

INTRODUÇÃO AO MONITORAMENTO E À AVALIAÇÃO DE POLÍTICAS PÚBLICAS

André Portela de Souza | Lycia Lima | Camila Soares
Gabriel Weber Costa | Juliana Camargo | Marina Ribeiro



AUTORIA:

FGV EESP CLEAR – Centro de Aprendizagem em Avaliação e Resultados para a África Lusófona e o Brasil

Diretor: André Portela de Souza

Vice-diretora: Lycia Lima

EQUIPE TÉCNICA:

Camila Mata Machado Soares

Gabriel Weber Costa

Juliana Camargo

Marina Mendes Rodrigues Ribeiro

INTRODUÇÃO

O curso de Introdução ao Monitoramento e à Avaliação de Políticas Públicas apresenta o ferramental teórico básico para que o aluno possa formular e analisar políticas públicas com base em evidências. Por meio dele, o aluno aprenderá como desenvolver a Teoria da Mudança de um programa e como desenvolver um plano de monitoramento e avaliação. Apresentaremos as ferramentas para que o aluno possa desenvolver indicadores para o plano de monitoramento da política em questão, bem como, de forma introdutória, os diferentes tipos de avaliação que podem ser realizados.

Este curso – que contribui com o eixo de atuação de formação de profissionais em habilidades para trabalhar com Monitoramento & Avaliação (M&A) da Iniciativa CLEAR – não pretende aprofundar-se nas metodologias de avaliação, entretanto, traz indicações de literatura para os alunos que desejam expandir os seus conhecimentos em cada tópico.

A Iniciativa CLEAR, criada em 2010, é um programa global de desenvolvimento de capacidades em M&A que congrega instituições acadêmicas e parceiros doadores de modo a contribuir para o uso de evidências nas tomadas de decisão em países em desenvolvimento. O CLEAR estabeleceu seis centros regionais, abrigados em instituições acadêmicas, com a missão de entregar e contribuir para replicar serviços de desenvolvimento de capacidades.

O Centro de Aprendizagem em Avaliação e Resultados para o Brasil e a África Lusófona (FGV EESP CLEAR) tem sede na Escola de Economia de São Paulo da Fundação Getulio Vargas (FGV EESP), conceituado *think tank* e instituição de ensino superior dedicada à promoção do desenvolvimento econômico e social brasileiro.

O objetivo geral é que o aluno desenvolva três conjuntos de habilidades complementares. Primeiro, compreender os conceitos de monitoramento e avaliação e a sua aplicação no ciclo da política pública. Segundo, estruturar e representar a lógica de uma intervenção, selecionando indicadores para o monitoramento. Terceiro, entender por que e quando avaliar uma política e as possibilidades metodológicas disponíveis.

De forma a cumprir o objetivo do curso, esperamos que o aluno atinja os seguintes objetivos específicos de aprendizagem:

- Compreender os conceitos de monitoramento e avaliação, bem como as suas aplicações.
- Identificar a diferença entre Teoria do Programa, Marco Lógico, Modelo Lógico e Teoria da Mudança.
- Identificar a diferença entre plano de monitoramento e plano de avaliação e saber como elaborar cada ferramenta.
- Compreender os critérios para a construção e a seleção de indicadores, assim como ter conhecimento de algumas fontes de dados.
- Compreender a diferença entre os principais tipos de avaliação de políticas.

SUMÁRIO

MÓDULO I – FUNDAMENTOS DE MONITORAMENTO E AVALIAÇÃO	7
MONITORAMENTO E AVALIAÇÃO: CONCEITOS	7
O que é monitoramento?	8
O que é avaliação?	8
Monitoramento e avaliação no ciclo da política pública	10
TRANSFORMAÇÃO DE INSUMOS EM RESULTADOS	12
Identificação de um problema com base em evidências	12
Representação da lógica de uma intervenção	14
Teoria da Mudança	16
Foco em resultados	18
PLANOS DE MONITORAMENTO E AVALIAÇÃO	20
Plano de monitoramento	20
Plano de avaliação	21
MÓDULO II – INDICADORES PARA MONITORAMENTO	23
INDICADORES E PLANO DE MONITORAMENTO	24
Seleção de indicadores para o plano de monitoramento	26
CONSTRUÇÃO DE INDICADORES PARA O PLANO DE MONITORAMENTO	29
Medidas de posição	29
Média	29
Mediana	31
Moda	32
Medidas de dispersão	32
Variância	32
Desvio padrão	33
Razão	33
Proporção	34
Percentual	34
Taxa	34
FONTES DE DADOS	36
Dados primários	37
Dados secundários	37
Ipeadata	38
Sidra	39
Inep	40
Datasus	40

MÓDULO III – INTRODUÇÃO À AVALIAÇÃO DE POLÍTICAS.....	41
POR QUE E QUANDO AVALIAR?	41
TIPOS DE AVALIAÇÃO DE POLÍTICAS	42
Avaliação <i>ex ante</i>	43
Avaliação de necessidades.....	43
Avaliação da Teoria do Programa	44
Avaliação <i>ex post</i>	45
Avaliação de desenho	46
Avaliação de processos.....	47
Avaliação de impacto	48
Avaliação de custo-benefício e avaliação de custo-efetividade	48
Avaliação executiva	50
METODOLOGIAS PARA AVALIAÇÃO DE POLÍTICAS.....	54
Métodos quantitativos	54
Métodos qualitativos.....	55
Métodos mistos.....	56
CONCLUSÃO	59
BIBLIOGRAFIA	60
FONTES DE DADOS	66
PROFESSORES-AUTORES	68
EQUIPE TÉCNICA.....	69



MÓDULO I – FUNDAMENTOS DE MONITORAMENTO E AVALIAÇÃO

O monitoramento e a avaliação são ferramentas essenciais para o planejamento e a gestão de políticas públicas baseadas em evidências. Com isso em mente, damos início a este curso apresentando essas ferramentas no arcabouço do ciclo da política pública. O objetivo deste módulo é introduzir ao leitor alguns conceitos importantes, com especial ênfase na ferramenta da Teoria da Mudança, que será utilizada ao longo de todo o curso como a base de todo o processo de monitoramento e avaliação.

Para tanto, este módulo está dividido em três unidades. Na unidade 1, vamos discutir as definições de monitoramento e avaliação, o ciclo da política e como esses componentes devem relacionar-se. A unidade 2 apresenta o conceito de Teoria do Programa e algumas das formas de representá-la, elaborando as etapas envolvidas nesse processo e a importância de se pensar em políticas públicas voltadas para resultados. Por fim, a unidade 3 demonstra brevemente a estrutura dos planos de monitoramento e avaliação, que servem como ponto de partida para os módulos seguintes.

Monitoramento e avaliação: conceitos

Tanto o monitoramento quanto a avaliação constituem funções de gestão indispensáveis que ajudam a medir e avaliar o desempenho da ação pública, mas o fazem de maneiras diferentes e em momentos distintos. O monitoramento é uma atividade de acompanhamento contínuo, que ocorre durante todo o curso de implementação da política pública. Por outro lado, a avaliação é conduzida, idealmente, de forma periódica e devidamente planejada, e é elaborada para que se obtenham respostas concretas a questões específicas, inclusive aquelas que surgem do acompanhamento contínuo dos indicadores de monitoramento.

A complementaridade dessas duas atividades é que permite as tomadas de decisão bem informadas para que uma política seja bem desenhada, planejada e executada, para que a sociedade e os atores interessados conheçam os resultados alcançados e para que seja garantida a eficiência na alocação de recursos públicos.

Uma observação sobre terminologia: ao longo deste curso, utilizaremos os termos *projeto*, *programa*, *política* e *intervenção* como sinônimos. Vale notar, entretanto, que eles são conceitos distintos. Segundo Imas e Rist (2009), um **projeto** é constituído por uma única intervenção implementada em uma ou mais localidades; um **programa** é formado por diversas atividades ou projetos relacionados, que contribuem para um objetivo em comum; e uma **política** é um conjunto de normas, diretrizes ou regras que orientam as decisões tomadas por uma organização ou um governo em uma determinada área, podendo ser composta de diversos programas. Neste curso, todos os conceitos e as ferramentas apresentados podem ser aplicados no nível de projetos, programas ou políticas, de forma que não faremos distinção entre os termos.

O que é monitoramento?

O monitoramento de uma intervenção, conforme discutido em OCDE (2002), pode ser definido como o processo contínuo e sistemático de coleta de dados e acompanhamento de indicadores específicos, que tem como objetivo informar os gestores e as demais partes interessadas sobre os progressos realizados, os objetivos atingidos e o uso dos recursos alocados.

Monitoramento é uma atividade contínua, rotineira e direcionada especialmente para o controle interno da intervenção (IMAS; RIST, 2009). Os dados coletados durante o processo de monitoramento provêm de indicadores selecionados na etapa de desenho da política que estão intimamente associados aos recursos – insumos – alocados, às atividades executadas, aos produtos e aos resultados da intervenção, bem como aos seus impactos pretendidos. Esses dados são utilizados para rastrear mudanças de desempenho ao longo do tempo, identificar tendências e padrões, adaptar estratégias e informar decisões. O seu objetivo principal é subsidiar a tomada de decisão e ajustar as linhas de atuação da ação pública de forma a torná-la mais eficiente.

Além de possibilitar a tomada de decisão rápida a partir da análise de evolução dos indicadores ao longo do período de implementação da política, a sistematização das informações coletadas no processo de monitoramento é um complemento vital do processo de avaliação, exatamente por prover muitos dos dados necessários para que uma avaliação possa ser conduzida (ROSSI; LIPSEY; FREEMAN, 2003).

O que é avaliação?

Seguindo ainda as definições de OCDE (2002), a avaliação, no contexto de políticas públicas, pode ser entendida como o processo de determinação de valor, de forma sistemática e objetiva, de uma atividade, um programa ou uma política específica, esteja ela em andamento, já finalizada ou ainda em fase de planejamento.

Avaliações são conduzidas por uma variedade de razões práticas: pode haver preocupações sobre a gestão e operação do programa, se o programa está atingindo o público-alvo desejado, se está tendo o impacto esperado e se os seus benefícios compensam os seus custos, entre diversas outras. Sob a perspectiva da gestão, o propósito de qualquer avaliação consiste em fornecer informações aos gestores de políticas públicas e habilitá-los a tomar as melhores decisões diante da concepção, da continuidade ou da reformulação de políticas públicas.

Dessa forma, uma avaliação deve ser elaborada a partir da definição das questões a que buscará responder. Várias questões de interesse podem surgir no decorrer do ciclo de uma política pública, a depender do seu estágio de maturidade: antes, durante ou depois da sua implementação. Entretanto, uma avaliação deve ser objetiva quanto ao conjunto de informações que pretende levantar, uma vez que a natureza das questões da avaliação influencia diretamente o tipo de avaliação a ser conduzida e a metodologia a ser aplicada.

Uma forma de classificar diferentes tipos de avaliação é considerar o momento em que a avaliação é conduzida (RUBIO, 2011). Uma avaliação *ex ante* é aquela que ocorre antes da implementação da política, partindo-se do diagnóstico do problema social a ser enfrentado até o momento de colocar em prática o planejamento realizado para a intervenção. O objetivo das avaliações conduzidas nessa etapa consiste em analisar a necessidade, a relevância, a coerência e a viabilidade da política a ser formulada, bem como garantir a definição clara dos objetivos a que a política se propõe e dos caminhos a serem trilhados para alcançar esses resultados esperados; e a consistência lógica desses objetivos com todas as etapas planejadas da política. Dessa forma, as informações levantadas nessa fase habilitam o gestor a tomar as decisões necessárias em relação à execução da política em questão.

