição; ou participativa, envolvendo os beneficiários em todas as etapas do ciclo (Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008). Admite-se, ainda, a avaliação mista, que combina características das modalidades interna e externa para mitigar vieses e maximizar potencialidades (Oliveira & Passador, 2019). Quanto à natureza, divide-se em formativa ou somativa. A formativa foca nos processos e na implementação, fornecendo informações para correção de rumos e aprimoramento procedimental (Lima & Mendes, 2021; Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008). A somativa, por sua vez, avalia o valor global do programa, verificando o alcance de metas e analisando a relação entre processos, resultados e impactos finais (Lima & Mendes, 2021; Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008).

Quanto ao objetivo da avaliação ou tipo de problema ao qual a avaliação responde, distingue-se a avaliação de processos (avaliação de implementação), realizada durante a implementação da política, relacionada à dimensão da gestão; e a avaliação de resultados e impacto (avaliação de metas e avaliação de impacto), mais abrangente e desafiadora do que a avaliação de processos, realizada com o objetivo de responder se a política funcionou ou não (Oliveira & Passador, 2019; Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008). Quanto à abordagem metodológica, classificam-se em qualitativas, quando focadas em narrativas (grupos focais) e análise documental, e em quantitativas, quando baseadas em dados numéricos e inferência estatística/econométrica. Existem também os métodos mistos (*mixed methods*), que combinam ambas as estratégias (Oliveira & Passador, 2019).

As avaliações também são classificadas de acordo com o uso ao qual se destinam, sendo possível identificar quatro formas básicas: instrumental, conceitual, persuasivo (simbólico) e de esclarecimento. O uso instrumental vincula-se à aplicação direta dos resultados na tomada de decisão, dependendo da qualidade técnica e da exequibilidade das recomendações (Oliveira & Passador, 2019; Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008). O uso conceitual refere-se ao aprendizado e à alteração da percepção dos gestores sobre a política. O uso persuasivo (ou simbólico) instrumentaliza a avaliação para legitimar posições políticas ou convencer *stakeholders* sobre a necessidade de mudanças (Oliveira & Passador, 2019; Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008). Por fim, o uso para esclarecimento incide sobre a comunidade ampliada e formadores de opinião, visando alterar convicções, qualificar o debate público e reorientar a agenda governamental (Oliveira & Passador, 2019; Ramos & Schabbach, 2012; Trevisan & Van Bellen, 2008).

2.2.2 Diversos caminhos possíveis em avaliação de políticas públicas

De forma geral, a avaliação de políticas públicas considera os critérios de eficiência, eficácia e efetividade. Costa e Castanhar (2003) propõem uma abordagem metodológica abrangente, defendendo a definição de critérios conforme o objetivo da avaliação. Dentre os indicadores mais comuns, destacam-se: eficiência, eficácia, impacto (efetividade), sustentabilidade, custo-efetividade, satisfação do usuário e equidade (Costa & Castanhar, 2003; Ipea, 2018b; Trevisan & Van Bellen, 2008).

Estabelecidos os critérios, a avaliação requer a comparação com padrões de referência, que podem ser: absolutos (metas predefinidas); históricos (evolução temporal); normativos (*benchmarking* com programas similares); teóricos (previsões causais do programa); e negociados (consensos entre gestores e formuladores) (Trevisan & Van Bellen, 2008).

Consideradas as diferentes metodologias existentes para avaliação de políticas públicas, reconhece-se a inexistência de um método único que possa ser considerado “padrão-ouro” em todos os contextos avaliativos (Jannuzzi, 2018). A literatura enfatiza que o essencial consiste na seleção de um desenho capaz de estabelecer argumentação causal consistente, adequada à precisão exigida pelo contexto e fundamentada em evidências robustas (Jannuzzi, 2021; Jannuzzi & Mello, 2023).

A robustez de uma avaliação transcende a consistência metodológica, englobando a solidez dos critérios e valores públicos subjacentes. A avaliação não é um ato neutro; suas definições de escopo e método são orientadas por motivações políticas e valores compartilhados pelos atores envolvidos no financiamento, controle e operação das políticas públicas (Boullosa et al., 2021; Jannuzzi, 2018, 2021, 2022; Tinoco et al., 2011).

2.2.3 Marcos teóricos

O presente estudo ancora-se em dois marcos teóricos contemporâneos que utilizam figuras metafóricas para identificar, categorizar e auxiliar na interpretação da diversidade de dados obtidos por meio de grupos focais realizados e relatórios examinados. Examina-se inicialmente a *Árvore da Teoria da Avaliação*, revisitada por Lemire et al. (2020), que apresenta uma taxonomia consolidada sobre os alicerces e as ramificações principais da avaliação (Lemire et al., 2020; Senra et al., 2024). Em seguida, o modelo do Jardim das Abordagens de Avaliação, proposto por Montrosse-Moorhead et al. (2024), que expande a análise da avaliação ao incorporar aspectos axiológicos emergentes e a complexidade dos contextos de aplicação. As duas sistematizações fornecem um arcabouço consistente para a interpretação das práticas dos Tribunais de Contas brasileiros.

2.2.3.1 A *Árvore da Teoria da Avaliação*

A ‘*Árvore de Avaliação*’, proposta por Christie e Alkin (2004), sistematiza o campo teórico em três ramos principais, conforme a ênfase de cada abordagem: Métodos (metodologia de pesquisa), Valoração (julgamento de mérito, atribuição de valor) e Uso (orientação para decisão e capacitação). Essa estrutura permite classificar os diversos teóricos e modelos avaliativos segundo seu foco preponderante (Carden & Alkin, 2012; Christie & Alkin, 2008; Heberger et al., 2010; Senra et al., 2024).

Revisões posteriores, como a de Christie e Alkin (2008), bifurcaram o ramo da valoração sob duas influências: a pós-positivista (que busca medir a verdade aproximada,

reconhecendo a falibilidade da ciência) e a construtivista/interpretativista (que assume realidades múltiplas e construídas socialmente). Essa distinção resultou em uma nova classificação dos teóricos na estrutura da árvore (Christie & Alkin, 2008).

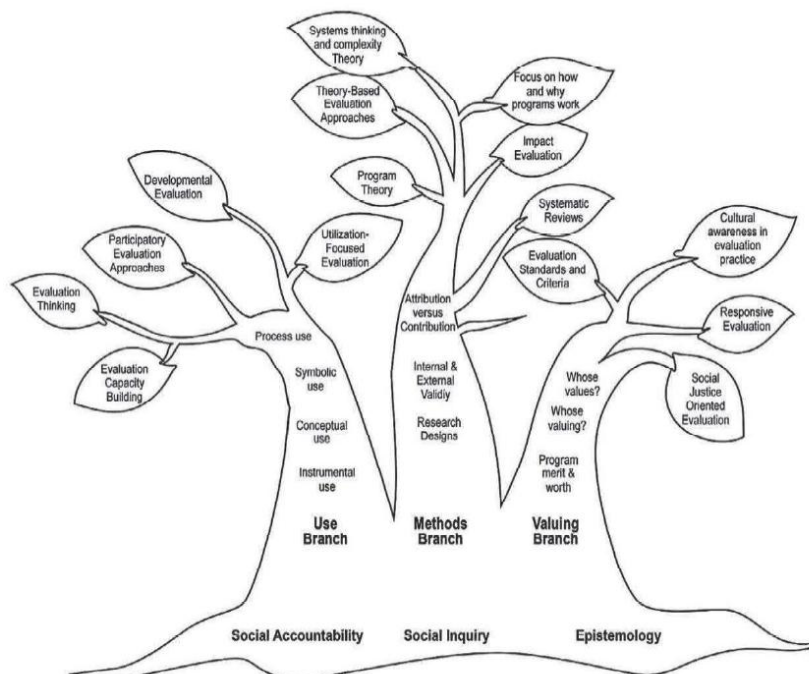
Senra (2024) cita que o modelo de Lemire et al. (2020) foi expandindo para integrar um quarto ramo: a Justiça Social. E que os ramos foram associados a quatro paradigmas filosóficos que orientam a escolha dos instrumentos: pós-positivista (métodos experimentais); pragmatismo (foco na utilidade); construtivismo (compreensão qualitativa); e transformativo (foco na equidade e direitos humanos). A escolha do paradigma orienta a definição dos instrumentos avaliativos: abordagens pragmáticas tendem a indicadores objetivos, enquanto as qualitativas priorizam narrativas e percepções dos participantes (Senra et al., 2024).

A luz das obras de referências, conclui-se que a Árvore de Christie e Alkin e os principais Paradigmas da Avaliação oferecem uma visão sistemática e abrangente dos componentes envolvidos no processo de avaliação, com destaque para a importância dos métodos, uso, valores e justiça social para a realização de avaliações relevantes (Senra et al., 2024).

Lemire et al. (2020) atualizaram o modelo para refletir o crescimento da árvore na floresta de análise de políticas públicas. Nesta versão, identificaram as raízes fundamentais da avaliação na responsabilidade social, na investigação social sistemática e na epistemologia, a partir das quais mapearam tendências contemporâneas nos três ramos clássicos (Lemire et al., 2020). A Figura 1 apresenta o ‘crescimento da árvore de avaliação na floresta de políticas públicas’, segundo Lemire et al. (2020):

Figura 1

A árvore de avaliação de Lemire et al. (2020).



Nota. Adaptado de Lemire et al. (2020).

As raízes sustentam o sistema: a responsabilidade social exige a demonstração de impactos; a investigação sistemática assegura o rigor científico; e a epistemologia valida a natureza do conhecimento produzido (Lemire et al., 2020). O modelo revisitado mantém os três ramos da avaliação: métodos, valoração e uso das avaliações.

O ramo principal e central da árvore teórica consiste nos ‘métodos’, referente à metodologia de pesquisa, incluindo projetos, métodos, abordagens e técnica para conduzir avaliações (Lemire et al., 2020). O ramo da ‘valoração’ trata da determinação do mérito da política. Engloba perspectivas normativas (baseadas em valores) e positivistas (baseadas em fatos), além de abordagens orientadas para a justiça social e responsivas às necessidades dos beneficiários. Também considera ‘como’, ‘por quem’ e ‘de que forma’ os julgamentos avaliativos sobre programas sociais são determinados. Este ramo também inclui abordagens orientadas para a justiça social e abordagens responsivas à avaliação (concentradas na diversidade social do programa, beneficiários e partes interessadas, e suas necessidades) (Lemire et al., 2020).

Por fim, o ramo do ‘uso’ investiga os fatores associados à utilização da avaliação. Além dos usos clássicos (instrumental, conceitual e simbólico), destaca-se contemporaneamente o

‘uso de processo’ (*process use*), referente ao aprendizado derivado da própria participação dos *stakeholders* no ciclo avaliativo (Lemire et al., 2020).

Lemire et al. (2020) identificaram três tendências proeminentes no ramo dos métodos: a) análise de *big data*: uso de novas fontes de dados e de estratégias analíticas, facilitadas por uma gama de tecnologias emergentes de informação e comunicação para coleta e análise de dados; b) compreensão de ‘como’ e ‘por que’ os programas funcionam, sob duas vertentes: métodos baseados em casos e métodos baseados em variância; e c) teoria da complexidade e pensamento sistêmico: programas complexos são fluidos e estão em constante desenvolvimento, reagem e respondem aos desafios e oportunidades emergentes, muitos dos quais são difíceis de prever e se encontram fora do controle do pessoal do programa e dos avaliadores, razão pela qual exigem a aplicação de diversas metodologias para sua avaliação e o uso de um pensamento sistêmico, partindo-se da premissa de que todos os programas e políticas estão incorporados em sistemas mais amplos (Lemire et al., 2020).

No ramo do ‘valor’, a tendência proeminente é a consciência e competência cultural. Exige-se que o avaliador reconheça vieses e considere o contexto cultural dos beneficiários, superando as limitações das abordagens epistemológicas estritamente ocidentais, especialmente em avaliações de desenvolvimento internacional e comunidades indígenas (Lemire et al., 2020).

A Associação Americana de Avaliação (2018) reforça essa diretriz, definindo a competência cultural como a capacidade de adequar teorias e métodos ao contexto, refletindo valores e perspectivas dos beneficiários para assegurar boas práticas avaliativas (Lemire et al., 2020).

No ramo do ‘uso’, as tendências convergem para a capacitação em avaliação (desenvolvimento do pensamento avaliativo organizacional) e para abordagens colaborativas/participativas, que integram os *stakeholders* no planejamento e monitoramento. A capacitação em avaliação (*evaluation capacity building*) refere-se ao desenvolvimento de estruturas, processos e do ‘pensamento avaliativo’ nas organizações, fomentando uma cultura crítica baseada em evidências e reflexão contínua para subsidiar decisões (Lemire et al., 2020).

As abordagens colaborativas enfatizam o envolvimento dos *stakeholders*. Incluem modalidades como a avaliação do desenvolvimento (foco em inovação e aprendizagem), a avaliação de ciclo rápido (*feedback* contínuo) e o mapeamento de resultados (*outcome mapping*), onde as metas e a aferição de resultados (intencionais ou não) são construídas participativamente (Lemire et al., 2020). Tais abordagens podem estruturar-se via pesquisa-

ação participativa, parcerias entre pesquisadores e profissionais, ou comunidades de prática, fortalecendo a colaboração entre pares (Lemire et al., 2020).

Em síntese, as seis tendências centrais da ‘árvore revisitada’ são: a) análise de *big data*; b) compreensão causal (‘como/por que’); c) teoria da complexidade; d) capacitação em avaliação; e) abordagens colaborativas; e f) consciência cultural na prática avaliativa (Lemire et al., 2020). A partir dessas tendências identificadas, Lemire et al. (2020) chegaram a três conclusões: i) a avaliação é um tipo de produção de conhecimento e uma forma de evidência; ii) os trabalhadores de conhecimento, os profissionais e os decisores políticos, são essenciais para a promoção de evidências; e iii) a resolução de problemas sociais exige a integração proposital de múltiplas disciplinas.

2.2.3.2 O Jardim das Abordagens de Avaliação

Montrosse-Moorhead et al. (2024) propuseram o ‘Jardim das Abordagens’, classificando diferentes abordagens de avaliação de acordo com oito dimensões para consideração prática: a) valores, b) valoração, c) justiça social, d) contexto, e) promoção do uso; f) envolvimento e g) profundidade no processo de avaliação, e h) dinâmica de poder na tomada de decisão. A classificação emprega a metáfora visual de flores para ilustrar as configurações filosóficas e metodológicas de cada vertente. Os valores se referem às crenças sobre o que é importante/válido; a valoração ao processo de determinação de mérito/valor; e o ativismo pela justiça social ao grau de compromisso da avaliação como instrumento de mudança social (Montrosse-Moorhead et al., 2024).

O contexto compreende aferir o quanto em uma abordagem de avaliação são considerados os contextos ou sistemas culturais, históricos e políticos em que a avaliação se insere, reconhecendo-se, é claro, que avaliações podem ser descontextualizadas. A promoção do uso, refere-se ao grau em que uma abordagem de avaliação orienta os avaliadores e facilita direta e ativamente o seu uso. O envolvimento no processo de avaliação tem por finalidade demonstrar ‘quem está envolvido’ na avaliação: se apenas os que detêm o poder formal sobre a avaliação e os que implementam as ações, ou se incluem outros grupos: público-alvo e partes beneficiárias, e avaliação participativa (Montrosse-Moorhead et al., 2024).

A profundidade do envolvimento no processo de avaliação, refere-se ao grau de envolvimento de diferentes grupos de pessoas ao longo de uma avaliação e em quais funções (se há envolvimento ou não, ou se este é limitado). Por fim, a dinâmica de poder na tomada de decisões de avaliação refere-se à orientação de uma abordagem de avaliação sobre quem está

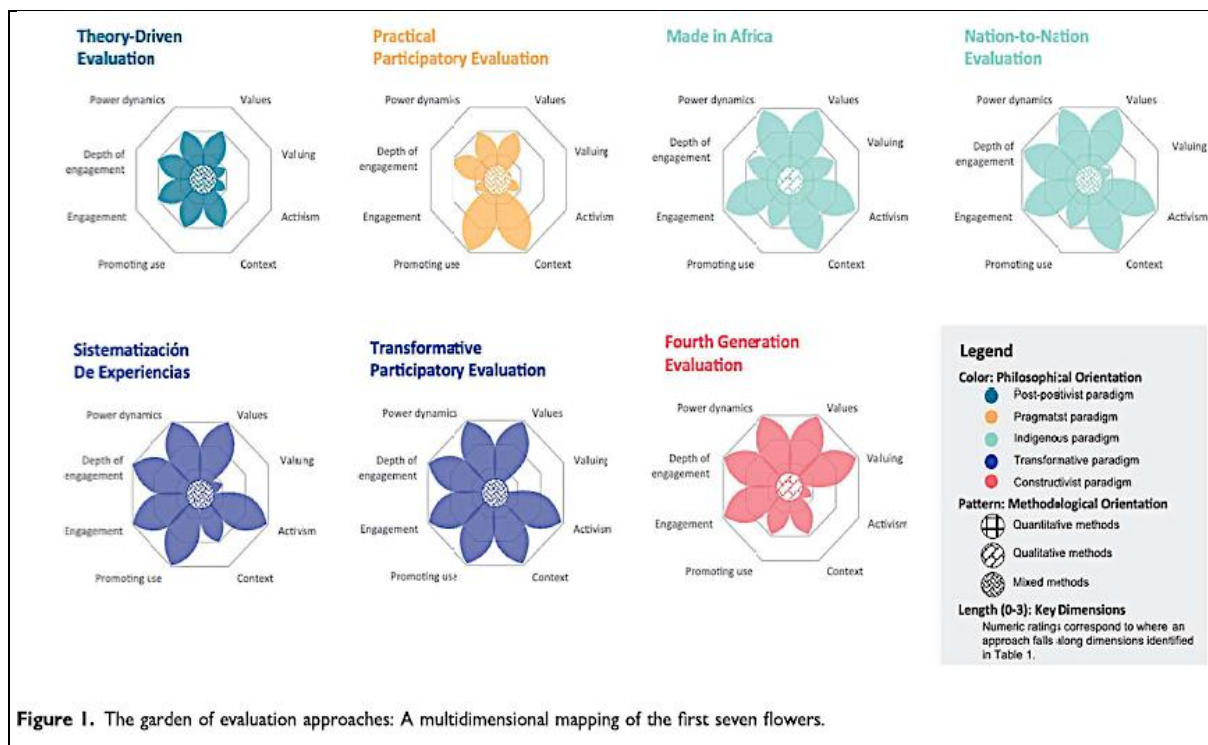
envolvido no processo de tomada de decisões, ou seja, o ‘que’ e ‘como’ é feito’ durante a avaliação (Montrosse-Moorhead et al., 2024).

Graficamente, o tamanho das pétalas representa a intensidade de cada dimensão na abordagem analisada, variando em uma escala de 0 (sem função) a 3 (papel decisivo/fundamental). Os critérios detalhados dessa gradação constam do Anexo A. A metáfora do jardim justifica-se pela possibilidade de mistura, evolução e descoberta de novas abordagens (‘flores’), cujas pétalas refletem a configuração única das oito dimensões em cada caso (Montrosse-Moorhead et al., 2024). Cada ‘flor’ possui três camadas: a orientação filosófica (cor), a disposição metodológica (padrão central: quantitativa, qualitativa, mista ou transformadora) e as dimensões-chave (comprimento das pétalas) (Montrosse-Moorhead et al., 2024).

A Figura 2 apresenta o mapa dimensional das sete abordagens avaliativas analisadas no estudo.

Figura 2

O jardim das abordagens de avaliação.



Nota: Montrosse-Moorhead et al. (2024).

Para o escopo deste estudo, a ‘Árvore Teórica’ e o ‘Jardim das Abordagens’ mostram-se promissores na medida que são utilizados para analisar as avaliações conduzidas pelos

Tribunais de Contas brasileiros. Até onde alcança o levantamento realizado, não se identificaram aplicações anteriores dos marcos teóricos da *Árvore* (Lemire et al., 2020) e do *Jardim* (Montrosse-Moorhead et al., 2024) no âmbito dos Tribunais de Contas. O trabalho visa preencher essa lacuna, contribuindo para o avanço teórico e o fortalecimento prático do Controle Externo nacional.

Foram adotados os ramos da *Árvore de Avaliação* de Lemire et al. (2020) - métodos, valoração e uso - como critérios a serem investigados nas avaliações de políticas públicas realizadas nos Tribunais de Contas. A fim de complementar os critérios e facilitar a interpretação dos dados coletados, adicionalmente foram utilizadas dimensões do *Jardim das Abordagens* de Montrosse-Moorhead (2024) - ativismo pela justiça social; contexto; promoção do uso; quem está envolvido no processo de avaliação e em que profundidade; e dinâmica de poder na tomada de decisões de avaliação.

2.3 Avaliação no âmbito dos Tribunais de Contas

A avaliação de políticas públicas vem sendo progressivamente reconhecida como uma prática estratégica para o aprimoramento da ação governamental e para a qualificação do processo decisório no setor público (Castro & Castro, 2021). No âmbito do Controle Externo, essa relevância se intensifica, uma vez que os Tribunais de Contas ocupam posição institucional singular, caracterizada pela independência funcional, pela imparcialidade na fiscalização e por competências constitucionalmente estabelecidas, que lhes permitem avaliar a atuação de todos os Poderes e órgãos públicos, sem subordinação hierárquica (Lima, 2023; Oliveira & Mendonça, 2020).

