TEXTO PARA DISCUSSÃO

2858

SISTEMAS E CICLOS DE MONITORAMENTO E AVALIAÇÃO: RECOMENDAÇÕES DA ANÁLISE *EX ANTE* DE POLÍTICAS PÚBLICAS E DE PROGRAMAS GOVERNAMENTAIS

ANTONIO LASSANCE

ipea

TEXTO PARA DISCUSSÃO

2858

Rio de Janeiro, março de 2023

SISTEMAS E CICLOS DE MONITORAMENTO E AVALIAÇÃO: RECOMENDAÇÕES DA ANÁLISE *EX ANTE* DE POLÍTICAS PÚBLICAS E DE PROGRAMAS GOVERNAMENTAIS[1]

ANTONIO LASSANCE[2]

1. Agradeço a Ivan Tiago Oliveira, Fábio Véras e José Eduardo Brandão pelo apoio institucional ao desenvolvimento deste estudo. A Martha Cassiolato, Ronaldo Garcia, Paulo Jannuzzi, Leandro Couto, Joana Alencar e André Pinelli agradeço pelos esclarecimentos importantes e textos elucidativos sobre algumas questões essenciais que precisavam receber uma explicação mais amigável, aplicada e nem sempre fácil de ser encontrada. A Kathleen Sousa Oliveira Machado e a toda a equipe do grupo de trabalho de análise *ex ante* da política de atenção especializada à saúde agradeço pela oportunidade de ter participado da experiência de maior amplitude e profundidade em análise *ex ante* do governo federal brasileiro, até o momento. Também sou grato a Guilherme Mansur Dias, Pedro Vilela, Diana Coutinho e Paula Mortari da Costa, da Escola Nacional de Administração Pública (Enap), e a Igor Vinicius de Souza Geracy, Rodrigo Leandro de Moura e Daniel Grabois, do Ministério da Economia, pelo aprendizado que me proporcionaram em trabalhos de análise *ex ante* dos mais diferentes escopos. Nenhuma das pessoas citadas tem qualquer responsabilidade sobre eventuais erros ou omissões neste trabalho.

2. Técnico de planejamento e pesquisa na Diretoria de Estudos Internacionais do Instituto de Pesquisa Econômica Aplicada (Dinte/Ipea); e doutor em ciência política.

Governo Federal

Ministério do Planejamento e Orçamento
Ministra Simone Nassar Tebet

ipea **Instituto de Pesquisa Econômica Aplicada**

Fundação pública vinculada ao Ministério do Planejamento e Orçamento, o Ipea fornece suporte técnico e institucional às ações governamentais – possibilitando a formulação de inúmeras políticas públicas e programas de desenvolvimento brasileiros – e disponibiliza, para a sociedade, pesquisas e estudos realizados por seus técnicos.

Presidenta
LUCIANA MENDES SANTOS SERVO

Diretor de Desenvolvimento Institucional (substituto)
FERNANDO GAIGER SILVEIRA

Diretora de Estudos e Políticas do Estado, das Instituições e da Democracia (substituta)
LUSENI MARIA CORDEIRO DE AQUINO

Diretor de Estudos e Políticas Macroeconômicas (substituto)
CLÁUDIO ROBERTO AMITRANO

Diretor de Estudos e Políticas Regionais, Urbanas e Ambientais (substituto)
ARISTIDES MONTEIRO NETO

Diretora de Estudos e Políticas Setoriais, de Inovação, Regulação e Infraestrutura (substituta)
FERNANDA DE NEGRI

Diretor de Estudos e Políticas Sociais (substituto)
CARLOS HENRIQUE LEITE CORSEUIL

Diretor de Estudos Internacionais (substituto)
FÁBIO VÉRAS SOARES

Coordenador-Geral de Imprensa e Comunicação Social
JOÃO CLÁUDIO GARCIA RODRIGUES LIMA

OUVIDORIA: http://www.ipea.gov.br/ouvidoria
URL: http://www.ipea.gov.br

Texto para Discussão

Publicação seriada que divulga resultados de estudos e pesquisas em desenvolvimento pelo Ipea com o objetivo de fomentar o debate e oferecer subsídios à formulação e avaliação de políticas públicas.

© Instituto de Pesquisa Econômica Aplicada – **ipea** 2023

Texto para discussão / Instituto de Pesquisa Econômica Aplicada.- Brasília : Rio de Janeiro : Ipea , 1990-

ISSN 1415-4765

1.Brasil. 2.Aspectos Econômicos. 3.Aspectos Sociais. I. Instituto de Pesquisa Econômica Aplicada.

CDD 330.908

As publicações do Ipea estão disponíveis para *download* gratuito nos formatos PDF (todas) e EPUB (livros e periódicos). Acesse: http://www.ipea.gov.br/portal/publicacoes

As opiniões emitidas nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do Instituto de Pesquisa Econômica Aplicada ou do Ministério do Planejamento e Orçamento.

É permitida a reprodução deste texto e dos dados nele contidos, desde que citada a fonte. Reproduções para fins comerciais são proibidas.

JEL: D04; H11; H43.
DOI: http://dx.doi.org/10.38116/td2858

SUMÁRIO

SINOPSE
ABSTRACT
1 INTRODUÇÃO 6
2 O QUE É MONITORAMENTO? PARA QUE SERVE? 7
2.1 O monitoramento fornece a matéria-prima para a avaliação .. 7
2.2 O fio de Ariadne dos indicadores 8
2.3 Já existem indicadores em uso? 11
2.4 Revise e, se precisar, construa novos indicadores 16
3 PREPARANDO O CAMINHO PARA A AVALIAÇÃO *EX POST* 30
3.1 Consistência *ex ante* e coerência *ex post* 30
3.2 Perguntas de avaliação 31
3.3 Tipos de avaliação *ex post* 32
3.4 Estabeleça o ciclo de monitoramento e avaliação 41
4 CONSIDERAÇÕES E RECOMENDAÇÕES FINAIS 42
REFERÊNCIAS 44

SINOPSE

Um dos propósitos da análise *ex ante* é justamente o de formular políticas públicas e modelar programas governamentais que já nasçam aptos ao monitoramento e avaliação. Para que possa ser baseada ou revisada a partir de evidências, cada passo dado pela ação pública deve deixar um rastro útil ao monitoramento de insumos, processos e produtos. Ao se concluir uma jornada, a avaliação, seguindo tais rastros captados pelo monitoramento, pode identificar a trilha de resultados e impactos que se sedimentaram ao longo do caminho. Este trabalho complementa o *Texto para Discussão* nº 2817, intitulado *Análise Ex Ante de Políticas Públicas: fundamentos teórico-conceituais e orientações metodológicas para a sua aplicação prática*. Enquanto aquele estudo detalha e orienta o uso de referenciais e serve de guia metodológico para a análise *ex ante*, este se dedica a mostrar como conceber e estruturar sistemas e ciclos de monitoramento e avaliação. O trabalho encoraja a formação de uma comunidade de especialistas em moderação e facilitação de atividades voltadas à análise *ex ante*. Tal qualificação é crucial para que as políticas e os programas estejam prontos para serem monitorados e avaliados de modo a evidenciar ao público sua razão de ser e seus maiores e melhores benefícios para a sociedade.

Palavras-chave: Análise *ex ante* de políticas públicas; formulação; implementação; monitoramento; avaliação; microeconomia do setor público.

ABSTRACT

The ex ante analysis aims, among other aspects, to formulate public policies and design government programs that are born ready to be monitored and evaluated. Evidence-based policies and programs may be improved only if each step taken by public action leaves some traces for the monitoring of inputs, processes, and outputs. In the long run, the monitoring system must leave clues that can be used to identify the results and impacts of the policy or program that have been sedimented along the way. This *Discussion Paper* complements the previous *Ex Ante Analysis of Public Policies: theoretical-conceptual foundations and methodological guidelines for practical purposes* (*Discussion Paper,* n. 2817). While that previous study details and guides the use of theoretical references and serves as a methodological guide for the ex ante analysis, this work is devoted to exploring how to design and implement monitoring and evaluation systems and cycles. This paper encourages the emergence of a community of experts on guidance and facilitating activities aimed at ex ante analysis. This qualification is essential for crafting policies and programs to demonstrate their *raison d'être* to the public and their biggest and best benefits to society.

Keywords: Ex ante analysis of public policies; formulation; implementation; monitoring; evaluation; microeconomic public policy.

1 INTRODUÇÃO

O propósito deste trabalho é oferecer orientações básicas sobre como aplicar fundamentos da análise *ex ante* para a formulação de políticas e o *design* de programas governamentais que nasçam prontos para serem monitorados e avaliados. Este estudo complementa o *Texto para Discussão* nº 2817, intitulado *Análise Ex Ante de Políticas Públicas: fundamentos teórico-conceituais e orientações metodológicas para a sua aplicação prática* (Lassance, 2022), e, por isso mesmo, é recomendável que seja lido após este último, que apresenta a teoria, os métodos e as técnicas que orientam os requisitos necessários e que precedem a definição e a construção de ciclos e sistemas de monitoramento e avaliação de políticas públicas e programas governamentais.

Uma das dificuldades mais comuns para que políticas e programas sejam baseados em evidências não é a incompreensão sobre a importância do monitoramento e da avaliação. A rigor, é ponto pacífico que a atividade e os resultados do setor público devam ser monitorados e avaliados. A dificuldade está no fato de que monitoramento e avaliação são normalmente pensados como fases a serem desenvolvidas *ex post*, e não *ex ante*. No entanto, quando a ação pública não é concebida para gerar informações (*ex ante*) que possam ser transformadas em dados, organizadas e depois interpretadas pelo trabalho de avaliação (*ex post*), os custos de remodelagem desse processo assoberbam-se.

Além de a maioria das políticas e programas ser concebida e estruturada por meio de processos muitas vezes caóticos, eivada de vieses e pressionada para responder a contingências críticas ou a grupos de interesse, boa parte dos que decidem e assessoram esses processos jamais foi apresentada a referenciais básicos de análise e avaliação que possam ser aplicados de modo a incluir requisitos de monitoramento e avaliação a essas fases iniciais de formulação e implementação. O mais comum é constatar a ausência de conhecimento sistematizado para a arquitetura de políticas e a engenharia de programas. O trajeto da formulação e *design* à implementação ainda é marcado pelo amadorismo, pelo voluntarismo e pelo abismo entre o que é previsto e o que é realizado.

Um primeiro passo para superar essa lacuna é disseminar conhecimento e permitir que as teorias, métodos e técnicas conhecidos ainda por poucas pessoas se transformem em aprendizado e em ferramentas de trabalho de amplo domínio público. De um assunto de burocratas ou especialistas, o ideal é que passe a ser um campo de estudos aplicado ao trabalho prático de pessoas dedicadas ao diálogo social para a solução de problemas coletivos – o que se convencionou chamar de políticas públicas.

Este trabalho consiste em um tópico sobre monitoramento e outro de avaliação. Parte-se do pressuposto básico de que o monitoramento ocorre intimamente ligado aos processos e atividades que geram os primeiros produtos, cujo propósito é realizar entregas sistemáticas que consolidem resultados aderentes aos objetivos da política. Explica-se o papel dos indicadores e como eles podem ser mais úteis e escolhidos de forma parcimoniosa. Em seguida, o tópico

ressalta a relação umbilical entre monitoramento e avaliação e a importância de que ambos sejam vistos não apenas como atividades, mas como um sistema integrado.

O tópico seguinte esmiúça a gramática da avaliação, expondo os tipos distintos de avaliação possíveis e o lugar de cada um deles. Diferentemente da propalada ideia de avaliações "padrão ouro", que reincidentemente são enaltecidas pela literatura e que mudam constantemente sobre a escolha de técnicas consideradas ideais, avaliações são efetivamente úteis quando são processos sistemáticos, cumulativos, com uma diversidade de métodos e técnicas que podem ser apropriadas para determinado tipo de problema e *timing*. Realizadas isoladamente e de forma descontextualizada, podem ser ferramentas completamente inapropriadas. A noção de que uma única técnica seja a "bola da vez", a melhor entre todas as demais, revela um reducionismo metodológico a ser superado, o que só pode ser feito com uma melhor compreensão sobre o que é o ofício de avaliação e como ele pode ser empreendido de modo integrado e com uma caixa de ferramentas que não seja composta apenas de martelos, por mais exímio que seja quem os manuseia.

As considerações finais oferecem recomendações básicas para tornar essa discussão mais amigável e melhor estruturada com base no referencial de análise *ex ante*.

2 O QUE É MONITORAMENTO? PARA QUE SERVE?

2.1 O monitoramento fornece a matéria-prima para a avaliação

O monitoramento é o processo que torna possível acompanhar se a implementação está ocorrendo e se os produtos, resultados e impactos estão aos poucos aparecendo, brotando. O monitoramento é responsável por coletar e organizar sistematicamente as informações quantitativas e qualitativas sobre o comportamento das ações e dos programas de uma política.

O binômio monitoramento e avaliação é um circuito lógico. Inclusive, a ordem dos fatores faz toda a diferença. O correto é falar em monitoramento e avaliação, nessa sequência, nunca o contrário, pois o que se avalia é aquilo que se monitorou ao longo do tempo. As duas dimensões e momentos estão diretamente conectados. Não há avaliação que possa ser feita a contento se o monitoramento não gerar as informações necessárias para se chegar a conclusões mais precisas e assertivas. O monitoramento antecede a avaliação e fornece a matéria-prima para o trabalho avaliativo.