Uma avaliação *ex post* é, então, aquela realizada após a implantação da política. Diversos tipos de avaliação *ex post* podem ser desenvolvidos, com finalidades diversas e complementares, a depender das questões para as quais se buscam respostas. Uma avaliação *ex post* de desenho, por exemplo, será responsável por analisar a consistência lógica da chamada Teoria do Programa, caracterizada pelas hipóteses de que a política em questão depende para alcançar os seus objetivos. Uma avaliação de processos diz respeito à caracterização da operação da política, com a análise dos seus insumos, atividades e produtos imediatos. Por sua vez, uma avaliação de impacto permitirá, ao fazer uso de técnicas estatísticas robustas, que se chegue a conclusões a respeito de relações de causa e efeito entre a implementação de uma política e mudanças observadas nos indicadores de resultado de interesse. Em todos os casos, as avaliações constituem uma fonte essencial de informações para subsidiar as tomadas de decisão no setor público, ao permitirem que sejam feitas reflexões quanto a possíveis melhorias de uma política, a sua expansão, replicação e continuidade.

Monitoramento e avaliação no ciclo da política pública

Em políticas públicas, tomar decisões baseadas em evidências significa utilizar os recursos da sociedade de forma responsável, com base no maior conjunto de evidências possível. É com este propósito que devemos enxergar o monitoramento e a avaliação: como ferramentas de tomada de decisão. Conforme visto anteriormente, as atividades de monitoramento e avaliação devem fazer-se presentes em todas as etapas de uma política pública, alimentando o processo decisório desde o momento de concepção de uma intervenção até depois da sua implementação. Em outras palavras, são atividades que perpassam todo o **ciclo da política pública**.

Podemos entender o ciclo da política pública como um processo de concepção, ação e gestão sobre um determinado problema que se deseja resolver ou atenuar, ou sobre um objetivo que se deseja alcançar. O ciclo se refere, portanto, a todo o processo decisório que envolve uma determinada política. É usual ilustrá-lo em etapas, como apresentado na figura 1, que adapta as discussões presentes em Secchi (2014) e Kraft e Furlong (2019).

Figura 1 – Ciclo da política pública



Fonte: Elaboração própria com base em Secchi (2014) e Kraft e Furlong (2019).

A ilustração do ciclo, ainda que constitua uma simplificação da realidade, auxilia na compreensão das etapas inerentes à implantação de uma política. Parte-se sempre da **identificação de um problema** que demande ação do governo para ser solucionado, ou ao menos mitigado. Um problema pode apresentar cunho social ou de gestão, relacionando-se, por exemplo, à desigualdade de acesso a serviços de saúde ou ao controle de gastos administrativos em uma determinada secretaria de governo. Identificar um problema consiste em defini-lo de forma precisa e caracterizá-lo com base em dados quantitativos e qualitativos para dimensionar a sua relevância e compreender as suas causas e consequências.

Na sequência, é necessário **desenhar a política**, definindo a sua forma de funcionamento lógico (a teoria por trás do seu funcionamento) e prático (como será implementada). A **implementação** trata do período em que a intervenção é colocada em prática, compreendendo todos os processos para que isso ocorra. A **avaliação de desempenho**, nesse contexto, se refere à análise do conjunto de evidências produzidas pelas atividades avaliativas e de monitoramento com o objetivo de se chegar a conclusões sobre a eficácia e eficiência da política. Realizar avaliações de impacto e de custo-benefício é particularmente importante nesse momento. Por fim, a **tomada de decisão** envolve emitir recomendações sobre o futuro da política, abrangendo possibilidades como a sua reformulação, expansão, continuidade ou interrupção.

Na prática, é claro, as complexidades inerentes às políticas públicas fazem o ciclo não se apresentar de forma tão clara quanto a ilustrada, com as diversas etapas naturalmente se sobrepondo. O monitoramento da política acontece principalmente durante a etapa de implementação, mas deve ser planejado desde o seu desenho.¹ As atividades de avaliação, por sua vez, devem acontecer ao longo de todo o ciclo, como será abordado no módulo 3. Em uma visão abrangente do ciclo, com um enfoque em resolução de problemas, o conceito de **gestão orientada para resultados** prevê o uso de evidências em todas as suas etapas, utilizando o monitoramento e a avaliação como ferramentas essenciais para tanto (KUSEK; RIST, 2004).

O uso de evidências toma formas distintas à medida que se avança no ciclo da política. Por exemplo, identificar o problema pode envolver o uso de evidências quantitativas e qualitativas para caracterizar a gravidade da situação enfrentada, delimitar a população afetada e entender as suas causas. Para formular a política, é necessário, entre diversas outras definições, delimitar os resultados esperados, selecionar a população-alvo e, frequentemente, comparar soluções alternativas de desenhos, cada uma idealmente com a sua lógica – ou teoria – baseada na literatura especializada e nas experiências de sucesso passadas.

As próximas unidades deste módulo tratam exatamente das etapas de identificação e formulação de uma política. Especificamente, a unidade 2 discute a identificação do problema e, na sequência, o encadeamento lógico que prevê a transformação dos recursos – insumos – previstos por uma determinada intervenção nos resultados de curto, médio e longo prazo desejados, sob o arcabouço da **Teoria da Mudança**. Em seguida, a unidade 3 apresenta uma visão geral de como as atividades de monitoramento e avaliação devem ser planejadas durante a própria fase de concepção da política e de como esses componentes se relacionam diretamente com a Teoria da Mudança.

¹ O módulo 2 abordará mais o monitoramento e a seleção dos indicadores.

Transformação de insumos em resultados

Identificação de um problema com base em evidências

No ciclo da política pública, ilustrado na unidade anterior, é possível perceber que o ponto de partida de toda política é a identificação de um problema. A importância dessa etapa está na própria razão de ser de uma política pública: ela deve ser pensada desde o começo como um meio para resolver o problema identificado.

Nesse sentido, é a identificação do problema que traz, quase imediatamente, os resultados almejados. A solução do problema identificado é exatamente o resultado que se espera alcançar com a política a ser formulada. É essa mentalidade orientada para resultados que deve embasar todo o processo de formulação que seguirá.

Identificar adequadamente um problema envolve utilizar evidências quantitativas e qualitativas para caracterizá-lo em detalhes. Esse processo envolve:

- **Caracterizar a população afetada.** Inclui perguntas como: “Quem são os indivíduos afetados direta ou indiretamente?”, “Qual é o tamanho dessa população?”, “Há alguns grupos que são mais afetados do que outros?”. Toda a população afetada compõe o grupo de potenciais beneficiários de uma política.
- **Caracterizar o problema temporal e regionalmente.** Envolve a utilização de dados para comparar a situação inicialmente identificada com o cenário do passado e de outras regiões, de forma a entender a evolução e a dimensão do problema.
- **Compreender as causas potenciais.** Inclui perguntas como: “Por que o problema existe?”, “Quais fatores o levaram a evoluir dessa forma?”, “Quais causas parecem ser mais importantes e podem ser endereçadas pela política?”. Esse mapa de causas e a compreensão das causas prioritárias auxilia na delimitação da forma de atuação da política, já que ela deve agir sobre as causas de um problema para que este possa ser resolvido.
- **Delimitar as consequências.** A relevância de um problema passa diretamente pelas consequências que ele gera. Delimitar essas consequências inclui mapear de que formas o problema afeta o bem-estar da população; entender como o problema é percebido pela sociedade, incluindo as populações afetadas; e definir como essas questões tornam o problema identificado mais ou menos relevante em face de outros problemas merecedores de atenção de políticas públicas.

Boxe 1: Programa de Infraestrutura Rodoviária – identificação do problema

Ao longo do curso, exemplificaremos os conteúdos abordados a partir do exemplo de um programa fictício, que chamaremos de Programa de Infraestrutura Rodoviária. Neste primeiro boxe, introduzimos esse exemplo sob a ótica da identificação do problema.

Suponha que em um determinado país, que denominaremos país X, o número de queixas recebidas pelo governo sobre a qualidade das estradas tem crescido muito nos últimos anos, em especial nas zonas rurais. Esse é o primeiro ponto de alerta para a possível existência de um problema social.

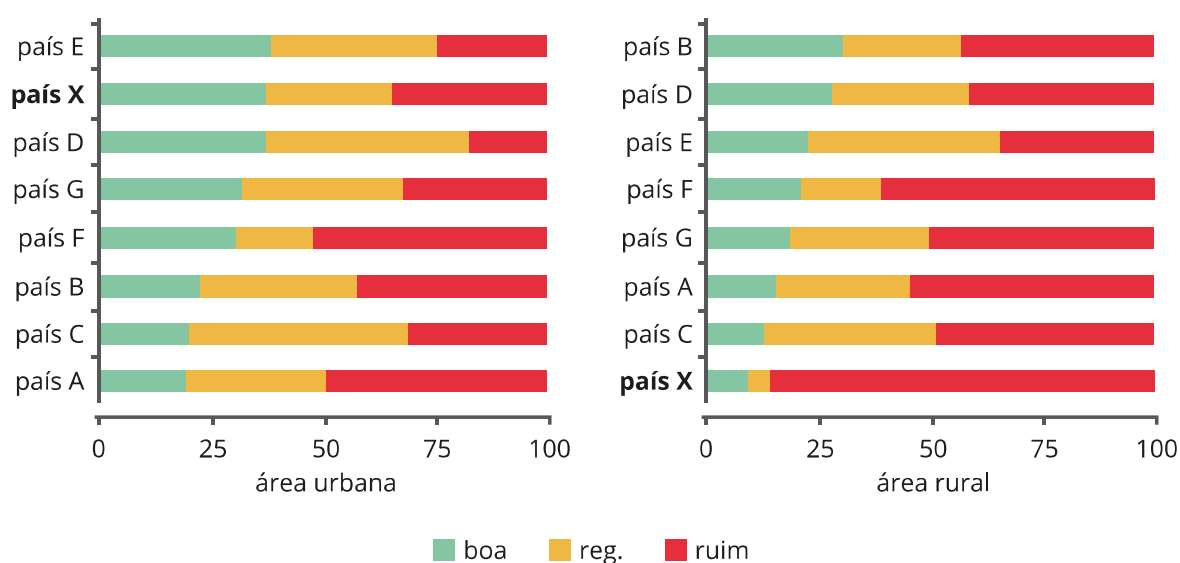
Com essa questão em mente, o governo do país X buscou compreender mais a fundo a natureza desse potencial problema a partir de um canal de comunicação direta com a população. Os moradores da zona rural informam estarem insatisfeitos com relação à facilidade de acesso e à qualidade das estradas, o que dificulta o acesso a serviços de educação e saúde e afeta também o escoamento da produção agropecuária, freando o desenvolvimento econômico regional.

Note que a percepção da própria população, nesse caso, compõe um ponto de partida importante – uma fonte de evidência – para que o problema seja compreendido. O passo seguinte envolve utilizar dados quantitativos para delimitar as questões de qualidade das rodovias e as consequências geradas.