2.3.1 Legitimidade e inovação no Controle Externo

A Constituição Federal de 1988 (Brasil, 1988) inaugurou uma nova ordem jurídica ao instituir um extenso rol de direitos fundamentais que passaram a orientar a definição das políticas públicas no Brasil (Cunda, 2011; Oliveira & Mendonça, 2020). Esse arcabouço constitucional impõe um gerenciamento das finanças públicas apto a assegurar a concretização de direitos sociais e individuais, exigindo, por conseguinte, que a aplicação dos recursos estatais seja otimizada e vocacionada à efetividade (Botelho & Albuquerque, 2024; Cunda, 2011; Leoni, 2023; Medeiros, 2023; Oliveira & Mendonça, 2020). Nesse cenário, o papel dos Tribunais de Contas transcende a verificação da mera legalidade estrita, estendendo-se, por força das competências fixadas no artigo 70 da Carta Magna, à análise de legitimidade, economicidade, eficiência, eficácia e efetividade do gasto público (Bitencourt & Tavares, 2022;

Brasil, 1988; Chudzij, 2020; Leoni, 2023; Lima, 2023; Medeiros, 2023). Historicamente, a atuação das Cortes de Contas brasileiras evoluiu de um controle eminentemente formalista e burocrático para um modelo indutor da qualidade da gestão pública. Contudo, essa transição não ocorreu sem turbulências. As instituições de controle externo têm sido alvo de críticas severas e crises de legitimidade, materializadas na tramitação de mais de 40 Propostas de Emenda à Constituição (PECs) entre 2000 e 2017, que visavam desde o aperfeiçoamento até a extinção desses órgãos, a exemplo do extinto Tribunal de Contas dos Municípios do Ceará (TCM-CE), sob argumentos de economia de recursos e eliminação de duplicidade de funções (Luz et al., 2019, 2021).

Tal movimento legislativo evidencia uma pressão social e política por maior efetividade no cumprimento do papel constitucional e pela modernização dos mecanismos organizacionais e processuais das Cortes (Luz et al., 2021). É nesse contexto desafiador que a avaliação de políticas públicas emerge como uma estratégia fundamental de legitimação institucional. Ao transcender o exame da conformidade financeira e incorporar dimensões substantivas de eficiência e efetividade, os Tribunais demonstram sua utilidade social, induzindo o aperfeiçoamento da gestão e entregando valor público à sociedade (Leoni, 2023; Luz et al., 2019; Macedo, 2023).

As críticas relativas à legitimidade democrática ou ao suposto ‘déficit eletivo’ dos membros dos Tribunais são contrapostas pelo duplo fundamento de sua atuação: a competência expressa nos artigos 70 e 71 da CF/88 e a capacidade de promover o diálogo institucional. Sob essa ótica, a avaliação qualifica os dados disponíveis para a gestão, fomenta o controle social e colabora para a consecução das políticas públicas, sem, contudo, substituir a discricionariedade do gestor eleito (Gomes, 2020; Leoni, 2023; Macedo, 2023).

A superação da crise de confiança institucional demanda, todavia, avaliações fundadas em parâmetros técnicos objetivos. O fortalecimento da *accountability* horizontal, via avaliação de políticas, visa assegurar a correta aplicação dos recursos na garantia do mínimo existencial e na efetividade dos direitos fundamentais, combatendo atitudes comissivas e omissivas do administrador que possam ofender os princípios da Administração Pública (Botelho & Albuquerque, 2024; Chudzij, 2020; Cunda, 2011; Lopes et al., 2023; Luz et al., 2019).

Essa atuação baseia-se precipuamente no exercício das funções pedagógica e colaborativa do Controle Externo, orientando o gestor a reavaliar decisões e promover ajustes necessários (Medeiros, 2023; Tavares & Bitencourt, 2022). Todavia, assegurando-se aos Tribunais a prerrogativa de exercer as funções sancionatória e cogente, no caso de ilegalidades

ou descumprimentos injustificados de deliberações, respeitando-se sempre os limites do mérito administrativo (Cunda, 2011; Willeman, 2020).

Não obstante os avanços, desafios persistem. O uso das avaliações muitas vezes ainda se restringe ao caráter informativo, com baixo impacto no julgamento das contas. A predominância de uma cultura burocrática e a tecnicidade excessiva dos relatórios podem limitar a transparência e dificultar o exercício do controle social vertical (Lopes et al., 2023; Medeiros, 2023). Por outro lado, a análise sugere que a adoção de práticas dialógicas pode potencializar os Tribunais como instâncias de articulação interinstitucional, induzindo um planejamento aderente às prioridades constitucionais e contribuindo para a efetiva garantia dos direitos fundamentais (Castro & Castro, 2021; Chudzij, 2020; Macedo, 2023; Ramalho & Graziane Pinto, 2024).

2.3.2 Auditorias Operacionais

O estudo parte da premissa de que os Tribunais de Contas têm o dever de avaliar sistematicamente as políticas públicas. Tal atuação, instrumentada por determinações, recomendações e sanções, configura-se como mecanismo legítimo de *accountability* e aprendizado institucional, visando o aprimoramento do valor público (Castro & Castro, 2021).

No Brasil, essa prática materializa-se predominantemente via Auditorias Operacionais (ANOP), ou Auditorias de Desempenho (AD), utilizadas por diversos Tribunais de Contas para examinar políticas, programas e organizações públicas em múltiplas áreas, como saúde, educação, segurança pública e infraestrutura (Carvalho et al., 2022; Lima, 2023; Sachsida, 2018; TCU, 2020). Essas auditorias têm produzido resultados considerados significativos, incluindo melhorias no desempenho administrativo, redução de desperdícios, estímulo a boas práticas, ampliação da transparência, fortalecimento do controle social e qualificação do debate público sobre políticas governamentais (Botelho & Albuquerque, 2024; Medeiros, 2023).

Contudo, a literatura aponta limites na equiparação entre avaliação de políticas públicas e auditoria operacional. Argumenta-se que as diretrizes tradicionais de ANOP, focadas nos ‘3 Es’ (economicidade, eficiência, eficácia), podem não contemplar plenamente as dimensões substantivas do valor público (Carvalho, 2021; Lima, 2023; Medeiros, 2023; Nunes et al., 2022).

Evidências empíricas reforçam essa distinção. Pesquisa conduzida no âmbito do Tribunal de Contas do Estado de Santa Catarina (TCE/SC) identificou que a avaliação de políticas públicas demanda a incorporação explícita da dimensão da ‘utilidade pública’, isto é, da capacidade da política de produzir efeitos socialmente relevantes para além de seu público-

alvo imediato, aspecto que, em regra, não é central nas auditorias operacionais tradicionais (Nunes et al., 2022).

Essa diferenciação também se expressa no plano normativo. A comparação entre a NBASP 300 (2013), voltada à auditoria operacional, e a NBASP 9020 (2021), específica para avaliação de políticas públicas, evidencia a ampliação do escopo avaliativo, ao incluir dimensões substantivas, como a utilidade e a relevância da política. Tais dimensões dizem respeito, respectivamente, ao alcance e impacto dos resultados na realidade social, econômica e ambiental e à coerência entre os objetivos da política e às necessidades sociais que se pretende atender (NBASP 300, 2013; NBASP 9020, 2021; Nelson, 2020).

A diversidade de práticas exige a sistematização de orientações metodológicas para assegurar qualidade e consistência. Tal necessidade deve ser balanceada com a crítica à manualização excessiva ('postura antimanual'), preservando o caráter reflexivo e contextual da avaliação (Boullosa et al., 2021).

Reconhecendo a inexistência de métodos universais, este trabalho defende a incorporação da dimensão normativo-valorativa como organizadora do método. O desenho metodológico deve estar ancorado em estrutura que permita explicitar juízos de valor e produzir sínteses socialmente relevantes, problematizando padrões de referência e argumentos (Boullosa et al., 2021). A avaliação é um processo interativo que exige dos auditores a articulação de saberes multidisciplinares e abordagens inovadoras. O objetivo é responder a desafios contemporâneos e subsidiar a tomada de decisão com evidências robustas para resultados sustentáveis (Senra et al., 2024).

2.3.3 Sistematização de categorias de acordo com o quadro teórico-conceitual da pesquisa

Em suma, o referencial teórico desta dissertação sustenta que a avaliação de políticas públicas no Controle Externo não pode ser compreendida como simples extensão da auditoria operacional, pois envolve escolhas metodológicas, julgamentos de valor e estratégias deliberadas de indução ao uso, condicionadas por capacidades institucionais e contextos de implementação. À luz da Árvore da Teoria da Avaliação, a análise organiza-se nos três eixos centrais (métodos, valoração e uso), que orientaram tanto a elaboração do roteiro dos grupos focais quanto a análise de conteúdo dos relatórios das fiscalizações examinados.

Em complemento, as dimensões do Jardim das Abordagens ampliam o olhar para elementos axiológicos e institucionais que incidem sobre o processo avaliativo, como valores públicos, justiça social, análise de contexto (institucional, político e ideacional), promoção do

uso, envolvimento e profundidade de participação de stakeholders e dinâmica de poder decisório.

O quadro analítico da Tabela 1 abaixo apresentada integra conceitos da teoria e da investigação empírica a fim de viabilizar a identificação e a interpretação de padrões e variações nas práticas avaliativas dos Tribunais de Contas, assim como a discussão sobre trajetórias institucionais e as implicações para o aprimoramento metodológico da avaliação de políticas públicas no âmbito do Controle Externo.

Tabela 1

Quadro Síntese da sistematização de categorias analíticas mobilizadas na pesquisa.

Categoria analítica (macro)	Ancoragem teórica (referenciais)	Dimensões / subcategorias mobilizadas (a priori)	Descritores operacionais (o que se observa / registra)
1. Métodos (como a avaliação é concebida e operacionalizada)	Árvore da Teoria da Avaliação (ramo “Métodos”) + tendências contemporâneas.	(i) instrumento de fiscalização predominante; (ii) aderência a referenciais normativos; (iii) modelos/abordagens adotados; (iv) adequação do desenho ao propósito (<i>ex ante/ex post</i> ; <i>avaliação executiva</i> etc.)	Como o TC “faz”: a) tipo de instrumento, passos do planejamento, tipo de evidência, técnicas e modelos, grau de formalização e padrões metodológicos declarados; b) existência de inovação metodológica e de parcerias com a "Academia".
2. Valoração (como se produz juízo de valor: critérios, valores e contexto)	Árvore (ramo “Valoração”) + expansão axiológica (valores emergentes) + Jardim (valores, valoração, justiça social, contexto)	(i) critérios clássicos de desempenho (economicidade, eficiência, eficácia e efetividade); (ii) dimensões substantivas NBASP 9020 (relevância e utilidade); (iii) valores públicos emergentes (<i>equidade, justiça social, sustentabilidade, satisfação do beneficiário</i>); (iv) análise de contexto (institucional, político e ideacional – CIPI); (v) padrões de referência (normativos, históricos, teóricos, <i>benchmarking</i> , negociados).	O que “vale” e “como” se julga: a) critérios utilizados; b) explicitação de valores e padrões de referência; c) emissão de valor quanto à relevância e/ou utilidade da PP; d) realização e registro formal de análise de contexto (CIPI); e) escolhas e prioridades.

3. Uso (para que serve e como se induz a utilização dos resultados)	Árvore (ramo “Uso”) + tipos de uso (instrumental, conceitual, simbólico e “uso de processo” etc.) + Jardim (<i>promoção do uso; envolvimento; profundidade; dinâmica de poder</i>)	(i) tipos de uso pretendidos (<i>instrumental, conceitual, persuasivo/simbólico, esclarecimento etc.</i>); (ii) estratégias de indução do uso das avaliações realizadas; (iii) uso de arranjos dialógicos e/ou consensuais; (iv) envolvimento de <i>stakeholders</i> (quem participa e em que profundidade); (v) dinâmica decisória (<i>quem decide “o quê” no planejamento, execução e divulgação; validação externa</i>).	Como se busca “fazer acontecer”: a) mecanismos e rotinas de engajamento; b) Monitoramento e comunicação; c) existência de arranjos participativos; grau de interação com gestores e/ou sociedade; d) transparência x controle social; e) “autoridade indutiva” vs. coercitiva.
4. Dimensões transversais (condicionantes institucionais e maturidade avaliativa)	Integração dos marcos + literatura sobre institucionalização e capacidade avaliativa	(i) capacidade avaliativa organizacional; (ii) cultura institucional (conformidade × desempenho/valor público); (iii) estrutura organizacional (especialização x difusão); (iv) governança e normatização; (v) limitações estruturais.	a) uniformidade ou heterogeneidade entre Tribunais de Contas; b) Trajetórias em APP; c) Limitações (dificuldades), oportunidades de melhoria e tendências em APP.

Nota: elaboração da autora.

3. MÉTODOS E TÉCNICAS

Para atingir o objetivo geral da pesquisa e descrever as percepções de auditores de controle externo sobre a avaliação de políticas públicas nos Tribunais de Contas brasileiros, com foco em métodos, critérios de valor ação e uso das avaliações, foi adotada uma abordagem qualitativa, de natureza exploratória e descritiva. A estratégia metodológica foi desenhada para atender aos objetivos específicos propostos, articulando evidências de natureza discursiva e documental.

O desenho da pesquisa estruturou-se em duas etapas complementares. A primeira consistiu na realização de grupos focais para a caracterização do panorama nacional das percepções e práticas avaliativas dos auditores de controle externo nos Tribunais de Contas brasileiros, mediante a análise de experiências consolidadas e trajetórias emergentes, com foco nos métodos, nos critérios de valoração e nos usos das avaliações (objetivos específicos ‘a’ e ‘b’). A segunda etapa compreendeu a análise de conteúdo de relatórios de fiscalização, permitindo comparar os achados discursivos com as práticas materializadas nos documentos oficiais (objetivo específico ‘b’). A triangulação desses dados foi realizada para a obtenção de subsídio empírico e teórico necessário para a construção de uma tipologia das trajetórias avaliativas a partir dos padrões e variações institucionais identificados, assim como para a discussão sobre as implicações para o aprimoramento metodológico da avaliação de políticas públicas no âmbito do Controle Externo e para a elaboração de um Guia de Maturidade em Avaliação de Políticas Pública para os Tribunais de Contas com base nas trajetórias institucionais de avaliação de políticas públicas (objetivos específicos ‘c’, “d” e “e”).

Na Tabela 2 a seguir explicita-se a relação entre os objetivos específicos e as dimensões de análise de forma sintética:

Tabela 2

Quadro Síntese dos objetivos específicos da pesquisa x dimensões de análise.

Objetivo específico	Dimensão de análise	Foco	Descritor
a) Caracterizar o panorama nacional das percepções e práticas avaliativas dos auditores de controle externo nos Tribunais de Contas brasileiros.	Descritiva	Concepções dos auditores e práticas avaliativas predominantes	Sistematização das percepções expressas nos grupos focais e das práticas identificadas nos documentos institucionais.

b) Analisar experiências consolidadas e trajetórias emergentes, com foco nos métodos, nos critérios de valoração e nos usos das avaliações.	Analítica	Métodos empregados, critérios de valoração e formas de uso institucional das avaliações	Exame crítico dos instrumentos metodológicos, critérios de julgamento e usos estratégicos das avaliações no Controle Externo.
c) Construir uma tipologia das trajetórias avaliativas a partir dos padrões e variações institucionais identificados.	Comparativa	Padrões recorrentes e variações institucionais nas práticas avaliativas	Triangulação entre percepções e evidências documentais para classificar e sistematizar trajetórias avaliativas nos TCs.
d) Discutir as implicações para o aprimoramento metodológico da avaliação de políticas públicas no âmbito do Controle Externo.	Propositiva	Implicações e diretrizes de aprimoramento metodológico	Integração dos achados para formulação de orientações voltadas ao aperfeiçoamento da avaliação no Controle Externo.

Nota: elaboração da autora.

Nas seções seguintes são apresentadas as duas etapas da pesquisa, iniciando com os grupos focais.

3.1 Grupos Focais (etapa 1)

3.1.1 Participantes da pesquisa e critérios de seleção

O universo da pesquisa abrangeu auditores de controle externo dos 33 Tribunais de Contas do Brasil. O critério de seleção dos participantes fundamentou-se na competência técnica e legal para o exercício da atividade (Distrito Federal, 2019). Por conseguinte, optou-se por restringir a amostra exclusivamente a auditores de controle externo, ou cargos equivalentes nas demais esferas, alocados na atividade finalística, na qualidade de profissionais detentores da competência legal para realizar as atividades finalísticas de caráter técnico, assegurando a qualificação das informações prestadas. No total, participaram dos grupos focais 126 auditores de controle externo.

Como critério eletivo, solicitou-se a participação apenas de auditores alocados em setores que desenvolvessem trabalhos relacionados à avaliação de políticas públicas ou que fossem reconhecidamente experientes na realização de fiscalizações dessa natureza, independentemente de tempo de serviço no tribunal, idade ou sexo. Em algumas Cortes, o atendimento do critério exigiu a realização de um segundo grupo focal, razão pela qual foram realizados 38 grupos focais.

3.1.2 Aspectos éticos e consentimento dos participantes

Os participantes dos grupos focais consentiram, de forma livre e esclarecida, com a gravação dos encontros exclusivamente para a consolidação das informações em formulário próprio, o qual foi submetido à validação pelos auditores após cada sessão. Para esta pesquisa, apenas as informações consolidadas nos formulários foram utilizadas, assegurando-se a anonimidade dos participantes e a confidencialidade dos dados. Foi comunicado previamente que não seria realizado qualquer *ranking* entre Tribunais de Contas e que não se pretendia criticar individualmente nenhuma Corte, mas mapear um panorama geral da avaliação de políticas públicas, destacando-se práticas avaliativas que se sobressaíssem.

3.1.3 Procedimentos de coleta de dados

A coleta de dados primários ocorreu entre 04/06/2025 e 23/08/2025, mediante a realização de 38 grupos focais, incluindo a realização de grupos focais adicionais em cinco Tribunais de Contas para atendimento do critério de elegibilidade dos participantes, explicitado no subitem anterior. A opção pelo grupo focal, ao invés de entrevistas, baseou-se na premissa de que a interação social no grupo propicia a emergência de diversos significados, crenças e culturas organizacionais que influenciam a prática fiscalizatória (Rabiee, 2004). Conforme Rabiee (2004) e a experiência desta pesquisa, os grupos focais permitem coletar múltiplas perspectivas, concordâncias e divergências espontâneas, revelando nuances culturais ou comportamentais.

Cada grupo focal foi composto por no mínimo dois e no máximo seis auditores. Considerado o tempo dispendido nos grupos focais por Tribunal, obtém-se um tempo médio de 106 minutos, sendo o mais curto de 57 minutos e o mais longo de 213 minutos. No total, a primeira etapa de coleta de dados consistiu na realização de aproximadamente 59 horas de discussões em grupo focal.

Para a composição dos grupos, adotou-se um protocolo de validação institucional. Inicialmente, solicitou-se a indicação de pontos focais aos Secretários-Gerais de Controle Externo, com nomenclaturas específicas para o cargo variando por instituição, na qualidade de responsáveis por coordenar e gerenciar as atividades de fiscalização nos Tribunais de Contas. Em seguida, divulgou-se o convite na Rede Integrar de Avaliação de Políticas Públicas, uma parceria entre IRB, ATRICON e Tribunais de Contas aderentes. A lista de participantes, compostas por indicados e voluntários, foi posteriormente divulgada entre os Secretários-Gerais

para fins de validação, com o objetivo de assegurar que os participantes do grupo focal detivessem notório conhecimento ou atuação na área referente à temática da pesquisa.

3.1.4 Roteiro dos grupos focais

O roteiro semiestruturado usado nos grupos focais (disponível no Apêndice A) foi elaborado conforme os referenciais teóricos da pesquisa. O roteiro seguiu a ‘classificação’ dos ramos da Árvore de Avaliação de Lemire et al. (2020), com os critérios ‘macro’ e escala de graduação das dimensões a serem avaliadas em relação às avaliações de políticas públicas realizadas pelos Tribunais de Contas: (a) metodologias de avaliação; (b) critérios para valoração; e (c) estratégias de fomento ao uso dos resultados. Adicionalmente, utilizou-se as dimensões propostas por Montrosse-Moorhead et al. (2024) para facilitar a parametrização do roteiro de discussão do grupo focal, assim como a interpretação e posterior apresentação dos resultados da pesquisa: (a) ativismo pela justiça social; (b) contexto; (c) promoção do uso; (d) quem está envolvido no processo de avaliação e (e) em que profundidade; e (f) dinâmica de poder na tomada de decisões de avaliação (ver a classificação completa no Anexo A).

3.1.5 Tratamento e análise dos dados do grupo focal

Todas as falas nos grupos focais foram gravadas e transcritas. Visando assegurar a fidedignidade dos dados, após cada grupo focal, o roteiro com a síntese dos registros foi encaminhada aos participantes para validação e eventuais retificações. Posteriormente, as informações foram sistematizadas em matrizes analíticas (planilhas), categorizadas em sete eixos temáticos: (a) nível de maturidade; setores encarregados; dimensões analisadas; normativos utilizados; abordagens avaliativas e metodologias; (b) estratégias para o uso efetivo das avaliações; (c) forma de valoração; (d) grau de validação externa das fiscalizações; (e) sugestões de melhoria e desafios; (f) tendências; e (g) observações adicionais.