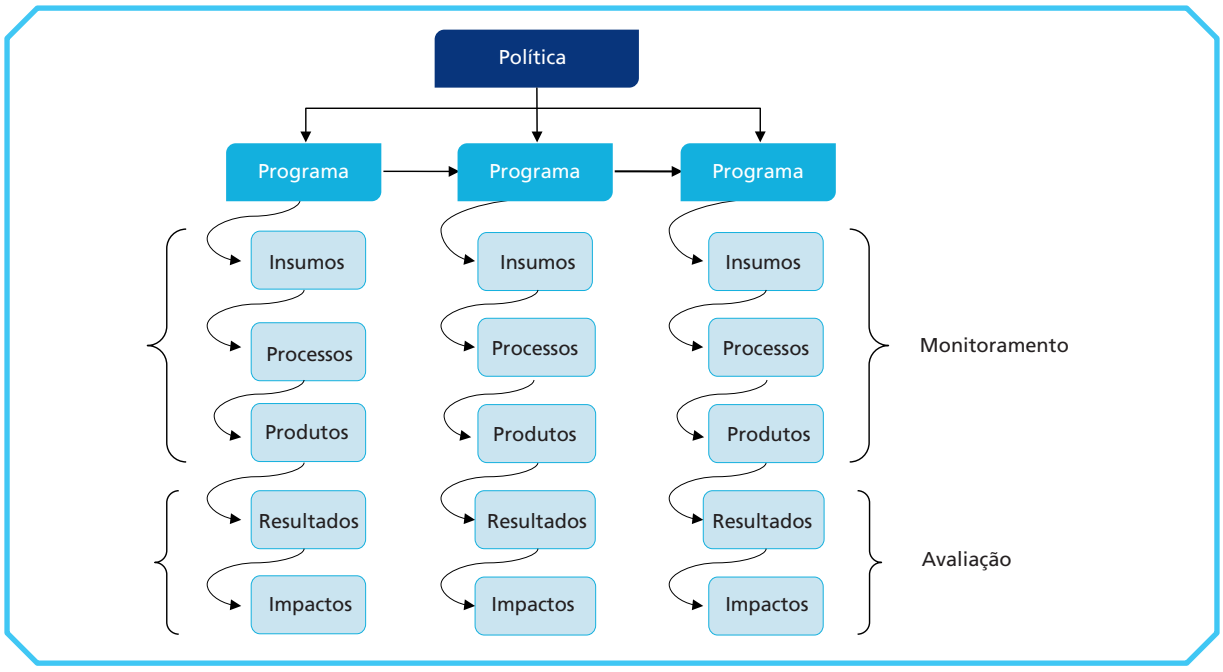
Como lembra Garcia (2015, p. 46),

> a condução de um programa não pode ser feita apenas com a avaliação de resultados finais, pois assim são restringidas as possibilidades da gestão. Os fatos geradores de resultados são os que devem ser monitorados, para que sobre eles seja possível exercer a crítica, a correção ou a confirmação. Avaliando-se apenas

> os resultados finais, será permitido averiguar os acertos ou os erros cometidos no final do programa ou das ações. E, quando diante desses últimos, chorar pelo leite derramado. (...) Os resultados finais de uma ação ou programa são construídos mediante a consecução de resultados parciais.

O monitoramento começa desde a chegada dos insumos e sua gestão pelos processos, antes mesmo de se ter o primeiro produto. Sua organização, no entanto, começa antes, na concepção da política e no *design* de programas que sejam criados para serem imediatamente monitorados e oportunamente avaliados.

FIGURA 1
Circuito do monitoramento à avaliação



Elaboração do autor.

2.2 O fio de Ariadne dos indicadores

Quando se faz uma árvore explicativa, um dos desafios é o de melhorar a formulação do problema, suas causas e consequências, especificadas em seus meandros. Essa especificação, desde os naipes da nuvem de problemas, serve para delimitar e garantir a desambiguação de seus significados, ou seja, escolher como um problema pode ser encaixado em uma grande explicação, tal e qual a peça de um quebra-cabeça.

Ao se ramificar a árvore explicativa, as causas e consequências em si já ajudam a especificar o problema, destrinchando as suas raízes e desfolhando seus efeitos. Separar esses blocos já é uma primeira categorização, refinada com a ramificação de conjuntos (naipes) de causas e consequências.

Como se viu no exemplo dos acidentes de trânsito em Lassance (2022),[1] o exercício explicativo foi, ao mesmo tempo, um trabalho de especificação de variáveis. Nesse processo, já se fez uma primeira aproximação à identificação de variáveis-chave e, quase que diretamente, de indicadores. Indicadores são construídos a partir das variáveis. Se desagregarmos os acidentes com vítimas fatais em acidentes decorrentes de colisões e os resultantes de atropelamentos, podemos ter indicadores da redução ou aumento de cada um deles, especificamente.

Passar dos problemas especificados aos indicadores é um passo essencial para o monitoramento e a avaliação. Se o problema fosse uma pessoa (e o problema sempre tem a ver diretamente com pessoas), a especificação significaria ter, além do nome e sobrenome do problema, a descrição sobre se ele é alto ou baixo, gordo ou magro, novo ou velho. Dito de outra forma, ao especificar o problema, queremos saber onde ele mora, seu número de telefone, se ele está em casa ou se saiu.

Só se pode monitorar aquilo que se consegue identificar e descrever, e só se pode avaliar aquilo que se monitorou. O caminho que vai do problema especificado e explicado ao indicador é um processo de aproximação, de *zoom*. Vamos paulatinamente nos acercando do problema, aprofundando e refinando nossa noção sobre ele e seu comportamento.

O que se quer é transformar o descritor em um indicador para que haja um bom sensor para o monitoramento. É como uma espécie de GPS da trajetória do problema e da intervenção. Gradualmente, o histórico dessa trajetória, feita pelo monitoramento, fornecerá os melhores dados para a avaliação. Este salto é que permite refinar o tipo de informação que será mais útil para bem medir a variação do problema e as consequências da intervenção.

Quando a discussão é bem estruturada, metodologicamente guiada, com pessoas experientes e em áreas que já tenham fundamentos teóricos mais sólidos e uma tradição de monitoramento, análise e avaliação, esse passo costuma ser bem mais rápido. Já se parte de algum consenso sobre os problemas centrais, as causas e consequências e seus indicadores. Uma comunidade epistêmica do assunto (a que reúne pessoas muito especializadas exatamente nisso) abrevia essa discussão com conclusões minimamente abalizadas. Áreas novas, problemas novos ou tendências desconhecidas exigirão um pouco mais de trabalho, assim como a construção de indicadores que talvez nem mesmo existam. Talvez os programas e as ações aconteçam sem que se tenha

1. Para mais informações, conferir página 65 em diante em Lassance (2022).

pensado na maneira mais simples e direta de coletar informações e produzir dados a partir da rotina diária de atividades (debate-se adiante no texto um exemplo de como isso pode ser pensado).

No processo de aproximação acerca do problema da política, as causas, a serem tratadas pelos programas, dirão quais são os indicadores de resultado. As consequências dirão quais são os indicadores de impacto. Pela lógica, se obtemos bons resultados nos programas, eles alteram positivamente os indicadores que descrevem o comportamento daquela causa. Se todos os programas apresentarem resultados positivos sistematicamente, eles irão contribuir diretamente para que o problema tenha uma alteração também positiva em suas mais graves consequências, isto é, em seus impactos.

Para que essa jornada aconteça e essas mudanças ocorram, é necessário preparar as sementes (insumos), plantar (processos), regar (alimentar os primeiros frutos) e colher (resultados e impactos). Cada um desses momentos precisa ser monitorado. Significa que é essencial acompanhar a chegada dos insumos, a execução dos processos, a entrega dos primeiros produtos, a geração de resultados e a consolidação dos impactos. Para cada momento, haverá um conjunto ou uma cesta de indicadores – como se diz tecnicamente, um sistema de indicadores.

Então, nessa longa caminhada, o sistema de indicadores funciona justamente como um fio de Ariadne capaz de nos guiar em um labirinto de informações, com grandes volumes de dados. Cada vez mais, há uma profusão de informações produzidas administrativamente, algumas delas úteis, muitas nem tanto. Os indicadores servem ao propósito de sintetizá-las. Uma escolha parcimoniosa desses indicadores permitirá seguir esse fio lógico dos insumos aos impactos sem nos perdermos em muitos números que podem até ser importantes administrativamente, mas nem sempre são significativos no monitoramento. Apesar da chegada da era dos *big data*, o mais importante não é ter muitos dados, mas os dados certos. Vale a velha regra de que nem tudo que pode ser contado conta, e nem tudo que conta pode ser contado.

Ademais, é bom evitar uma longa lista de indicadores, pois conseguir dados é algo custoso e nem todo número existente é justificável como indicador do fenômeno que se quer medir e acompanhar. O número disponível pode não ser representativo do problema. Ter poucos indicadores, mas que sejam robustos e sistematicamente atualizados, é melhor do que ter muitos indicadores precários e pouco representativos.

O equilíbrio está em combinar um conjunto razoável de indicadores, mas apenas os que sejam absolutamente necessários para entendermos o que está acontecendo, em cada dimensão e a cada momento: nas dimensões distintas – a da política, dos programas e das ações –; nas fases de implementação, da geração dos primeiros produtos, no momento em que começam a aparecer

os primeiros resultados e, no longo prazo, quando serão colhidos os impactos; no monitoramento da implementação ou na avaliação de resultados e impactos. É praticamente uma maratona. Por isso que a ideia de sistema de indicadores é bastante coerente.

Ao mesmo tempo, quanto mais se puder economizar fôlego no uso de indicadores, melhor. Assim, para que tenhamos pernas para essa maratona, é preciso muita parcimônia. Deve-se reduzir o número de indicadores de cada fase para escolher sempre os melhores e mais módicos de serem colhidos. Só assim é possível caminhar no labirinto de problemas e encontrar saídas objetivas, sem se perder por meandros cheios de confusões, algumas delas até sedutoras, como sereias à espera de quem se desvie do rumo e opte por caminhos que enganosamente parecem ser mais agradáveis.

2.3 Já existem indicadores em uso?

Na fase inicial, quando basicamente se estabelecem os objetivos da política, os indicadores serão essencialmente os de medição do problema – para ver se ele está se tornando maior ou menor, mais grave ou mais mitigado – e de impacto – nesse caso, mais relacionados às consequências, e não às causas dos problemas.

Quem de fato deve produzir uma boa cesta de indicadores a respeito das causas são os programas. É na fase de modelagem dos programas que temos a chance de melhor especificá-los e dali extrair os melhores indicadores, de modo mais preciso e realista. São os programas, como se viu, que basicamente precisam se preocupar e indicar quais são as causas do problema central da política.

Na dimensão dos programas, os resultados esperados são consequências de melhoria ou superação de parte do problema. É para isso que serve um programa: atacar um problema em um dos flancos e gerar consequências positivas, mesmo que parciais. Da mesma forma acontece com as ações. Para cada grande problema em que o programa incide, as ações atacam causas mais microscópicas e, sobre elas, produzem ações benéficas.

Se olharmos bem esse fio lógico que une causas e consequências quase indefinidamente, veremos que a relação entre política, programas e ações é quase como uma matriosca, aquele brinquedo inventado pelos russos formado por bonecas iguais, com diferenças de escala e de hierarquia, indo sempre da maior para a menor (uma boneca menor não consegue "engolir" uma maior). No caso, o que se quer é que a ação na ponta tenha a cara do programa, e que este esteja à imagem e semelhança da política. Essa sequência também vale para os

indicadores. Seu processo de refinamento é sucessivo, sequencial, meticuloso, de refinamento após refinamento, do macro para o micro. É preciso debulhar a matriosca e deduzir os indicadores sistematicamente.

No entanto, há um problema prático e corriqueiro: nem sempre se trabalha sob condições ideais. Há urgências que se impõem e às vezes não é possível dar tempo ao tempo e aguardar que grupos de trabalho e unidades de monitoramento e avaliação façam esse trabalho meticuloso. Sendo assim, nem sempre é viável especificar que os indicadores possam decorrer de uma análise *ex ante* completa, com redesenho de programas e ações calcadas em mapeamentos de processo.

Nesse caso, é comum que o monitoramento se faça com os dados disponíveis, os reais, e não os ideais, e com variáveis *proxies*, substitutos que não se ligam às causas diretas do problema. A alternativa é usar imediatamente o que se tem à mão. À medida que a análise *ex ante* reformular a política, que os programas forem sendo remodelados e que as ações forem mapeadas, em seus insumos e processos, os indicadores podem ser redefinidos e melhor especificados. Os antigos indicadores podem ser confirmados ou substituídos por outros, numa lista maior ou menor, mais parcimoniosa.

O bom senso recomenda que é melhor ter algum do que nenhum sistema de indicadores. Só tendo algo à mão é possível chegar à conclusão de que se trata de algo insuficiente e que precisa de aprimoramentos.

2.3.1 Alguns exemplos

Digamos que um dos problemas tidos como centrais em uma política pública de saúde seja a existência de doenças, comorbidades e condições de saúde consideradas muito graves e que exijam atenção especializada, e não apenas atenção básica. Suponhamos ainda que isso requeira uma política específica, como um dos braços da política de saúde mais ampla. Detalhar os tipos de doenças, comorbidades e condições de saúde, de modo mais específico, já é um bom começo. Isso será a especificação das variáveis do problema. Câncer, doenças coronárias, doenças congênitas ou adquiridas, transtornos psicológicos, entre outras, são as componentes ou variáveis do que chamamos de problemas de saúde que requerem atenção especializada.

Daí já se explica que o público dessa política é aquele afetado por qualquer dessas doenças ou por uma combinação delas, o que as ciências da saúde chamam de comorbidades. Também se supõe, em uma teoria que é a mais simples, mas muito importante, que qualquer alteração em uma das componentes, para mais ou para menos, irá alterar o problema central.

Quanto menos pessoas manifestarem os problemas de saúde indicados, menor será o problema central, e vice-versa, quando o agravamento de alguma de suas componentes o torna maior.

Essa especificação ramificada é o primeiro passo para construirmos indicadores. Se o problema ganhar escala, será ou por alguma de suas componentes ou todas elas ao mesmo tempo. No caso específico, o indicador será a leitura do problema pelas taxas de cada doença.

O que se espera, em uma política de saúde, pode ser medido por indicadores. Um deles seria a diminuição da incidência de casos de doenças desse tipo. O sucesso dessa política depende sobretudo do braço especializado na atenção básica. A atenção especializada será bem-sucedida em duas frentes de atuação, a do diagnóstico e a do tratamento. A teoria existente é de que quanto mais cedo ocorrer essa detecção, menos complexo e custoso é o tratamento e maiores as chances de cura do paciente.

Indicadores como o número de exames preventivos realizados no país, se diminuírem muito em relação à média histórica, nos permitem supor, pela teoria, de que haverá um aumento de casos graves não detectados precocemente. Indicadores de diminuição do número de equipamentos em operação capazes de realizar exames mais sofisticados (como mamógrafos) nos levam a crer que o aumento do número de casos pode demorar a ser detectado. Um aumento no tempo médio de espera na fila, entre o diagnóstico e o tratamento, indica que a taxa de mortalidade dessas doenças tende a crescer pelo fato de as pessoas obterem tratamentos meramente paliativos ou nem mesmo isso.