Ao buscar evidências quantitativas, a equipe do governo responsável pelas questões de mobilidade interurbana se deparou com dados de um organismo internacional sobre a qualidade da condição de uso de rodovias em diversos países, conforme apresentado na figura 2. Esse é um bom exemplo de dado quantitativo que pode ser usado para indicar a dimensão do problema e realizar uma comparação, nesse caso, entre países.

Figura 2 – Exemplo de dado quantitativo para a identificação do problema

Distribuição das rodovias por condição de uso em 2014, por país (% do total)



Fonte: Elaborado pelos autores com base em Foster e Briceño-Garmendia (2010).

Os países de A até G, cujos dados também estão na figura 2, são da mesma região do país X, com cultura e características socioeconômicas similares, portanto, comparar os cenários entre eles pode ser útil. Os dados mostram que 35% das vias do país X foram avaliadas como em situação “boa”, 30% como “regular” e 35% como “ruim”. Já nas zonas rurais, a situação é bem mais crítica: 10% das vias foram avaliadas como em condição “boa”, 5% como “regular” e 85% como “ruim”. Comparada com a dos países vizinhos, a situação das zonas rurais é ainda mais preocupante: ainda que não haja muita variação na qualidade das rodovias urbanas, o país X está claramente muito atrás dos demais com relação à qualidade de estradas rurais.

O uso de dados durante a identificação do problema não deve ser limitado a uma única dimensão. Indicadores adicionais a serem analisados para o país X poderiam incluir, por exemplo, taxas de mortalidade por acidentes de trânsito, comparando o indicador ao longo do tempo e entre países, índices de acesso a serviços de educação e saúde, entre outros. Esses dados auxiliariam já na caracterização da relevância do problema, com o entendimento das suas consequências.

Uma vez caracterizadas as dimensões quantitativas, conforme discutido acima, o governo do país X se preocupou com o mapeamento das causas potenciais da deterioração do acesso e qualidade da malha rodoviária. Algumas possíveis causas a serem levantadas, por exemplo, a partir do conhecimento de profissionais da área, incluem: (i) o processo de urbanização não foi acompanhado por reforços à malha rodoviária, gerando congestionamentos; (ii) inexistência de calendário para reparos periódicos; (iii) aumento não previsto do fluxo de veículos motorizados, em especial veículos pesados; (iv) quantidade insuficiente de agentes e equipamentos de fiscalização; (v) pouca capilaridade de programas de educação para o trânsito, principalmente na zona rural.

Algumas dessas causas, entendidas como as mais importantes no contexto do país X a partir da análise dos dados quantitativos e qualitativos levantados, incluindo por exemplo a percepção das populações locais, serão o alvo do Programa de Infraestrutura Rodoviária, conforme veremos nas videoaulas e no restante do curso.

Representação da lógica de uma intervenção

Conforme discutido anteriormente, a primeira etapa no ciclo de uma política pública consiste em identificar o problema que será alvo da ação a ser desenvolvida. O uso de evidências nessa fase de diagnóstico é importante para que o problema identificado seja caracterizado da maneira mais adequada possível, com um detalhamento da sua relevância, das suas causas potenciais e da população atingida. Esse percurso é essencial para que se possam definir com clareza os resultados esperados da política pública. Em outras palavras, definir aonde se quer chegar em curto, médio e longo prazo.

Dessa forma, para planejar uma política é necessário prever o desenvolvimento de atividades que façam sentido para alcançar os objetivos, ou seja, é preciso que haja uma sequência lógica interna, com relações de causa e efeito entre cada fase da intervenção. A essa lógica subjacente, que garante a coerência de uma política, dá-se o nome de **Teoria do Programa**.

A Teoria do Programa é, assim, o conceito abstrato de toda a teoria que está por trás do funcionamento lógico de uma determinada intervenção. Ela engloba as relações de causalidade entre as diversas etapas da implementação que culminarão nos resultados desejados, considerando inclusive as complexidades do mundo real relacionadas aos motivos pelos quais essas relações fazem sentido e às condições e hipóteses necessárias para tanto.

Para visualizar de forma prática e sucinta a teoria subjacente de uma política pública, faz-se uso de ferramentas de monitoramento e avaliação. Há diversas maneiras de se fazer isso, cada uma com as suas vantagens e desvantagens. Em alguns casos, opta-se por uma representação em formato de matriz, como no caso do **Marco Lógico**, cujo modelo é usualmente similar ao encontrado na tabela 1. Essa abordagem lista os componentes da intervenção em uma única coluna, separando-os nas categorias de insumos, atividades, produtos, resultados e impactos.² Cada componente é associado a um indicador, a uma fonte de dados e à hipótese necessária para que os componentes imediatamente acima possam acontecer, na chamada “lógica vertical”.³

Tabela 1 – Exemplo de estrutura de um Marco Lógico

etapa	componentes	indicadores	fonte de dados	hipóteses
impactos				
resultados				
produtos				
atividades				
insumos				

Alternativamente, a ferramenta conhecida como **Modelo Lógico** foca apenas os próprios componentes listados, apresentando-os em colunas e omitindo as demais informações presentes no Marco Lógico, conforme ilustrado na figura 3, a seguir.

Figura 3 – Exemplo de estrutura de um Modelo Lógico



Fonte: Elaborado pelos autores.

² Essas etapas serão discutidas em mais detalhe no tópico 1.2.3 – Teoria da Mudança.

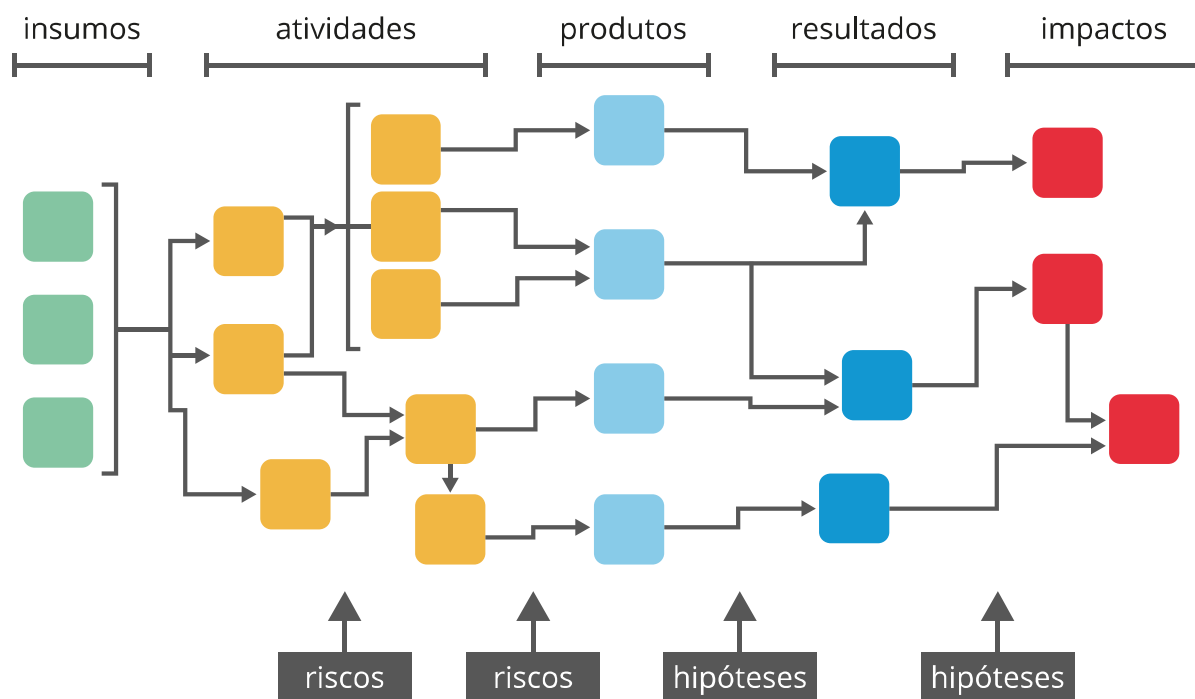
³ Para mais informações sobre Marco Lógico, é possível consultar, por exemplo, Ortegón, Pacheco e Prieto (2005).

Uma crítica frequente às abordagens do Marco Lógico e do Modelo Lógico é a sua linearidade excessiva, que ignora a complexidade de situações do mundo real, no qual as ações não se desenrolam da forma mecânica que parece ser prevista por essas ferramentas (BAKEWELL; GARBUTT, 2005). Como possível solução, surge a chamada **Teoria da Mudança**, discutida no próximo tópico, que consiste em uma representação bastante visual dos componentes da política pública. Ela abrange um diagrama que explicita as relações de causalidade entre cada um dos seus itens utilizando setas e, assim, permite representar intervenções complexas, isto é, cujos diversos componentes podem relacionar-se de forma não linear.

Teoria da Mudança

A Teoria da Mudança é uma representação de como uma intervenção deve funcionar para atingir os resultados desejados. A sua construção envolve a identificação das relações de causa e efeito e dos pressupostos envolvidos nessas relações, proporcionando assim um caminho lógico entre o problema real, as soluções propostas, e os impactos de médio e longo prazo. Em outras palavras, a Teoria da Mudança é um mapa, tal como o representado na figura 4, a seguir, no qual é possível visualizar o encadeamento lógico entre os *insumos* que serão utilizados – financeiros, humanos ou físicos –, as *atividades* desenvolvidas, que geram os *produtos* ou serviços oferecidos, os quais, por sua vez, devem levar aos *resultados* e, em última instância, aos *impactos* de longo prazo.

Figura 4 – Exemplo de estrutura de uma Teoria da Mudança



Fonte: Elaborado pelos autores.

Conforme discutido por Kusek e Rist (2004), as etapas da Teoria da Mudança podem ser definidas da seguinte forma:

- **Insumos** – São todos os recursos utilizados pela intervenção para que as atividades previstas sejam desenvolvidas, incluindo recursos financeiros, equipamentos, equipe e infraestrutura, por exemplo.
- **Atividades** – Todas as ações que o implementador do programa precisa executar para transformar os insumos em bens ou serviços. Incluem atividades diretas, que recaem sobre os beneficiários do programa; e indiretas, necessárias para a articulação interna da intervenção, como o treinamento da equipe de implementação. Em geral, as atividades são descritas por verbos no infinitivo, por exemplo, entregar, construir, treinar e fornecer.
- **Produtos** – São os resultados imediatos e tangíveis das atividades, por exemplo, os profissionais treinados e os hospitais construídos.
- **Resultados** – São os resultados intermediários da política, que representam os benefícios gerados aos beneficiários. Em geral, são relacionados a mudanças de comportamento, conhecimentos e habilidades, que só se mostram possíveis com a conclusão dos produtos, por exemplo, os profissionais treinados aumentam a sua produtividade, a população de uma região utiliza mais serviços de saúde.
- **Impactos** – São as mudanças esperadas no longo prazo, que culminam do conjunto de resultados alcançados da política; os efeitos de longo prazo alcançados após a devida maturação dos resultados intermediários. Os impactos precisam responder ao problema social para o qual a política foi desenhada. Por exemplo, um dos impactos esperados por um programa que aumenta o acesso à rede de serviços de saúde poderia ser a redução da mortalidade daquela região.

Essa estrutura, que cobre desde os insumos aplicados por uma política até os seus efeitos de mais longo prazo, evidencia a importância de compreender o encadeamento lógico entre cada etapa para que os resultados esperados sejam alcançados. Nesse sentido, a Teoria da Mudança pode ser entendida como uma ferramenta de *planejamento* de uma política pública. Analisar a coerência interna da intervenção, com o bom uso de evidências, auxilia no processo de identificar possíveis falhas na teoria, como redundâncias, inconsistências e expectativas irrealistas.