Essa organização permitiu a identificação de padrões e a interpretação dos dados à luz dos objetivos específicos da pesquisa. Para apoiar a verificação de consistência, empregaram-se ferramentas de inteligência artificial, embora a análise interpretativa tenha sido mantida sob responsabilidade exclusiva da pesquisadora.

3.2 Relatórios de fiscalizações (etapa 2)

3.2.1 Seleção dos relatórios de fiscalizações

A etapa documental da pesquisa consistiu no exame qualitativo de 16 relatórios de auditoria e de outros instrumentos de fiscalização equivalentes de oito Tribunais de Contas, classificados como ‘mais avançados’ na avaliação de políticas públicas após a análise dos dados obtidos dos grupos focais. Nessa categoria, foram incluídos: (a) os tribunais cujos participantes dos grupos focais classificaram a atuação das Cortes como ‘intermediária’ (TCU; TCE/PE; TCE/PI e TCM/RJ); (b) tribunais que iniciaram avaliações denominadas *stricto sensu*, assim consideradas as realizadas mediante observância plena da NBASP 9020 (TCE/ES e TCE/GO); e (c) tribunais em que foi identificado o uso de metodologias avaliativas singulares (TCE/RN e TCE/SC) em relação às demais Cortes de Contas, não obstante também realizem ANOP.

Dentre os relatórios que compuseram a amostra, priorizou-se a seleção dos relacionados às áreas de saúde, educação, assistência social, segurança pública e justiça e cultura, esporte e lazer em face da recorrência temática identificada nos grupos focais. No âmbito do Tribunal de Contas da União (TCU), a amostra compreendeu relatórios de três das nove auditorias integrantes do Relatório de Fiscalizações em Políticas Públicas e Programas de Governo (REPP/2025), selecionados e acessados mediante consulta ao respectivo site institucional. Dos demais tribunais, foram analisados os relatórios selecionados dentre os disponibilizados ou indicados pelos participantes de cada grupo focal. A relação dos relatórios examinados encontra-se detalhada no Apêndice B.

De forma geral, os relatórios examinados se referiram a fiscalizações realizadas em período anterior à realização dos grupos focais, excepcionados os relatórios das primeiras avaliações *stricto sensu* do TCE/ES e TCE/GO, disponibilizados nos meses de outubro e dezembro de 2025, respectivamente, considerados essenciais para a pesquisa.

3.2.2. Análise dos dados

O tratamento dos relatórios selecionados seguiu o rito de análise de conteúdo (Bardin, 2016), percorrendo três fases: a) pré-análise, com leitura flutuante do relatório integral, b) identificação das unidades de registro; e c) categorização, com auxílio do uso do software ATLAS.TI. Após o tratamento dos dados, procedeu-se a análise e interpretação dos textos dos relatórios.

Ao longo da leitura do material transcrito foi realizada a codificação das unidades de registo recorrendo à marcação do texto sem ajustes. Para as análises, foram consideradas na dimensão da ‘metodologia’ as abordagens avaliativas, modelos, métodos e procedimentos avaliativos. Na dimensão de ‘valor’, os critérios utilizados foram economicidade, eficácia, eficiência, relevância, utilidade, equidade, justiça social, sustentabilidade, espécies de análise de contexto etc. Em relação à dimensão do ‘uso’, não foram identificadas transcrições nos textos examinados, confirmando-se o entendimento de que se trata de uma ‘estratégia’, *a priori* não explicitamente declarada nos relatórios.

3.3 Integração dos dados dos grupos focais e dos relatórios

A etapa de integração dos dados operou-se mediante a comparação dialética entre as percepções dos auditores obtidas durante os grupos focais e a materialidade técnica evidenciada nos relatórios das fiscalizações. Essa abordagem permitiu confirmar e complementar a análise meramente declaratória, possibilitando o exame da coerência entre a intenção avaliativa manifestada pelos participantes do grupo focal e as práticas efetivamente registradas nos documentos oficiais.

Para operacionalizar essa integração, utilizou-se como eixo estruturante a "Árvore de Avaliação" de Lemire et al. (2020), complementada pelas dimensões do "Jardim das Abordagens" de Montrosse-Moorhead et al. (2024). As categorias analíticas definidas *a priori* (métodos, valoração e uso) serviram de chaves interpretativas para o mapeamento de convergências e complementação de informações, não tendo sido identificadas dissonâncias. Nesse processo, buscou-se confirmar o uso de abordagens/modelos avaliativos (ANOP, Teoria da Mudança, Modelo Lógico, Diagnóstico do Problema, modelos econométricos de avaliação de impacto etc.); dimensões de análise para valoração da PP avaliada (relevância, utilidade, eficiência, eficácia, justiça social etc.); assim como a existência e tipos de análise de contexto (institucional, político e ideacional).

Adicionalmente, o cruzamento dos dados permitiu confirmar o grau de institucionalização das práticas, distinguindo diferentes ‘formas’ ou ‘estratégias’ avaliativas incorporadas à cultura organizacional das Cortes. A síntese interpretativa resultante desse cruzamento constituiu a base empírica para a elaboração da tipologia das trajetórias institucionais apresentada na discussão dos resultados, permitindo classificar os Tribunais de Contas pela evidência material de suas práticas avaliativas em incipientes, incrementais, normativo-estruturantes e disruptivos (científicos).

3.4. Limitações

Apesar do rigor metodológico empregado, a pesquisa apresenta limitações inerentes ao seu desenho. Primeiramente, ao basear-se fundamentalmente na percepção de auditores, por meio de grupos focais, e em relatórios selecionados como ‘casos de sucesso’, os resultados podem refletir um viés de desajustabilidade social ou uma visão otimizada que não corresponde integralmente à prática cotidiana média de todas as fiscalizações. Ademais, o estudo restringiu-se à ótica do ‘avaliador técnico’ (auditor de controle externo), não incorporando a percepção do Ministério Público de Contas ou dos Membros dos Tribunais (Ministros e Conselheiros), tampouco dos ‘avaliados’ (gestores públicos) ou dos ‘beneficiários’ (parte da sociedade atendida).

4. RESULTADOS E DISCUSSÃO

A apresentação dos resultados foi organizada de modo a oferecer, primeiro, um panorama geral e nacional das percepções e práticas relatadas pelos auditores e, em seguida, um exame aprofundado de experiências consideradas ‘mais avançadas’ ou ‘inovadoras’. Isso permite identificar trajetórias emergentes, gargalos estruturais e diretrizes para o aprimoramento das avaliações de políticas públicas no Controle Externo. Por fim, apresenta-se uma Tipologia dos Tribunais de Contas do Brasil com base nas trajetórias avaliativas de políticas públicas das Cortes de Contas nacionais.

4.1 Concepções e práticas avaliativas dos auditores: caracterização do panorama nacional

A análise dos dados coletados nos grupos focais revelou que a valoração das políticas públicas é uma prática disseminada nos Tribunais de Contas do país. Não obstante, observa-se uma dissonância conceitual significativa, uma vez que a maioria dos auditores não reconhece as ações de fiscalização como avaliação de políticas públicas *stricto sensu*, especialmente pela ausência de aderência aos parâmetros da NBASP 9020. Predomina a percepção de que as fiscalizações são de forma geral fortemente orientadas pela lógica da conformidade e as avaliações de políticas públicas realizadas por meio de auditorias operacionais com adaptações, carecendo de metodologias específicas para aferir sobretudo o impacto e a efetividade das políticas públicas fiscalizadas. Esse cenário é agravado por barreiras culturais e institucionais que privilegiam a análise de contas de governo e de processos de conformidade, em detrimento de avaliações de desempenho e de políticas públicas.

Em síntese, embora a maioria das Cortes de Contas realize fiscalizações envolvendo valoração de políticas públicas, há ampla convergência de que tais práticas, em geral, não configuram avaliação de políticas públicas em sentido estrito, sobretudo à luz da NBASP 9020. Predomina, assim, o entendimento da avaliação como extensão da auditoria operacional, com ênfase em critérios clássicos de desempenho, como economicidade, eficiência e eficácia. Enquanto isso, dimensões substantivas como relevância e utilidade permanecem pouco usadas e institucionalizadas.

Mesmo nos tribunais considerados em estágio inicial da avaliação de políticas públicas, com base nas falas dos participantes do grupo focal, foi possível identificar níveis distintos de maturidade e inovação na realização das avaliações nos Tribunais. Essa diversidade em termos de maturidade e inovação ocorre com ou sem a observância, ainda que parcial, da NBASP 9020.

Com base nos casos descritos, observam-se práticas avaliativas consistentes em quatro Cortes de Contas (TCU, TCE/PE, TCE/PI e TCM/RJ), ainda que não totalmente alinhadas à NBASP 9020, nos quais os participantes dos grupos focais classificaram como ‘intermediária’ a atuação das Cortes. Também se observou a existência de avaliações de políticas públicas ‘pioneiras’, com a observância integral da NBASP 9020 no TCE/ES e no TCE/GO. Adicionalmente, observaram-se inovações metodológicas no TCE/RN e no TCE/SC, que adotaram abordagens bem distintas em relação aos demais Tribunais de Contas, embora também realizem auditorias operacionais para avaliar políticas públicas.

Do ponto de vista organizacional, identifica-se uma tendência à especialização, devido à instituição em 11 Tribunais de setores específicos dedicados à avaliação de políticas públicas. Entretanto, modelos alternativos coexistem, variando desde a execução difusa por diferentes setores da estrutura institucional, a exemplo do TCU (especialização temática), resultando no intercâmbio de experiências entre diferentes unidades técnicas e relevantes avanços nas fiscalizações, até o modelo de supervisão centralizada (por um setor específico) de trabalhos executados por setores diversos do tribunal, adotado pelo TCM/RJ.

Diante desse panorama heterogêneo, a apresentação dos resultados a seguir traz uma análise transversal das percepções coletadas nos grupos focais com auditores dos 33 Tribunais de Contas, bem como análises realizadas por meio do exame documental dos relatórios produzidos pelos Tribunais identificados como mais avançados ou com iniciativas inovadoras, permitindo confrontar o discurso institucional com a materialidade técnica das avaliações.

4.2 Métodos, critérios e usos das avaliações: experiências avançadas e trajetórias emergentes nos TCs

4.2.1 Métodos: instrumentos predominantes, adesão à NBASP 9020 e inovação metodológica

No que se refere à metodologia (o ‘como fazer’) utilizada pelos Tribunais de Contas nas fiscalizações para avaliar políticas públicas, o que inclui projetos, métodos, abordagens e técnicas para conduzir as avaliações (Lemire et al., 2020; Montrosse-Moorhead et al., 2024), a análise dos dados coletados nos grupos focais revela que a Auditoria Operacional (ANOP) é o principal instrumento de fiscalização utilizado. No entanto, os tribunais também utilizam outros meios, como levantamentos, acompanhamentos e monitoramentos. Alguns tribunais implementaram novos instrumentos ou metodologias de fiscalização, como é o caso do TCE/ES, TCE/GO, TCE/RN e TCE/SC.

Nessa linha, participantes do grupo focal de 12 Tribunais (36,4%) consideram a ANOP suficiente para avaliar políticas públicas. Nos demais 18 tribunais (54,5%), foi reconhecida sua adequação, mas desde que com adaptações metodológicas, complementações ou uso de instrumentos adicionais. Participantes do grupo focal de apenas um tribunal informaram não possuir, naquele momento, condições de opinar sobre a suficiência da ANOP, por se tratar da primeira fiscalização realizada. Em outro tribunal, a percepção do grupo focal foi de ausência de interesse institucional na avaliação de políticas públicas. E em um terceiro tribunal, a argumentação foi de que efetivamente ‘não avaliam, mas fiscalizam’ políticas públicas, embora os participantes também considerem a ANOP inadequada para este fim.

Instados a se manifestarem quanto à utilização da NBASP 9020, os participantes dos grupos focais foram unânimes ao reconhecer sua importância no âmbito dos Tribunais de Contas. Contudo, em 28 (84,8%) os participantes dos grupos focais afirmaram que ainda não seguem os preceitos do normativo nas fiscalizações, embora seu uso se encontre em estudo e debate em cinco destes tribunais. Participantes do grupo focal de outros três tribunais declararam que usam alguns conceitos ou que existem ações individuais de auditores neste sentido. Na Tabela 3 encontra-se detalhado o panorama dos Tribunais de Contas quanto ao uso da NBASP 9020.

Tabela 3

Proporção de Tribunais de Contas que usam a NBASP 9020, segundo os participantes do grupo focal.

Uso da NBASP 9020	% TCs
Não implementaram	60,6%
Implementaram, ainda que parcialmente	15,2%
Não implementaram, mas está em estudo e discussão	15,1%
Não implementaram, mas usam alguns conceitos	9,1%

Nota: elaborado pela autora.

Dentre as metodologias inovadoras, distintas da ANOP, incorporadas e adotadas por alguns Tribunais de Contas, apontadas pelos participantes do grupo focal, foram identificadas as seguintes: a) abordagem avaliativa da Teoria da Mudança, por meio de avaliação executiva (TCE/ES); b) uso do Modelo Lógico de Programa, com avaliação da existência do ciclo lógico no desenho da política pública (TCE/GO); c) adoção de metodologia do Banco Mundial para elaboração de levantamentos econométricos (TCE/SC); d) Síntese de Evidências (TCE/RN); e e) Análises com utilização de *big data* (TCDF, TCE/MG, TCE/RS e TCE/PI).

A análise das práticas nos Tribunais de Contas, conforme os participantes dos grupos focais, revela a emergência de trajetórias inovadoras na avaliação de políticas públicas, caracterizadas pela adoção de referências metodológicas robustas e pela celebração de parcerias institucionais estratégicas. Embora a Auditoria Operacional permaneça como instrumento predominante, observa-se, em Cortes de Contas específicas, um movimento de sofisticação técnica alinhado a padrões internacionais.

No âmbito das avaliações de desenho e implementação, destacam-se as experiências do TCE/ES e do TCE/GO, ambas alinhadas à NBASP 9020 e suportadas por parcerias com a academia. O TCE/ES realizou sua avaliação piloto sobre a Política Pública Estadual de Cofinanciamento dos Serviços Especializados em Reabilitação (SERDIA), adotando a Teoria da Mudança e a modalidade de Avaliação Executiva. A escolha, fundamentada na celeridade e no potencial de indução de melhorias em curto prazo, contou com a parceria técnica da Fundação Getúlio Vargas (FGV/EESP CLEAR). Em movimento análogo, o TCE/GO avaliou a Política de Atenção à Saúde Materno Infantil utilizando a metodologia do Modelo Lógico. O trabalho amparou-se nos guias práticos publicados pelo IPEA e no Referencial de Controle de Políticas Públicas do TCU, além de contar com suporte de especialistas da Universidade Federal de Goiás (UFG).

Em uma vertente metodológica distinta, focada na inferência causal e na análise quantitativa, verificou-se que o TCE/SC desponta como caso singular na realização sistemática de avaliações de impacto (efetividade) por meio de levantamentos econométricos, seguindo as diretrizes do *Impact Evaluation in Practice*, do Banco Mundial (2016). A viabilidade dessas avaliações, segundo os participantes do grupo focal, decorre diretamente do perfil de profissionais em seu quadro técnico, com formação acadêmica em economia e econometria.

Inovando no suporte à decisão, o TCE/RN institucionalizou a Síntese de Evidências, processo de análise sistemática voltado à identificação das melhores evidências disponíveis (literatura e boas práticas) para subsidiar decisões dos gestores públicos. O primeiro trabalho concluído sob essa metodologia no exercício de 2025 analisou a oferta de vagas em creches municipais. Transversalmente a essas experiências, nota-se a crescente incorporação de análises de *Big Data* e uso de Inteligência Artificial pelas Cortes, como, por exemplo, TCDF, TCE/MG, TCE/RS e TCE/PI em suas fiscalizações, visando à identificação de padrões, irregularidades e geração de *insights* considerados ‘valiosos’ para orientar os trabalhos.

Também se observou uma forte influência normativa do TCU em relação às metodologias avaliativas. Conforme relatos, observa-se um isomorfismo normativo, uma vez

que auditores participantes do grupo focal de 78,8% das Cortes indicaram o uso de manuais do TCU, especialmente o Referencial de Controle de Políticas Públicas, como guias subsidiários no planejamento de suas ações.

Apesar dos avanços, a institucionalização da avaliação *stricto sensu* colide com barreiras estruturais. Na Tabela 4 são apresentadas as principais limitações metodológicas dos tribunais, segundo os participantes do grupo focal. A principal restrição, apontada em 69,7% dos tribunais, reside na dificuldade de avaliar a efetividade (impacto). Esse gargalo, segundo os participantes do grupo focal, encontra-se intrinsecamente ligado à escassez ou baixa qualidade dos dados governamentais disponíveis e à ausência de indicadores adequados. O cenário é agravado pela carência de normativos internos específicos e pela incompatibilidade entre o tempo exigido para fiscalizações complexas e os prazos processuais vigentes.

Tabela 4

Limitações metodológicas nos TCs, segundo os participantes do grupo focal.

Limitações metodológicas	% TCs
Dificuldade em avaliar impacto (efetividade)	69,7%
Dificuldade de acesso ou baixa qualidade dos dados disponíveis	57,6%
Ausência de indicadores adequados de resultado e de impacto	21,2%
Baixa qualidade do desenho das políticas públicas	18,2%
Ausência de normativos específicos orientando a realização das avaliações	27,3%
Insuficiência de tempo ou prazos incompatíveis para a realização das avaliações, que são mais complexas que as ANOP	27,3%
Limitações tecnológicas como deficiência de infraestrutura de TIC, de ferramentas de análise e integração de dados	12,1%
Incertezas quanto ao papel do Tribunal na avaliação, especialmente sobre os limites entre avaliação e auditoria	12,1%

Nota: elaborado pela autora.

A superação dos obstáculos demanda, segundo a percepção dos auditores de 66,6% dos tribunais, a priorização institucional da avaliação de políticas públicas como estratégia de legitimação do Controle Externo. Na maioria dos tribunais (81,8%) defende-se a delimitação formal do *modus operandi* avaliativo, mediante o estabelecimento de diretrizes metodológicas claras que evitem o engessamento da fiscalização. Em 33% dos tribunais foi apontada a celebração de parcerias com universidades e centros de pesquisa como vetor de qualificação técnica. Complementarmente, participantes do grupo focal sugerem o investimento contínuo em ferramentas de TI e Inteligência Artificial (em 45,4% dos tribunais), o foco inicial na

avaliação do desenho e implementação, dimensão da relevância (em 15,1% dos tribunais), e a introdução de práticas de meta-avaliação (em 9,1% dos tribunais).

4.2.2 Valoração: critérios usados, análise de contexto e valores emergentes

A valoração das políticas públicas pelos Tribunais de Contas brasileiros decorre do exame do caso concreto, condicionado ao objeto e ao escopo de cada fiscalização. Os critérios de avaliação são formalizados nas Matrizes de Planejamento e fundamentam-se em parâmetros normativos, princípios constitucionais, e guias de boas práticas, dentre outros, visando conferir suporte técnico e legal às deliberações emanadas pelas Cortes de Contas. Em consonância com a metodologia de Auditoria Operacional (ANOP), o julgamento avaliativo concentra-se nas dimensões da economicidade, eficiência e eficácia. A análise da efetividade (impacto), embora reconhecida como prioritária, permanece um desafio metodológico para a maioria das Cortes, excepcionalizado o TCE/SC, devido à realização de levantamentos econométricos.

A comparação entre as falas e os documentos examinados sugere um descompasso entre a prática vigente e as diretrizes da NBASP 9020 no que tange às dimensões substantivas da avaliação. A análise da dimensão da relevância com a observância do normativo ainda é incipiente, tendo sido iniciada, segundo a percepção dos participantes dos grupos focais, efetivamente em apenas 18,2% dos tribunais. Não se identificou a incorporação sistemática da dimensão da ‘utilidade’ por nenhuma Corte, muito devido a suscitadas limitações metodológicas para sua aferição, semelhantes às aplicáveis para avaliação da dimensão da efetividade. Apesar disso, os participantes do grupo focal do TCE/ES e TCE/GO declararam que a avaliação da ‘utilidade’ das políticas públicas constitui uma meta, um segundo passo após a consolidação da avaliação da relevância, iniciada nas mencionadas avaliações ‘pilotos’.

Paralelamente, nota-se a emergência de novos vetores axiológicos, visto que entrevistados em 36,3% dos tribunais declararam a incorporação de valores como justiça social, equidade, sustentabilidade e satisfação dos beneficiários como parâmetros orientadores das avaliações de políticas públicas. Além disso, relataram o uso de critérios baseados em informações (declaração de objetivos, metas e prioridades) constantes dos principais instrumentos orçamentários: PPA, LDO e LOA.