Note que a leitura dos indicadores deve ser orientada pela teoria da mudança (Lassance, 2022, p. 69-75) e pela teoria do programa (Lassance, 2022, p. 78-81). Por isso se pode dizer que a política e os programas têm base em evidência: o indicador denota uma interpretação que confirma o sucesso ou o fracasso das ações em curso ou expõe se as ações previstas estão realmente em curso ou não. Esses indicadores podem ainda ser granularizados até níveis mais microscópicos de modo a entender por que determinados problemas macro estão ocorrendo. Nesse sentido, funcionam como sensores do monitoramento da implementação. Quanto mais isso for possível, mais antecipadamente os problemas poderão ser detectados.

Suponha o exemplo também hipotético de uma prefeitura que passou a exigir que o atendimento de saúde fosse feito por meio de prontuários digitais. A médica adentra a unidade de saúde e registra sua entrada com sua impressão digital. O registro de ponto naquele momento marca sua chegada ao local de trabalho e significa o início da cronometragem de sua atuação. O monitoramento tem início.

Em seguida, a médica abre o aplicativo de atendimento, pelo seu celular ou no computador da unidade. Ali já está lançado o primeiro atendimento a ser feito. A primeira paciente já aguarda na sala de espera. O cadastro foi preenchido assim que ela deu entrada na unidade de saúde. Logo depois, a paciente passou pela triagem e foi encaminhada a uma médica especialista, no caso, aquela que acabou de chegar ao hospital. Quando a médica chama, pelo aplicativo, essa primeira pessoa a ser atendida, é feito o registro de que o atendimento foi iniciado. A paciente é atendida e seu prontuário é preenchido eletronicamente, no aplicativo. Encerrada a consulta, a médica chama a próxima paciente.

Esse processo simples de registro gera automaticamente inúmeros dados de monitoramento. Entre outros:

- total de pacientes atendidos por ano, mês, semana, dia e por horários específicos;
- perfil dos pacientes atendidos, por gênero, idade, tipo de comorbidade ou condição de saúde; e
- tipos de atendimento e tratamento a que cada paciente foi submetido.

É possível obter daí um rol importante de indicadores:

- porcentagem de pacientes atendidos por cada médico;
- tempo médio de atendimento de cada paciente por cada médico;
- tempo médio de atraso nos agendamentos;
- porcentagem de exames solicitados por médicos;
- registros do atendimento da ouvidoria, com reclamações; e
- porcentagem de sucesso no tratamento por doenças ou comorbidades.

É muito viável obter boas conclusões do que anda bem, do que anda mal e do que há de pior na unidade de saúde com esse conjunto de informações. Daí é possível tomar decisões e expedir orientações que permitam correções e uma atenção particularizada da gestão a cada profissional.

O monitoramento de um dia, ao se consolidar ao longo de um ou mais anos, pode ensejar um ciclo de avaliação, ou seja, um momento em que se pode analisar mais profundamente o que está indo bem ou mal e tomar decisões gerenciais importantes, conforme exemplificado a seguir.

- A gerência do hospital pode verificar quando há um maior afluxo de pacientes e se elas podem ser bem atendidas pelo total de profissionais da unidade. Esse dado deveria ser importante, entre outras decisões, para montar as escalas de férias e folgas dos profissionais.

- Pode-se verificar o excesso ou a deficiência de profissionais como um todo ou em determinadas especialidades. Isso pode justificar o pedido de reforço na equipe, para que venham mais profissionais, ou a realocação interna.

- Em nível macro, os dados de todas as unidades de saúde podem indicar aquelas de melhor *performance* (*benchmarks*).

- Pode-se medir o grau de eficiência e efetividade de cada profissional, não apenas pelo tempo de atendimento prestado por cada qual, considerando especificidades. Por exemplo, idosos ou pessoas com determinadas condições de saúde mais graves ou complexas podem exigir um tempo maior de atendimento do que pacientes apenas gripados.

As diferenças de eficiência e efetividade (resolubilidade e satisfação de pacientes com o atendimento) podem motivar reuniões gerenciais e aprendizados para estimular que alguns profissionais contem suas experiências e ensinem a aprimorar o atendimento. Profissionais que são muito "fora da curva", em termos de sua *performance* mais positiva ou negativa, podem ser chamados para conversas reservadas com suas chefias, seja para ensinar, seja para aprender. A particularização das evidências do problema não deve ser usada para constranger, mas, antes, para ver o que está acontecendo, por que problemas essas pessoas estão passando e de que maneira podem ser ajudadas.

É preciso exaltar e ao mesmo tempo ter muita prudência com o que são ou parecem ser as melhores práticas (*benchmarks*). A avaliação precisa investigar a fundo e evitar conclusões precipitadas, afastando problemas de viés. Por exemplo: as melhores unidades têm melhor desempenho porque estão melhor providas de profissionais, recursos e equipamentos? Ou porque a unidade conta com profissionais mais experientes? Ou porque a gestão realmente inovou em algo? Que inovações foram essas? Elas podem ser reproduzidas por outras unidades? Em suma, há vários segredos a serem decifrados.

A melhora ou piora agregada nesses indicadores explica e antecipa parte significativa da melhora ou piora dos indicadores de saúde do município. Lembrando que há determinantes sociais e econômicos que vão além da saúde e que podem ocasionar situações de piora e crise, como é o caso de pandemias; desastres; poluição do ar, da água, das ruas; desemprego; criminalidade,

entre tantos outros. Os determinantes são fatores exógenos, ou seja, estão além do alcance direto da política de saúde, mas afetam diretamente as questões relacionadas a ela.

Os indicadores de monitoramento da implementação, portanto, são instrumentos para checar se o que foi previsto de ser feito está sendo bem desenvolvido. A expectativa é a de que, se a implementação de fato acontecer, os resultados e impactos positivos aparecerão. Eles aparecerão nos ciclos de avaliação. Se algo que havia sido previsto não está sendo feito, não há como se observar melhorias de resultados. Ao contrário, é possível estimar uma piora significativa deles. E, finalmente, se tudo foi feito e bem feito e não produziu resultados, é sinal de que a teoria está errada e a intervenção não tratou o problema certo da forma correta.

A avaliação não só detecta problemas, mas tira conclusões e recomenda decisões mais transparentes e fundamentadas (baseadas em evidências). Ela pode, por exemplo, indicar e quantificar a necessidade de abertura de concursos, sugerir a realocação de pessoal e recursos entre unidades, transformar inovações localizadas em programa ou mesmo em diretriz da política, entre tantas outras possibilidades.

Esse é o papel mais nobre do monitoramento e da avaliação: gerar aprendizado, entender a lógica dos processos, identificar seus gargalos e orientar decisões.

2.4 Revise e, se precisar, construa novos indicadores

2.4.1 Antes de mais nada, evite o fetichismo dos dados

Construir indicadores é um dos objetivos essenciais da análise *ex ante*. Indicadores escolhidos apenas *ex post* podem ser um sinal de que talvez se tenha jogado uma flecha e só depois se levou o alvo ao local mais conveniente para dizer que acertou na mosca. Ao contrário, o indicador deve nascer intimamente associado aos objetivos da política e dos programas para depois ser o ponto focal da avaliação *ex post*. Esses indicadores serão a prova de que o que foi previsto, na teoria da mudança e do programa, foi ou está sendo alcançado.

Como lembra Paulo Jannuzzi, é bom começar por um alerta que parece uma obviedade: o de saber que o indicador apenas indica. A preocupação é a de evitar uma prática corrente de substituir o conceito sobre o qual o indicador é construído, e que é a chave para interpretar seu significado, pelas medidas em si, como se os dados tivessem vida própria. Como exemplo, diz Jannuzzi (2002, p. 56):

> embora definidos muitas vezes de forma bastante abrangente, os conceitos são operacionalmente banalizados, como se os indicadores e índices criados fossem a expressão exata, mais válida ou ideal dos conceitos indicados. Assim, por exemplo, a avaliação da melhoria das condições de vida ou desenvolvimento humano em países, regiões e municípios reduz-se a uma apreciação da variação do indicador construído. Não tendo havido modificação no indicador, não haveria eventuais avanços ou retrocessos das condições de vida ou desenvolvimento humano, ainda que fossem realizados (ou não) esforços de políticas para mudança social em uma dimensão não contemplada pela medida.
>
> Há situações em que, inclusive, primeiro se define a medida para depois inferir o conceito indicado, uma deturpação flagrante da boa prática metodológica no campo das ciências sociais aplicadas.
>
> Imagina-se que, pela combinação de diversos indicadores sociais, referidos a múltiplas dimensões da realidade e agregados através de algum processo estatístico sofisticado, pode-se chegar a um indicador final com conteúdo social substantivo, de pronta identificação com o mundo empírico. Esta mitificação das técnicas quantitativas nas ciências sociais hoje em dia está por trás de muitas das propostas de novos indicadores.

Tal mitificação das técnicas e o fetichismo dos dados ganharam impulso ainda maior na era dos *big data*. Se imagina, erroneamente, que os dados falam por si e que um grande volume de dados é o que proporciona resultados mais confiáveis e interpretações atentas aos problemas de viés. Uma grande empresa de *software* de análise de dados chega a dizer em sua propaganda: "Os dados podem fazer coisas incríveis". Na verdade, os dados não fazem nada, não dizem nada, não mostram nada. Ao contrário, a análise de dados é orientada por perguntas direcionadas a testar suposições fundamentadas por teorias. Eles devem ser confrontados com expectativas para que se veja em que medida elas estão sendo confirmadas ou frustradas e por quê.

Quando uma paciente se consulta, os sintomas relatados a uma médica são indicadores de um problema. É a formação e a teoria que a médica tem sobre cada tipo de problema que a leva a submeter a paciente a determinados tipos de exame, e não a todo e qualquer exame possível e desnecessário. A investigação que se faz é para testar, por meio de exames, se as suspeitas de diagnóstico se confirmam ou não. Mesmo exames periódicos compõem uma prescrição básica para o monitoramento das condições de saúde e servem ao mesmo propósito: verificar (*check-up*) se determinadas suspeitas sobre problemas de saúde se confirmam ou não.

Então, suposições sobre avaliação orientada por indicadores, tomada de decisões orientadas por dados, governança orientada por resultados ou gestão por resultados são uma inversão completa do sentido do uso de indicadores e da análise de dados. É o equivalente a falar em medicina baseada em exames.

Nas palavras de Jannuzzi (2002, p. 56): "enfim, um indicador consistente deve estar referido a um modelo teórico ou a um modelo de intervenção social mais geral, em que estejam explicitados as variáveis e categorias analíticas relevantes e o encadeamento causal ou lógico que as relaciona".

2.4.2 Pendure seus indicadores no modelo lógico

Retome a teoria do programa, conforme o *design* feito pelo modelo lógico (Lassance, 2022, p. 78-81). A cada insumo, processo, produto, resultado e impacto, pense quais os indicadores correspondentes e necessários. Os insumos precisam ser estimados em quantidades necessárias conforme a necessidade dos processos, a previsão de produtos demandados, a expectativa de resultados e a suposição dos impactos.

Muitas vezes, esse é um trabalho de engenharia reversa. Se começa dos impactos, mesmo que eles sejam mais amplos e genéricos. Em seguida, se pergunta que resultados e em que medida (a que taxas, quantidades, valores, número de ocorrências) eles são minimamente necessários para gerar um impacto em determinado tempo. Adiante, se pergunta que produtos geram aqueles resultados, mensurados também conforme taxas, quantidades, valores ou número de ocorrências suficientes para se levar àqueles resultados. A mesma coisa é feita para processos e, finalmente, os insumos. Dessa maneira, mais uma vez se reforça a lógica do modelo e são previstos processos e insumos apenas na medida necessária para se chegar a resultados do programa e a impactos condizentes com as expectativas de impacto da política.

O grande problema, certamente, é que aquilo que o público espera de um programa e da política supera em muito a proporção de insumos disponível por aquela organização e sua capacidade de processamento e entrega. Em geral, o Estado conta com menos recursos do que poderia; eles nem sempre chegam no tempo exato, muitas vezes são desperdiçados no meio do caminho e não são entregues às pessoas certas.

Essa insuficiência dos insumos e as falhas do processo precisam ser detectadas e corrigidas, mas também devem ser levadas em consideração para refazer as contas das promessas, com base nas expectativas. Os produtos aparecerão em menor quantidade. Os resultados serão mais modestos. Os impactos não serão tão exuberantes. O tempo será maior.

Portanto, a engenharia reversa das expectativas é agora revisada com base no realismo das circunstâncias. Os recursos disponíveis e a capacidade de processá-los corretamente são fatores limitadores e também desafios. São um convite para que dirigentes, diretores e gestores garantam mais recursos ou parcerias (contar com os recursos de outros), ao mesmo tempo em que precisam urgentemente reunir sua organização para melhorar a eficiência e a produtividade dos processos.

Com os indicadores escolhidos e dimensionados conforme curvas de eficiência, ou seja, a percepção do que é possível melhorar se forem feitos aprimoramentos, é possível rever as metas inicialmente previstas. Com menos recursos e processos não tão eficientes, o tempo para se entregar produtos, alcançar resultados e consolidar impactos será maior ou o público beneficiado será mais restrito. A cada vez que as previsões de recurso ou a capacidade de gerenciamento passar por abalos (cortes orçamentários, redução da equipe, saída de parceiros dos projetos), se tal quadro não sofrer reversão, as metas precisarão ser revistas e anunciadas às autoridades dirigentes, e por essas (ou com autorização dessas) ao público.

A ausência desse requisito de transparência faz com que o ônus do que não será feito seja atribuído a fatores que abalam a reputação da organização. Ela será certamente acusada de gerir mal os recursos que tem (sendo que esses recursos podem ter sido reduzidos à metade, por exemplo), de ser incompetente e, obviamente, de não ter cumprido uma promessa feita, anunciada e injustificadamente frustrada.

Comece com indicadores já existentes. Veja em que medida eles já atendem bem ao que se precisa. Se houver problemas, tente saná-los. A última opção é a de construir um novo indicador.

Indicadores, como já dissemos, funcionarão no monitoramento como sensores. Eles demandam um grande esforço de registro, tratamento (sistematização, organização, compilação) e processamento (transformação do registro em dado disponível para a análise).