Além da sequência de insumos, atividades, produtos, resultados e impactos, e da representação visual ilustrada na figura 4, acima, Imas e Rist (2009) discutem alguns aspectos adicionais que devem fazer parte de uma Teoria da Mudança:

- Descrição das condições necessárias e dos eventos – riscos – que podem afetar o sucesso da intervenção em alcançar os resultados esperados. Por exemplo, condições macroeconômicas podem afetar a capacidade de beneficiários prosperarem nos seus negócios; condições climáticas não favoráveis podem diminuir a produtividade agrícola, etc.

- Identificação das hipóteses necessárias para que as relações de causa e efeito sejam cumpridas. Por exemplo, para que os profissionais que participaram de um treinamento aumentem a produtividade, é preciso que o conteúdo desse treinamento seja relevante ao seu trabalho. As hipóteses devem idealmente partir de embasamento teórico na literatura especializada da área, levando em conta considerações acerca do contexto em que a política está inserida. Algumas das relações de causa e efeito por trás das hipóteses podem ser testadas no arcabouço de avaliações, como veremos no módulo 3.

Para abranger esse detalhamento teórico, que vai além do que pode ser representado em um diagrama (especialmente para intervenções mais complexas, que preveem muitos componentes), é frequentemente uma boa ideia que a Teoria da Mudança seja também descrita de forma narrativa, em um **relatório de desenho** (*design report*, em inglês). Esse documento será o anexo principal da representação visual do funcionamento da política, elaborando os aspectos teóricos e de planejamento e detalhando as condições e hipóteses necessárias para alcançar o sucesso da intervenção.

A Teoria da Mudança, além de trabalhar com todo o embasamento em evidências disponíveis, é fruto de um processo analítico, buscando explicitar, de forma coerente e realista, como determinados insumos, quando mobilizados e implementados conforme as atividades desenhadas, vão produzir os efeitos desejados, isto é, os resultados esperados em curto, médio e longo prazo. Além de uma ferramenta de planejamento, ela também é importante para a comunicação entre todos os envolvidos sobre o que a política de fato representa e o que pretende alcançar. Nesse sentido, a Teoria da Mudança também é uma ferramenta de *gestão*, pois guia a implementação da política, levando a processos efetivos que aumentarão as chances de sucesso e orientando todo o processo de tomada de decisão na direção de alcançar resultados.

Foco em resultados

O arcabouço de Teoria da Mudança discutido acima prevê três etapas que, em uma primeira leitura, podem não parecer tão diferentes: produtos, resultados e impactos. É importante, então, enfatizar as diferenças entre essas camadas para entender de que forma a Teoria da Mudança pode ser aplicada como ferramenta de gestão orientada para resultados.

Conforme discutido, os *produtos* são imediatamente gerados pelas *atividades* de uma política. Assim, eles refletem uma mudança apenas na oferta de bens e serviços. Em outras palavras, os produtos nada mais são do que itens orçamentários de uma intervenção, que estão sob o controle da equipe que a implementa. Os *resultados*, por sua vez, vão além do escopo direto da implementação da política, relacionando-se a mudanças de atitudes, motivação, habilidades e comportamentos. São os resultados que refletem de fato os benefícios que um programa deseja prover. Por exemplo, a construção de um hospital, por si só, não representa um benefício, mas apenas um item orçamentário a ser cumprido por uma política de infraestrutura de saúde; o acesso aumentado a serviços de saúde é que configura um *resultado* real. Pensar em política públicas voltadas para resultados significa compreender que produtos não são suficientes.

Ainda além dos *resultados* previstos por uma política estão os seus *impactos*. Esses efeitos esperados são alinhados aos objetivos de última instância de determinada intervenção, mas frequentemente estão fora da sua alçada. Os impactos são os aspectos que dizem respeito a mudanças de longo prazo mais abrangentes, geradas pelo conjunto de *resultados*. Por exemplo, o aumento do acesso a serviços de saúde pode ser um *resultado* da construção de um hospital, cujo *impacto* desejado de longo prazo é a redução da mortalidade na população-alvo da política.

Compreender essas diferenças é essencial para analisar a Teoria da Mudança como um todo e distinguir sucessos de falhas. Uma **política bem-sucedida** é aquela que consegue transformar os insumos em resultados e impactos, conforme ilustrado na figura 5, a seguir.

Figura 5 – Roteiro de uma intervenção de sucesso



Fonte: Elaborado pelos autores.

Por outro lado, quando os resultados não são alcançados, há essencialmente duas possíveis razões. A sequência entre *insumos*, *atividades* e *produtos* diz respeito aos processos de uma intervenção. Diz-se que há uma **falha de implementação** quando essa sequência é quebrada, conforme apresentado na figura 6, a seguir. Essa situação implica que os resultados não poderão ser alcançados, mesmo que a teoria subjacente à política apresente sentido lógico, porque a equipe de implementação não conseguiu concluir as suas responsabilidades.

Figura 6 – Roteiro de uma intervenção com falha de implementação



Fonte: PRIME (2018).

O segundo caso é o de uma **falha de teoria**, representado na figura 7, a seguir. Diferente de antes, nessa situação os resultados não serão alcançados mesmo que todos os processos previstos sejam cumpridos perfeitamente, devido a um problema fundamental com o desenho da política. Uma Teoria da Mudança bem construída, com as suas implicações tanto de *planejamento* como de *gestão*, é uma ferramenta essencial para que isso não aconteça e, em última instância, que os objetivos pretendidos possam ser atingidos.

Figura 7 – Roteiro de uma intervenção com falha de teoria



Fonte: PRIME (2018).

Planos de monitoramento e avaliação

Como vimos anteriormente, a Teoria da Mudança ajuda a visualizar as cadeias que interligam os resultados esperados de uma determinada intervenção. Ela é a ferramenta que traz clareza sobre os componentes da política e as suas inter-relações, buscando garantir uma sequência lógica entre eles e guiar a implementação da política com base em evidências, porém há ainda outros aspectos que ressaltam a importância da elaboração da Teoria da Mudança. É ela que embasa todo o processo de monitoramento e avaliação a ser conduzido ao longo do ciclo da política pública, servindo como um guia primordial para a seleção dos indicadores a serem monitorados e a definição das perguntas de avaliação que serão respondidas.

Plano de monitoramento

Um plano de monitoramento é uma ferramenta utilizada para sistematizar as dimensões a serem monitoradas em uma política pública, listando os indicadores relacionados aos componentes de cada etapa da Teoria da Mudança: insumos, atividades, produtos, resultados e impactos. Dessa forma, a seleção dos indicadores adequados depende diretamente da construção da própria Teoria da Mudança, conforme será discutido em mais detalhes no módulo 2 deste curso.

O plano de monitoramento pode ter estrutura de tabela, como a tabela 2, a seguir, detalhando aspectos como a frequência da coleta do indicador, a fonte de dados que será utilizada, o responsável por essa coleta, os valores de linha de base e das metas definidas, além da própria descrição do indicador e a indicação da etapa da Teoria da Mudança à qual ele se refere.

Tabela 2 – Plano de monitoramento

indicador	etapa					valor de linha de base	frequência da coleta	meta			órgão responsável pela coleta de dados
	insumos	atividades	produtos	resultados	impactos			ano 1	ano 2	ano 3	

Fonte: Adaptada de World Bank Institute (2009).

Plano de avaliação

De forma similar ao plano de monitoramento, durante o período de formulação da política pública também se deve elaborar o plano de avaliação. Essa é a ferramenta que especifica de antemão todas as avaliações que deverão ser conduzidas ao longo do ciclo da política, definindo as perguntas a serem respondidas, as metodologias que serão utilizadas e quem será o responsável por conduzir a avaliação. Assim como antes, as perguntas de avaliação surgem da Teoria da Mudança, podendo ser relacionadas, por exemplo, aos processos da intervenção (As atividades foram realizadas conforme o cronograma planejado?), aos efeitos (O impacto foi alcançado?), ao desenho (A Teoria da Mudança prevista foi cumprida?) ou aos custos (Os recursos gastos compensaram os benefícios gerados?).⁴

O plano de avaliação também pode ser estruturado em formato de tabela, como exemplificado na tabela 3, a seguir.

Tabela 2 – Plano de Avaliação

pergunta a ser respondida	subperguntas	tipo de avaliação	metodologia	fonte de dados	órgão responsável pela análise de dados

Fonte: Adaptada de World Bank Institute (2009).

⁴ Os tipos de avaliação serão discutidos no módulo 3 deste curso.



MÓDULO II – INDICADORES PARA MONITORAMENTO

Conforme discutido no módulo anterior, para que seja possível formular uma política consistente, é preciso delimitar muito bem o problema que se deseja solucionar. Além disso, é fundamental que se tenha clareza da lógica por trás da intervenção, do que é necessário para que as atividades se transformem em produtos e de quais são os resultados e impactos esperados. Mas como saber se a política está sendo executada conforme o planejado? Como saber se os resultados esperados estão sendo atingidos?

A definição de indicadores envolve identificar parâmetros que permitam descrever eventuais avanços ou retrocessos ao longo da implementação do programa, desde insumos e atividades até resultados e impactos. Munido de um conjunto de **indicadores**, variáveis que permitem verificar mudanças decorrentes de uma intervenção e comparar o desempenho dos resultados obtidos aos resultados planejados (OCDE, 2002), o gestor terá uma ferramenta objetiva para acompanhar o andamento da execução do programa.

Para que o gestor tenha uma ferramenta que seja útil para acompanhar todas as etapas da implementação do programa, é fundamental que todas as etapas da Teoria da Mudança – insumos, atividades, produtos, resultados e impactos – sejam associadas a indicadores. É a partir da Teoria da Mudança da política que se devem selecionar os indicadores que comporão o seu plano de monitoramento, ferramenta essencial para a gestão voltada para resultados.

O objetivo deste módulo é apresentar os passos a serem seguidos na seleção e construção de indicadores para o acompanhamento de uma política. A unidade 1 expõe de que forma se relacionam os indicadores e o plano de monitoramento, bem como discute critérios para a seleção de indicadores para o plano de monitoramento; a unidade 2 apresenta conceitos e formas de construção de indicadores; por fim, a unidade 3 discute possíveis fontes de dados para a construção de indicadores.

Indicadores e plano de monitoramento

O uso de indicadores na gestão pública é amplo. Ao se formularem políticas públicas baseadas em evidências, indicadores são essenciais ao longo de todo ciclo da política: identificação do problema, formulação, implementação, avaliação e tomada de decisão (SECCHI, 2014; KRAFT; FURLONG, 2019).

Mas o que define um indicador? De acordo com OCDE (2002), indicador é uma variável quantitativa ou qualitativa que permite, de forma simples e confiável, medir resultados, exibir mudanças, ou avaliar desempenho. Indicadores podem ser quantitativos, utilizados para descrever um fenômeno por meio de números; ou qualitativos, utilizados para descrever características intangíveis, por meio de variáveis nominais, ordinais ou categóricas, por exemplo (PNUD, 2005). Além disso, indicadores podem ser unidimensionais, referindo-se a uma única dimensão – por exemplo, indicadores de pobreza que utilizam apenas a variável de renda para mensurar a pobreza (FIELDS, 2002) –, ou multidimensionais – também conhecidos como indicadores sintéticos, ou índices –, combinações de um conjunto de medidas de diferentes dimensões em um único indicador. Um exemplo de indicador multidimensional é o Índice de Desenvolvimento Humano (IDH), que é uma medida resumida de três dimensões básicas do desenvolvimento humano: renda, educação e saúde (PNUD, 2019).