Constatou-se que participantes do grupo focal em 81,8% dos tribunais reconhecem a relevância da qualidade da análise de contexto (ideacional, institucional e política) para um diagnóstico preciso e para a compreensão dos valores que orientam a política pública, todavia, sua execução carece de sistematização. De forma geral, a análise de contexto institucional e política costuma ser realizada, ainda que nem sempre seja objeto de registro nos relatórios de

fiscalização. Participantes do grupo focal de apenas 39,4% dos tribunais declararam que realizam a análise de contexto ideacional de forma sistemática, embora também nem sempre a documentem nos relatórios ou nos demais registros relacionados às fiscalizações (papéis de trabalho). A Tabela 5 traz o panorama dos TC em relação à análise de contexto ideacional.

Tabela 5

Panorama da realização da análise de contexto ideacional nos TCs.

Análise de contexto ideacional (valores que orientaram a PP)	% TCs
Realizada de forma sistemática	39,4%
Realizada de forma difusa, em menor grau ou foi iniciada	36,3%
Não é realizada	24,2%

Nota: elaborado pela autora.

A ausência de registro formal da análise de contexto nos relatórios ou papéis de trabalho, verificada na maioria dos casos, tende a gerar implicações negativas, na medida que pode comprometer a transparência do processo avaliativo, dificultar a verificação da profundidade analítica e, ainda, fragilizar a fundamentação dos achados (constatações significativas, positivas ou negativas). Isso ocorre porque o avaliador pode não captar adequadamente as disputas e valores subjacentes à formulação da política. Ademais, a concentração dessa análise apenas na fase de planejamento, realidade declarada por participantes dos grupos focais em 45,1% dos tribunais, tem o potencial de limitar a capacidade do auditor de perceber mudanças contextuais durante a avaliação da execução da política.

A consolidação de uma prática avaliativa substantiva esbarra em barreiras estruturais e culturais significativas, como mostra a Tabela 6. A principal limitação, apontada por participantes dos grupos focais em 69,7% dos tribunais, é a ausência de uma cultura e consciência institucional voltada para a avaliação de políticas públicas *stricto sensu*. Segundo depoimento de participantes dos grupos focais em 57,6% dos tribunais, predomina uma cultura organizacional de conformidade e responsabilização, que por vezes drena recursos e prioriza a sanção em detrimento da melhoria da gestão. Essa lógica reflete-se na escassez crítica de pessoal, com casos extremos de tribunais operando com quadros mínimos para todo o estado, e em sistemas de produtividade que, ao privilegiarem a celeridade processual, desincentivam a realização de avaliações complexas e de longo curso.

Tabela 6

Limitações na valoração de avaliações de políticas públicas nos TCs.

Limitações na valoração de PP	% TCs
Ausência de uma cultura e consciência cultural de avaliação de PP	69,7%
Foco em conformidade e na responsabilização	57,6%
Insuficiência de pessoal	54,5%
Dificuldades para avaliar produtividade dos ACE e dos TCs de forma adequada	21,2%

Nota: elaborado pela autora.

Para superar as limitações apresentadas na Tabela 6, e contribuir para o aprimoramento da emissão de juízos de valor (valoração), os caminhos indicados pelos participantes dos grupos focais convergem quase unisonamente para a educação corporativa. Em 78,8% dos tribunais, foi apontada a implementação de programas de capacitação específica como a estratégia central. Tais ações formativas são percebidas não apenas como meio de transferência técnica, mas como vetores essenciais para a formação de uma nova consciência cultural, capaz de ampliar a realização de avaliações de políticas públicas como estratégia para a legitimação das Cortes de Contas e para viabilizar maiores entregas efetivas de valor público à sociedade.

4.2.3 Uso: estratégias de indução, consensualismo e gargalos do monitoramento

Os grupos focais convergiram no sentido de que a finalidade última da avaliação de políticas públicas reside na promoção da melhoria da gestão e no atendimento efetivo das demandas sociais prioritárias. Embora os tipos clássicos de utilização (instrumental, conceitual e simbólico) sejam identificáveis dependendo do caso concreto, a pesquisa revela que o esforço institucional das Cortes se concentra na criação de mecanismos para induzir a apropriação dos resultados pelos gestores.

Nos relatos, destacam-se estratégias de natureza dialógica e pedagógica. A realização de reuniões com gestores para divulgação, discussão ou mediação de soluções destaca-se como a prática mais disseminada, adotada em 66,6% dos tribunais. Em segundo plano, mas com relevância expressiva (citado em 51,5% dos tribunais), situa-se o monitoramento tempestivo do cumprimento das decisões, evidenciando uma dualidade na atuação do Controle Externo, que oscila entre a orientação e a verificação do *enforcement*. Outras estratégias de fomento incluem a divulgação midiática e a realização de reuniões de validação dos achados, além da capacitação

de gestores e publicização das fiscalizações em sites institucionais, conforme demonstrado na Tabela 7 a seguir.

Tabela 7

Ações adotadas nos para fomentar o uso das avaliações de políticas públicas nos TCs.

Ações adotadas para fomentar o uso das avaliações realizadas	% TCs
Reuniões com os gestores, seja para divulgar, discutir os resultados ou mediar soluções	66,6%
Realização de monitoramentos tempestivos do cumprimento de Decisões da Corte de Contas em processo de avaliação	51,5%
Divulgação dos resultados das fiscalizações na mídia, incluindo as redes sociais	36,3%
Reuniões de encerramento para validação dos resultados antes da conclusão dos relatórios	36,3%
Realização de eventos de capacitação dos gestores	33,3%
Divulgação dos resultados das fiscalizações no site institucional	33,3%
Exigência de um Plano de Ação elaborado pelos gestores detalhando ações e prazos para o atendimento das decisões	24,2%
Participação dos <i>stakeholders</i> (participação social) nas avaliações	21,2%

Nota: elaborado pela autora.

A análise qualitativa das iniciativas, resumida na Tabela 8, permitiu categorizar as estratégias de indução em três eixos principais: (1) articulação dialógica, (2) transparência social e (3) indução por emulação (*benchmarking*) e cooperação. No eixo da articulação dialógica, destacam-se o TCE/GO e o TCM/GO, que institucionalizaram a busca por soluções conjuntas mediante Comitês Consultores e Mesas de Discussão Técnica, respectivamente, integrados por gestores de diferentes órgãos envolvidos na política pública avaliada. Similarmente, a atuação regionalizada do TCE/RS promove maior proximidade com os jurisdicionados, potencializando a implementação das decisões.

Tabela 8

Ações singulares relevante nos TCs para fomentar o uso das avaliações.

Tribunal	Ações adotadas para fomentar o uso das avaliações realizadas
TCE/GO	Divulgação dos resultados das fiscalizações para um Comitê Consultor, formado por diferentes órgãos envolvidos na PP fiscalizada.
TCE/PE	Emissão de Alertas de Responsabilização: na fase do Procedimento Interno de Fiscalização - PIP (atuação prévia, sem atuação de processo específico); Distribuição de Cartilhas para a sociedade divulgando os trabalhos realizados.

TCE/PI	Forte parceria com outros órgãos de controle, especialmente com o Ministério Público, por meio de Acordos de Colaboração; Bom relacionamento com órgãos colegiados (UNDIME, COSEMS, COEGEMAS etc.)
TCE/PR	Ampla divulgação no site institucional dos dados do Programa de Avaliação de Contas Municipais de Governo (ProGov), com o objetivo de incentivar o controle social e induzir a melhoria da gestão a partir da comparação dos resultados entre os Municípios.
TCE/RS	Atuação articulada das unidades centrais e das regionais do Tribunal junto aos jurisdicionados.
TCE/SC	Ampla divulgação dos resultados dos Levantamentos Econométricos.
TCM/GO	Realização de mesas de discussão e construção de soluções para as inconformidades identificadas
TCM/SP	O TCM/SP instituiu em 2021 o Observatório de Políticas Públicas que funciona junto à Escola de Contas-ESCON e conta com a parceria de diferentes instituições acadêmicas, públicas e privadas, e da sociedade civil. Comunicação dos resultados das fiscalizações para a sociedade de forma mais inteligível, com uso de recursos como infográficos e painéis interativos no site institucional.

Nota: elaborado pela autora.

No eixo da transparência social e controle democrático, tribunais como o TCM/SP e o TCE/PE inovam ao traduzir a complexidade técnica para a linguagem cidadã. O primeiro, por meio de seu Observatório de Políticas Públicas e pela comunicação dos resultados das fiscalizações com o uso de infográficos e painéis interativos no site institucional; o segundo, recorrendo à distribuição de cartilhas para a sociedade também em linguagem cidadã com o mesmo objetivo. O TCE/SC segue diretriz análoga ao promover conferências para ampla divulgação de seus levantamentos econométricos. Participantes do grupo focal do TCM/SP e do TCE/GO, por exemplo, destacaram a institucionalização da participação social e técnica por meio, respectivamente, do Observatório de Políticas Públicas e de um Comitê Consultor, formado por diversos órgãos envolvidos com a política pública fiscalizada.

Já no eixo de indução por cooperação, o TCE/PI tem atuado em forte parceria com outros órgãos de controle, como o Ministério Público Estadual, e o por meio de parcerias estratégicas com conselhos e órgãos deliberativos (UNDIME, COSEMS, COEGEMAS etc.) para induzir o uso de suas avaliações de políticas públicas pelos gestores estaduais e municipais.

Na indução por emulação (*benchmarking*) destaca-se a estratégia adotada no TCE/PR para a modernização da fiscalização e o aprimoramento da qualidade das políticas públicas, embora não decorrente de ANOP. A estratégia consiste na divulgação, em painel disponibilizado no portal institucional do Tribunal, dos índices do Programa de Avaliação de Contas Municipais do Governo (ProGov), gerados a partir da análise de dados autodeclaratórios das prefeituras do estado para consulta pública. Segundo os participantes do grupo focal, os

resultados positivos observados decorrem do interesse dos gestores o desempenho nos indicadores avaliados, bem como sua posição no ‘*ranking*’ dos municípios paranaenses.

Instados a se manifestarem sobre eventuais limitações, apenas uma dificuldade com representatividade significativa foi indicada por participantes dos grupos focais em 18,2% dos tribunais: a ausência de monitoramentos tempestivos. A intempestividade no acompanhamento das deliberações é percebida como fator crítico para o uso instrumental das avaliações; na sua falta, o potencial de correção de rumos se dilui.

Diante desse cenário, observa-se uma mudança de paradigma em direção ao consensualismo. Apontada como tendência em 30,3% dos tribunais, a institucionalização de práticas consensuais, materializada em Mesas Técnicas e Termos de Ajustamento de Gestão (TAGs), revela-se como instrumento potente para superar a lógica adversarial. Para os participantes do grupo focal, a otimização do uso das avaliações depende desse tríplice movimento: (a) implementação e/ou consolidação do consensualismo; (b) ampliação da participação dos gestores no desenho da avaliação; e (c) investimento contínuo na capacitação dos gestores envolvidos no planejamento, execução, avaliação e retroalimentação da política pública.

4.3 Padrões e variações institucionais: proposição de tipologia das trajetórias avaliativas

A análise integrada das evidências empíricas, oriundas do exame das percepções dos participantes dos grupos focais e do material dos relatórios de fiscalização, revela um cenário de acentuada heterogeneidade no panorama do Controle Externo brasileiro. Diante disso, apresenta-se uma tipologia de Tribunais de Contas baseada nas distintas trajetórias evolutivas identificadas: (1) tribunais incipientes, marcados pela predominância do uso de ANOP, com observância de sua metodologia própria; (2) tribunais incrementais, que reconhecem o potencial do uso da ANOP nas avaliações, desde que condicionada à realização de adaptações; (3) tribunais normativos-estruturantes, caracterizados pela adesão plena às diretrizes da NBASP 9020; e (4) tribunais disruptivos, que introduziram inovações metodológicas singulares baseadas em ciência de dados e inferência causal.

Importa ressaltar que a tipologia proposta para os Tribunais de Contas (incipientes, incrementais, normativo-estruturantes e disruptivos/científicos) não deve ser interpretada como uma hierarquia de mérito institucional, nem como uma proposta de ‘evolução linear’ com um ‘padrão-ouro’ a ser alcançada. Trata-se, antes, de uma classificação descritiva de arquiteturas institucionais e escolhas avaliativas identificadas empiricamente, que coexistem no sistema de

controle externo e respondem a combinações distintas de capacidades avaliativas, ambiente e cultura institucionais, bem como escolhas estratégicas de governança.

Nessa perspectiva, as trajetórias incremental, normativo-estruturante e disruptiva/científica não estão condicionadas a uma progressão sequencial: podem ser acionadas a partir de diferentes pontos de partida, combinadas em uma mesma Corte e mobilizadas seletivamente conforme o propósito avaliativo e o contexto analisado. Essa leitura reforça o caráter não normativo da tipologia e sustenta a tese central de que inexistem um ‘padrão-ouro’ para avaliações de políticas públicas.

4.3.1 Tribunais incipientes

O primeiro grupo, denominado de ‘incipiente’, é formado pelos 25 Tribunais de Contas em que os participantes do grupo focal classificaram a atuação da Corte como ‘inicial’ em avaliação de políticas públicas. As percepções dos participantes dos grupos focais revelaram a predominância nestas Cortes de fiscalizações com foco no exame de conformidade legal, financeira e orçamentária, e na análise das contas de governo, setores nos quais se encontra alocada a maior parte da força de trabalho.

Nesse grupo de tribunais, as avaliações de políticas públicas são realizadas por meio de ANOP, com foco no exame das dimensões de economicidade, eficiência e eficácia, sem adentrar na efetividade, dadas as limitações já indicadas, nem no exame da relevância ou da utilidade das políticas públicas, consoante dispõe a NBASP 9020. Embora reconheçam a importância e significância da avaliação de políticas públicas, destacam a falta de instrumentalização técnica para de fato realizá-las.

Participantes dos grupos focais de alguns tribunais destacaram a percepção quanto à ausência de priorização na realização de ANOP de forma geral, considerada a reduzida alocação de auditores, especialmente com o objetivo de obter um resultado positivo na avaliação de desempenho levada a efeito pela Associação de Tribunais de Contas do Brasil (ATRICON), por meio do Marco de Medição de Desempenho dos Tribunais de Contas (MMD-TC), que afinal se destaca como um agente indutor de excelência nas fiscalizações realizadas no âmbito Controle Externo.

Em apenas um tribunal o grupo focal apontou que, embora o MMD-TC assegure a realização de fiscalizações em temas relevantes, pode também limitar a atuação das Cortes no caso da definição de temáticas não prioritárias para a sociedade de determinada unidade da federação. Situação que poderia ser mitigada, segundo percepção do grupo focal, mediante a

ampliação da participação dos auditores alocados na área fim de cada Tribunal de Contas nos fóruns de discussão e planejamento do MMD-TC, e a possibilidade de ‘adequações’ das fiscalizações de acordo com a realidade e das demandas prioritárias de cada sociedade ‘local’.

Um ponto amplamente discutido no âmbito dos tribunais do grupo ‘incipiente’ consiste na percepção dos participantes de que a especialização temática (e.g. saúde ou educação) tem o potencial de ampliar a qualidade das fiscalizações de forma geral, incluindo as avaliações de políticas públicas, uma vez que o acúmulo de conhecimento e experiências em áreas específicas tem o potencial de contribuir para avaliações cada vez mais densas e profundas. Nessa mesma direção, também foi reconhecida como relevante a criação de setores específicos para a avaliação de políticas públicas: tanto para assegurar a realização efetiva de avaliações, quanto para o fomento de novas metodologias, critérios e estratégias para induzir os gestores a se apropriarem de fato dos produtos dos trabalhos realizados (uso das avaliações dos TCs).

Outro ponto recorrentemente suscitado nos grupos focais refere-se aos questionamentos quanto aos encaminhamentos das avaliações. Há um consenso quanto às formas usuais de deliberações, determinações e recomendações não serem as mais apropriadas, ainda que constituam alternativas possíveis em determinadas situações, como na ‘ausência’ deliberada de ações ‘saneadoras’ por parte dos gestores em face dos ‘achados’ das avaliações. A percepção dos auditores, conforme já mencionado, é que a adoção de estratégias de natureza dialógica e pedagógica têm o potencial de ampliar de forma significativa o efetivo uso das avaliações pelos gestores públicos.

4.3.2 Tribunais incrementais

O grupo ‘incremental’ reúne Tribunais que mantêm a Auditoria Operacional (ANOP) como plataforma central de trabalho, mas ampliam progressivamente o repertório metodológico, os critérios de valoração e as estratégias de uso dos achados para além do tripé clássico de desempenho (economicidade, eficácia e eficiência); fazem-no por meio de planejamentos mais densos (matrizes e levantamentos), interação continuada com gestores e dispositivos de engajamento (painéis, reuniões de encerramento, monitoramentos, consensualismo), aproximando-se de uma avaliação *stricto sensu* mesmo sem incorporar de forma plena a NBASP 9020, nem reconhecer o efetivo exame das dimensões de ‘relevância’ e ‘utilidade’ das políticas públicas.

Em relação aos arranjos institucionais e processos, no TCU a estrutura temática da Segecex viabiliza a difusão das avaliações entre unidades, com autorização prévia do relator,

encaminhamento de Relatório Prévio aos gestores para manifestação e envio do Relatório de Fiscalização (RFA) ao Ministro Relator designado, nos termos da Resolução 315/2020; há validação do planejamento via painéis de referência e consolidação anual no Relatório de Fiscalizações em Políticas e Programas de Governo (RePP).

No TCE/PE, a Diretoria de Controle Externo opera por departamentos temáticos, com Plano Bienal de Fiscalização e um Procedimento Interno Prévio (PIP): levantamento célere que pode resultar em auditoria, inspeção, alerta de responsabilização ou arquivamento, antes da autuação de processo formal.

No TCE/PI, a criação da Diretoria de Fiscalização de Políticas Públicas em 2019 e a especialização temática na estrutura do Corpo Técnico do Tribunal deslocaram o foco do controle *a posteriori* para a atuação concomitante, com reuniões desde o planejamento e reunião de encerramento obrigatória; o Relatório Final segue para deliberação Plenária após contraditório.

No TCM/RJ, existe desde 2021 uma Coordenação de Avaliação de Políticas Públicas que acompanha o planejamento (matrizes), apoia quando demandada e revisa relatórios, enquanto todas as Inspeções realizam avaliações (arranjo difuso por jurisdicionado); a análise de contexto (CIPI) tornou-se obrigatória no planejamento por memorando e está em vias de normatização interna.

Em termos metodológicos, os quatro Tribunais ancoram-se na ANOP e, conforme o objeto, incorporam ferramentas avaliativas: no TCU, Teoria da Mudança e Modelo Lógico são acionados para estruturar perguntas e critérios, com acompanhamentos e monitoramentos usados para dar oportunidade a correções durante a implementação; há investimento em TIC/IA; uso de *Objectives and Key Results* (OKRs); e aplicação de questionários a gestores (IESGO) para robustecer governança e aprendizado. No TCE/PE, o PIP opera como instrumento indutivo e célere, uma ‘porta’ que abre trilhas diferenciadas (auditoria/inspeção/alerta/arquivamento), e a parceria com a UFPE busca formatar um modelo e construir indicadores alinhados ao uso decisório; a divulgação ocorre em plataformas como o ‘Tome Conta’.

O TCE/PI combina ANOP com densidade analítica (satisfação do usuário, recortes de equidade quando pertinentes), cruzamentos de dados (e.g. CADÚnico) e levantamentos para mapear políticas e orientar a seleção do objeto, com a Resolução nº 32/2022, os Manuais do IPEA (*ex ante/ex post*) e o Referencial de Avaliação de PP do TCU como balizas.

No TCM/RJ, as avaliações, historicamente por ANOP e acompanhamentos, concentram-se no desenho e implementação (resultados imediatos/eficácia), incluindo consultas a especialistas (e.g. política de saúde mental), uso de materiais do IPEA e do TCU e referências OCDE (via Rede Integrar); além de experiências setoriais com RAPPAM e IndiMapa em fiscalizações ambientais.

Predomina, nos quatro casos, o julgamento por economicidade, eficiência e eficácia, porém com alargamento para dimensões transversais. No TCU, há atenção explícita a equidade, governança e sustentabilidade, além de esforço para analisar pertinência/relevância do desenho das políticas públicas, quando viável, reconhecendo que lacunas de indicadores e de arquitetura dificultam medir efetividade/impacto.

No TCE/PE, ampliam-se recortes de qualidade do serviço, governança e, em alguns objetos, justiça social/equidade, embora relevância e utilidade ainda não estejam sistematizadas como dimensões avaliativas institucionais. No TCE/PI, a visão finalística (resultados para o público-alvo) e a satisfação do usuário tornam-se lentes frequentes, com equidade/justiça social aparecendo em determinados casos; relevância/utilidade permanecem em construção.

No TCM/RJ, a valoração se centra no desenho (existência/suficiência) e na adequação aos planos e normas programáticas (PPA, planos setoriais), com ênfase em induzir gestores a justificarem escolhas e qualificar o planejamento, um passo que tangencia a análise de pertinência do desenho às necessidades identificadas.

Quanto às estratégias para o uso das avaliações, o TCU investe em monitoramentos e acompanhamentos continuados, painéis de referência, oitiva prévia (Res. 315/2020), RePP como vitrine anual e uso do IESGO para induzir respostas gerenciais, um conjunto que combina uso instrumental (mudanças concretas) e uso processual (aprendizagem/ajuste em curso). O TCE/PE emite alertas por meio dos PIPs, investe em TAGs/planos de ação, palestras e painéis com gestores e divulgação pública (Tome Conta) como estratégia dialógica e indutiva para acelerar correções e apropriação dos achados.