2.4.3 Indicadores são essenciais, imperfeitos e nunca dizem tudo

Um filme clássico de faroeste, dirigido pelo italiano Sérgio Leone, *The Good, the Bad and the Ugly* (*O Bom, o Mau e o Feio*, em tradução literal),[2] mostra um dilema entre pessoas que competem entre si, mas, embora rivais, precisam um do outro para alcançar um mesmo objetivo: encontrar um tesouro enterrado em um cemitério. O trunfo e o infortúnio de cada um era saber apenas uma parte do segredo. Um deles sabia a localização do cemitério, mas não da cova, e vice-versa.

2. No Brasil, o filme ganhou o título *Três Homens em Conflito*.

Essa situação é parecida com a escolha que temos que fazer em relação a indicadores. Um bom indicador pode não estar disponível, e aquele que existe talvez não seja o ideal. Procurar indefinidamente pelo indicador ideal pode envolver um alto custo e ser uma busca inútil. Além disso, assim como no filme, em que o trio está em busca de um tesouro perdido, os dados não caem do céu. Eles precisam ser localizados com precisão e desenterrados meticulosamente.

Os critérios para a escolha ou construção dos melhores indicadores são aqueles que estão colados ao que se quer medir, seja quanto ao objetivo final, seja quanto aos esforços para a sua conquista. O mesmo raciocínio que pode ser feito quanto aos objetivos deve ser seguido pelos indicadores, que serão como uma sombra que acompanha a caminhada da política e dos programas.

Objetivos bem definidos, que tradicionalmente podem ser checados pelo acrônimo SMART (*specific, measurable, attainable, realistic, time bound*) (Brasil, 2010), são o primeiro passo para a escolha de indicadores bem definidos. Estes também devem fazer jus aos mesmos critérios e ser:

- específicos: o indicador deve se ater ao que pretende representar. Evite indicadores muito amplos e genéricos e que não tenham nexo causal estabelecido com os demais (entre a política e os programas e no sequenciamento lógico da modelagem, de insumos a impactos);

- mensuráveis: o que pode ser aferido e dimensionado, conforme um dado, seja ele um valor ou uma ocorrência;

- alcançáveis: o indicador deve estar expresso em um dado que esteja ao alcance da organização, que seja viável de ser colhido em face dos recursos (humanos, materiais, financeiros, tecnológicos);

- relevantes: o indicador deve estar relacionado ao problema, ao ponto de monitoramento (Insumo? Processo? Produto? Resultado? Impacto?) e ser suficientemente representativo do que visa expressar. Melhor ainda se ele for considerado relevante por uma comunidade epistêmica que aprofunde conclusões sobre até que ponto se pode chegar a conclusões mais robustas sobre um fenômeno com aquele indicador e em que medida ele nos fornece dados comparáveis nacional ou internacionalmente; e

- tempestivos: o indicador deve fornecer a informação no tempo certo e de modo regular, sem interrupções, valendo-se de ou colaborando para uma série histórica. Isso permitirá análises de trajetória e avaliações com maior potencial comparativo.

As recomendações são boas, mas como pensar de modo prático? Como escolher ou construir indicadores que atendam bem a todos esses quesitos?

2.4.4 Comece a montar o sistema de monitoramento e avaliação

Registrar e sistematizar informações e produzir dados que possam alimentar indicadores exigem alguém que os faça. Ou esses dados já são produzidos por alguma unidade ou processo automatizado confiável ou isso precisará ser feito. É preciso que haja alguém ou uma equipe que se especialize em dar tratamento e fazer análise de dados.

Em geral, o que há nos governos é a escassez ou raquitismo de estruturas mínimas e sistemática de registro e tratamento de informações, produção e atualização de dados e divulgação na forma de dados abertos. Nesse caso, é preciso criar secretarias ou assessorias de monitoramento e avaliação, nem que seja para saber montar processos de contratação de ajuda externa. Mas a ajuda externa é inútil se não há equipe minimamente treinada para até mesmo saber o que contratar, que perguntas fazer, que soluções considerar adequadas. Do contrário, a consultoria externa apresentará uma proposta que interessa mais a ela do que ao próprio órgão público.

Quando há tais unidades ou equipes, em geral, elas são bastante reduzidas. As bases de dados normalmente estão dispersas, cheias de falhas e sem critérios e regularidade na entrada e saída de dados e na checagem de inconsistências. A falta ou precariedade de muitos dados é um indicador desse histórico de negligências com algo muito precioso: a produção de informações sobre a ação governamental e seus resultados.

O mais curioso é que não faltam informações. A pessoa da portaria que faz o registro de entrada de usuários de um serviço; a atendente telefônica; aquela que abre e encaminha um processo administrativo; a unidade que faz uma contratação de um prestador de serviços; a empresa que realiza a entrega de um produto; a ouvidoria que recebe reclamações, dúvidas ou elogios; cada uma delas gera muita informação, do início ao fim do processo. Todavia, as informações vão para um cemitério. Os dados estão lá, debaixo dessa camada de informações, se decompondo aos poucos enquanto esperam em silêncio para serem desenterrados por alguém que saiba como encontrá-los.

Organizar os fluxos de informação conforme um sistema de entrada de informações e saída de dados é, na prática, a construção do que pomposamente se chama de sistema de informação ou sistema de monitoramento e avaliação. É importante lembrar que esse sistema não é o computador, o *software* ou o aplicativo que processa esses dados. Hoje em dia, se confunde o sistema com a

solução de processamento e de geração de painéis de dados (*business intelligence* – BI). Sistema é o resultado de uma sistemática. É tudo aquilo que é recolhido, sistematizado (compilado, organizado, tratado), processado e publicado em dados abertos. O que as ferramentas digitais permitem é apenas automatizar, acelerar o processamento e visualizar melhor os dados.

Corrigir, treinar e valorizar quem faz o registro de informações, combinar essas sistemáticas entre diferentes unidades, revisar, limpar, corrigir e recuperar informações faltantes são fatores que também devem começar com um mapeamento das ações.

Um sistema de informações deve se espelhar na teoria do programa e estar colado ao modelo lógico. O monitoramento e a avaliação também devem ser pensados segundo seus insumos (recursos financeiros e tecnológicos, equipe, normas), processos (registros, cadastros, entrevistas, pesquisas), produtos (relatórios, pareceres, análises), resultados (avaliações) e impactos (recomendações de revisão da política).

Portanto, essa construção deve ser guiada pelos objetivos da política e do programa. Ou seja, o dado que se quer desenterrar é aquele de que mais se precisa para aferir se os insumos estão chegando, se os processos estão rodando, se os produtos estão saindo, se os resultados estão aparecendo e se os impactos podem ainda ser esperados. Sendo assim, o que se pergunta à organização é onde estão os dados de:

- linha de base da situação atual do problema e do público (ponto zero da política e dos programas);
- componentes e volume de entrada e saída de recursos;
- projetos e processos desenvolvidos pelas equipes e seu mapeamento de ações;
- quantificação dos produtos entregues;
- certificação de que o público efetivamente recebeu os produtos e serviços e de que esse é de fato o público desejado;
- resultados de alteração intermediária (médio prazo) da situação da linha de base; e
- efeitos de longo prazo (impactos).

Note que algumas dessas informações não são exclusivamente quantitativas. Projetos e processos desenvolvidos pelas equipes e seu mapeamento de ações são dados objetivos, mas sua

checagem é do tipo: há ou não há projetos, registro de processos desenvolvidos e mapeamento de ações? Essas informações são uma *proxy* do grau de maturidade e dizem muito sobre se os processos foram minimamente pensados e preparados e se foram executados de forma condizente com o esperado.

Informações qualitativas também podem ser quantificadas e depois transformadas em dados, se isso for útil para sumarizar a informação. Por exemplo: de todos os processos desenvolvidos, quantos o foram por equipes que planejaram projetos e mapearam ações? O monitoramento pode obter essas informações por meio de um repositório (se possível) ou por entrevistas.

Todas essas informações, para serem obtidas, precisam ser minimamente planejadas. A título de exemplo, pode-se exigir que as organizações tenham, para a sua política, planos de monitoramento e avaliação que digam e pactuem, entre várias unidades, quem fará o que, quando e como para registrar, sistematizar e processar as informações que comporão o sistema de monitoramento e avaliação. O desenlace desse plano seria justamente firmar protocolos operacionais mínimos de compromisso, que podem ser graficamente representados por fluxogramas com o mapeamento das ações de entrada e saída de dados em cada unidade responsável. Uma portaria pode selar esse pacto e firmar a governança (quem pode o que e quem faz o que) e a gestão (como se faz) do sistema.

O monitoramento começa por gerar informações mais descritivas e, em seguida, mais sintéticas (dados), para finalmente serem cobertas por uma segunda nova camada que já entra no campo avaliativo (como se verá na seção 3) e se debruça sobre: economicidade; eficácia; eficiência (em termos da quantidade e qualidade, o que inclui a satisfação do público com os produtos e serviços); equidade (perfil dos beneficiários); efetividade dos resultados dos programas; efetividade dos impactos dos programas para um dos impactos da política; e efetividade dos impactos de todos os programas para todos os impactos da política.

2.4.5 Revise regularmente as metas

Uma das funções do monitoramento é ajudar a, de novo, revisar e fixar metas mais realistas. A recomendação, desde o início, foi a de uma meta modesta na previsão do que era possível fazer, pelo pouco tempo que dura uma gestão, e ambiciosa o suficiente em sua implementação para ampliar seu arco de apoio, desenvolver parcerias e convencer sobre a importância dos resultados e impactos finais, de maneira que essas expectativas sejam alcançadas antes do tempo e, melhor ainda, que se possa ir além do prometido.

Metas não são imutáveis. Se elas foram feitas com base em pressupostos que se alteraram sobre recursos financeiros, equipes, determinações legais, parcerias, entre outras condições, podem e devem ser atualizadas sempre que necessário, justificadamente. O público deveria ser o primeiro e não o último a saber que um compromisso não será cumprido no prazo inicialmente previsto. Muitas vezes nem mesmo o dirigente é alertado previamente. Simplesmente, algo que deveria acontecer não acontece e ninguém se deu ao trabalho de explicar por quê.

A frustração de prometer e não entregar afeta não apenas a imagem do dirigente, mas a reputação do órgão público prestador de serviços. Em contrapartida, a saída simples de desenvolver ações sobre as quais não se tem qualquer expectativa e as quais não se sabe antecipar minimamente é um sinal de amadorismo e totalmente contrário à lógica da administração pública, que exige que as ações tenham justificativa, previsões (toda a previsão orçamentária é baseada nisso) e metas. Começar algo sem saber que produto, resultado ou impacto isso proporcionará é inadmissível, um indicador de mediocridade da governança e um cheque em branco para a gestão.

A melhor surpresa que o público pode ter, porém, é a de receber resultados maiores e antes que o previsto. A regra da administração pública deveria ser a de poder festejar produtos entregues com rapidez e em quantidades superiores ao anunciado, antecipar resultados e, sobretudo, criar a sensação positiva de que grandes mudanças começam a se concretizar com pequenas entregas que acontecem a cada dia. A diferença entre uma governança competente e com gestão de excelência e outras que são medíocres, muitas vezes, não está no volume de realizações, mas no tamanho das promessas. As mais bem-sucedidas são as mais parcimoniosas nas metas e mais exuberantes nas entregas.

A seleção ou construção de indicadores é um processo não apenas técnico, mas político. Envolve escolhas, pactuação, análise situacional. A técnica anda de mãos dadas com a política para protegê-la de apostas erradas. Para isso, a burocracia precisa demonstrar não o quanto ela entende de detalhes, mas o quanto vale a pena atuar de modo mais profissional, competente e mais bem organizado. Essa é a conversa a ser tecida com quem quer obter reconhecimento político mais amplo, algo justo e necessário a qualquer processo democrático.

2.4.6 Exemplificando a diferença entre variáveis, parâmetros e indicadores

Pense no exemplo simples da preparação de um bolo. Cada insumo, como farinha (*f*), açúcar (*a*), levedura química (fermento) (*q*), óleo (*o*) etc., pode ser considerado uma variável do bolo. Lembrando sempre que este exemplo é uma simplificação, uma analogia – não uma

representação, pois os estatísticos são os primeiros a nos alertar que modelos estatísticos não são receita de bolo. De todo modo, modelos estatísticos também são simplificações e bons bolos não são tão simples de serem feitos como muitos acreditam – principalmente as pessoas que já os compram feitos.

Feitas as ressalvas, voltemos ao bolo. Ele em si é uma variável. No caso, uma variável dependente, explicada por outras (farinha, açúcar, levedura). Cada variável deve ser usada conforme medidas definidas. Essas medidas são o que chamamos de parâmetros. Os parâmetros são tão fundamentais no resultado do bolo quanto as suas variáveis, pois eles determinam o quanto de cada variável é necessário para obtermos um resultado replicável – o mesmo bolo, muitas vezes, ou o dobro da quantidade de bolo, se multiplicamos por dois toda a equação.

Digamos que, nessa receita, o bolo precise de 3 xícaras (0,6 kg) como quantidade ideal de farinha de trigo, 2 xícaras (0,4 kg) de açúcar, 1 colher de sopa (0,0075 kg) de levedura química e assim por diante. A equação de um bolo ficaria como algo assim:

$B = 0{,}6\,f + 0{,}4\,a + 0{,}0075\,q$ etc.

Para evitar a confusão entre parâmetros e indicadores, é importante diferenciar esses conceitos. Primeiro, para com isso entender que o monitoramento pode se valer de algumas variáveis ou de alguns de seus atributos indiretos para chegar a indicadores. Segundo, porque, em geral, a lista de indicadores é bem menor que a de variáveis.

A primeira tarefa do monitoramento, muito antes de se chegar aos indicadores, é mais básica. Começa por verificar se os ingredientes estão todos presentes e se eles aparecem na medida indicada. Ou seja, o monitoramento não observa só os indicadores. Boa parte de seu tempo é dedicada a checar a presença das componentes de um modelo.

Mesmo essa tarefa aparentemente simples é também fundamentada na teoria do programa. A teoria pressupõe que uma eventual ausência de ingredientes ou a insuficiência em seus parâmetros (pouca farinha, por exemplo) já compromete ou limita o processo. A ausência total já prenuncia que o produto não pode ser feito, o resultado não será gerado e os impactos jamais virão.

A ausência dos insumos em quantidades ideais revela de antemão que o produto estará aquém do necessário para atender às necessidades estimadas. Não haverá bolo suficiente para a quantidade de pessoas que se espera. Não havendo produto, o resultado é inexistente ou insuficiente para atender às expectativas. O impacto é uma frustração parcial ou generalizada.