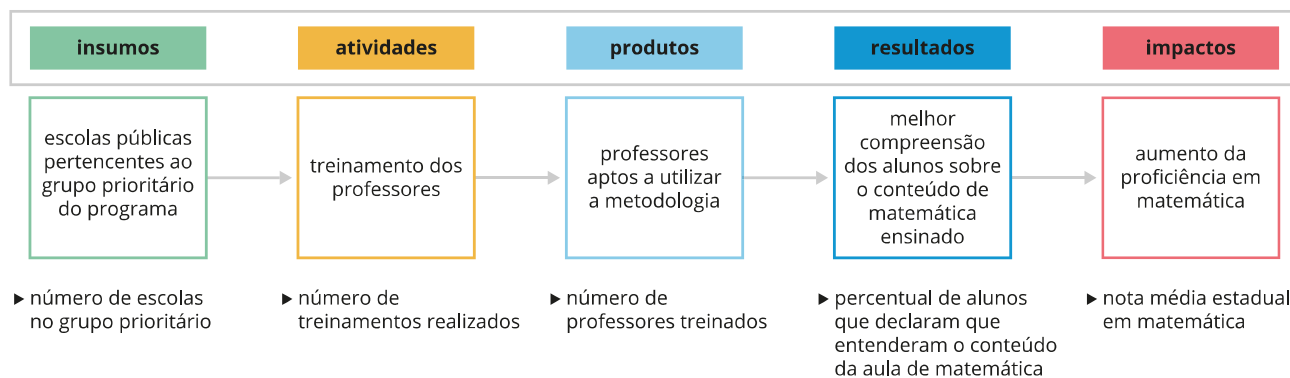
Um processo de gestão voltada para resultados geralmente utiliza como uma das suas principais ferramentas os **indicadores**, que medem se os objetivos de uma política ou um projeto estão sendo atingidos (OCDE, 2002). Entretanto, para o monitoramento do progresso de uma política, os indicadores não devem restringir-se apenas a medir os seus resultados e impactos, mas também os produtos, as atividades e os insumos.

Considere, como exemplo, o monitoramento da implementação de um programa estadual que tem como objetivo melhorar o desempenho acadêmico dos alunos em matemática por meio da aplicação de uma metodologia que incorpora jogos e problemas do dia a dia no ensino de matemática. Como parte do programa, são realizadas capacitações com os professores sobre essa metodologia. Nesse caso, pode-se definir como **atividade** o treinamento dos professores. Um **indicador** para essa atividade seria o número de treinamentos realizados. Já um impacto pretendido pelo programa pode ser o aumento da proficiência dos alunos em matemática. Um indicador para ele seria a nota média em matemática dos alunos da rede estadual. Assim, é possível acompanhar de forma rápida e objetiva a execução das atividades implementadas e o progresso em direção aos efeitos desejados.

O monitoramento de um projeto ou de uma política é importante para identificar áreas nas quais se está no caminho para alcançar os resultados esperados e aquelas nas quais aprimoramentos serão necessários. O monitoramento do programa deve contemplar tanto indicadores finais – resultados e impactos – quanto indicadores intermediários – insumos, atividades e produtos –, pois estes tendem a sinalizar mais rapidamente mudanças em decorrência de uma política, sendo importantes ferramentas de gestão, que auxiliam o acompanhamento do programa durante a etapa de implementação (PNUD, 2005; KUSEK; RIST, 2004; ROSSI; LIPSEY; FREEMAN, 2003).

Retomando o exemplo do programa para melhorar o desempenho dos alunos em matemática, a figura 8, a seguir, traz um exemplo de uma parte da Teoria da Mudança desse programa. Além disso, a figura 8 ilustra exemplos de indicadores para as etapas da Teoria da Mudança, que poderiam fazer parte do plano de monitoramento do programa.

Figura 8 – Exemplo de teoria da mudança e indicadores



Fonte: Elaboração própria.

Em resumo, como foi apresentado no módulo anterior, políticas e projetos são implementados para promover mudanças de modo a atingir os resultados desejados. Para isso, eles utilizam insumos e atividades que se convertem em produtos. Estes devem transformar-se em resultados que, no longo prazo, resultarão em impactos (KUSEK; RIST, 2004). A Teoria da Mudança permite estruturar a política com todos os seus elementos-chave, mostrando o papel de cada um na lógica do programa e as relações de causalidade entre eles.

Após identificar e selecionar os elementos que integram cada etapa da Teoria da Mudança, o próximo passo para a construção do plano de monitoramento é fazer com que eles sejam mensuráveis: transformá-los em indicadores. Assim, os indicadores podem ser classificados de acordo com a etapa da Teoria da Mudança de que fazem parte.

Conforme a definição utilizada por Mosse e Sontheimer (1996):

- **Insumos** – Medem a quantidade (algumas vezes, a qualidade) dos recursos existentes para desenvolver as atividades ligadas ao programa.
- **Atividades** – Medem o esforço e as ações realizados para obter os resultados planejados para o programa, utilizando os insumos disponíveis. Em outras palavras, medem a utilização dos insumos.
- **Produtos** – Medem a quantidade, ou qualidade, dos bens e serviços criados ou ofertados em decorrência do uso dos insumos.
- **Resultados** – Medem a quantidade e a qualidade dos resultados alcançados pelo programa para o seu público-alvo.
- **Impactos** – Medem quantitativa e qualitativamente os efeitos do programa.

Após a seleção e a construção de um indicador, são coletados os valores iniciais, os chamados valores de linha de base – medidos antes do início do programa – e é estabelecida a meta para ele, valor a ser alcançado para esse indicador em um determinado período de tempo. Para se estabelecer uma meta realista para o indicador, sugere-se considerar os seguintes pontos (IMAS; RIST, 2009):

- o ponto de partida, ou seja, valores do indicador no passado, incluindo o valor de linha de base e, por exemplo, a média dos anos anteriores;
- o tamanho da mudança que se deseja promover, a magnitude do resultado que se deseja atingir, desagregando-o em um conjunto de metas intermediárias a serem periodicamente alcançadas pelo indicador em direção à sua meta final;
- recursos financeiros e humanos disponíveis para o programa no período de tempo previsto para o alcance da meta e
- questões políticas relevantes.

Uma vez estabelecida a meta, define-se a frequência com que o indicador será atualizado para o monitoramento do seu progresso em direção à meta. A frequência da atualização pode variar para cada indicador. Idealmente, conforme apresentado na tabela 2 do módulo 1, o plano de monitoramento deve ser capaz de indicar os valores de linha de base, e como está a evolução dos indicadores em relação às metas estabelecidas. Assim, ao passo que os indicadores são periodicamente atualizados, esse monitoramento permite o uso dessas informações para a gestão e a tomada de decisão em momentos-chave da implementação da política (ROBERTS; KHATTTRI, 2012).

O plano de monitoramento é composto de indicadores que refletem todas as etapas da Teoria da Mudança. Entretanto, é possível que existam vários indicadores para cada etapa, sendo, portanto, fundamental selecionar os melhores indicadores. A próxima unidade discutirá critérios utilizados para a seleção de indicadores para o monitoramento.

Seleção de indicadores para o plano de monitoramento

Como apresentado no começo desta unidade, o plano de monitoramento deve partir da Teoria da Mudança do programa. O passo seguinte para a sua construção é o processo de seleção dos indicadores. Segundo Kusek e Rist (2004), essa seleção deve ser feita de forma participativa, considerando as preocupações e os interesses dos atores-chave do programa. Mas como selecionar os indicadores? Quantos indicadores devem ser selecionados?

Em uma primeira análise, pode-se pensar que quanto mais indicadores selecionados, maior o número de informações sobre o andamento do programa. Portanto, quanto mais indicadores, melhor. Entretanto, a seleção deve considerar não só os benefícios, mas também os custos associados à adição de um indicador. O gestor deve levar em conta se existem custos de coleta, atualização, análise e disseminação da informação. Esses custos são compensados pelo benefício dessa informação adicional?

Para selecionar indicadores de boa qualidade, sugere-se que sejam seguidos oito passos (KUSEK; RIST, 2004; GÖRGENS; KUSEK, 2009), conforme tabela 4, abaixo:

Tabela 3 – Oito passos para selecionar indicadores de boa qualidade

1)	Identifique o que precisa ser medido.
2)	Faça uma lista de possíveis indicadores, preferencialmente utilizando processos participativos com os atores-chave do programa.
3)	Avalie cada possível indicador.
4)	Selecione os melhores indicadores.
5)	Desenvolva uma proposta de protocolo para a construção dos indicadores.
5.1)	Defina: i. fonte de dados; ii. método para a coleta de dados; iii. responsáveis pela coleta dos dados; iv. frequência com que os dados serão coletados; v. custo e dificuldade para se coletarem os dados; vi. responsáveis por analisar os dados e divulgar os relatórios e vii. usuários dos dados.
6)	Colete dados de linha de base.
7)	Defina metas.
8)	Refine os indicadores e protocolos e conclua a sua seleção.

Fonte: Elaboração própria com base em Kusek e Rist (2004) e Görgens e Kusek (2009).

Para a avaliação e a seleção dos indicadores, os passos 3 e 4 da tabela 4, acima, existem alguns critérios que podem ser aplicados, incluindo a metodologia “SMART”, que é amplamente utilizada.

De acordo com World Bank (2013), um indicador “SMART” deve apresentar as seguintes propriedades:

- **e**specífico – O indicador deve medir de forma simples apenas o objeto a ser mensurado, e não outros elementos do programa.
- **M**ensurável – Deve haver uma forma clara e objetiva para atribuir um valor ao indicador.
- **A**tribuível – O indicador deve ser uma medida válida do objeto que se quer mensurar.

- **Realista** – O indicador deve ser realista em relação à capacidade de se coletarem periodicamente as informações necessárias para a sua construção, considerando restrições existentes no programa, tais como recursos humanos, materiais e financeiros, entre outros.
- **Tempo** – Devem existir intervalos regulares de tempo para medir os indicadores e para verificar se houve mudanças ao longo do tempo.

Boxe 2: Critérios complementares para a seleção de indicadores

Em complementariedade aos critérios SMART, Görgens e Kusek (2009) apresentam 10 critérios a serem utilizados para a avaliação e a seleção de bons indicadores:

1. **Mensurabilidade** – O indicador deve poder ser quantificado e medido por meio de alguma escala.
2. **Praticidade** – As informações necessárias para construir o indicador devem poder ser coletadas periodicamente e com um baixo custo.
3. **Confiabilidade** – Os dados devem poder ser mensurados, com precisão, por diferentes pessoas.
4. **Relevância** – Existe uma relação clara e significativa entre o objeto a ser medido e o indicador.
5. **Utilidade para a tomada de decisão** – A informação medida pelo indicador é importante para tomada de decisão? Se a informação não será utilizada para o gerenciamento e monitoramento do projeto, o indicador não deve ser selecionado.
6. **Medida direta** – O indicador deve ser uma medida mais próxima possível do objeto a ser mensurado.
7. **Sensível a mudanças** – Um indicador sensível deve ser capaz de refletir proporcionalmente e na mesma direção mudanças que estejam ocorrendo no objeto medido.
8. **Responsivo** – Se o objeto medido sofre mudanças como consequência das ações do programa, o indicador que o mede deve refletir mudanças responsivas às ações do programa.
9. **Objetivo** – A medida é precisa e mede apenas um fenômeno. Não há ambiguidade na medida.
10. **Desagregável** – Os dados devem poder ser desagregados: gênero, idade e região, por exemplo.