O TCE/PI formaliza planos de ação, realiza reunião de encerramento obrigatória, mobiliza redes setoriais (UNDIME, COSEMS, COEGEMAS) para difundir boas práticas e aciona cooperações com o Ministério Público para ampliar o alcance das recomendações. E no TCM/RJ, a Coordenação atua como instância de qualidade (revisão das matrizes e relatórios), há envolvimento de especialistas (e.g. saúde mental) e crescimento da exigência de planejamento robusto (análise CIPI obrigatória) com vistas a elevar a credibilidade científica e, por consequência, o uso das avaliações realizadas.

Os quatro Tribunais convergem quanto a limitações de dados (qualidade, disponibilidade e integração), deficiências no desenho das políticas (que inviabilizam análises de efetividade/impacto) e necessidades de padronização e capacitação (inclusive em dados/IA). Desafios específicos incluem: no TCU, infraestrutura para dados sensíveis e regionalização de objetos sociais; no TCE/PE, padronização metodológica, priorização institucional e bases de dados confiáveis; no TCE/PI, governança multinível, intersetorialidade e avanço para impacto; no TCM/RJ, dissipação do viés de conformidade, delimitação de papéis (controle × gestão) e capacidade para análises quantitativas mais sofisticadas (estatística/econometria).

Assim, pode-se apresentar, em linhas gerais, as seguintes trajetórias dos Tribunais de Contas do grupo incremental: i) TCU: “portfólio com monitoramento”: consolidação anual (RePP), acompanhamentos e painéis para calibrar desenho/implementação e estimular equidade e governança nos objetos; ii) TCE/PE: “celeridade indutiva”: PIP como porta de entrada, consensualismo (TAGs) e divulgação pública para acelerar a apropriação e a correção de falhas; (iii) TCE/PI: “concomitância e redes”: especialização temática, interação obrigatória com gestores, planos de ação e difusão por colegiados para ampliar alcance; e (iv) TCM/RJ: “coordenação revisora e densificação”: coordenação que supervisiona planejamento e relatórios, CIPI obrigatório e consulta a especialistas em objetos complexos (e.g. saúde mental, PSR, hospital).

4.3.3 Tribunais normativos-estruturantes

No grupo ‘normativos-estruturantes’ encontram-se dois Tribunais de Contas que vêm institucionalizando arranjos organizacionais, normativos e procedimentais específicos para avaliações de políticas públicas e internalizando diretrizes contemporâneas de avaliação (métodos, valoração e uso) de forma progressiva e sistemática: TCE/ES e TCE/GO. O grupo se distingue por três características combinadas: (i) existência de setores específicos para realizar as avaliações; (ii) mecanismos normativos e operacionais: normas e instituição de ‘processos’ ou ‘sistemáticas’ próprias; e (iii) estratégias de uso que transcendem a lógica sancionatória, priorizando aprendizagem institucional, indução de melhorias e diálogo com *stakeholders*. Tais características reforçam a competência dos Tribunais de Contas para emitir juízos de valor (valoração) e propor encaminhamentos com base em evidências, atributo que diferencia e qualifica o Controle Externo.

No TCE/GO, a criação do Serviço de Inovação de Avaliação de Políticas Públicas (Resolução Administrativa nº 19/2022), a edição da Resolução Normativa nº 02/2024 (que

estabelece diretrizes e procedimentos de APP) e o lançamento do Observatório de Políticas Públicas (2024) compõem um pacote organizacional-normativo que formaliza o instrumento ‘Avaliação de Política Pública’, define governança do processo e cria capacidade analítica para diagnóstico e monitoramento de intervenções de saúde, educação e segurança, com previsão de ampliação para outras áreas. No Espírito Santo, a SECEX Social desponta no âmbito das avaliações, articulando piloto metodológico com assessoria acadêmica (FGV - CLEAR) e processo de implementação da NBASP 9020, o que sinaliza um movimento deliberado de convergência entre prática de controle e referenciais contemporâneos de avaliação.

No plano metodológico, esses Tribunais usam abordagens que já superam a matriz clássica dos ‘3 Es’ (economicidade, eficiência e eficácia) ao incorporar modelos lógicos, diagnóstico de problemas, árvore de problemas e Teoria da Mudança, nessa primeira fase no exame do desenho das políticas públicas (*ex ante*), quando políticas estão em (re)formulação.

O TCE-GO, por exemplo, estrutura avaliações com Diagnóstico Situacional e Modelo Lógico para verificar a coerência do desenho, a identificação adequada do problema público e o alinhamento entre demandas sociais e instrumentos implementados, o que incluiu na avaliação ‘piloto’ o exame do desenho da política pública em saúde materno-infantil com participação de comitê consultivo, distinção de papéis decisórios e ênfase no critério de relevância previstos na NBASP 9020.

Já o TCE-ES combinou na sua primeira avaliação *stricto sensu* uma Auditoria Operacional com um Relatório de Avaliação como produto específico, adotando a Avaliação Executiva (mais célere, orientada à tomada de decisão) e a Teoria da Mudança para explicitar pressupostos, riscos e cadeias causais, com vistas a preparar terreno para futuras avaliações de impacto quando houver dados e maturidade de implementação.

Em termos de valoração, identifica-se ampliação dos critérios de julgamento: além de desempenho, emergem dimensões como equidade e sustentabilidade da intervenção. No TCE-ES, a parceria com a FGV CLEAR conduz à declaração de riscos e à explicitação de pressupostos causais, enquanto no TCE-GO o comitê consultivo aporta diversidade de perspectivas e validações interativas do planejamento e de versões preliminares do relatório, preservada a independência do Tribunal na formação de juízo técnico. Esse deslocamento valorativo está alinhado aos marcos conceituais de avaliação (como os guias CLEAR/FGV), que privilegiam o uso qualificado de evidências e a adequação metodológica ao propósito decisório, elementos que vêm sendo apropriados pelos TCs em suas rotinas.

O uso dos resultados é induzido por estratégias de engajamento e arranjos de governança específicos. No TCE/GO, destaca-se a proposta de Mesa Técnica para pactuar planos de ação com gestores (consensualismo) e o uso do Observatório de Políticas Públicas para dar publicidade e perenidade às séries de indicadores e diagnósticos; há, ainda, a previsão de monitoramento das recomendações e a intenção de institucionalizar um Sistema Permanente de M&A no âmbito do Estado, envolvendo diferentes poderes. No TCE/ES, a preocupação explícita com acessibilidade e divulgação dos produtos e a referência a experiências do TCU na síntese anual das avaliações (REPP) sinalizam caminhos para interlocução com o Legislativo e maior impacto orçamentário-programático, em linha com boas práticas já consolidadas no plano federal.

Do ponto de vista procedimental, esses Tribunais avançam na padronização e na qualidade dos processos. O TCE-GO combina NBASP 9020, NBASP 100/300 e manuais IPEA (*ex ante/ex post*) com o Referencial de Avaliação de Políticas Públicas do TCU, produzindo Procedimento Operacional (PO) específico para avaliação, com ciclo PDCA para revisão contínua, movimento sustentado por um sistema de certificações ISO (9001, 14001, 37001, 37301) e por um catálogo de POPs para instrumentos de fiscalização. O TCE-ES, por sua vez, ancora suas práticas especialmente na NBASP 9020 e no referencial de Avaliação de Políticas Públicas do TCU, além de seguir a metodologia do Guia de Avaliação de Políticas Públicas do Clear Brasil (FGV).

Essas escolhas normativas e estruturantes não eliminam desafios. Persistem limitações de dados para avaliações de impacto e dificuldades na mensuração orçamentária do gasto por política em contexto subnacional, o que afeta análises de custo-efetividade e utilidade; há, ainda, o desafio cultural de consolidar uma visão de avaliação distinta (e complementar) à auditoria operacional, com prazos e técnicas próprios e com coprodução de informação junto a gestores e usuários. Em resposta, esses Tribunais têm apostado em avaliações *ex ante* (quando políticas estão em fase de (re)desenho, avaliações executivas (para ganho de oportunidade), e modelos lógicos como passos intermediários, além de parcerias acadêmicas e de observatórios como infraestrutura de dados públicos.

De forma geral, o padrão ‘normativo-estruturante’ configura-se como uma trajetória de inovação incremental com base institucional: fortalece governança e qualidade metodológica das avaliações; amplia o escopo valorativo (relevância, com pretensão de avançar futuramente para a dimensão da ‘utilidade’); e cria mecanismos de uso (planos de ação pactuados, monitoramento e interfaces com observatórios e Poder Legislativo). Ao fazê-lo, esses Tribunais

aprofundam a capacidade de emitir juízos de valor baseados em evidências e de induzir aperfeiçoamentos em políticas públicas, em consonância com o papel constitucional do Controle Externo e com a literatura que situa os Tribunais de Contas como atores estratégicos na entrega de valor público.

4.3.4 Tribunais disruptivos

Neste agrupamento, situam-se Tribunais de Contas que desenvolvem avaliações ‘singulares’, consideradas as demais Cortes de Contas, ao realizarem trabalhos com o uso de distintos dispositivos metodológicos de fronteira. Nesse grupo, duas vertentes metodológicas distintas foram identificadas, com efeitos distintos sobre a dimensão da valoração.

No recorte analisado, o caráter ‘disruptivo’ decorre menos de um arcabouço normativo consolidado e mais de trajetórias experimentais e tecnicamente sofisticadas, ainda em processo de institucionalização, mas com forte potencial indutivo sobre a formulação e o aprimoramento de políticas públicas.

Do ponto de vista institucional, os dois casos ilustram o padrão da especialização. No TCE/SC, as avaliações de políticas públicas se concentram na Diretoria de Atividades Especiais (DAE), que, além de auditorias operacionais e levantamentos, desenvolve ‘Levantamentos Econométricos’ voltados à efetividade e ao impacto de políticas, com uma equipe especializada (dois auditores e um assessor técnico, com formação em economia/econometria), e fluxo decisório que envolve autorização da direção.

Já no TCE/RN, houve reorganização da estrutura do Controle Externo com a criação, em jan/2025, da Diretoria de Avaliação de Políticas Públicas (DPP), com foco em políticas sociais, coexistindo com a Diretoria de Infraestrutura e Meio Ambiente (DIA), que combina auditorias operacionais e levantamentos (e.g. em agendas climáticas coordenadas com o TCU). O TCE/RN aprovou normas de qualidade e ciclo de auditoria (SisDGF; controle de qualidade por amostragem), aderiu às NBASPs, e passou a instituir consensualismo como instrumento de indução de melhorias, o que compõe um ambiente organizacional propício à inovação avaliativa.

No eixo metodológico, os ‘disruptivos’ introduzem abordagens fora do repertório padrão. O TCE-SC realizando avaliações de impacto por meio dos mencionados Levantamentos Econométricos, aplicadas a políticas como o ProEMI (Ensino Médio Inovador), ICMS-Educação e Iluminação Pública, com parceria acadêmica (UFSC) e uso de marco metodológico internacional (Referencial do Banco Mundial) como bibliografia orientadora para

o desenho e a inferência causal. O TCE-RN, por sua vez, institucionalizou a ‘Síntese de Evidências’ como instrumento científico (norma própria: Instrução Técnica 02/2025-TCE/RN), voltado a intermediar evidências para gestores e sociedade. O piloto sobre vagas em creches combinou revisão da literatura (nacional/internacional), boas práticas e notas técnicas para informar alternativas de solução e alimentar futuras fiscalizações.

Quanto à valoração, observa-se alargamento dos critérios: no TCE/SC, embora os levantamentos se orientem por critérios normativos ligados ao objeto, a ênfase no mérito substantivo (efeitos/impactos) termina por tangenciar dimensões como relevância e equidade, ainda que a NBASP 9020 não seja aplicada de modo explícito como guia institucional.

No TCE/RN, constatou-se que na Síntese de Evidências não são emitidos juízos de valor prescritivos, tampouco são hierarquizadas as alternativas apresentadas; o produto limita-se a apresentar opções existentes e seus níveis de evidência, preservando a decisão às instâncias competentes. Todavia, nas demais fiscalizações realizadas envolvendo análises de políticas públicas são realizados exames ancorados em critérios definidos em matrizes de planejamento (como a maioria das demais Cortes de Contas) e nas dimensões selecionadas por objeto, havendo referência à participação de partes interessadas como princípio a ser perseguido, coerente com a orientação da NBASP 9020, sobretudo nos processos de auditoria operacional e nos espaços de validação com gestores.

Quanto ao uso dos resultados, os ‘disruptivos’ investem em estratégias de difusão e engajamento que fortalecem a oportunidade e a utilidade das análises. No TCE/SC, destaca-se a combinação de conferências, reuniões e palestras com jurisdicionados, divulgação em portal e revista do Tribunal e o papel ativo de Relatores temáticos na comunicação pública dos achados; além disso, há acompanhamentos e experiências de monitoramento, como o *follow-up* anual da política de ICMS-Educação, que retroalimentam o ciclo decisório. No TCE/RN, a Síntese de Evidências foi publicada no site e difundida na mídia e em painéis, havendo planejamento de ‘mesas técnicas’ (consensualismo) para pactuação de planos de ação e maior interação com gestores ao longo do processo, com o Tribunal posicionado como ‘instituição intermediadora de evidências’.

No plano procedimental, há arranjos diversos. No TCE/SC, não há ainda um instrumento normativo específico para ‘avaliação de políticas públicas’; os trabalhos seguem regramento de levantamentos e referências metodológicas (Banco Mundial), com fluxo decisório interno (autorização pela direção técnica; eventual relatoria temática) e prospecção de parceria formal com universidade para inaugurar nova fase de avaliações. No TCE/RN, as

auditorias seguem Manual do TCU e NBASPs, com uso de sistema estabelecendo o fluxo de procedimentos para as fiscalizações (SISDGF), para padronizar o ciclo e manual próprio para Síntese de Evidências.

Apesar dos avanços, persistem desafios críticos. No TCE/SC, a capacidade operacional é limitada para a ambição de avaliações econométricas (necessidade de equipes maiores e acesso a bases de dados confiáveis), além de exigências de TI/IA para tratamento de dados em escala e da necessidade de maior aproximação com gestores para potencializar o uso dos resultados. No TCE/RN, os entraves incluem recursos humanos e tecnológicos, especialização temática, articulação com atores da política e sistematização/normatização de um fluxo avaliativo próprio, além de dificuldades com bases de dados dos jurisdicionados e complexidades das políticas multiníveis, o que reforça a importância de parcerias acadêmicas e soluções digitais.

Em resumo, os ‘disruptivos’ indicam trajetórias de fronteira: no TCE-SC, a sofisticação técnica das avaliações de impacto: Levantamentos Econométricos que tendem a posicionar o Tribunal como produtor de evidências causais com potencial de reorientar políticas; no TCE-RN, a Síntese de Evidências e a aposta na implementação do consensualismo consolidam um papel intermediador entre ciência e decisão, com produtos de rápida circulação e alto potencial de uso. Em ambos, a institucionalização incremental (normas, manuais, parcerias) e os arranjos de engajamento (relatores temáticos, painéis, mesas técnicas) induzem o uso das avaliações e ampliam a capacidade de emitir juízos de valor baseados em evidências, delineando um padrão inovador e replicável de atuação das Cortes no sistema nacional de avaliação.

Em termos interpretativos, propõe-se a seguinte síntese comparativa entre os quatro grupos de tribunais, conforme a classificação apresentada na Tabela 9 a seguir.

Tabela 9

Síntese comparativa da avaliação de políticas públicas entre os quatro grupos de Tribunais de Contas.

Grupo de Tipologia	Trajatória/ Perfil	Instrumentos Predominantes	Referenciais Normativos e Metodológicos	Dimensões de Desempenho Avaliadas	Características Centrais	TCs por grupo
Tribunais Incipientes	Atuação inicial em APP; Falta de priorização da avaliação de PP; Autopercepção dos ACE: fase inicial em APP.	Auditoria Operacional (ANOP) ‘tradicional’, sem integração sistemática de dimensões de relevância ou utilidade.	Manual de ANOP; Ausência de referenciais específicos para avaliação de PP; Papel indutor do MMD-TC (ATRICON).	Economicidade; Eficiência; e Eficácia.	Priorização de fiscalizações de conformidade e exame de contas de governo; Reconhecimento da relevância da avaliação, mas com baixa instrumentalização técnica;	Maioria dos TCs estaduais e municipais; 25 Tribunais de Contas brasileiros.
Tribunais Incrementais	Consolidação gradual de práticas avaliativas no âmbito da auditoria operacional; Autopercepção dos ACE: fase ‘intermediária’ em APP	ANOP com adaptações para avaliar PP; Procedimento de Investigação Prévio (TCE/PE).	Manual de ANOP (próprios e do TCU); Referencial de Controle de PP do TCU; Guias do IPEA;	Foco no exame da ‘qualidade’ da PP; Economicidade; Eficiência; Eficácia; Efetividade restrita à verificação de indicadores; Incorporação de critérios transversais: equidade, sustentabilidade e governança.	Auditorias mais analíticas, com exame do desenho e da implementação das políticas (qualidade das PP); Análises de contexto (político, institucional e ideológico); Tangenciamento da relevância da política pública sem adesão integral à NBASP 9020	<ul style="list-style-type: none"> • TCU: liderança normativa; e subsidiam o Congresso (REPP). • TCM/RJ: destaca-se pelo uso de Modelos Lógicos na ANOP; foco na Equidade; e existência de uma Coordenação revisora das APP; • TCE/PI: ênfase na avaliação <i>Ex Ante</i> (qualidade da formulação) baseada em manuais do IPEA. • TCE/PE: foco na celeridade e <i>indução à correção de ‘impropriedades’</i> (emissão de alertas de responsabilização via PIP).

Tribunais Normativos-Estruturantes	Transição deliberada para avaliação de políticas públicas <i>stricto sensu</i>	Instituição de um instrumento próprio de Avaliação de PP. Teoria da Mudança; Modelo Lógico do IPEA.	NBASP 9020 (adesão plena); Guias do IPEA; Referencial de Controle de PP do TCU; Suporte técnico-científico da Academia (FGV e UFG).	'Relevância' (diagnóstico e desenho da PP); Pretensão de avançar para o exame da 'utilidade' da PP em um segundo momento.	Institucionalização planejada da avaliação de PP; Base metodológica específica da avaliação de PP; Elaboração colaborativa do diagnóstico do "problema" (Árvore de Problemas"); Foco na qualidade do desenho da política e em sua lógica interventiva; Parcerias acadêmicas (FGV e UFG).	TCE/GO: Res. Norm. 02/2024; Observatório de PP; Modelo Lógico do IPEA; TCE/ES: SECEX Social; piloto com FGV; Avaliação Executiva; Teoria da Mudança.
Tribunais Disruptivos	Expansão metodológica para além das práticas avaliativas dos demais TC (<i>dos modelos usuais de controle</i>); Autopercepção: 'avançada' ou 'inovadora' em nichos específicos.	Levantamentos Econométricos (avaliação de impacto); Síntese de evidências (revisão sistemática).	Métodos de inferência causal (<i>PSM, Neyman-Rubin, DEA etc.</i>); Evidências científicas; Ausência de referência explícita à NBASP 9020	Efetividade (impacto); Evidências para orientar a tomada de decisão pelos gestores.	Inovação metodológica singular; Uso intensivo de dados e ciência aplicada; Deslocamento do controle da verificação para a produção e curadoria de evidências; Apoio estratégico ao gestor público	TCE/SC: único tribunal a realizar sistematicamente Avaliações de Impacto (causalidade) usando métodos econométricos (contrafactual, <i>propensity score matching</i>). Responde à pergunta: 'A política mudou a realidade?'" TCE/RN: Inovou com a Síntese de Evidências, que consiste em uma revisão sistemática da literatura para oferecer um 'menu de soluções' a serem avaliadas pelos gestores.

Nota: elaborada pela autora.

4.4 Implicações metodológicas para o Controle Externo: discussão integradora dos achados

A análise integrada dos dados empíricos com o quadro teórico-conceitual da pesquisa visa extrair implicações analíticas e institucionais para o Controle Externo. A discussão tem por objetivo articular especialmente os ‘ramos’ (métodos, valoração e uso) da Árvore da Teoria da Avaliação (Lemire et al., 2020) com as dimensões do Jardim das Abordagens (Montrosse-Moorhead e com os achados relatados no capítulo dos ‘resultados e discussão’.

Importante ressaltar que a compreensão de avaliação de políticas públicas como exame sistemático, baseado em evidências, orientado à emissão de juízo de valor de políticas públicas com base no exame de diferentes dimensões (incluindo relevância e utilidade) explica a predominância do uso da Auditoria Operacional como principal instrumento utilizado pela maioria das Cortes de Contas. Observa-se, entretanto, uma assimetria entre tribunais que restringem à avaliação ao exame dos ‘3 Es’ e os que ‘avançam’ nas práticas avaliativas ao incorporarem adaptações à ANOP (TCU, TCE/PE; TCE/PI e TCM/RJ), além de dois outros tribunais nos quais emergem gradualmente práticas avaliativas *stricto sensu* (TCE/ES e TCE/GO) duas trajetórias consideradas singulares (TCE/RN e TCE/SC).