E quais seriam os indicadores? Para os insumos, se pode imaginar que um indicador de qualidade é fundamental. A teoria, dessa vez, é a de que ingredientes de má qualidade irão gerar produtos ruins, resultados sofríveis e impactos contraproducentes. A cozinheira pode considerar como indicadores de qualidade dos insumos de seu bolo: a marca ou procedência do produto, a validade, a aparência e o odor de cada ingrediente. Na ocorrência de qualquer desvio nesses indicadores, o insumo deve ser descartado e substituído antes do processo se iniciar.

Note que, nesse caso, o indicador não é exatamente nenhuma variável ou parâmetro, mas atributos dessas variáveis. Ele sumariza ou sintetiza uma ou várias informações ao mesmo tempo. Também está fundamentado em teoria, mesmo que elementar, o que permite levantar uma expectativa de resultado, e a sua função é sempre a de orientar a tomada de decisão, como ocorreu nessa fase de monitoramento dos insumos. O indicador serviu para decidir se os ingredientes podiam ser usados.

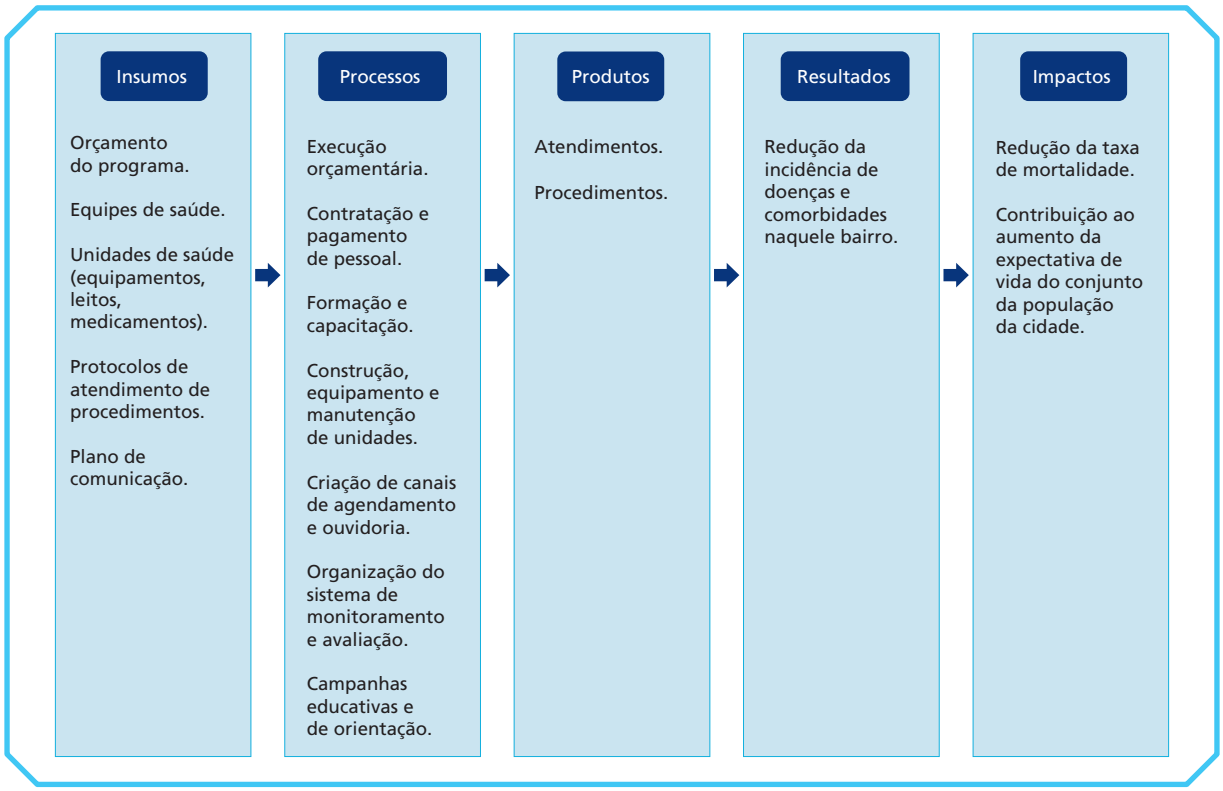
Indicadores que não cumprem essas funções são desnecessários e não se justificam como dignos de maior atenção.

2.4.7 Exemplo de como os indicadores devem estar encadeados

A mesma lógica de encadeamento que orientou a modelagem do programa deve se mostrar consistente o suficiente para costurar uma relação coerente entre os indicadores. Sendo estes a sombra da trajetória de cada componente (dos insumos aos impactos), nada mais natural do que exigir que essa relação se estabeleça fortemente ao longo de toda a implementação. É como se um indicador prenunciasse que os próximos números estão a caminho de apresentar bons resultados.

No exemplo representado pela figura 2, em um programa de atenção à saúde, os insumos são: i) orçamento; ii) equipes; iii) unidades de atendimento; iv) protocolos de atendimento; e v) plano de comunicação.

FIGURA 2
Exemplo de encadeamento de indicadores: programa de atenção à saúde



Elaboração do autor.

Note que, a partir dos processos, há indicadores que podem ser pensados para o monitoramento da implementação. Por exemplo, a porcentagem de execução orçamentária pode indicar uma *proxy* da medida em que as ações já estão ocorrendo e de quão perto ou longe elas estão de serem concluídas. O andamento da formação e capacitação, que é uma variável qualitativa (basicamente, o que se quer medir não é exatamente a quantidade de pessoas que estão sendo capacitadas, mas o prazo de conclusão da atividade), pode antecipar quando as equipes estarão prontas para dar início aos trabalhos ou a partir de quanto (de quais parâmetros) se podem medir melhorias de qualidade do atendimento, se considerarmos processos de aperfeiçoamento profissional.

Essas equipes, enquanto são formadas, aguardam a construção e instalação de equipamentos e o início da entrada das unidades em operação para poderem dar início ao trabalho. O cronograma de obras e a contratação e formação de pessoal dirão em que momento se dá o início da gestão do programa para a entrega dos primeiros produtos (os primeiros atendimentos e procedimentos). Por sua vez, sistemas de monitoramento e avaliação darão métricas de fluxo de agendamento, avaliação de satisfação das pessoas com a prestação de serviços, reclamações sobre problemas e profissionais. Esses indicadores irão consolidar avaliações sobre a eficiência e a efetividade dos serviços prestados.

À medida que forem surgindo resultados no médio prazo, se saberá qual o impacto das ações na redução da incidência de determinadas doenças naquela região. Comparando com o Estado ou o país (tomando cuidado com comparações muito amplas pela dificuldade de se lidar com outros fatores intervenientes que são bastante distintos), mas sobretudo comparando-se com o passado daquela cidade ou localidade onde as unidades passaram a atuar, se poderá medir ou aproximar a avaliação do impacto daquelas ações em consequências mais profundas do problema.

Fica claro que, quando os processos administrativos geram os registros adequados, se torna mais fácil, quase automático, o monitoramento da implementação das ações, do resultado do programa e de sua contribuição para os impactos da política. A todo instante, um bom conjunto de indicadores oferece aos dirigentes e diretores a chance de monitorar a gestão e de produzir bons anúncios de conquistas parciais a serem comemoradas ou de alertas a serem levados em conta.

A produção sistemática de informações administrativas passa a produzir dados de monitoramento quase automaticamente, desde que os protocolos sejam estabelecidos de maneira que qualquer atividade em andamento, ao ter início, registre ligações, agendamentos, impressões digitais, prontuários, pedidos e resultados de exames, listas de medicamentos, protocolos de internação e alta etc. Esse tipo de abordagem deve se dar *ex ante*. As providências que tornam possível fazer o *design* das organizações para que ajam em conformidade com esse novo tipo de visão devem ser tomadas previamente à implementação.

Um passo necessário é convencer e motivar a governança superior, responsável pelo comando e controle, a não só permitir como estimular que as unidades operacionais revisem seus procedimentos. O pacto a ser feito é o de criar procedimentos menos regulados com o objetivo de produzir registros de comando e controle e melhor formatados a produzir dados. Em troca, ganha-se transparência interna e pública e mostra-se mais exatamente o que se está fazendo. Do contrário, o registro torna-se um mecanismo meramente burocrático, de registro de autorizações e de cumprimento de obrigações, e não um sistema de monitoramento e avaliação. O prejuízo dessa visão arcaica é imenso. Gasta-se muito tempo juntando documentos, e não realizando ações.

O que se deve ter em mente é que um bom sistema de monitoramento e avaliação cumpre plenamente o papel de ser um bom mecanismo de controle. Nesse pacto, a governança sai do dilema contraproducente entre ter mais ou ter menos comando e controle e assume como desafio ter melhor monitoramento e avaliação. As autorizações e os registros de atividades continuarão sendo feitos. A única e mais importante diferença é que sua motivação é acompanhar e impulsionar a geração de resultados, e não formalizar procedimentos para atender a requisitos de comando e controle. Ações governamentais são realizadas não para satisfazer normas e atender

a condições de órgãos de controle, mas para gerar produtos, resultados e impactos que sejam concretos, benéficos e efetivos. A ação dos órgãos de controle, porém, não será mitigada ou barrada sem que em troca se mostre um avanço no mapeamento de processos e na construção de sistemas de monitoramento que atendam a requisitos mais inteligentes.

A todo momento, e dessa vez não será diferente, o esforço é de que cada passo se torne explícito e gere dados que serão públicos e que possam ser interpretados de modo bem contextualizado, conforme uma lógica inteligível por qualquer pessoa. Mantendo a consistência de uma teoria integrada da política e dos programas, que interligue uma concepção a resultados, é plenamente possível demonstrar a relação entre o que se planta no dia a dia e os sinais do que se espera colher, paulatinamente.

2.4.8 Exemplo do papel decisivo do monitoramento: saúde durante a pandemia de covid-19

Quando eclodiu a pandemia de covid-19, em 2020, a primeira e mais importante providência adotada por muitos governos, seguindo orientações da comunidade médica, foi a de permanecer em casa. O Serviço Nacional de Saúde (National Health Service – NHS) inglês transformou a recomendação em norma, sujeita a multa para quem a desobedecesse, e acrescentou a ela outras medidas profiláticas: usar máscaras e lavar sempre as mãos ou higienizá-las com álcool em gel. Um *slogan* publicitário deixou claro que a regra tinha por trás uma teoria da ação muito clara e bem fundamentada (figura 3).

FIGURA 3
***Slogan* do NHS durante a pandemia de covid-19**

Fonte: NHS.
Obs.: Em tradução livre: Fique em Casa, Proteja o NHS, Salve Vidas.

Cartazes e adesivos foram espalhados por todo o Reino Unido. Todas as conferências e coletivas de imprensa das autoridades que tratavam do assunto traziam o *slogan*. A sequência lógica fazia exatamente o encadeamento de três teorias: a teoria da ação individual (ficar em casa), do programa (proteger o NHS, permitindo que ele tivesse condições de atender à demanda de pacientes) e da política (salvar vidas, que é o grande objetivo de uma política de saúde).

Com esses objetivos em mente e com uma diretriz de saúde bem específica (manter as pessoas em casa), vários países utilizaram indicadores de isolamento médio nacional, por região, estado, cidade e bairro. Esse indicador foi criado por meio de dados de geolocalização gerados pelo uso dos aparelhos telefônicos celulares. Cada vez que as pessoas, com seus aparelhos, saíam de casa, seu deslocamento pela cidade era automaticamente alimentado e georreferenciado. Sem revelar a identidade de cada usuário, mas analisando o comportamento coletivo, o indicador expressava aproximadamente a porcentagem de moradores que estavam cumprindo ou descumprindo essa diretriz em cada área delimitada e em que horários do dia. Com isso, era possível, pela teoria, supor reduções ou aumentos do número de contaminados e da superlotação de unidades de saúde com pacientes suspeitos de terem contraído a doença. Antes disso, os governos poderiam intensificar a fiscalização sobre essas localidades.

Note, mais uma vez, como um indicador cumpre a função de ser uma síntese informativa totalmente aderente à teoria da ação, do programa e da política, útil para orientar decisões e capaz de antecipar tendências.

3 PREPARANDO O CAMINHO PARA A AVALIAÇÃO *EX POST*

3.1 Consistência *ex ante* e coerência *ex post*

Avaliar é dar valor, conforme pesos, medidas, critérios. Como lembra Weiss (1998, p. 4), para isso precisamos de referenciais que definam melhor o que queremos medir. Não se pode fazer a avaliação certa sem réguas, termômetros, barômetros. Contudo, a própria escolha do instrumento de medida correto depende do propósito e da natureza do objeto que se quer avaliar.

Avaliar é um trabalho que só pode ser bem feito *ex post* se for bem preparado *ex ante*. Como dito anteriormente, o que se avalia é se o que foi feito está bem ou mal feito, o que só pode ser devidamente respondido com a compreensão do propósito com que se fez aquilo. A avaliação *ex post* só faz sentido retomando os requisitos da análise *ex ante* e confrontando esses pressupostos com os resultados.

O propósito da avaliação é confrontar os efeitos de uma política e seus programas com os objetivos que eles pretendiam alcançar, conforme critérios explícitos de julgamento. Por sua vez, a matéria-prima para que os avaliadores façam seu trabalho são informações fornecidas pelo monitoramento.

O que se quer na avaliação *ex post* é verificar se os efeitos da ação governamental são coerentes com a concepção que motivou a intervenção. Se não forem, o que a avaliação fará

é investigar se isso se deveu a problemas na implementação ou a falhas de consistência da concepção.

3.2 Perguntas de avaliação

Toda avaliação é motivada por perguntas avaliativas. Todas elas têm relação direta ou no mínimo indireta com a análise *ex ante*, como podemos observar nos exemplos a seguir.

- Qual o objetivo da política e dos programas? Eles foram atingidos? Parcial ou completamente?
- O público alcançado corresponde ao que se pretendia?
- As ofertas (produtos) previstas estão sendo entregues às pessoas certas, da forma certa, na quantidade e no tempo certo?
- A qualidade do que é realizado é boa? Como isso pode ser atestado?
- O público está satisfeito?
- Houve melhoria na situação das pessoas, em relação ao problema principal escolhido para tratamento?
- Houve mudanças na condição de vida das pessoas? Essas mudanças mais substantivas foram proporcionadas de alguma maneira pela contribuição dada pela política e seus programas?
- Ocorreram problemas? Onde eles se localizam? Nos insumos? Na implementação (processos e entrega dos produtos) ou na teoria da mudança ou do programa?
- Esses problemas de implementação dizem respeito à governança, à gestão ou a ambas?
- Como melhorar ou superar essas falhas?