Fonte: Traduzido de Görgens e Kusek (2009).

Os critérios apresentados no boxe 2 apontam na direção de se avaliar a importância e a finalidade do indicador que será selecionado. Além disso, é preciso considerar se as informações a serem utilizadas satisfazem o propósito imaginado, se elas são passíveis de atualização, se estão disponíveis e bem documentadas, ou se será necessário coletar dados.

É importante que o plano de monitoramento identifique quem ficará responsável pela coleta e atualização dos dados, aponte as fontes de dados e estime quais serão os custos associados à construção e à atualização do indicador, considerando recursos humanos e financeiros. Portanto, tanto para a seleção dos indicadores quanto para a elaboração do plano de monitoramento, é de grande utilidade o conhecimento acerca de formas de construção dos indicadores e possíveis fontes de dados. As próximas unidades deste módulo abordarão esses dois temas.

Construção de indicadores para o plano de monitoramento

Uma vez selecionados os indicadores, o próximo passo é desenvolver o protocolo para a sua construção. Para isso, é preciso definir a sua fórmula, que descreve de forma objetiva como o cálculo do indicador deve ser realizado. A seguir, são apresentadas algumas fórmulas e conceitos de estatística amplamente utilizados para a construção de indicadores, baseados em Bussab e Morettin (2017) e PNUD (2005).

Medidas de posição

Média

A **média** é a soma dos valores observados ($X_1 + X_2 + \dots + X_n$) dividida pelo número total de observações (n). A média aritmética, ou apenas média, pode ser expressa por meio da fórmula abaixo:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Onde:

- $\sum$ é o operador soma; i corresponde à observação ($i = 1, 2, \dots, n$);
- X_i é o valor da variável principal, X , para a observação i e
- n refere-se ao número total de observações.

Para facilitar a compreensão, a tabela 5 apresenta um exemplo utilizando notas de matemática de um grupo de alunas.

Tabela 4 – Exemplo: Notas de matemática das alunas (1)

aluna	nota
Juliana	5
Valentina	4
Laura	7
Camila	9
Valéria	5

Fonte: Elaborado pelos autores.

Nesse caso, $i = 1, 2, \dots, 5$, X_i é a nota da aluna i , $n = 5$. Portanto, a média das notas das alunas será:

$$\bar{X} = \frac{5+4+7+9+5}{5} = 6.$$

Assim, generalizando, um indicador da nota média estadual em matemática seria construído da seguinte forma:

$$\text{nota média estadual em matemática} = \frac{1}{n} \sum_{i=1}^n \text{nota de matemática do aluno}_i$$

Mediana

A **mediana** é a realização que ocupa a posição central em um conjunto de observações, quando estas estão ordenados em ordem crescente. Supondo que se têm os seguintes valores observados em ordem crescente: $X_{(1)}, X_{(2)}, \dots, X_{(n)}$, a mediana será dada por:

$$Med(X) = \begin{cases} X_{(\frac{n+1}{2})}, & \text{se } n \text{ é ímpar} \\ \frac{X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}}{2}, & \text{se } n \text{ é par} \end{cases}$$

Considerando-se o exemplo anterior, na tabela 5, em ordem crescente, as notas das alunas foram: 4, 5, 5, 7 e 9. Nesse caso, o número de observações (n) é ímpar, e é igual a 5. A mediana será o valor que ocupa a posição central, dividindo o conjunto de dados em duas partes iguais $((5+1)/2)$. Dessa forma, a mediana será o valor da terceira observação, que no caso é igual a 5.

No caso em que o número de observações é par, como na tabela 6, a seguir, a mediana é calculada da seguinte forma:

$$X_{(\frac{n}{2})} = 5; X_{(\frac{n}{2}+1)} = 7; Med(X) = ((5+7)/2) = 6.$$

Tabela 5 – Exemplo: Notas das Alunas (2)

aluna	nota
Juliana	5
Valentina	5
Laura	7
Camila	9

Fonte: Elaborado pelos autores.

Moda

Por fim, a **moda** é o valor mais frequente em um conjunto de observações. A tabela 7, a seguir, apresenta, para o exemplo disponível na tabela 5, a lista dos valores observados – notas – e a frequência com a qual eles ocorrem.

Tabela 6 – Exemplo: Notas das Alunas (1) (continuação)

valores observados	frequência
4	1
5	2
7	1
9	1

Fonte: Elaborado pelos autores.

Dessa forma, a nota 4 ocorreu uma única vez (nota da Valentina), assim como as notas 7 e 9. Já a nota 5 ocorreu duas vezes (nota da Juliana e da Valéria). Assim, a moda será igual a 5, valor correspondente à realização mais frequente.

Medidas de dispersão

As medidas de dispersão são importantes para informar sobre a distribuição de um conjunto de dados.

Variância

A **variância** mede a dispersão dos dados em torno da média. É definida como a soma dos quadrados dos desvios, dividida pelo número de observações. A sua fórmula é dada por:

$$Var(X) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

Onde:

- i corresponde à observação;
- X_i é o valor da variável X para a observação i ;
- $\bar{X}$ corresponde à média de X e
- n refere-se ao número total de observações.

Considerando novamente o exemplo da tabela 5, a variância será dada por:

$$\begin{aligned} \text{Variância} &= \left(\frac{1}{5}\right) * [(5 - 6)^2 + (4 - 6)^2 + (7 - 6)^2 + (9 - 6)^2 + (5 - 6)^2] \\ &= \left(\frac{1}{5}\right) * [1 + 4 + 1 + 9 + 1] \\ &= \left(\frac{1}{5}\right) * 16 \\ &= 3,2 \end{aligned}$$

Desvio padrão

O **desvio padrão** corresponde à raiz quadrada positiva da variância. Ele mensura qual será o "erro" cometido ao tentar substituir cada observação pela média. A fórmula do desvio padrão é dada por:

$$Dp(X) = \sqrt{Var(X)}$$

Novamente, considerando o exemplo da tabela 5, tem-se que o desvio padrão é calculado da seguinte forma:

$$Dp(x) = \sqrt{3,2} \cong 1,79$$

Razão

A **razão** é a divisão de duas quantidades que são medidas na mesma unidade. O resultado dessa operação não possui unidade, pois a do numerador e a do denominador se cancelam. Por exemplo, razão de meninas matriculadas na creche em relação aos meninos:

Razão de meninas matriculadas na creche em relação aos meninos matriculados =
(número de crianças do sexo feminino matriculadas em creches/
número de crianças do sexo masculino matriculadas em creches).

Proporção

A **proporção** é um tipo especial de razão, na qual o denominador é a quantidade que representa o total de um determinado objeto de análise e o numerador é um subgrupo desse objeto. Assim como a razão, a proporção não possui unidade. Por exemplo, proporção de meninas matriculadas na creche em relação ao total de crianças matriculados na creche:

$$\text{Proporção de meninas matriculadas na creche} = \frac{\text{(número de crianças do sexo feminino matriculadas em creches)}}{\text{número total de crianças matriculadas em creches}}.$$

Percentual

O **percentual** é a proporção multiplicada por 100%. Ainda utilizando o exemplo anterior, para obtermos o percentual de meninas matriculadas na creche em relação ao total de crianças:

$$\text{Percentual de meninas matriculadas na creche} = 100\% \times \frac{\text{(número de crianças do sexo feminino matriculadas em creches)}}{\text{número total de crianças matriculadas em creches}}.$$

Taxa

A **taxa** é a razão entre duas quantidades que são medidas em unidades diferentes. Normalmente a unidade é expressa pela unidade do numerador em relação à do denominador:

$$\text{PIB per capita} = \text{PIB} / \text{população}.$$

Boxe 3: Medidas de pobreza

A literatura sobre pobreza é repleta de controvérsias, que vão desde divergências éticas sobre o conceito de pobreza até as diferentes abordagens empregadas para a sua mensuração, passando pela discussão acerca de quais dimensões, variáveis e indicadores utilizar para medir pobreza (COMIM; BORSI; VALERIO MENDOZA, 2020; RAVALLION, 2016; ANAND; SEGAL; STIGLITZ, 2010; SEN, 1983; ALKIRE; FOSTER, 2011).

Em relação às dimensões, existem os indicadores unidimensionais de pobreza, cuja abordagem considera pobreza enquanto insuficiência de renda, ou seja, considera apenas a dimensão renda para medir a pobreza de uma determinada população (FIELDS, 2002; FOSTER; GREER; THORBECKE, 1984) e os indicadores multidimensionais de pobreza. Estes consideram que a pobreza deve ser medida nas suas múltiplas dimensões, enquanto

insuficiência de realizações – privações – em diferentes dimensões do bem-estar, tais como educação, saúde e acesso a serviços básicos, entre outras (ATKINSON; BOURGUIGNON, 1982; CHAKRAVARTY; SILBER, 2008; SEN, 2000; ALKIRE; SANTOS, 2010).

Dessa forma, abaixo são apresentados indicadores unidimensionais e multidimensionais de pobreza amplamente utilizados na literatura.

Indicadores unidimensionais de pobreza

Os indicadores unidimensionais consideram pobreza a insuficiência em apenas uma dimensão: renda. Para isso, eles comparam a renda de um indivíduo (y_i) com uma linha de pobreza (z). A medida de pobreza da população como um todo é obtida por meio da agregação das funções de pobreza individuais, que dependem da renda de cada indivíduo e da linha de pobreza considerada (FIELDS, 2002).

Uma medida comumente utilizada na literatura é a proposta por Foster, Greer e Thorbecke (1984), os chamados indicadores da classe P_α :

$$P_\alpha \equiv \frac{1}{n} \sum_{i=1}^q \left(\frac{z - y_i}{z} \right)^\alpha$$

Onde:

- q é o número de pessoas pobres na população;
- n é o número total de pessoas na população;
- z é a linha de pobreza estabelecida e
- y_i é a renda do indivíduo i .

Os diferentes valores assumidos por α resultam em diversos indicadores de pobreza, que variam entre 0 e 1. Quanto mais próximo de 1, pior a situação de pobreza apontada. Quando $\alpha = 0$, tem-se o chamado P_0 , a **taxa de pobreza**, que mede a incidência de pobreza na população. Isto é:

$$P_0 = \frac{q}{n}$$

Quando $\alpha = 1$, tem-se o **hiato da pobreza** (P_1), que mede a intensidade da pobreza em uma população. O indicador do hiato da pobreza é calculado da seguinte forma:

$$P_1 = \frac{1}{n} \sum_{i=1}^q \left(\frac{z - y_i}{z} \right)$$

Por fim, outro indicador amplamente utilizado na literatura é o de **severidade da pobreza** (P_2), quando $\alpha = 2$. Esta medida atribui um peso maior aos indivíduos mais afastados da linha de pobreza. Portanto:

$$P_2 = \frac{1}{n} \sum_{i=1}^q \left(\frac{z - y_i}{z} \right)^2$$

Indicadores multidimensionais de pobreza

Existe uma crescente literatura que enfatiza o caráter multidimensional da pobreza. Assim, as medidas de pobreza devem considerar privações dos indivíduos não apenas na renda, mas nas múltiplas dimensões que compõem o desenvolvimento humano, tais como acesso a serviços básicos, cidadania e direitos, educação e saúde, entre outros (ATKINSON; BOURGUIGNON, 1982; CHAKRAVARTY; SILBER, 2008; SEN, 2000; ALKIRE; SANTOS, 2010). Para isso, são utilizados os indicadores multidimensionais de pobreza.