Na perspectiva evolutiva, as quatro ondas (ciência, diálogo, neoliberal/NGP e evidências) conversam com o panorama dos Tribunais de Contas: diversidade metodológica e valoração de evidências convivem com a preocupação com o desempenho (NGP) e com arranjos dialógicos que buscam a legitimidade e a adesão dos gestores, e a preocupação com a justiça social: o ‘quarto’ ramo da Arvore de Avaliação, mencionado por Senra (2024).

À luz da Árvore revisitada de Lemire et al. (2020) e dos achados, a robustez no ramo dos métodos no Controle Externo tende a decorrer do gerenciamento de trilhas metodológicas conforme a capacidade avaliativa institucional, a maturidade da política, e a disponibilidade/qualidade de dados.

O Jardim de Montrosse-Moorhead et al. (2024) complementa que a escolha do método não é apenas técnica: ela envolve valores, contexto e dinâmica de poder. Assim, a decisão entre um diagnóstico célere de relevância, um estudo de implementação, uma avaliação de impacto ou uma ‘revisão sistemática’ deve ser articulada com mecanismos de indução do uso (diálogo com os *stakeholders*; construção colaborativa de ‘soluções’; divulgação de resultados para a sociedade etc.) que alteram a autoridade do Controle Externo de meramente responsabilizadora

para cognitiva-indutiva, sem afastar, indubitavelmente, o poder sancionatório quando necessário.

No ramo da valoração, a emissão de juízo de valor permanece ancorada predominantemente em economicidade, eficiência e eficácia e tende a produzir achados consistentes sobre desempenho imediato, avançando em alguns tribunais no exame da ‘relevância’, e nenhum especificamente na ‘utilidade’ das políticas públicas, embora no TCE/SC já tenha se consolidado no exame de impacto (efetividade).

Conforme Montrosse-Moorhead et al. (2024), o Jardim oferece os instrumentos, que combinados com os achados, permitem qualificar esse julgamento, tornando explícitos os valores públicos adotados (equidade, satisfação do beneficiário, sustentabilidade) e exigindo a realização de análises do contexto (institucional, político e ideacional), que devidamente documentados (nem sempre o são nas Cortes que as realizam), aliada ao desenvolvimento de competência cultural, tem o potencial de elevar a legitimidade e a transparência do ‘julgamento’ das políticas e melhorar a qualidade das avaliações.

No ramo do uso, a literatura e os dados sugerem que maximizar o uso implica planejar, desde o início, quem participa, como decide e de que maneira se traduzem achados para os diversos públicos interessados (gestores, parlamento, sociedade). Vislumbra-se, assim, aumentar a probabilidade dos diferentes ‘usos’ indicados pela doutrina: de uso instrumental (ajustes tempestivos), de uso conceitual (aprendizado) e de esclarecimento (qualificação do debate público), enquanto o controle conserva o *enforcement* para faltas graves ou omissões, e mantém o foco no seu objetivo maior: assegurar a correta (adequada) aplicação dos recursos públicos para maiores entregas à sociedade, ou seja, para otimizar o atendimento das demandas sociais prioritárias. Segundo as percepções dos auditores de controle externo, isso demanda rotinas de monitoramento tempestivo, estratégias de comunicação cidadã e práticas consensuais (mesas técnicas e TAGs) que reduzam assimetrias de informação e resistência à implementação.

Como síntese, os achados empíricos dialogam coerentemente com os dois referenciais: (i) a pluralização metodológica ventilada por Lemire et al. (2020) se revela necessária e gradual, dadas as restrições de dados e tempo processual; (ii) a qualificação da valoração proposta pelo Jardim (com valores, contexto e justiça social) responde a lacunas recorrentes de fundamentação explícita do mérito; e (iii) a centralidade do uso, nos dois modelos, legitima a adoção de arranjos consensuais e pedagógicos que ampliam a efetividade, sem abdicar da função sancionatória, nem do exame de conformidade. Como implicação prática, propõe-se

migrar de ‘projetos de auditoria’ para ‘cadeias de valor avaliativas’ em que método, valoração e indução ao uso devem ser planejadas de forma articulada.

5. CONCLUSÕES E RECOMENDAÇÕES

A pesquisa atingiu o objetivo geral proposto de analisar como os Tribunais de Contas brasileiros concebem e operacionalizam a avaliação de políticas públicas, na perspectiva dos auditores de controle externo, identificando padrões e implicações para o Controle Externo.

Nesse intuito, os objetivos específicos do estudo também foram alcançados. Apresentou-se o panorama nacional das percepções e práticas avaliativas dos auditores de controle externo dos 33 Tribunais de Contas brasileiros. Com base no exame das evidências coletadas por meio dos 38 grupos focais realizados, totalizando 126 auditores de controle externo, e exame documental dos 16 relatórios de fiscalizações selecionados dos tribunais considerados ‘mais avançados’ na avaliação de políticas públicas, foi possível analisar experiências consolidadas e trajetórias emergentes, com foco nos métodos, nos critérios de valoração e nos usos das avaliações.

A partir dos exames realizados foi possível propor uma tipologia das trajetórias avaliativas a partir dos padrões e variações identificadas, discutir as implicações para o aprimoramento da avaliação de políticas públicas no âmbito do Controle Externo e propor como Produto Técnico-Tecnológico, como se segue, um Guia de Maturidade em Avaliação de Políticas Pública para os Tribunais de Contas com indicação de metodologias, critérios para valoração e estratégias para incentivar o uso pelos gestores públicos dos trabalhos realizados.

A investigação empírica demonstrou uma dissonância conceitual estrutural: embora a avaliação de políticas públicas seja uma prática disseminada nas Cortes, predomina a compreensão de que os trabalhos se configuram como extensão das auditorias operacionais, ainda que condicionada à incorporação de adequações, e não avaliações em sentido estrito, conforme parâmetros da NBASP 9020. Por conseguinte, a valoração se concentra na análise das dimensões dos ‘três Es’ (economicidade, eficiência e eficácia), não alcançando as dimensões substantivas associadas à avaliação de políticas públicas: relevância e utilidade.

No ‘**ramo dos métodos**’, a Auditoria Operacional é o instrumento predominante para avaliar políticas públicas, embora a percepção dos auditores indique a necessidade de adaptações metodológicas, complementações e/ou o uso de instrumentos adicionais para o alcance de dimensões substantivas. Da mesma forma, apesar do reconhecimento da importância da NBASP 9020, sua efetiva institucionalização e aplicação ainda é restrita: auditores da maioria dos Tribunais de Contas relataram não seguir seus preceitos, ainda que iniciativas neste sentido se encontrem em discussão e análise.

A despeito da predominância da ANOP, o panorama é heterogêneo. Foram identificadas trajetórias emergentes de inovação, nas quais Tribunais de Contas por isso considerados ‘mais avançados’ em avaliação de políticas públicas adotam abordagens avaliativas específicas e/ou ‘métodos de fronteira’, tais como a Teoria da Mudança, Avaliação Executiva, Modelo Lógico, Árvore de Problemas, avaliações de impacto por meio de Levantamentos Econométricos, Síntese de Evidências e análise Big Data/IA em fiscalizações. Além disso, foi identificada uma liderança normativa do TCU, cujos referenciais têm sido utilizados por diferentes Cortes como fonte subsidiária para a realização dos trabalhos, especialmente o Referencial de Controle de Políticas Públicas da Corte de Contas federal.

Quanto às restrições, os auditores convergem em relação às limitações estruturais que limitam o avanço das avaliações: dificuldade para avaliar efetividade/impacto e, por conseguinte, a dimensão da ‘utilidade; deficiência ou ausência de bases de dados e de indicadores das políticas públicas; incompatibilidade de tempo para a realização de fiscalizações mais complexas e os prazos processuais, dentre outros mencionados. Os obstáculos indicam que a transição para avaliações de políticas públicas depende de infraestrutura de dados, capacidade técnica e condições organizacionais, mais do que a simples adesão ao novo normativo (NBASP 9020).

Em relação ao ‘**ramo da valoração**’, a pesquisa evidencia um descompasso entre a norma e a cultura institucional das Cortes de Contas. Não obstante valores emergentes (justiça social, equidade, sustentabilidade e satisfação do beneficiário) tenham sido apontados, persiste uma cultura orientada à conformidade e responsabilização. Ademais, embora reconhecida a relevância da análise de contexto (institucional, político e ideacional) para orientar a emissão de juízo de valor (julgamento), tem-se que a prática ainda não é sistematizada de forma geral, nem documentada, com maior fragilidade da análise de contexto ideacional. A principal estratégia apontada pela maioria dos auditores consiste na educação corporativa como meio de transferência técnica e formação de uma nova consciência cultural, hábil para incentivar a ampliação de avaliações de políticas públicas como estratégia para a legitimação das Cortes de Contas, por meio de maior entrega de valor público para a sociedade, mediante o aprimoramento das políticas públicas, concebidas como o meio por excelência para o atendimento das demandas sociais consideradas prioritárias.

Quanto ao ‘**ramo do uso**’, verificou-se uma preocupação com a efetiva apropriação dos resultados das avaliações pelos gestores públicos como caminho para a promoção de melhorias nas políticas públicas. Nesse intuito, identificou-se a superação de uma lógica centrada na

responsabilização para uma lógica dialógica e pedagógica (reuniões com gestores, validação de achados, divulgação e capacitação), ao lado de mecanismos de monitoramento e uso de Planos de Ação, e a institucionalização de práticas de consensualismo (Mesas Técnicas e TAGs) com a finalidade de aumentar a efetividade do controle, mediante ampliação do uso instrumental das avaliações, promoção da transparência e levantamento de evidências para o processo decisório.

As análises e discussões promovidas quanto aos resultados da pesquisa permitiram a proposição de uma tipologia das trajetórias avaliativas dos Tribunais de Contas (incipientes, incrementais, normativos-estruturantes e disruptivos/científicos) que organiza e categoriza a heterogeneidade das Cortes de Contas conforme as diferentes ‘abordagens’ avaliativas adotadas, segundo as opções metodológicas, critérios de valoração e estratégias para fomentar o uso dos trabalhos realizados.

Em síntese, o quadro empírico identificado junto aos auditores de controle externo sugere que a incorporação da avaliação de políticas públicas pode reconfigurar paulatinamente a base de legitimidade do Controle Externo, deslocando o foco no controle de conformidade e a autoridade centrada na responsabilização para o de geração de valor público, na autoridade técnica, dialógica e orientada por evidências, sem afastar, sempre que necessário, a função sancionatória e o exame de conformidade.

A luz do diagnóstico empírico e dos obstáculos identificados, recomendam-se as ações estruturantes para os Tribunais de Contas organizadas em três níveis: (1) nível institucional: recomenda-se a priorização estratégica da avaliação de políticas públicas na agenda institucional, acompanhada da especialização temática dos órgãos técnicos (e.g. saúde, educação, segurança) e a implementação de programas contínuos de educação corporativa voltados para a formação de uma cultura avaliativa, mitigando o viés excessivo de conformidade; (2) nível metodológico: adoção de estratégias para a incorporação gradual da NBASP 9020, inicialmente pela avaliação sistemática da dimensão da relevância (desenho e diagnóstico), utilizando metodologias consolidadas como a Teoria da Mudança e Modelos Lógicos, com a pretensão de avançar para a análise de utilidade em um segundo momento; além disso, considera-se necessária a sistematização e documentação da análise de contexto (institucional, político e ideacional) nos relatórios de fiscalização para garantir a transparência dos juízos de valor; (3) nível de relacionamento interinstitucional e indução ao uso das avaliações produzidas: implementação do consensualismo (Mesas Técnicas e TAGs), fomento à participação social, e fortalecimento de monitoramentos tempestivos como caminhos para induzir o uso efetivo dos resultados.

Por fim, recomenda-se aprofundar a agenda de pesquisa a fim de examinar e consolidar os caminhos identificados: (1) **efetividade do consensualismo**: investigar comparativamente se o uso de estratégias de indução consensual (Mesas Técnicas e TAGs) gera, de fato, mais resultados concretos e tempestivos do que o tradicional modelo de responsabilização; (2) **impacto da Síntese de Evidências**: analisar o impacto real da metodologia de ‘Síntese de Evidências’ no processo decisório dos gestores, verificando se a oferta de um ‘menu de soluções’ altera a formulação das políticas públicas locais; (3) **avaliação das parcerias acadêmicas**: comparar a qualidade técnica e a profundidade analítica de avaliações realizadas em parceria com universidades *versus* aquelas produzidas exclusivamente por equipes internas, testando a hipótese de que a cooperação externa proporciona ganhos exponenciais; (4) **meta-avaliação**: realizar uma meta-avaliação dos levantamentos econométricos de impacto produzidos pelo TCE/SC, visando validar essas metodologias como ‘padrão-ouro’ para a análise de efetividade (talvez da ‘utilidade’ oportunamente) e facilitar sua difusão; e (5) **evolução dos Tribunais de Contas**: realização de estudo longitudinal para acompanhar a evolução e investigar os fatores condicionantes de mudança institucional no Controle Externo.

Em suma, presente dissertação defende que futuro da avaliação de políticas públicas nos Tribunais de Contas depende da articulação entre rigor científico (método), sensibilidade social na incorporação das dimensões de valor público (valoração) e diálogo interinstitucional voltado ao incentivo ao uso das avaliações (uso). Espera-se, a luz das evidências e recomendações apresentadas, que as Cortes de Contas, ao investirem nas avaliações de políticas públicas, contribuam de fato para a geração de resultados socialmente relevantes e capazes de promover transformações na realidade da sociedade brasileira.

A leitura integrada dos achados empíricos evidencia que a evolução das práticas avaliativas nos Tribunais de Contas brasileiros se expressa sobretudo por meio de arranjos institucionais distintos de organização dos métodos, da valoração e das estratégias de indução ao uso. As diferenças observadas revelam trajetórias alternativas e, por vezes, combinadas de institucionalização da avaliação, condicionadas por capacidades organizacionais e contexto decisório. Essa constatação empírica indica que o aprimoramento da Avaliação de Políticas Públicas no Controle Externo demanda instrumentos capazes de acomodar múltiplos caminhos de transição, oferecendo referências práticas, entregáveis mínimos e orientações metodológicas compatíveis com diferentes cenários institucionais, fundamento a partir do qual se estrutura o Produto Técnico-Tecnológico apresentado no capítulo subsequente.

6. PRODUTO TÉCNICO-TECNOLÓGICO (PTT): Guia de Arquiteturas e Trilhas para Avaliação de Políticas Públicas (APP) em Tribunais de Contas. Modelo modular, não hierárquico, com autodiagnóstico, entregáveis mínimos e indicadores por trilha.

6.1 Introdução

O presente Produto Técnico-Tecnológico (PTT) constitui um processo de gestão documentado, concebido como um framework de apoio à estruturação e ao fortalecimento das práticas de Avaliação de Políticas Públicas (APP) no âmbito do Controle Externo brasileiro. O Guia deriva diretamente dos achados empíricos desta dissertação, que evidenciaram heterogeneidade institucional entre os 33 Tribunais de Contas quanto às concepções, às metodologias adotadas, aos critérios de valoração e às estratégias de indução ao uso das avaliações.

Em consonância com a literatura contemporânea e com os achados desta pesquisa, o Guia foi redesenhado sob o entendimento de que inexiste um ‘padrão-ouro’ para APP. As arquiteturas identificadas empiricamente (incremental, normativo-estruturante e disruptiva/científica) não configuram estágios ou níveis hierárquicos de maturidade, mas trilhas alternativas, plenamente viáveis e combináveis. A partir de um ponto de partida comum (perfil base de capacidades e estratégia organizacional), uma Corte pode optar por qualquer trilha; combinar trilhas; alterná-las ao longo do tempo; manter arranjos distintos por unidade temática; ou inovar por meio de uma nova trilha (alternativa/híbrida) construída a partir de suas prioridades, contexto e capacidade instalada.

6.2 Descrição Geral

O PTT enquadra-se na categoria Processo de Gestão Documentado (Grupo 3), conforme a classificação da CAPES, e estrutura-se como um Guia de Arquiteturas e Trilhas. Seu objetivo central é apoiar Tribunais de Contas na definição e implementação de arranjos avaliativos coerentes com suas capacidades institucionais e com o propósito de cada avaliação, fortalecendo, de forma integrada, as dimensões de métodos, valoração e uso.

O Guia é composto por quatro elementos estruturantes:

- Matriz de autodiagnóstico em duas etapas (capacidade instalada × escolha de trilha/módulos);
- Perfil base de capacidades (ponto de partida), com entregáveis mínimos e indicadores iniciais;

- Trilhas (arquiteturas) não hierárquicas, com módulos combináveis, entregáveis mínimos e indicadores por trilha;
- Diretrizes transversais para indução ao uso, qualidade, transparência e governança ética de dados.

Importante ressaltar que o Guia preserva a autonomia de cada Corte para promover adaptações conforme suas prioridades, objetivos estratégicos e capacidade instalada (recursos humanos, estrutura organizacional, qualidade dos bancos de dados e estrutura de TIC). Visualiza-se inclusive a possibilidade de intercâmbio e a combinação de módulos entre trilhas, inclusive por portfólio, sem ordem obrigatória de adoção.

6.3 Base teórica utilizada

A fundamentação teórica do Guia reflete o quadro teórico-conceitual desta dissertação. Sua estrutura adota como eixo central os três ramos da Árvore da Teoria da Avaliação (métodos, valoração e uso), conforme revisitada por Lemire et al. (2020), incorporando a compreensão de que a avaliação constitui um processo integrado de produção de conhecimento e indução de decisões públicas.

De forma complementar, o Guia dialoga com o Jardim das Abordagens de Avaliação (Montrosse-Moorhead et al., 2024), ao reconhecer que decisões metodológicas e avaliativas são condicionadas por valores, contexto institucional, dinâmicas de poder e estratégias deliberadas de promoção do uso das evidências produzidas.

No plano empírico, o Guia ancora-se na tipologia construída a partir dos grupos focais com auditores dos 33 Tribunais de Contas brasileiros e da análise documental de relatórios de fiscalização, a partir da qual emergiram arquiteturas institucionais recorrentes. O PTT, entretanto, reorganiza tais padrões em trilhas não hierárquicas e módulos combináveis, para que possam ser adotados em diferentes configurações, inclusive por Tribunais já considerados “avançados”, em estratégia de portfólio e inovação.

6.4 Relevância e impacto

Sob a perspectiva acadêmica e profissional, o Guia apresenta elevada aderência aos objetivos do Mestrado Profissional em Administração Pública, ao converter evidências empíricas qualificadas em instrumento de gestão aplicável à realidade institucional do Controle Externo. Do ponto de vista institucional, o Guia: (i) apoia a delimitação do *modus operandi* avaliativo em APP; (ii) favorece o aprendizado interinstitucional entre Cortes com repertórios

distintos; (iii) contribui para a qualificação das avaliações e para a ampliação de sua utilidade social e legitimidade; e (iv) estimula inovação responsável, ao permitir combinações de trilhas e a construção de trilhas alternativas/híbridas.

6.5 Documentos comprobatórios e evidências

São documentos e evidências que fundamentam o PTT:

- Roteiro do grupo focal (Apêndice A).
- Relatórios das fiscalizações e avaliações de políticas públicas analisados (Apêndice B), disponíveis nos portais institucionais e/ou acessíveis mediante solicitação, quando cabível.

6.6 Como implementar o Guia: autodiagnóstico e seleção de trilhas

O Guia recomenda implementação em quatro passos, com autodiagnóstico em duas etapas:

Passo 1: Diagnóstico do ponto de partida (capacidade instalada): identificar capacidades atuais (pessoas, dados, governança, rotinas, cultura avaliativa e comunicação) e lacunas críticas;

Passo 2: Definição de propósito e audiência da avaliação: definir a pergunta avaliativa e o tipo de decisão a apoiar (aperfeiçoar desenho, qualificar implementação, estimar efeitos/impactos, curar evidências para decisão, fortalecer controle social etc.);

Passo 3: Seleção de trilha(s) e módulos: escolher uma ou mais trilhas (incremental, normativo-estruturante, disruptiva/científica ou alternativa/híbrida) e os módulos coerentes com o propósito e com a capacidade instalada;

Passo 4: Execução, aprendizagem e ajuste: executar o ciclo avaliativo com mecanismos de qualidade, comunicação e monitoramento; registrar lições aprendidas e ajustar o arranjo ao final do ciclo (melhoria contínua).

6.6.1 Perfil base (ponto de partida) — capacidades e entregáveis mínimos

O perfil base descreve condições típicas de Cortes em início de institucionalização de APP. Não representa inferioridade, mas um retrato de capacidades e rotinas mínimas necessárias para sustentar qualquer trilha escolhida (inclusive trilhas avançadas em nichos específicos).

Entregáveis mínimos (MVP do perfil base):

- Registro de ao menos um trabalho com recorte de desempenho (mesmo via ANOP tradicional), com perguntas avaliativas explícitas.

- Matriz de planejamento com critérios e padrões de referência definidos e rastreáveis (ao menos normativo/histórico).
- Plano mínimo de evidências (fontes, técnicas, limitações) e estratégia inicial de comunicação (sumário executivo).
- Rotina piloto de validação (reunião de encerramento) e previsão de acompanhamento/monitoramento.