Essas e tantas outras questões ainda mais detalhadas podem ser resumidas a duas perguntas apropriadas a cada momento do desenvolvimento de uma política ou programa, cada qual com uma subpergunta diretamente associada. São elas, em suma, as destacadas a seguir.

- O que está acontecendo (e por quê)? O que se pode corrigir?
- O que aconteceu (e por quê)? É possível aprimorar?

Um erro comum é supor que a avaliação deve escolher que programas manter e premiar (com mais recursos, por exemplo) e quais extinguir ou reduzir e penalizar. A concepção por trás desse equívoco é que isso levaria a uma competição dos gestores entre si e a um aumento da eficiência dos programas.

Ao contrário, o principal objetivo da avaliação é compreender a política e a interconexão que há entre os programas, identificar problemas que possam ser corrigidos e aprender como aperfeiçoar esses programas. Recomendar a sua extinção ou a de ações pode ser uma das conclusões da avaliação quando isso contribui com os objetivos da política e promove a otimização da modelagem dos programas.

3.3 Tipos de avaliação *ex post*

3.3.1 Avaliação formativa

A primeira pergunta básica de avaliação ("O que está acontecendo?", no gerúndio mesmo) é própria de avaliações de processo. É feita enquanto os programas estão em fase de implementação e começam a gerar seus primeiros produtos. Essas avaliações servem para detectar falhas que ocorrem de imediato ou ao longo da implementação, de modo a corrigi-las o quanto antes, evitando o desperdício de recursos (insumos) e esforços (processos). São avaliações do *making of* dos programas, por isso são chamadas de formativas. Sua realização, porém, já deixa de pertencer ao momento de análise (*ex ante*) e inaugura o da avaliação (*ex post*). Ainda assim, sem se descolar do que foi previamente formulado.

O momento da implementação é aquele que vai da chegada dos insumos e seu processamento até sua transformação em produtos. Quando ela completa um ciclo inteiro de entrega de produtos, o que se espera é que comecem a surgir os primeiros resultados. Esse momento encerra um primeiro ciclo de implementação. É como se o programa tivesse superado sua fase inicial de aprendizado formativo e alcançado um momento de maturidade.

Considera-se que, nessa fase, o programa corrigiu suas falhas e estabilizou suas incoerências e travamentos. Antes que isso aconteça, é temerário fazer avaliações mais conclusivas sobre os programas e ações. Cabe à análise *ex ante* formalizar os ciclos de avaliação somativa e formativa, delimitando especificamente a partir de que momento as avaliações mais conclusivas podem ser feitas.

Programas têm um ciclo de vida que precisa ser respeitado para evitar conclusões apressadas. Com indicadores parcimoniosos e metas realistas, é possível estabelecer marcos temporais

dos ciclos de avaliação que devam ser respeitados também pelos órgãos de controle, evitando que conclusões precipitadas comprometam a continuidade de ações que ainda precisam ser estabilizadas para produzir resultados efetivos.

3.3.2 Avaliação somativa

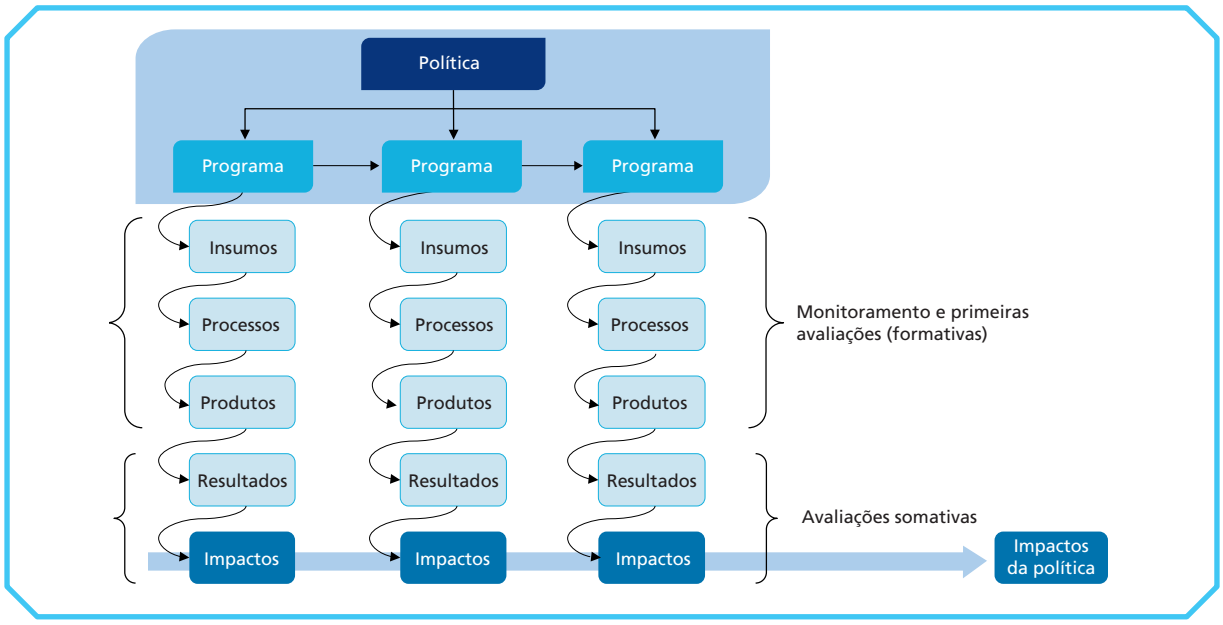
Michael Scriven, criador da tipologia que distingue a avaliação formativa da somativa (Scriven, 1967), explicava que "quando o cozinheiro prova a sopa, esta é a avaliação formativa; quando o convidado a degusta, esta é a avaliação somativa" (Scriven, 1991, p. 19, tradução nossa).[3] Quando a avaliação se debruça sobre a implementação, é formativa. Quando os resultados e impactos são seu objeto, a avaliação é somativa.

Avaliações somativas podem ser consideradas como sinônimos de avaliações de efetividade ou avaliações de impacto. Quando se debruçam sobre os resultados de um programa, as avaliações somativas podem ser especificadas como de efetividade dos resultados do programa, enquanto as classificadas como de impacto alcançam justamente os impactos daquele programa para a política. A diferença entre uma e outra está no fato de que o ciclo de maturação ou de vida de um programa, no médio prazo, pode ter apresentado apenas resultados, mas não ainda os impactos. Avaliações somativas de longo prazo, essas, sim, podem alcançar os impactos para a efetividade da política.

A lógica que preside esse ciclo é de que, havendo uma política integral e programas integrados, cada programa contribui para viabilizar pelo menos um impacto da política. Somados os impactos de todos os programas, se pode fazer uma avaliação somativa da efetividade dos impactos da política.

3. "*When the cook tastes the soup, that's formative evaluation; when the guest tastes it, that's summative evaluation*."

FIGURA 4
Avaliação somativa da política e de seus programas



Elaboração do autor.

3.3.3 Avaliação de conformidade

A avaliação de conformidade verifica e atesta se programas e ações estão sendo implementados conforme as regras. Nesse caso, o critério mais importante é a aderência à norma. A pergunta de avaliação é: foram cumpridos à risca os preceitos normativos estabelecidos pela lei e pelo próprio órgão gestor dos programas?

Quanto mais genéricas e dúbias as normas e menos específicos os planos, mais a gestão abre espaço para que essas lacunas e ambiguidades fiquem abertas à interpretação dos órgãos de controle interno e externo. Com isso, os órgãos de controle podem suprir essas dubiedades com recomendações e acórdãos e firmar entendimento diverso dos propósitos dos que formularam ou que são responsáveis pela governança superior das políticas e programas.

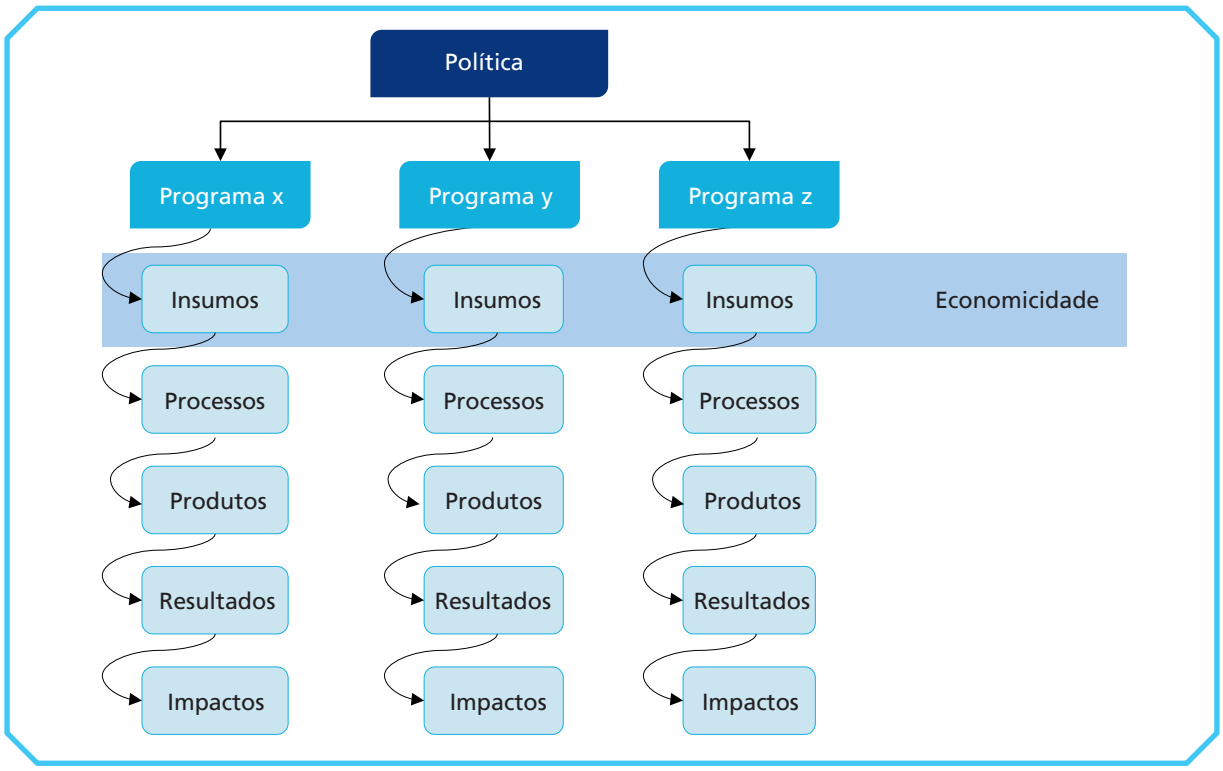
Nem todos os detalhes do planejamento e da implementação precisam se tornar normas. As normas podem e devem remeter esses detalhes a planos, análises, manuais. Esses instrumentos podem fornecer esses detalhes de modo flexível e atualizado.

Deve-se lembrar que os órgãos de controle têm expandido sua atuação e destinado suas recomendações e determinações a campos cada vez mais abrangentes da governança e gestão públicas. Portanto, também nesse aspecto, a análise *ex ante* é providencial às autoridades dirigentes das políticas, diretoras de programas e gestoras das ações.

3.3.4 Avaliação de economicidade

A avaliação de economicidade é a que verifica a aquisição dos insumos ao menor custo possível. A pergunta dessa avaliação é: os preços referentes aos insumos adquiridos são compatíveis com os de mercado? Eles resultaram em economia para a administração pública?

FIGURA 5
Avaliação de economicidade

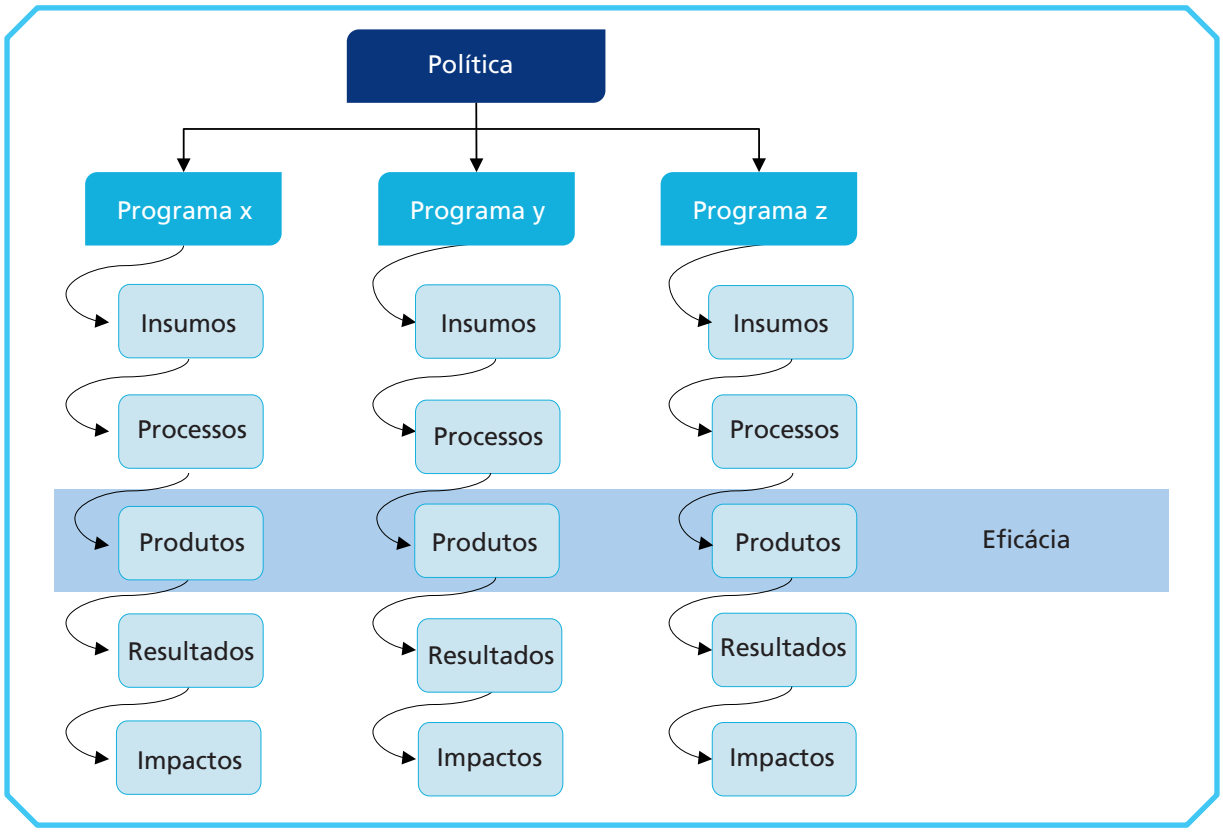


Elaboração do autor.