O Programa das Nações Unidas para o Desenvolvimento (PNUD) possui um Índice de Pobreza Multidimensional (IPM) global, fortemente associado ao Objetivo de Desenvolvimento Sustentável (ODS) 1: erradicação da pobreza (PNUD, 2020). O IPM é um índice sintético composto de 10 indicadores que fazem parte de três dimensões nas quais as pessoas vivenciam pobreza: saúde, educação e padrão de vida. Cada dimensão possui o mesmo peso no IPM.

Esse indicador multidimensional de pobreza é complementar ao indicador unidimensional de pobreza, também utilizado pelo ODS, que considera pobreza enquanto insuficiência de renda.

Para ambos os indicadores de pobreza – unidimensional e multidimensional – é calculada a taxa de pobreza, a qual, conforme apresentado anteriormente, mede a incidência de pobreza na população. A diferença entre os dois indicadores é o critério utilizado para considerar um indivíduo pobre. Para o ODS 1, na abordagem unidimensional, um indivíduo é considerado pobre se a sua renda é menor do que US\$ 1,90 por dia. A abordagem multidimensional ajuda a compreender a natureza das privações de cada pessoa em diferentes dimensões, além da renda. No IPM, um indivíduo é considerado multidimensionalmente pobre se ele está em situação de privação em um terço ou mais dos 10 indicadores (PNUD, 2020).

Fontes de dados

Por fim, após selecionar os indicadores e desenvolver os protocolos para a sua construção, o plano de monitoramento deve indicar se as informações necessárias para a construção dos indicadores já existem ou se deverão ser construídas. Existem duas possíveis fontes de dados: (a) primários e (b) secundários.

Dados primários

Dados primários são aqueles produzidos diretamente pelas organizações envolvidas no programa. A sua construção pode ser realizada por meio de diferentes metodologias e instrumentos: coleta de dados administrativos, entrevistas, observação, análise de documentos, grupos focais e análise de diários, entre outros.

As metodologias de coleta de dados podem ser classificadas como estruturada ou semiestruturada (KUSEK; RIST, 2004; IMAS; RIST, 2009). Segundo Imas e Rist (2009), uma metodologia estruturada requer que os dados sejam coletados por meio de instrumentos padronizados. Podem-se citar como exemplo de metodologia estruturada entrevistas realizadas por meio de questionários, que sempre contêm as mesmas perguntas, aplicadas na mesma ordem e de forma padronizada. Já as metodologias semiestruturadas exigem um menor grau de padronização. Podem-se citar como um exemplo de metodologia semiestruturada entrevistas que utilizam um roteiro de perguntas, mas a ordem e a forma como elas são aplicadas varia de acordo com o contexto da entrevista. Os dados coletados por essas diferentes metodologias podem ser tanto quantitativos quanto qualitativos. Dados quantitativos são dados em forma numérica, dados qualitativos são dados em forma não numérica.

A qualidade dos dados primários coletados é um ponto muito importante a ser considerado. Alguns fatores que podem comprometer a qualidade dos dados são: i) erros na codificação/programação dos instrumentos de coleta; ii) instrumento incapaz de captar de forma acurada as informações desejadas, em geral, resultado de pré-teste insuficiente do instrumento; iii) longo tempo de duração total da entrevista, o que pode levar o entrevistado a responder às perguntas de forma aleatória; iv) treinamento insuficiente dos entrevistadores; v) falta de supervisão e falta de controle de qualidade da coleta de dados em campo (IMAS; RIST, 2009).

Dados secundários

Dados secundários são informações coletadas por outras organizações, que não estão diretamente relacionadas ao programa. Exemplos de dados secundários são censos, pesquisas realizadas por institutos ou órgãos de governos. O uso de dados secundários possui vantagens e desvantagens. Por um lado, pode ser custo-eficiente utilizar esse tipo de fonte de dados já que, por serem dados pré-existent, não há a necessidade de recursos adicionais de coleta para aquele propósito específico. Entretanto, o uso de dados secundários requer alguns cuidados. É preciso investigar se os dados são confiáveis e se os instrumentos de coleta são periodicamente validados (IMAS; RIST, 2009).

Além disso, muitas vezes não é possível medir um objeto de forma direta a partir de dados já existentes. Nesse caso, pode-se optar pelo uso de um indicador *proxy*, isto é, uma variável utilizada para medir outra de difícil mensuração ou para a qual não existem dados, supondo que ambas possuam uma relação estreita entre si. Entretanto, de acordo com Kusek e Rist (2004), só se deve optar pelo uso de indicadores *proxy* quando realmente não há possibilidade de medir os indicadores diretos. Por exemplo, o uso de *proxies* pode ser justificado nos casos em que a coleta tem um elevado custo financeiro, ou em que não é factível coletar os dados para os indicadores diretos periodicamente.

Para diversos temas, é possível encontrar séries temporais de indicadores já calculados para diferentes países:

- Demographic and Health Surveys (DHS);
- Biblioteca de Microdados do Banco Mundial;
- UNdata e
- Global SDG Indicators Database.

Além disso, é possível encontrar séries temporais de indicadores já calculados para diferentes países da África lusófona:

- Instituto Nacional de Estatística (Moçambique);
- Instituto Nacional de Estatística (Cabo Verde);
- Instituto Nacional de Estatística (Guiné-Bissau);
- Instituto Nacional de Estatística (São Tomé e Príncipe);
- Instituto Nacional de Estatística (Angola) e
- Banco Africano de Desenvolvimento.

Para diversos temas, é possível encontrar séries temporais de indicadores já calculados para o Brasil, as unidades da Federação e os municípios. Algumas fontes importantes de informação existentes no Brasil são:

- Ipeadata;
- Sistema IBGE de Recuperação Automática (Sidra);
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) e
- Departamento de Informática do Sistema Único de Saúde (Datasus).

A seguir, serão apresentadas informações adicionais sobre as fontes de dados secundários para o Brasil, mencionadas acima.⁵

Ipeadata

O Ipeadata é uma plataforma mantida pelo Instituto de Pesquisa Econômica Aplicada (Ipea) e disponibiliza dados de diversas fontes organizados em três grupos: macroeconômico, regional e social. Para cada grupo, é possível encontrar diferentes séries de dados disponíveis para *download*, sendo possível selecionar o nível geográfico (p. ex., municípios), a abrangência (p. ex.: região Sudeste) e o período desejado (p. ex.: jan/2000 a dez/2014).

⁵ Biblioteca de Microdados do Banco Mundial. Disponível em: <<https://microdata.worldbank.org/index.php/catalog>>.

No grupo macroeconômico, estão disponíveis dados econômicos e financeiros (p. ex.: séries estatísticas da economia brasileira e internacional), tendo como principais temas: balanço de pagamentos, câmbio, comércio exterior, contas nacionais, correção monetária, finanças públicas, preços, produção e projeções. Algumas fontes disponíveis são: Banco Central do Brasil (Bacen), Banco Mundial, Banco Nacional de Desenvolvimento Econômico e Social (BNDES), Bolsa de Mercadorias & Futuros (BM&FBovespa), Confederação Nacional da Indústria (CNI), Fundação Getulio Vargas (FGV), Fundação Instituto de Pesquisas Econômicas (Fipe), Instituto Brasileiro de Geografia e Estatística (IBGE), Ipea, Ministério da Fazenda, Organização das Nações Unidas (ONU), Centralização de Serviços Bancários (Serasa).

No grupo regional, estão disponíveis dados demográficos, econômicos e geográficos para as regiões, as unidades da Federação e os municípios brasileiros, tendo como principais temas: agropecuária, consumo e vendas, contas regionais, eleições, emprego, finanças públicas, população, produção, renda, segurança pública e transporte. Algumas fontes disponíveis são: Datasus, IBGE, Ipea, Ministério da Fazenda/Secretaria do Tesouro Nacional, Ministério do Trabalho e Emprego e Tribunal Superior Eleitoral (TSE).

Por fim, no grupo social, estão disponíveis dados e indicadores sociais abrangendo temas diversos, como nível de renda *per capita*, desigualdade na distribuição de renda dos indivíduos e domicílios, desempenho educacional, condições de saúde e habitação e inserção no mercado de trabalho, assistência social, demografia, desenvolvimento humano, educação, habitação, mercado de trabalho, renda e saúde. Algumas fontes disponíveis são: IBGE, Ipea, Ministério do Desenvolvimento Social e TSE.

Sidra

O Sidra é disponibilizado pelo IBGE e contém dados já processados de diversas pesquisas, entre elas: Censo Demográfico, Pesquisa de Orçamentos Familiares (POF), Pesquisa Industrial Anual (PIA), Pesquisa Mensal de Emprego (PME), Pesquisa Nacional por Amostra de Domicílios (PNAD) e Pesquisa Nacional por Amostra de Domicílios Contínua Trimestral (PNAD Contínua).

Para uma determinada pesquisa, é possível fazer o *download* de informações em formato de tabela, sendo viável escolher diversas opções, como grupo (p. ex., mulheres), ano (p. ex., 2012) e nível territorial (p. ex., grandes regiões).

Exemplo: PNAD

- população residente, por situação, sexo e grupos de idade;
- pessoas de 10 anos de idade ou mais e valor do rendimento médio mensal, por sexo, situação e classes de rendimento mensal;
- número de filhos tidos, nascidos vivos, no período de referência de 12 meses, tidos pelas mulheres de 15 anos de idade ou mais, por situação do domicílio e grupos de idade das mulheres;
- domicílios particulares permanentes e moradores em domicílios particulares permanentes, por classes de rendimento mensal domiciliar, situação do domicílio e esgotamento sanitário.

Inep

O Inep é uma autarquia federal vinculada ao Ministério da Educação (MEC). Produz informações que subsidiam a formulação de políticas públicas educacionais dos diferentes níveis de governo. Além disso, esses dados podem subsidiar avaliações de programas cujo objetivo é promover o desenvolvimento educacional. Por meio do Inep, é possível acessar dados sobre:

- avaliações, exames e indicadores da educação básica, tais como Sistema Nacional da Educação Básica (Saeb), Índice de Desenvolvimento da Educação Básica (Ideb), Exame Nacional do Ensino Médio (Enem), Exame Nacional para Certificação de Competências de Jovens e Adultos (Encceja), Indicadores Educacionais;
- avaliações, exames e indicadores da educação superior, entre eles, Sistema Nacional de Avaliação da Educação Superior (Sinaes), Exame Nacional de Desempenho de Estudantes (Enade), Exame Nacional de Revalidação de Diplomas Médicos (Revalida), Indicadores de Qualidade da Educação Superior, Sistema de Avaliação de Escolas de Governo (Saeg);
- estatísticas educacionais – censos educacionais (Censo da Educação Superior, Censo Escolar).

Datasus

O Datasus produz informações que subsidiam políticas públicas voltadas para a promoção da saúde. Além disso, essas informações podem ser utilizadas para avaliações de programas cujo objetivo é a promoção da saúde. No Datasus é possível acessar dados sobre indicadores de saúde, tais como morbidade, incapacidade, acesso a serviços, qualidade da atenção, condições de vida e fatores ambientais.