Indicadores sugeridos (perfil base):

- Percentual de trabalhos com perguntas avaliativas explícitas no planejamento;
- Percentual de relatórios com critérios e padrões de referência declarados;
- Realização de ao menos uma (01) reunião de validação/encerramento por avaliação;
- Existência de monitoramento planejado (sim/não) e prazo de início do monitoramento.

6.7 Trilhas (arquiteturas) não hierárquicas: módulos, entregáveis e indicadores

As trilhas a seguir representam arquiteturas institucionais e repertórios avaliativos. Elas podem ser adotadas por Tribunais em qualquer ponto da tipologia, inclusive por aqueles já classificados como incrementais, normativo-estruturantes ou disruptivos/científicos, seja para avançar em outra arquitetura, seja para compor um portfólio com unidades especializadas. Não há ordem obrigatória nem superioridade intrínseca entre trilhas.

6.7.1 Trilha Incremental (ANOP qualificada e valoração ampliada)

Foco: qualificar auditorias operacionais e instrumentos correlatos para incorporar planejamento mais denso, critérios transversais (equidade, sustentabilidade, governança, qualidade) e estratégias de uso, sem exigir adesão plena a um único normativo.

Módulos e entregáveis mínimos:

M1 — Planejamento enriquecido

- Perguntas avaliativas e teoria implícita do programa explicitadas (mesmo que de forma simplificada).
- Matriz de planejamento com critérios clássicos + pelo menos um (01) critério transversal.

- Plano de evidências com triangulação mínima (documentos + dados + entrevistas/visitas *in loco*/observação direta).

M2 — Contexto (CIPI) pragmático

- Registro sintético de contexto institucional, político e ideacional, sempre que possível e pertinente;
- Justificativa de como o contexto afeta achados e recomendações.

M3 — Uso e indução

- Reunião de abertura e/ou encerramento com gestores (ou equivalente), preservada a independência técnica.
- Plano de ação solicitado/pactuado e/ou previsão de monitoramento.
- Produto de comunicação cidadã (sumário executivo, infográfico, painel, cartilha etc.).

Indicadores sugeridos:

- Proporção de avaliações com critérios transversais explícitos (equidade/sustentabilidade/governança).
- Percentual de avaliações com registro de contexto (CIPI) no relatório.
- Percentual de avaliações com plano de ação e monitoramento iniciado em prazo definido.
- Número de produtos de comunicação cidadã publicados por ciclo avaliativo.

6.7.2 Trilha Normativo-Estruturante (APP *stricto sensu* e foco em desenho/relevância)

Foco: institucionalizar um processo de APP com governança e procedimentos próprios, com ênfase inicial em diagnóstico do problema, desenho e relevância da política (*ex ante* e/ou avaliação executiva), utilizando modelos como Teoria da Mudança e Modelo Lógico, e incorporação efetiva da NBASP 9020 como baliza. Essa trilha pode coexistir com rotinas incrementais e com núcleos disruptivos, a depender do portfólio institucional.

Módulos e entregáveis mínimos:

N1 — Diagnóstico e desenho

- Diagnóstico situacional e/ou árvore de problemas.
- Modelo Lógico e/ou Teoria da Mudança com pressupostos e riscos.
- Perguntas avaliativas alinhadas à dimensão de relevância (coerência problema-objetivos-meios).

N2 — Governança e qualidade

- Ato interno definindo APP, papéis e fluxo (planejamento-execução-reporte-monitoramento).
- Rotina de revisão técnica/controle de qualidade (*checkpoints*) sem engessamento (flexibilidade metodológica).
- Parceria técnica/científica (quando viável) para reforço metodológico.

N3 — Estratégia de uso e aprendizagem

- Arranjo de engajamento (comitê consultivo, mesa técnica, painéis de referência), preservada a independência do Tribunal.
- Plano de comunicação para públicos distintos (gestão, parlamento, sociedade).
- Monitoramento com ciclos curtos de retroalimentação e transparência dos encaminhamentos.

Indicadores sugeridos:

- Percentual de avaliações com uso declarado do Modelo Lógico/Teoria da Mudança ou outra abordagem própria da APP.
- Existência de ato interno e POP/roteiro de APP (sim/não) e grau de uso (percentual de trabalhos).
- Número de avaliações com instâncias formais de engajamento e registro de decisões.
- Taxa de implementação de encaminhamentos pactuados em planos de ação e/ou consensualismo, como celebração de TAGs (em prazo definido).

6.7.3 Trilha Disruptiva/Científica (impacto, inferência causal e/ou síntese de evidências)

Foco: produzir evidências avançadas para decisões públicas por meio de duas vertentes metodológicas complementares, que podem coexistir no mesmo portfólio institucional:

(i) Levantamentos Econométricos: avaliações de impacto/efetividade, apoiadas em inferência causal, com valoração substantiva baseada em efeitos observados e suas limitações; e/ou

(ii) Síntese de Evidências: destinada à curadoria e organização do conhecimento científico, sem emissão de juízo avaliativo prescritivo e sem hierarquização de alternativas como ‘melhor/pior’, limitando-se a apresentar opções existentes, seus níveis de evidência e condicionantes contextuais.

Essa trilha pode ser adotada como núcleo especializado em temas selecionados, mesmo quando a Corte mantém, no restante do portfólio, rotinas incrementais ou normativo-estruturantes.

Módulos e entregáveis mínimos:

D1 — Avaliação de impacto (quando viável)

- Protocolo de desenho causal (pergunta, estratégia de identificação, grupo de comparação, suposições e ameaças à validade).
- Plano de dados (fontes, qualidade, tratamento, privacidade e segurança).
- Relatório com magnitude do efeito, heterogeneidades e limites interpretativos, explicitando implicações para a decisão pública.

D2 — Síntese de Evidências (curadoria científica para decisão, sem valoração prescritiva)

- Pergunta de decisão e protocolo de busca/seleção de estudos (critérios de inclusão/exclusão e rastreabilidade).
- Critérios de qualidade da evidência e síntese de achados, com explicitação de lacunas e incertezas.
- Produto de decisão (*policy brief* / “menu de opções”) que apresente alternativas existentes e níveis de evidência, sem recomendar hierarquicamente uma alternativa como “a melhor” (a decisão permanece com as instâncias competentes).

D3 — Inteligência analítica (*Big Data/IA*) como suporte

- Integração de bases e trilhas de fiscalização orientadas por dados.
- Painéis/*dashboards* para monitoramento contínuo de indicadores.
- Rotina de governança, segurança e mitigação de vieses em sistemas analíticos/algorítmicos, quando aplicável.

Indicadores sugeridos:

- Número de avaliações com protocolo causal documentado e reproduzível (D1).
- Percentual de estudos com plano de governança de dados e avaliação de risco (privacidade/ética) (D1/D3).
- Número de Sínteses de Evidências produzidas com protocolo rastreável e critérios de qualidade explicitados (D2).
- Evidências de uso por gestores e/ou instâncias deliberativas (citações, planos de ação, atos normativos, decisões) quando rastreável (D1/D2).

6.7.4 Trilha Alternativa/Híbrida (terceira via e inovação orientada ao contexto)

Foco: permitir que a Corte desenvolva uma arquitetura própria quando a natureza do problema, o ciclo da política ou as condições institucionais recomendarem arranjos não plenamente contemplados pelas trilhas anteriores. A trilha híbrida pode combinar módulos das demais trilhas e incorporar abordagens como avaliação participativa, avaliação do desenvolvimento, ciclos rápidos de *feedback*, mapeamento de resultados e outras estratégias compatíveis com a independência e com as competências do Controle Externo.

Entregáveis mínimos:

- Definição explícita da abordagem escolhida e justificativa de adequação ao contexto.
- Mapa de *stakeholders* e desenho de participação (quem participa, em que profundidade, com quais salvaguardas).
- Produto de aprendizagem (lições, hipóteses testadas, recomendações adaptativas) e mecanismo de acompanhamento.

Indicadores sugeridos:

- Existência de mapa de *stakeholders* e registro de decisões participativas (sim/não).
- Número de ciclos de *feedback* implementados e ajustes realizados durante a execução.
- Grau de satisfação/adesão dos atores envolvidos (quando aplicável) e qualidade percebida do aprendizado.

6.7.5 Regras de combinação e portfólio (aplicáveis a qualquer trilha)

- **Combinação por ciclo da política:** desenho/relevância pode favorecer módulos normativo-estruturantes; implementação pode favorecer incremental; maturidade de dados pode habilitar *impact evaluation* e síntese de evidências.
- **Combinação por capacidade:** com dados limitados, priorizar diagnóstico, contexto e indução ao uso; com dados robustos, acrescentar desenho causal e reprodutibilidade.
- **Combinação por risco e legitimidade:** em temas sensíveis, fortalecer análise de contexto (CIPi), participação e transparência, além de comunicação cidadã.
- **Portfólio institucional:** uma Corte pode manter simultaneamente unidades/rotinas em diferentes trilhas (incremental, normativo-estruturante e disruptiva), por área temática, complexidade e prioridade social.
- **Inovação:** o intercâmbio de módulos deve ser registrado e avaliado (lições aprendidas), reforçando melhoria contínua.

Os quadros a seguir descrevem **perfis institucionais típicos** de Avaliação de Políticas Públicas observados empiricamente nos Tribunais de Contas brasileiros, organizados conforme os **focos predominantes** de métodos, valoração e uso das avaliações.

Esses perfis **não configuram níveis sequenciais de maturidade**, tampouco trajetórias obrigatórias de evolução institucional. Um mesmo Tribunal pode apresentar características de mais de um perfil, seja por área temática, unidade organizacional, tipo de política avaliada ou por estratégia institucional deliberada.

Assim, os quadros devem ser lidos como **referenciais comparativos e instrumentos analíticos**, aptos a apoiar escolhas de **arquitetura avaliativa**, composição de **portfólio institucional** e definição de **trilhas metodológicas**, em consonância com o propósito avaliativo, o contexto e a capacidade instalada.

Quadro-Síntese — Perfil Incipiente: foco na conformidade.

Dimensão	Síntese (caracterização típica)	Capacidades/condições associadas	Entregáveis mínimos esperados	Evidências de institucionalização	Riscos e gargalos recorrentes
Perfil institucional	Predominância de fiscalizações voltadas à conformidade legal, financeira e orçamentária; Avaliação de políticas públicas ocorre de forma esparsa e não sistematizada.	Baixa capacidade avaliativa institucional; Pouca especialização temática; Cultura predominantemente orientada à conformidade.	Registro de ao menos um trabalho com recorte de desempenho (ainda que via ANOP tradicional); Matriz de planejamento tradicional.	Relatórios com foco nos “3 Es”, sem referência explícita à NBASP 9020.	Viés de conformidade; Limitação de dados; Ausência de normativos e de tempo compatível com avaliações mais densas; Dificuldade para examinar efetividade/impacto.
Métodos	Utilização de instrumentos tradicionais de fiscalização, com baixa pluralização metodológica.	Rotinas de planejamento pouco densas para avaliação; Infraestrutura limitada de dados.	Matriz básica com objetivos, critérios e fontes de evidência.	Descrição sumária de método e escopo no relatório.	Engessamento por prazos processuais; Baixa integração de bases de dados.
Valoração	Ênfase quase exclusiva em economicidade, eficiência e eficácia; Relevância e utilidade não explicitadas.	Critérios e padrões de referência pouco sistematizados; Análise de contexto pouco documentada.	Critérios clássicos explícitos + padrão normativo ou histórico.	Achados vinculados a critérios tradicionais de desempenho.	Fragilidade do juízo de valor pela ausência de análise contextual estruturada.
Uso	Uso predominantemente incidental; Monitoramentos nem sempre tempestivos.	Comunicação técnica; Baixa interação com gestores.	Reunião piloto de encerramento/validação; Recomendações formalmente viáveis.	Registro de encaminhamentos e eventual acompanhamento.	Baixo uso instrumental; Resistência dos gestores; Perda de oportunidade de correção.

Quadro-Síntese — Perfil Incremental: foco no desempenho e governança.

Dimensão	Síntese (caracterização típica)	Capacidades/condições associadas	Entregáveis mínimos esperados	Evidências de institucionalização	Riscos e gargalos recorrentes
Perfil institucional	Auditoria Operacional funciona como plataforma central, com adaptações avaliativas ; Aproxima-se da APP, sem aplicação plena na NBASP 9020.	Especialização temática, ainda que parcial; Maior densidade analítica; Início de práticas dialógicas.	Matriz de planejamento com adaptações para APP; Registro, ainda que parcial, de contexto; Previsão de estratégia inicial de uso.	Relatórios mais analíticos; Validação com gestores; Divulgação e monitoramento, nem sempre tempestivos.	Dados insuficientes; Produtividade e prazos incompatíveis com a complexidade da APP; Cultura institucional orientada para a responsabilização.
Métodos	ANOP qualificada por ferramentas avaliativas (<i>modelo lógico, ToC, levantamentos, TIC/IA</i>), conforme o objeto.	Planejamento mais robusto; Uso crescente de dados.	Perguntas avaliativas explícitas; Critérios transversais; Plano de evidências.	Referenciais do TCU/IPEA citados no planejamento/relatório.	Heterogeneidade metodológica sem normatização interna.
Valoração	Ampliação dos “3 Es” com valores emergentes e análise de contexto institucional e político.	Capacidade de explicitar critérios; Contexto ideacional ainda não sistematizado.	Seção de contexto + critérios e padrões de referência justificáveis.	Relatórios com contextualização e fundamentação mais densa.	Possibilidade de fragilização do julgamento quando o contexto não é documentado.
Uso	Estratégias dialógicas e pedagógicas combinadas com monitoramento tempestivo.	Rotinas de interação; Planos de ação e produtos de comunicação.	Plano de ação pactuado ou solicitado; Sumário executivo ou infográfico.	Registros de reuniões, monitoramentos e divulgação pública.	Limitação quanto ao uso instrumental da APP quando o monitoramento não é tempestivo.

Quadro-Síntese — Perfil Normativo-Estruturante: foco no desenho e relevância (NBASP 9020).

Dimensão	Síntese (caracterização típica)	Capacidades/condições associadas	Entregáveis mínimos esperados	Evidências de institucionalização	Riscos e gargalos recorrentes
Perfil institucional	Transição deliberada para APP <i>stricto sensu</i> , com governança, normas e arranjos próprios.	Unidade especializada em APP; Existência de POPs ou normativos disciplinando a APP; Parcerias acadêmicas; Implementação de Observatórios.	Ato normativo de APP; Produto específico de avaliação; Realização de diagnóstico e desenho da política.	Resoluções; POPs; Comitês consultivos e cooperação acadêmica.	Limitações de dados para avançar da relevância à utilidade; Desafio cultural para ampliação da APP (considerada a predominância da ANOP).
Métodos	Uso de metodologias de desenho (Modelo Lógico, Teoria da Mudança, Árvore de problemas) e Avaliações Executivas.	Competência técnica em desenho avaliativo; Padronização flexível.	Pergunta avaliativa + ToC/Modelo Lógico + Matriz de Evidências.	Relatórios ancorados na NBASP 9020 e Análise do Desenho da PP (relevância), com pretensão de avançar para o exame da utilidade.	Risco de formalização excessiva ('manualização').
Valoração	Ênfase na relevância (coerência problema-objetivos-meios), com perspectiva de avançar para a utilidade.	Capacidade de explicitar critérios substantivos e contexto CIPI.	Seção de valoração com critérios NBASP 9020 e análise de contexto registrada.	Julgamento fundamentado e transparente.	Limitações para mensuração custo-efetividade/utilidade.
Uso	Arranjos estruturados de engajamento e governança para indução do uso.	Capacidade de pactuação, monitoramento e comunicação acessível.	Plano de ação pactuado; Monitoramento tempestivo; Produto de transparência.	Mesas técnicas; Observatórios; e Monitoramentos tempestivos.	Avaliação perde impacto se o uso não for planejado desde o início.

Quadro-Síntese — Perfil Disruptivo/Científico: foco na utilidade, impacto e evidências para decisão.

Dimensão	Síntese (caracterização típica)	Capacidades/condições associadas	Entregáveis mínimos esperados	Evidências de institucionalização	Riscos e gargalos recorrentes
Perfil institucional	Produção de evidências avançadas (impacto, inferência causal ou síntese científica) para subsidiar decisões públicas.	Equipes especializadas; Acesso a bases de dados; Parcerias acadêmicas.	Estudo de impacto ou síntese de evidências institucionalizada; Estratégia de difusão.	Produção regular de estudos e divulgação ampliada.	Alta dependência de dados; Capacidade operacional limitada.
Métodos	Inferência causal, revisão sistemática/síntese de evidências e <i>Big Data</i> /IA.	Infraestrutura analítica; Metodólogos; Governança de dados.	Protocolo metodológico reprodutível.	Métodos explícitos e replicáveis; Padrões internacionais citados.	Sofisticação técnica sem indução adequada ao uso ('método sem uso').
Valoração	(i) Na Avaliação de Impacto: valoração substantiva (efeitos; impacto; utilidade), com discussão de magnitude, distribuição e limites; (ii) Na Síntese de Evidências: ausência de valoração prescritiva (apresentação de opções e evidências para deliberação informada).	Capacidade de traduzir achados técnicos em juízo útil.	Seção interpretativa com efeitos, distribuição e implicações.	Relatórios orientados à decisão.	Fragilidade se a causalidade for mal especificada.
Uso	Tribunal como produtor/intermediador de evidências; Ampla difusão e governança para implementação.	Comunicação cidadã + Diálogo + Monitoramento.	“Menu de opções” + Pactuação + Acompanhamento.	Conferências; Mesas Técnicas; Difusão pública.	Contestação política se faltarem salvaguardas de legitimidade e diálogo.

Quadro-Síntese transversal — “o que muda por perfil” (visão comparativa rápida)

Eixo	Perfil Incipiente	Perfil Incremental	Perfil Normativo-Estruturante	Perfil Disruptivo
Métodos	ANOP tradicional	ANOP com adaptações (qualificada) + ferramentas	APP NBASP 9020 (Avaliação de desenho)	Impacto causal / síntese científica
Valoração	“3 Es”	“3 Es” + Valores/contexto	Relevância (→ utilidade)	Impacto/efetividade (quando há inferência causal), e/ou curadoria de evidências (Síntese de Evidências), sem valoração prescritiva — apresentação de alternativas e níveis de evidência para deliberação informada.
Uso	Não faz parte do planejamento da ANOP	Dialógico + Monitoramento	Governança do uso	Evidência para decisão + Difusão
Normativos predominantes	Manuais de ANOP	Manuais de ANOP – Referencial de Controle de Avaliação de PP do TCU	NBASP 9020 Normativos Internos POPs	Norma da Síntese de Evidências Metodologia do Banco Mundial
Tribunais referenciais	25 TCs	TCE/PE; TCE/PI; TCM/RJ e TCU	TCE/ES e TCE/GO	TCE/RN e TCE/SC

REFERÊNCIAS

- Bardin, L. (2016). *Análise de conteúdo* (1a). Almedina.
- Bitencourt, C. M., & Tavares, A. A. (2022). Avaliação de políticas públicas no contexto do federalismo cooperativo brasileiro. *A&C - Revista de Direito Administrativo & Constitucional*, 22(90), 171. <https://doi.org/10.21056/aec.v22i90.1736>
- Botelho, Marcos & Albuquerque, Larisse. (2024). Ministério Público e Tribunal de Contas como competente na fiscalização das políticas públicas. *Revista do Direito Público*. 19. 50-65. DOI:10.5433/1980-511X.2024.v19. n.2.44598.
- Boullosa, R. de F., Breyner R. de O., Araújo, E. T. de, & Guss, A. F. (2021). Por um antimanual de avaliação de políticas públicas. *Revista Brasileira de Avaliação*, 10(1). e100521. <https://doi: 10.4322/rbaval202110005>.
- Bourgeois, I., Maltais, S. (2023). Translating evaluation policy into practice in government organizations. *American Journal of Evaluation*, Vol. 44 (3) 353-373. <https://doi: 10.1177/10982140221079837>.
- Brasil. (1988). *Constituição da República Federativa do Brasil*. http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm
- Carden, F., & Alkin, M. C. (2012). *Raízes da avaliação: uma perspectiva internacional*. 8.
- Carvalho, A. (2021). Auditoria operacional como método de avaliação de políticas públicas. [Dissertação de mestrado (Programa *Stricto Sensu* em Políticas Públicas)] - Universidade Católica de Brasília. <https://bdtd.ucb.br:8443/jspui/handle/tede/2775>
- Carvalho, A., Tófoli, P.V., & Moreira, T. B. S. (2022). Auditoria Operacional como método de avaliação de políticas públicas. *Revista Razão Contábeis e Finanças*, volume 13, n. 2.
- Castro, S. H. R., & Castro, R. R. (2021). Avaliação de políticas públicas: uma nova fronteira para o Controle Externo e pilar estruturante da democracia. *Revista Controle - Doutrina e Artigos*, 19(1), 21–38. <https://doi.org/10.32586/rcda.v19i1.676>
- Christie, C. A., & Alkin, M. C. (2008). Evaluation theory tree re-examined. *Studies in Educational Evaluation*, 34(3), 131–135. <https://doi.org/10.1016/j.stueduc.2008.07.001>
- Chudzij, L. F. (2020). Controle das políticas públicas: uma apreciação sob a ótica do Poder Legislativo, Poder Judiciário e Tribunal de Contas. *Revista Digital de Direito Administrativo*. 7. 275-296. [10.11606/issn.2319-0558.v7i1p275;296](https://doi.org/10.11606/issn.2319-0558.v7i1p275;296).