3.3.5 Avaliação de eficácia

A avaliação de eficácia constata se os produtos foram gerados e entregues a contento. A pergunta dessa avaliação é: os produtos são exatamente os previstos? Eles foram entregues ao seu destino e público corretos e no tempo previsto?

FIGURA 6
Avaliação de eficácia

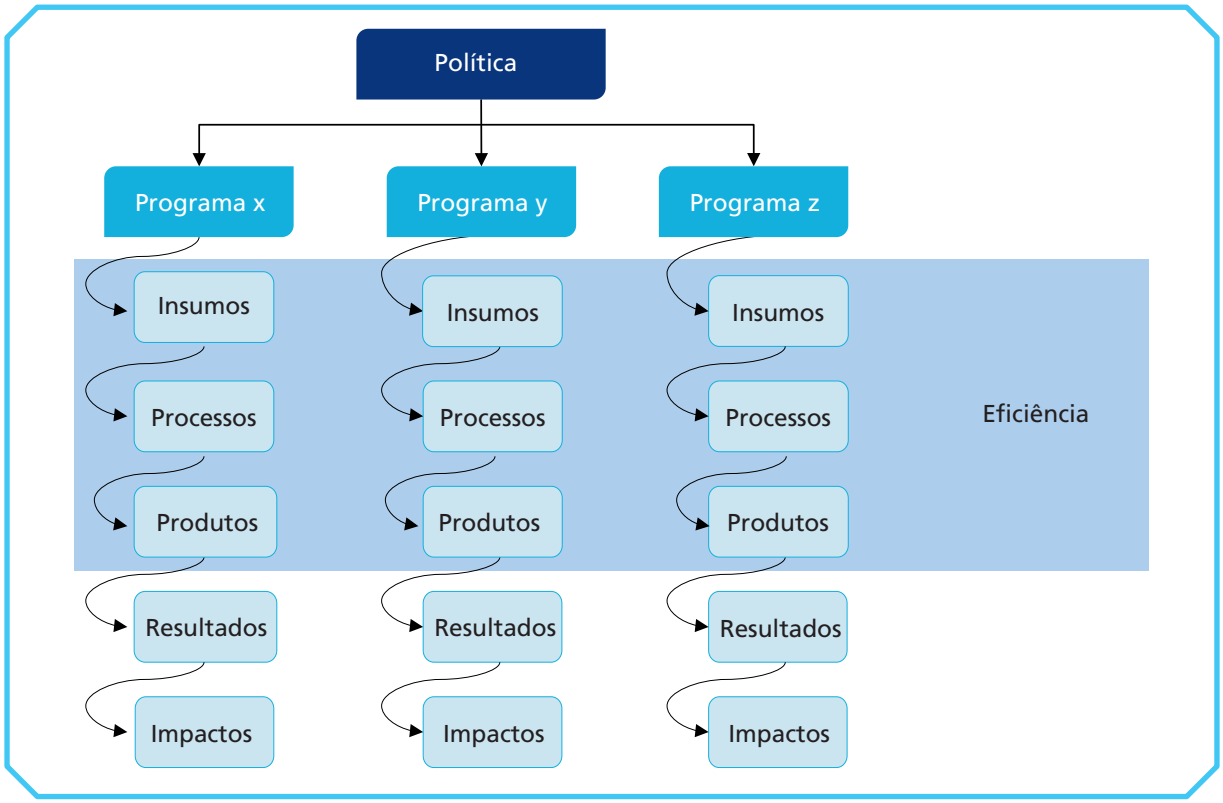


Elaboração do autor.

3.3.6 Avaliação de eficiência

A avaliação de eficiência verifica se os processos empregados aproveitam e otimizam ao máximo o uso de insumos na geração de cada produto. A pergunta dessa avaliação é: os processos, procedimentos, tecnologias, usos ou manuseio dos insumos empregados proporcionam a melhor oferta de produtos, em quantidade e qualidade, no tempo adequado?

FIGURA 7
Avaliação de eficiência



Elaboração do autor.

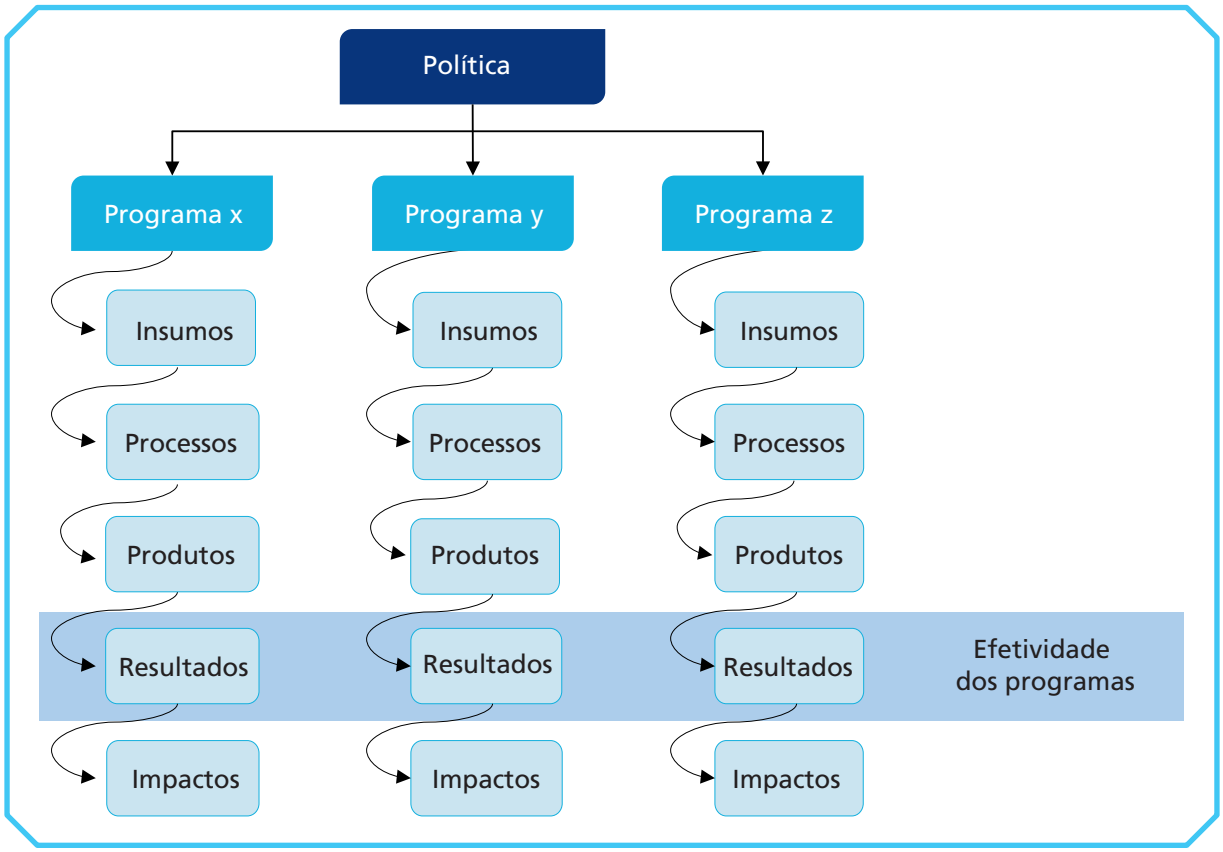
3.3.7 Avaliações de efetividade

A avaliação de efetividade busca encontrar os resultados dos programas ou da política, a partir do momento em que a implementação foi estabilizada, os programas alcançaram sua maturidade, seus resultados estão consolidados e os impactos podem ser aferidos.

Portanto, a avaliação de efetividade de um programa pergunta se há resultados de médio prazo previstos e alcançados e se já existem impactos. O impacto de um programa é um impacto parcial da política. Todo programa deve ser capaz de gerar, no longo prazo, pelo menos um impacto mais amplo que contribua com a política, revertendo consequências mais profundas de um problema. Por isso, só faz sentido e só é de fato factível fazer análises definitivas de impacto de uma política quando seu conjunto de programas cumpriu seus ciclos de resultados e foi capaz de proporcionar pelo menos seus primeiros impactos.

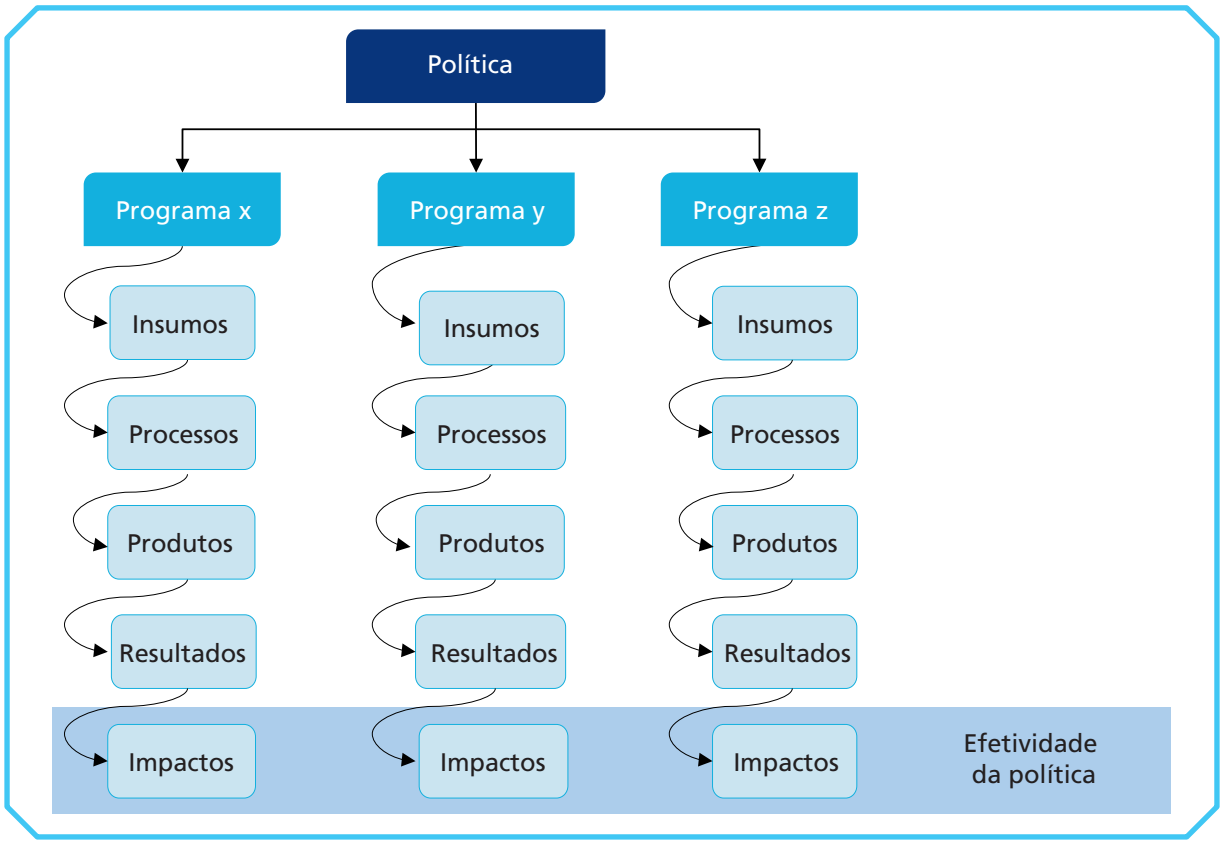
Dessa forma, são possíveis e necessárias pelo menos três avaliações distintas e sequenciadas de efetividade: i) avaliações de efetividade dos resultados de um programa; ii) avaliações de efetividade dos impactos de um programa; e iii) avaliações de efetividade dos impactos da política.

FIGURA 8
Avaliação de efetividade dos programas



Elaboração do autor.

FIGURA 9
Avaliação de efetividade da política



Elaboração do autor.

3.3.8 Avaliações somativas robustas dependem de uma sequência cumulativa de avaliações preliminares

A sequência das avaliações de economicidade, eficácia e eficiência deixa claro que as três são avaliações de processo, portanto, formativas. As avaliações de efetividade são de médio prazo, para programas, e de longo prazo, para a política, sendo elas somativas.

A noção de somatividade revela seu caráter cumulativo. Apenas uma sequência de avaliações sistemáticas permite consolidar um entendimento mais preciso sobre a real capacidade demonstrada pela política e seus programas de debelar um problema grande e complexo.

Como sempre, essas avaliações não acontecem descoladas e alienadas em relação ao *ex ante*, mas rigorosamente dedicadas a responder às perguntas: a teoria da mudança está sendo confirmada? A teoria do programa funciona? Se não, o que ocorreu? O problema está na teoria?

Se escolheu o problema errado? Se adotou a solução errada? Ou o problema está na implementação? Se fez erradamente? Em um ou em todos os programas? Ou se fez a coisa certa para um público errado? Essas são perguntas típicas de avaliação que também precisam revisitar a teoria para serem respondidas.

3.3.9 Avaliação de equidade

A avaliação de equidade busca atestar os públicos abrangidos pela política e o grau de prioridade dado pelos programas àquelas pessoas que mais precisam dos serviços públicos. O princípio da equidade parte da noção de que, sendo as pessoas submetidas a condições desiguais e a situações distintas, cabe ao Estado reconhecer essas diferenças e agir para corrigi-las.

Um exemplo do quanto o princípio da equidade tem efeito prático ocorre em qualquer unidade de saúde que siga, em sua triagem de pacientes, uma classificação de risco.[4] Nesse caso, a prioridade de atendimento considera a gravidade, em seguida, a urgência, e, por último, a ordem de chegada. Pelo princípio da igualdade, o atendimento deveria ser prestado primeiro a quem chegou antes, mas isso seria iníquo. Afinal, para casos graves, o tempo é um fator primordial e o objetivo maior da política, que é salvar vidas, estaria comprometido pela falta de um critério de equidade.