- Costa, F. L., & Castanhar, J. C. (2003). Avaliação de programas públicos: desafios conceituais e metodológicos. *Revista De Administração Pública*, 37(5), 969 a 992. Recuperado de <https://periodicos.fgv.br/rap/article/view/6509>
- Cunda, D. Z. G. (2011). Controle de políticas públicas pelos Tribunais de Contas: tutela da efetividade dos direitos e deveres fundamentais. *Revista Brasileira de Políticas Públicas*. DOI: 10.5102/rbpp.v1i2.1270
- Distrito Federal (2009). *Institui o plano de carreira, cargos e remunerações dos serviços Auxiliares do Tribunal de Contas do Distrito Federal e dá outras providências*. https://www.sinj.df.gov.br/sinj/Norma/60794/Lei_4356.html#capII_art6_incII_incIII_add
- Gomes, S. R. S. (2020). Controle de políticas públicas: fundamentos e diretrizes para atuação dos Tribunais de Contas. *Caderno de Direito e Políticas Públicas*, 2(1), pp. 174-196
- Heberger, A. E., Christie, C. A. & Alkin, M. C. A. (2010). Bibliometric analysis of the academic influences of and on evaluation theorists published works. *American Journal of Evaluation* (Sage Journals), 31 (1) 24-44. <https://doi.org/10.1177/1098214009354120>
- Instituto de Pesquisas Avançadas – IPEA. (2018a). *Avaliação de políticas públicas: guia prático de análise ex post*, volume 2. Brasília, DF.
- Instituto de Pesquisas Avançadas – Ipea. (2018b). *Avaliação de políticas públicas: guia prático de análise ex ante*, volume 1. Brasília, DF.
- Jannuzzi, P. de M. (2018). Mitos do desenho quase-experimental na avaliação de programas. *NAU Social*, 9(16). <https://doi.org/10.9771/ns.v9i16.31419>
- Jannuzzi, P. de M. (2021). Economia política e avaliação em políticas públicas no Brasil pós-2014. *Cadernos de Saúde Coletiva* 29(spe):103–14. <https://doi.org/10.1590/1414-462x202199010366>.
- Jannuzzi, P. de M. (2022). A importância do contexto institucional, político e ideacional na avaliação de políticas públicas. *Revista Brasileira de Avaliação* 11(2):e113722. <https://doi.org/10.4322/rbaval202211037>.
- Jannuzzi, P. de M. & Mello, J.M.P. (2023). Delineamentos experimentais na avaliação de políticas públicas: Usos e abusos. *Estudos em Avaliação Educacional*, 34, e09956. <https://doi.org/10.18222/ea.v34.9956>

- Lemire, S., Peck, L. R., & Porowski, A. (2020). The growth of the evaluation tree in the policy analysis forest: recent developments in evaluation. *Policy Studies Journal*. V. 48, Issue S1.
- Leoni, F. (2023). A legitimidade democrática e o controle de políticas públicas: o que dizer sobre os Tribunais de Contas? *Revista Controle: Doutrina e Artigos*, 21(1), 15-39.
<https://doi.org/10.32586/rcda.v21i1.804>
- Lima, E. C. P. (2023). *Os Tribunais de Contas e as Políticas Públicas: uma introdução*. Fórum, p. 19-30. ISBN 978=-65-5518-596-6.
- Lima, W. A. S., & Mendes, V. L. P. S. (2021). Avaliação da implementação de políticas e programas públicos no Brasil: uma discussão das dimensões analíticas. *Revista Meta: Avaliação*, 13(40), 674. <https://doi.org/10.22347/2175-2753v13i40.3480>
- Lopes, G. A. F., Grin, E. J., Teixeira, M. A. C., & Rehner, M. (2023). Tribunais de Contas Estaduais: controle ou avaliação de políticas públicas? Em P. Palotti, E. C. Licio, S. Gomes, C. I. Segatto, & A. L. N. Silva (Orgs.). *E os Estados? Federalismo, relações intergovernamentais e políticas públicas no Brasil contemporâneo* (pp. 351-454). Ipea.
<https://doi.org/10.38116/9786556350509cap16>
- Luz, B. B. C., Guarido Filho, E. R., & Sousa, M. M. (2021). Legal, normativo e responsivo: frames no debate sobre a reforma dos Tribunais de Contas brasileiros. *Revista de Administração Pública*, 55(4), 861-880. <https://doi.org/10.1590/0034-761220200394>.
- Luz, B. B. C., Sousa, M. M., & Guarido Filho, E. R. (2019). Legitimacy accounts no debate sobre a reforma dos Tribunais de Contas brasileiros. *Revista de Administração Contemporânea*, 23(3), 311-330. <https://doi.org/10.1590/1982-7849rac2019180300>.
Disponível em: <https://www.portaldeperiodicos.idp.edu.br/cadernovirtual/article/view/7040>
- Macedo, J. P. L. (2023). O Tribunal de Contas como instância de articulação interinstitucional na implementação de políticas públicas. *Revista de Direito Administrativo e Gestão Pública*, 9(1), 50-71
- Medeiros, J. L. (2023). A atuação do Tribunal de Contas e as políticas públicas de saúde. *Revista Caderno Virtual*, 1, 1-19
- Montrosse-Moorhead, B., Schröter, D., & Becho, L. W. (2024). The garden of evaluation approaches. *American Journal of Evaluation*, 45(2), 166–185.
<https://doi.org/10.1177/10982140231216667>

- Nelson, N. G. N. (2020). Território: análise sobre a fiscalização de políticas públicas. *IRB - Instituto Rui Barbosa*. <https://irbcontas.org.br/irb-territorio-analise-da-fiscalizacao-de-politicas-publicas/>
- Normas Brasileiras de Auditoria do Setor Público (2013). *Princípios fundamentais de auditoria operacional*. NBASP 300. Instituto Rui Barbosa.
- Normas Brasileiras de Auditoria do Setor Público (2021). *Avaliação de políticas públicas*. NBASP 9020. Instituto Rui Barbosa.
- Nunes, G. S. de F., Monteiro, M. A. A., & Secchi, L. (2022). Perspectivas sobre a atuação do Tribunal de Contas do Estado de Santa Catarina na avaliação de políticas públicas. *IX Encontro Brasileiro de Administração Pública*. São Paulo, SP, Brasil. <https://sbap.org.br/ebap/index.php/home/article/view/702>
- OCDE. *Organização para a Cooperação e o Desenvolvimento Econômico*. (2022). Recommendation of the Council on Public Policy Evaluation. <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0478>
- Oliveira, L. R. D., & Passador, C. S. (2019). Ensaio teórico sobre as avaliações de políticas públicas. *Cadernos EBAPE.BR*, 17(2), 324–337. <https://doi.org/10.1590/1679-395169657>
- Oliveira, T. C., & Mendonça, M. L. C. (2020). Os Tribunais de Contas na Constituição Federal e as interações do controle com as políticas públicas. *Inovação & Tecnologia Social*, 2(6), 23–38. <https://doi.org/10.47455/2675-0090.2020.2.6.7744>
- Patton, Michael. (2023). *Public policy evaluation: origins and evolution*. Doi: 10.4337/9781800884892.00008.
- Queiroz, L. de F. N. de, & Capelari, M. G. M. (2020). Condições para avaliação de resultados em políticas públicas: um debate sobre o papel da institucionalidade. *Revista de Administração Pública* 54(3):433–47. <https://doi:10.1590/0034-761220190258>.
- Rabiee, F. (2004). Focus-group interview and data analysis. *Proceedings of the Nutrition Society*, 63(4), 655–660. doi:10.1079/PNS2004399.
- Ramalho, R. & Graziane Pinto, E. (2024). Tribunais de Contas e planejamento estatal: uma concepção do controle de legitimidade para o fortalecimento de políticas públicas. *Caderno da Escola Paulista de Contas Públicas*. Cadernos, [S.l.], v. 1, n. 14, p. 93-120, jun. 2025. ISSN 2595-2412. <<https://www.tce.sp.gov.br/epcp/cadernos/index.php/CM/article/view/355>>

- Ramos, M. P., & Schabbach, L. M. (2012). O estado da arte da avaliação de políticas públicas: conceituação e exemplos de avaliação no Brasil. *Revista de Administração Pública*, 46(5), 1271–1294. <https://doi.org/10.1590/S0034-76122012000500005>.
- Sachsida, A. (2018). Políticas públicas: avaliando mais de meio trilhão de reais em gastos públicos (Org.). *Ipea*. 595 p. il.; gráfs. ISBN: 9788578113339.
<https://repositorio.ipea.gov.br/handle/11058/8750>
- Salama, C., & Picalarga, S. (2024). Promoting policy evaluation across government: The contribution of the OECD recommendation to public policy evaluation. *Evaluation*, 30(3), 327-337. <https://doi.org/10.1177/13563890241234699>
- Senra, E. M. D. A., Elliot, L. G., Da Silva, L. R. B., & Bastos, T. M. (2024). Bases epistemológicas da avaliação: Entre saberes e práticas. *Revista Meta: Avaliação*, 16(51), 463. <https://doi.org/10.22347/2175-2753v16i51.4984>
- Tribunal de Contas da União. (2020). Manual de auditoria operacional (4ª ed.). Secretaria-Geral de Controle Externo.
https://portal.tcu.gov.br/data/files/F2/73/02/68/7335671023455957E18818A8/Manual_auditoria_operacional_4_edicao.pdf
- Tinoco, D. dos S., Souza, L. M., & Oliveira, A.B. (2011). Avaliação de políticas públicas: modelos tradicional e pluralista. *Revista Políticas Públicas*. V.15, n.2, p. 305-313.
- Trevisan, A. P., & Van Bellen, H. M. (2008). Avaliação de políticas públicas: uma revisão teórica de um campo em construção. *Revista de Administração Pública*, 42(3), 529–550. <https://doi.org/10.1590/S0034-76122008000300005>.
- Tribunal de Contas da União. (2020). *Referencial de controle de políticas públicas*. Brasília, DF.
- Vedung, E. (2010). Four waves of evaluation diffusion. *Evaluation*, 16(3), 263–277. <https://doi.org/10.1177/1356389010372452>.
- Willeman, M. M. (2020). *Accountability democrática e o desenho institucional dos Tribunais de Contas no Brasil* (2ª ed.). Fórum.
- Wollmann, H. ([s.d.]). 26 *Avaliação de políticas e pesquisa de avaliação*.

Apêndice A

Roteiro dos grupos focais

Orientação geral: o termo “Políticas Públicas” engloba a avaliação de Programas e Atividades governamentais.

I – QUALIFICAÇÃO DOS PARTICIPANTES DO GRUPO FOCAL

TC:

Data:

Participantes do grupo focal:

II – QUESTÕES DO GRUPO FOCAL

A – Tema 1: Particularidade das Avaliações de Políticas Públicas

- 1)** Como e em que setor (s) têm sido realizadas as avaliações de Políticas Públicas no TC (setor específico ou vários setores, de forma difusa)? Quem realiza o “controle/acompanhamento/revisão” destas avaliações (especificar os níveis de revisão dentro da Unidade Técnica)? Qual o nível de maturidade deste tipo de fiscalização no TC (inicial/intermediário/avançado - justificar)?
- 2)** Quais particularidades diferenciam, ou não, as avaliações de Políticas Públicas das avaliações realizadas como “Auditorias Operacionais? (Eficácia? Eficiência? Economicidade? Efetividade? Pertinência? Relevância? Sustentabilidade? Análise custo-benefício? Satisfação do cliente? Equidade? Outros?)
- 3)** Quais auditorias mais recentes sobre avaliações de políticas públicas podem ser indicadas como “paradigmáticas” no âmbito do TC (citar de 3 a 5 documentos representativos deste tipo de fiscalização no TC, como, por exemplo, Relatórios de Auditoria)?
- 4)** Quais normativos são utilizados para a realização da avaliação de políticas públicas no seu TC? (Manual de Auditoria Operacional? Normativos específicos? Detalhar)
- 5)** Quais normativos, orientações e organismos internacionais são (ou podem) ser considerados como “guias” ou “fontes de consulta” para a realização de avaliações? (além da NBASP 9020 e normativos do órgão)
- 6)** Em alguma fase da avaliação de PP são realizadas análises e/ou ponderações quanto ao contexto institucional, político e ideacional que cercam a política/programa a ser avaliado? (Em caso afirmativo, especificar em quais contextos e em que medida impactam na avaliação)

B – Tema 2: Abordagens utilizadas nas avaliações

- 7)** Quais “abordagens avaliativas” ou “modelos de avaliação” têm sido utilizados (ou inspirado) na/a realização de avaliações de políticas públicas no seu TC?
- 8)** Quais ações/procedimentos têm sido adotados, ou podem ser adotados, a fim de assegurar a observância de um Método/Metodologia na realização das avaliações?
- 9)** Qual uso das avaliações se pretende alcançar com a sua realização? (para que servem as avaliações?)
- 10)** Quais ações/procedimentos têm sido adotados, ou podem ser adotados a fim de fomentar o uso efetivo das avaliações realizadas? (justiça social? Participação e envolvimento dos stakeholders, público-alvo e partes envolvidas? Outros?)
- 11)** Quais critérios ou padrões têm sido adotados, ou podem ser adotados, para a emissão de juízo de valor sobre o objeto de avaliações de Políticas Públicas (Mérito e Valor)? (justiça social? Outros?)
- 12)** No processo de avaliação, quem decide e em que grau/profundidade nas fases de planejamento, execução e consolidação/divulgação de resultados? (validação externa? Feedback com gestores? Processo participativo?)
- 13)** É possível indicar “tendências” (em abordagens, em metodologias, para fomentar o uso, para a delimitação do valor e mérito) para a realização de avaliações de políticas públicas?
- 14)** Quais os principais “desafios” (metodológicos, valoração e uso)” e/ou sugestões de melhoria vocês apontam para a realização de avaliações em Avaliações de Políticas Públicas no âmbito dos TCs? (limitações, dificuldades e sugestões quanto aos métodos, valoração e uso).

Apêndice B

Relação dos 16 relatórios de fiscalizações analisados no estudo

Tribunal	Tipo de avaliação e nome da PP ou equivalente examinada
TCU	ANOP - Programa Bolsa Família (2025)
	ANOP - Programa Mais Médicos (2025)
	ANOP - Política Nacional Aldir Blanc de Fomento à Cultura (2025)
TCE/ES	Avaliação Executiva da Política Estadual de Cofinanciamento dos Serviços Especializados em reabilitação para pessoas com deficiência intelectual e transtorno do espectro autista (SERDIA) (2025)
TCE/GO	Avaliação da Política Estadual de Atenção à Saúde Materno Infantil (2025)
TCE/PE	Levantamento Prévio - Educação inclusiva - Neuropatias (2025)
	Levantamento Prévio - Operação transporte escolar seguro (2023)
TCE/PI	ANOP - Governança e Gestão da SESAPI nos Hospitais geridos por OSS, período 2023/2024 (2024)
	ANOP - Educação em tempo integral (2024)
	ANOP - Políticas de ressocialização da pessoa privada de liberdade e do egresso do sistema prisional (2024)
TCE/RN	Síntese de Evidências - Oferta de vagas em creches municipais (2025)
TCM/RJ	ANOP - População de rua (2024)
	ANOP - Eficiência na gestão de leitos do Hospital Municipal Miguel Couto (HMMC)
	ANOP - Eficiência da Rede de Atenção Psicossocial (RAPS)
TCE/SC	Levantamento Econométrico - Eficiência dos hospitais estaduais de SC (2019)
	Levantamento Econométrico - Avaliação de Impacto do Programa Ensino Médio Inovador (ProEMI) em SC (2019)

Nota: Fonte: elaborado pela autora.

Anexo A

Critérios e escala de graduação das dimensões de avaliação

Graduação (do ausente ao plenamente existente) / dimensões	0	1	2	3
Valores <i>Refere-se à extensão em que uma avaliação inclui o surgimento e o uso de valores em uma avaliação. Valores incluem as crenças, atitudes e ideias das pessoas envolvidas na avaliação sobre o que é valioso, bom, importante, útil, desejado, necessário ou preferido. Os valores orientam, implícita ou explicitamente, o que acontece em cada estágio do processo e como o trabalho em cada estágio é realizado.</i>	Desconhecido/ausente	Os valores não têm função na avaliação. A abordagem argumenta que os avaliadores minimizam a influência dos valores.	Os valores são importantes, mas a abordagem de avaliação não desempenha um papel ativo na divulgação desses valores.	Os valores são fundamentais, e a abordagem de avaliação desempenha um papel ativo na divulgação e no uso desses valores para orientar a avaliação.
Valorização <i>A valorização refere-se medida em que uma avaliação inclui um processo implícito ou explícito de determinação do mérito, do valor ou da importância de algo.</i>	Desconhecido/ausente	A valorização não tem lugar na avaliação, e não são gerados julgamentos avaliativos.	A valorização é importante, mas a abordagem de avaliação não explica o papel dos avaliadores de facilitar um julgamento sobre mérito, valor ou significância.	A valorização é importante, e é função da avaliação envolver ou facilitar um julgamento sobre o mérito, o valor ou a importância.
Ativismo pela justiça social <i>Ativismo pela justiça social refere-se a extensão em que uma avaliação reflete a adoção de medidas claras em apoio a uma causa, e seu posicionamento de defesa ou ativismo como o objetivo principal das atividades de avaliação.</i>	Desconhecido/ausente	A defesa ou o ativismo não é central na avaliação.	A defesa da mudança social é usada para estruturar a avaliação.	O quadro ativismo pela justiça é usado e posiciona a avaliação como um impulsionador ou ferramenta para a mudança social.
Contexto <i>O contexto refere-se à extensão em que as avaliações atendem direta e ativamente aos contextos ou sistemas culturais, históricos e/ou políticos que as cercam.</i>	Desconhecido/ausente	O contexto não é levado em consideração. A avaliação é, em sua maior parte, descontextualizada.	O contexto é uma consideração e é inventariado, mas não direciona as atividades de avaliação (por exemplo, o contexto é usado principalmente no estágio de planejamento e pode ou não ser essencial na interpretação dos dados).	O contexto direciona a conceitualização e as atividades da avaliação. A abordagem considera essencial o atender e integrar o contexto e inclui processos para fazer isso durante toda a avaliação.

Gradação (do ausente ao plenamente existente)/ dimensões	0	1	2	3
Promoção do uso <i>A promoção do uso refere-se à medida em que a avaliação orienta os avaliadores a facilitar direta e ativamente o uso. Esse uso pode ser o uso dos resultados da avaliação ou do conhecimento adquirido por meio do processo de participação em uma avaliação. O uso pode ser imediato e amplo, ou lento e constante, ocorrendo ao longo do tempo.</i>	Desconhecido/ ausente	O uso não é levado em consideração; a avaliação fornece apenas resultados avaliativos.	O uso é importante, mas a avaliação não tem um papel ativo na promoção de seu uso.	O uso é essencial, e a avaliação toma medidas ativas para maximizar ou influenciar o uso.
Envolvimento no processo de avaliação <i>Engajamento refere-se a quem está envolvido no planejamento, na interpretação, no relatório e na tomada de decisões da avaliação. Esses grupos de pessoas podem incluir aqueles que trabalham na concepção, implementação e/ou gerenciamento de uma avaliação (por exemplo, doadores, financiadores, contribuintes), aqueles que são os destinatários imediatos de um programa (por exemplo, participantes do programa ou aqueles que recebem serviços) e aqueles que não são destinatários diretos, mas que se beneficiam mesmo assim (por exemplo, famílias de pessoas que participaram do programa, outros que realizam atividades semelhantes).</i>	Desconhecido/ ausente	Apenas os grupos que detêm poder formal sobre a avaliação, como os financiadores, patrocinadores ou a equipe de avaliação, estão envolvidos na tomada de decisões sobre a avaliação.	Aqueles que detêm o poder formal, juntamente com os grupos que implementam as atividades do programa, estão envolvidos na tomada de decisões da avaliação.	Uma variedade de grupos está envolvida na tomada de decisões sobre a avaliação, incluindo aqueles com poder formal, implementadores de programas, usuários pretendidos, beneficiários de programas, participantes, membros da comunidade ou organizações parceiras.
Profundidade do envolvimento <i>A profundidade do envolvimento refere-se ao grau de envolvimento de diferentes grupos de pessoas ao longo de uma avaliação e em quais funções.</i>	Desconhecido/ ausente	Ninguém limita o envolvimento nas fases de avaliação.	Os envolvidos são consultados em algumas ou em todas as fases da avaliação.	Os envolvidos são sócios, codiretores ou diretores em algumas ou em todas as fases de avaliação.
Dinâmica de poder na tomada de decisões de avaliação <i>A dinâmica de poder na tomada de decisões de avaliação é sobre quem está envolvido na tomada de decisões e como.</i>	Desconhecido/ ausente	Tomada de decisão de cima para baixo por aqueles com poder formal (por exemplo, financiadores, doadores, comissários).	Tomada de decisão cooperativa entre a equipe de avaliação e os grupos envolvidos.	Tomada de decisão colaborativa e inclusiva envolvendo uma ampla gama de grupos engajados.

Nota. Adaptado de Montrosse-Moorhead et al. (2024).