Nesse caso, a avaliação vai além de simplesmente contabilizar atendimentos, tratamentos e curas. O importante será analisar por que determinados públicos talvez não estejam sendo devidamente priorizados e recomendar a inclusão de critérios práticos que traduzam o princípio da equidade para situações muito concretas e que fazem toda a diferença.

A equidade é um princípio basilar. Deve estar presente desde a escolha do problema até a definição da estratégia da política e estender-se à formatação dos processos e à verificação da entrega dos produtos. Portanto, faz parte da avaliação formativa. No ciclo de avaliações somativas, novamente se confere a aderência dos resultados ao público-alvo definido como prioritário dos programas e o impacto final da política para um maior grau de equidade social, no âmbito da política pública em questão (saúde, assistência social, educação, emprego, mobilidade urbana etc.).

3.3.10 Métodos e técnicas de avaliação

Não confunda os tipos de avaliação com seus métodos e técnicas. As avaliações são orientadas por propósitos e por perguntas, e não por técnicas. O fetichismo de dados tem induzido ao equívoco

4. Disponível em: <https://pensesus.fiocruz.br/equidade>.

de achar que a avaliação é o método. Essa é uma completa inversão da lógica da avaliação. É o propósito e a pergunta de avaliação que levam à escolha do método, e não este que conduz as definições sobre a avaliação. Dependendo do tipo de problema a ser avaliado, alguns métodos podem ser mais adequados que outros e alguns podem se mostrar completamente inapropriados.

Métodos como o de diferenças em diferenças, pareamento, uso de variáveis instrumentais, análise envoltória de dados, regressão descontínua e as técnicas de experimentos aleatorizados, pesquisas com grupos focais ou *surveys* com amostras representativas, entre outras, são maneiras de operacionalizar uma avaliação.

Os centros e institutos de pesquisa, como o Instituto de Pesquisa Econômica Aplicada (Ipea), têm pessoal especializado na realização de avaliações utilizando as metodologias e técnicas mais adequadas. No entanto, a primeira pergunta que uma analista (econometrista, estatística ou outra) faz é justamente de que tipo de avaliação se trata, qual o propósito, qual o grau de maturidade atingido pela política e pelo programa. Só depois se perguntará: onde estão os dados? Eles são confiáveis? São completos? Foram bem colhidos e tratados? Nenhuma especialista em análise de dados, por mais habilitada que seja no uso dos métodos mais sofisticados que existem, será capaz de executar um bom trabalho de avaliação sem que a pergunta de pesquisa avaliativa esteja clara e a matéria-prima trazida pelo monitoramento seja razoável.

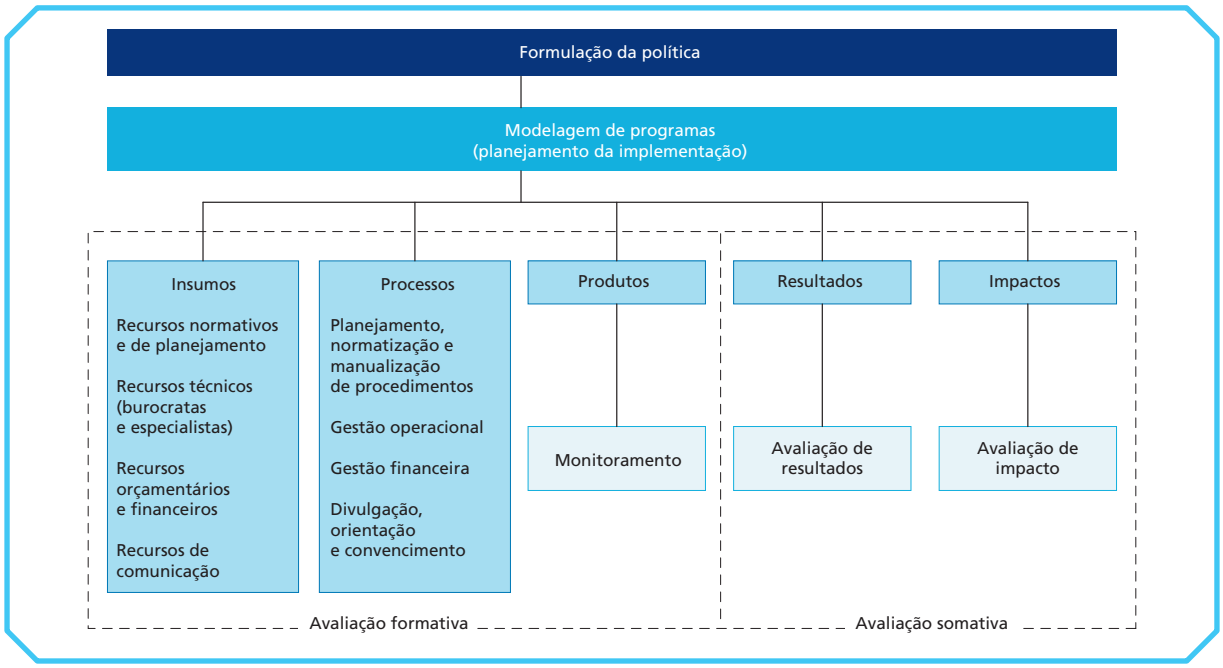
Outra recomendação importante é buscar a chamada triangulação de métodos, ou seja, combinar métodos e técnicas quantitativas e qualitativas para obter conclusões mais abrangentes e aprofundadas e também entender melhor a lógica que incide sobre o problema, assim como a maneira como as pessoas se comportam diante dele. Metodologicamente, informações quantitativas (dados) e qualitativas são complementares, e não rivais (Weiss, 1998, p. 14). Além disso, é bom levar em conta que métodos complexos precisam ser explicados aos dirigentes e ao público de forma lógica e inteligível. Sendo assim, triangular métodos pode oferecer narrativas que melhor expliquem os números.

3.4 Estabeleça o ciclo de monitoramento e avaliação

O ciclo de monitoramento e avaliação é a previsão de fases ou momentos de avaliação sequenciados. Toda análise *ex ante* precisa prever seu ciclo, estabelecendo prazos em que cada tipo de avaliação será efetuado.

Note que as avaliações são encadeadas e acompanham a modelagem dos programas (figura 10).

FIGURA 10
Ciclo de monitoramento e avaliação



Elaboração do autor.

O ciclo possibilita que o monitoramento e as avaliações obtenham um caráter cumulativo – ou somativo, como diz Scriven (1967). Assim, se não há economicidade na aquisição de insumos, a eficiência estará comprometida, pois produtos caros demais têm menor oferta – sendo, nesse sentido, menos eficientes. Não se fará mais com menos, mas menos com mais. Por sua vez, se não houver eficácia e o produto for mal feito ou não for entregue devidamente, também se compromete a eficiência. Ou seja, economicidade e eficácia são componentes da eficiência. Consequentemente, dependendo do grau de ineficiência, compromete-se parcial ou completamente a efetividade dos programas. Havendo baixa efetividade dos programas, compromete-se a efetividade da política.

4 CONSIDERAÇÕES E RECOMENDAÇÕES FINAIS

Os fundamentos teóricos, as orientações metodológicas e os exercícios de aplicação prática da análise *ex ante* são o mapa que orienta uma longa caminhada de uma política e seus programas. Quando dados os primeiros passos, o monitoramento busca identificar os sinais de que realmente se está no caminho certo previsto pelo mapa ou se são necessárias correções de rota. A avaliação, por sua vez, é a conclusão parcial ou final sobre o caminho percorrido, se foi alcançado o objetivo final, se havia trilhas alternativas mais promissoras que não foram previstas pelo mapa e se a própria viagem valeu a pena e merece novas jornadas.

A análise *ex ante* é um requisito essencial para a montagem de sistemas de monitoramento e avaliação que sejam o desdobramento de ciclos sistemáticos e cumulativos de coleta, organização e análise de dados. A consistência *ex ante* só pode ser aferida, baseada em evidências, se houver coerência *ex post*. Quem atesta tal coerência são os dados de monitoramento e os estudos de avaliação.

Como monitorar e avaliar são processos que envolvem custos financeiros e organizacionais não triviais, é possível preliminarmente se pensar em formas de otimizar ao máximo as tarefas administrativas regulares como fontes para a geração, se possível, automatizada de informações que possam ser transformadas em dados e que alimentem indicadores que serão úteis apenas se tiverem sido construídos para tal propósito, teórica e metodologicamente fundamentados. Ao mesmo tempo, esse é um bom mote para rever tarefas burocráticas que, se não servirem minimamente para fornecer dados relevantes a avaliações de economicidade, eficácia, eficiência e efetividade, talvez não façam sentido de existir.

Monitoramento e avaliação são duas dimensões indissociáveis. A avaliação tira conclusões a partir da matéria-prima fornecida pelo monitoramento. O monitoramento, que antecede a avaliação, torna possível acompanhar desde o início da implementação até os resultados que só serão gerados no médio e longo prazo.

O monitoramento e a avaliação não se prestam a apontar culpados. Esse papel cabe aos dirigentes, aos órgãos de controle e ao público. Da mesma maneira, não servem para disseminar a cizânia e a competição por recursos, mas para orientar a cooperação e a interligação de unidades de uma organização que vivem compartimentadas e brigam entre si como se fossem adversários, e não parceiros com objetivos comuns.

A presença ou a ausência de sistemas de monitoramento e de ciclos regulares de avaliação diz muito sobre a consistência e maturidade de políticas e programas e sobre a capacidade de suas equipes. Também diz muito a respeito da possibilidade ou não de se chegar a alguma conclusão razoável sobre se o que foi feito gerou resultados e se valeu ou não a pena, em termos de seu valor público. Como as jornadas de avaliação precisam sempre seguir os rastros captados desde o início pelo monitoramento, é crucial que as políticas e os programas nasçam bem concebidos e prontos (*ex ante*) para serem monitorados e avaliados. Do contrário, parafraseando uma clássica analogia de Scriven (1967), é como se houvesse um grande caldeirão borbulhante de sopa à nossa frente e não tivéssemos uma única colher para prová-la.

REFERÊNCIAS

BRASIL. Ministério do Planejamento, Orçamento e Gestão. **Indicadores de programas**: guia metodológico. Brasília: MP, 2010. Disponível em: <https://bibliotecadigital.economia.gov.br/bitstream/777/84/1/Indicadores_programas-guia_metodologico.pdf>.

CASSIOLATO, Martha; GUERESI, Simone. Como elaborar modelo lógico: roteiro para formular programas e organizar avaliação. *In*: CARDOSO JUNIOR, José Celso Cardoso; CUNHA, Alexandre dos Santos (Org.). **Planejamento e avaliação de políticas públicas**. Brasília: Ipea, 2015. p. 297-332. Disponível em: <https://bit.ly/3pL23KE>.

GARCIA, Ronaldo Coutinho. A reorganização do processo de planejamento do governo federal: o PPA 2000-2003. *In*: CARDOSO JUNIOR, José Celso Cardoso; CUNHA, Alexandre dos Santos (Org.). **Planejamento e avaliação de políticas públicas**. Brasília: Ipea, 2015. p. 17-54. Disponível em: <https://bit.ly/3QVj7JY>.

JANNUZZI, Paulo de Martino. Considerações sobre o uso, mau uso e abuso dos indicadores sociais na formulação e avaliação de políticas públicas municipais. **Revista de Administração Pública**, Rio de Janeiro, v. 36, n. 1, p. 51-72, jan./fev. 2002. Disponível em: <https://bibliotecadigital.fgv.br/ojs/index.php/rap/article/download/6427/5011>.

LASSANCE, Antonio. **Análise *ex ante* de políticas públicas**: fundamentos teórico-conceituais e orientações metodológicas para a sua aplicação prática. Rio de Janeiro: Ipea, 2022. (Texto para Discussão, n. 2817). Disponível em: <https://repositorio.ipea.gov.br/bitstream/11058/11399/2/TD_2817_Analise_ExAnte.pdf>.

SCRIVEN, Michael. The methodology of evaluation. *In*: TYLER, Ralph W.; GAGNÉ, Robert M.; SCRIVEN, Michael (Ed.). **Perspectives of curriculum evaluation**. Chicago: Rand McNally and Company, 1967.

_____. Beyond formative and summative evaluation. *In*: MCLAUGHLIN, Milbrey Wallin; PHILLIPS, Denis Charles. **Evaluation and education**: at quarter century. Chicago: University of Chicago Press, 1991. p. 19-64.

WEISS, Carol. **Evaluation**: methods for studying programs and policies. 2nd ed. Upper Saddle River: Prentice Hall, 1998.

WILDAVSKY, Aaron. **The politics of the budgetary process**. Boston: Little, Brown and Company, 1964.

Ipea – Instituto de Pesquisa Econômica Aplicada

EDITORIAL

Coordenação
Aeromilson Trajano de Mesquita

Assistentes da Coordenação
Rafael Augusto Ferreira Cardoso
Samuel Elias de Souza

Supervisão
Aline Cristine Torres da Silva Martins

Revisão
Bruna Neves de Souza da Cruz
Bruna Oliveira Ranquine da Rocha
Carlos Eduardo Gonçalves de Melo
Elaine Oliveira Couto
Luciana Bastos Dias
Rebeca Raimundo Cardoso dos Santos
Vivian Barros Volotão Santos
Débora Mello Lopes (estagiária)
Maria Eduarda Mendes Laguardia (estagiária)

Editoração
Aline Cristine Torres da Silva Martins
Mayana Mendes de Mattos
Mayara Barros da Mota

Capa
Aline Cristine Torres da Silva Martins

Projeto Gráfico
Aline Cristine Torres da Silva Martins

The manuscripts in languages other than Portuguese published herein have not been proofread.

Logotipo/slogan de governo removido em observancia ao disposto no art. 40 da Lei n 9.504/1997

Missão do Ipea

Aprimorar as políticas públicas essenciais ao desenvolvimento brasileiro por meio da produção e disseminação de conhecimentos e da assessoria ao Estado nas suas decisões estratégicas.

ISSN 1415-4765

ipea Instituto de Pesquisa Econômica Aplicada

MINISTÉRIO DO PLANEJAMENTO E ORÇAMENTO