Além disso, o Datasus também possui informações sobre assistência à saúde da população, os cadastros – Rede Assistencial – das redes hospitalares e ambulatoriais, o cadastro dos estabelecimentos de saúde, recursos financeiros e informações demográficas e socioeconômicas da população.



MÓDULO III – INTRODUÇÃO À AVALIAÇÃO DE POLÍTICAS

Avaliar é analisar de forma sistemática uma atividade – que pode ser um projeto ou uma política pública – em planejamento, em andamento ou até mesmo finalizada. Avaliações permitem, por exemplo, investigar as necessidades do público-alvo, aperfeiçoar a gestão e operação do projeto, mensurar a eficácia da política, analisar se os seus benefícios compensam os seus custos e investigar o seu impacto.

Este módulo tem como objetivo, por um lado, discutir a importância e a utilidade de avaliações – temática abordada na unidade 1 – e, por outro, apresentar diferentes tipos de avaliação e diferentes metodologias, permitindo que o leitor saiba quando lançar mão de cada uma delas – temáticas abordadas nas unidades 2 e 3.

Este módulo é de nível introdutório, não tendo como objetivo apresentar o passo a passo da realização de avaliações, tampouco se aprofundar quanto às diferentes metodologias que podem ser utilizadas ao longo do processo avaliativo. Entretanto, ao longo do texto, serão identificadas referências que possibilitem o aprofundamento do aluno nos estudos de áreas específicas de interesse.

Por que e quando avaliar?

Antes de entrarmos em detalhes quanto aos diferentes tipos, e metodologias, de avaliação, é importante discutirmos por que avaliar. Em geral, avaliações são realizadas para gerar evidências que possam contribuir para o processo de tomada de decisão, por parte do gestor, em qualquer etapa do ciclo de uma política pública.

No início do ciclo, a avaliação pode ser utilizada, por exemplo, para decidir se o desenho planejado para uma dada política ou projeto é adequado. Já durante a implementação de uma política, a avaliação permite ao gestor entender se essa política está sendo implementada da forma adequada.

Após a implementação, avaliações de impacto permitem ao gestor compreender se o seu projeto, já implementado, está gerando os impactos esperados. Estes são apenas alguns exemplos de como avaliações respondem a diversos questionamentos dos gestores em diferentes etapas do ciclo da política a depender do seu estágio de maturidade: antes, durante ou depois da sua implementação.

É, então, possível, e desejável, realizar avaliações ao longo de todo o ciclo de política pública; em cada etapa, a avaliação trará um tipo de informação diferente que contribuirá para a tomada de decisão por parte do gestor. Independentemente de quando sejam realizadas, avaliações são um instrumento poderoso para o guiar o gestor ao longo do ciclo de um projeto e devem ser compreendidas como essenciais em um processo de gestão com base em evidências.

Para se aprofundar mais em gestão com base em evidências, com um foco especial no setor público, recomendamos o guia *“Evidence-Based Policymaking: a guide for effective government”* elaborado pela Pew-MacArthur Results First Initiative e o guia *“Ten steps to a results-based monitoring and evaluation system: a handbook for development practitioners”*, editado pelo Banco Mundial.

Antes de prosseguirmos, cabe destacar que, se por um lado, é a etapa do ciclo de política que define o tipo de avaliação a ser realizada, por outro, todas as avaliações devem ser sempre guiadas por questionamentos. São esses questionamentos, denominados perguntas de avaliação, que devem orientar a realização de avaliações ao longo de todo o ciclo e as escolhas metodológicas, que serão discutidas na unidade 3 deste módulo.

As perguntas avaliativas, por sua vez, podem ser organizadas em um plano de avaliação, apresentado no módulo 1. Essa é uma importante ferramenta da gestão com base em evidências, pois contribui para o planejamento e a organização das avaliações que, por sua vez, trarão evidências necessárias para guiar a tomada de decisão.

Idealmente, o plano deve ser elaborado durante a fase de formulação da política, conjuntamente com o plano de monitoramento, e deve considerar a realização de avaliações ao longo de todo o ciclo. A elaboração do plano antes da implementação da política torna possível que não apenas as avaliações sejam inseridas ao longo de todo o ciclo, mas também que essas avaliações sejam pensadas e planejadas com antecedência e junto com o ciclo da política.

Tipos de avaliação de políticas

Diferentes tipos de avaliação respondem a diferentes perguntas avaliativas ao longo das etapas do ciclo de política pública. Nesta unidade, faremos uma breve apresentação de cada um desses diferentes tipos. Para tal, vamos partir de duas categorias de tipos de avaliação: a *ex ante* e a *ex post*.

A **avaliação *ex ante*** é aquela que ocorre antes da implementação da política; já a **avaliação *ex post*** é aquela realizada durante ou após a implementação da política, para analisar o seu desempenho de forma global. Cada um desses grandes grupos engloba diferentes tipos de avaliação, como será discutido posteriormente.

É importante destacar que, neste curso, optamos por categorizar as avaliações de acordo com o momento que ocorrem no ciclo do projeto e, assim, classificamos os tipos de avaliação em *ex ante* e *ex post*. Entretanto, outras categorizações são possíveis. O PNUD (2019), por exemplo, define, nas suas recomendações, que dois tipos de avaliação sejam conduzidos: a de meio-termo e a final.

Enquanto a avaliação de meio-termo permite analisar o andamento do projeto e propor modificações, uma vez que ocorre ao longo do processo de implementação; a avaliação final é um marco importante de *accountability* do projeto, ocorrendo após a sua finalização. Essa categorização é especialmente relevante no contexto de projetos que têm um prazo para o seu término, como é o caso de muitos projetos financiados por organismos internacionais.

Outra categorização possível é entre avaliações formativas e somativas. As formativas ocorrem ao longo da implementação e têm como foco a melhoria do projeto que está sendo avaliado; já as avaliações somativas ocorrem ao final da implementação e têm como objetivo avaliar a efetividade e o impacto de um projeto (IFRC, 2011).

Avaliação *ex ante*

Avaliações *ex ante* contribuem para que a política seja desenhada de forma robusta e tenha capacidade potencial de cumprir os objetivos definidos. Para realizar uma avaliação *ex ante*, é preciso avaliar de forma crítica todas as etapas da pré-implementação da política, incluindo os planos para a implementação e também para o monitoramento e a avaliação. O principal objetivo de realizar uma avaliação *ex ante* é antecipar problemas e riscos que possam vir a ocorrer após o início da implementação e incluir no planejamento soluções para tais problemas e formas de mitigar os possíveis riscos. Pretende-se, então, garantir que a política a ser implementada seja a mais robusta possível, aumentando, assim, a sua chance de ser bem implementada e, potencialmente, atingir os resultados esperados.

Apresentamos a seguir alguns tipos de avaliações que se encaixam nesta categoria. É importante destacar que não pretendemos esgotar os tipos de avaliação *ex ante* que podem ser realizados e que nem sempre será possível realizar todas as etapas apresentadas antes da implementação do projeto, seja por limitações de tempo, seja por escassez de informações ou de recursos.

Avaliação de necessidades

A **avaliação de necessidades** corresponde, muitas vezes, à primeira etapa de uma avaliação *ex ante*. Essa etapa tem como objetivo identificar a natureza e o escopo do problema que motivou a criação da política, além das suas causas e consequências. Para tal, esse tipo de avaliação inclui a definição e a delimitação clara desse problema, que também pode ser entendido como uma necessidade que precisa ser atendida. Essa avaliação inclui a caracterização do problema por meio do uso de indicadores que permitam compreender a sua magnitude e a sua evolução no tempo, além de possibilitar comparações com outros territórios.

Além disso, esse tipo de avaliação inclui a análise das causas e das consequências desse problema – o que permitirá entender, por um lado, as causas prioritárias, sobre as quais a política deve atuar para solucionar o problema, e, por outro, porque é tão importante agir nesse problema – via análise das suas consequências.

Para saber mais sobre este tipo de avaliação, uma boa fonte é o livro “*A guide to assessing needs: essential tools for collecting information, making decisions, and achieving development results*” de Watkins, West Meiers e Visser (2012).

Avaliação da Teoria do Programa

A segunda etapa de uma avaliação *ex ante* corresponde à análise do desenho do projeto que deve ser feita por meio da **avaliação da Teoria do Programa**. Nesta etapa, pretende-se compreender se a política representa um caminho lógico e plausível para solucionar o problema social já diagnosticado; ou seja, se as atividades e os produtos que compõem a política podem, potencialmente, contribuir, de forma relevante, para a solução do problema social. Para isso, é preciso analisar criticamente a Teoria do Programa – que, para tal, precisa estar sistematizada por meio, por exemplo, de uma Teoria da Mudança. Caso essa sistematização não tenha sido feita, essa etapa, então, deve incluir a sua elaboração – abordada em detalhes no módulo 1.

É importante que, nesta etapa, seja considerado o contexto em que o projeto será implementado. Afinal, determinadas hipóteses que dão suporte à Teoria do Programa podem ser verdadeiras em determinados contextos, mas não em outros, e essas particularidades devem ser consideradas nessa avaliação.

Para realizar a avaliação da Teoria do Programa, experiências de políticas semelhantes implementadas em outros territórios ou com outro público-alvo são muito úteis. É possível que tais referências já tenham sido identificadas e contribuído para o seu desenho; em caso afirmativo, essas devem ser analisadas aqui e, em caso negativo, esse esforço de identificação e análise deve ser empreendido. A consulta a especialistas também pode contribuir muito com essa etapa, já que esses, com o seu conhecimento quanto à temática, podem contribuir para a validação da Teoria do Programa.

Nesse ponto, uma fonte interessante de metodologia é o guia do governo canadense, “*Guide to rapid impact evaluation*”.⁶

Além desses dois tipos de avaliação, um processo de avaliação *ex ante* pode ser composto de outros tipos de avaliação. Por exemplo, pode-se analisar criticamente todo o processo de formulação da política, de modo a compreender se a alternativa escolhida – ou seja, a política que será implementada – era, de fato, a melhor opção. Também é possível incluir uma avaliação *ex ante* do plano de implementação – ou seja, do planejamento do uso dos insumos, da execução das atividades

⁶ Disponível em: <<https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/centre-excellence-evaluation/guide-rapid-impact-evaluation.html>>.

e da geração dos produtos – de modo a analisar a sua robustez. Também poderia ser feita uma análise crítica dos planos de monitoramento e de avaliação – discutidos no módulo 1 – quanto à sua robustez e consonância com a Teoria do Programa da política.

É possível, ainda, realizar uma avaliação dos riscos que podem potencialmente prejudicar a execução de uma política. Isso incluiria a identificação dos potenciais riscos e das suas consequências, além da elaboração de formas de mitigar tais riscos, que seriam, então, incluídas no plano de implementação da política. Outro tipo de avaliação *ex ante* é a avaliação dos potenciais impactos da política, sejam esses impactos orçamentários e fiscais, sejam os impactos sobre os resultados esperados.

Em todos esses tipos de avaliação *ex ante*, será importante contar com a experiência dos gestores da política, e de outros especialistas, além de utilizar experiências similares, que já tenham passado por processos de avaliação *ex post*, como comparação.

Para conhecer mais sobre avaliação *ex ante*, boas referências são o guia do governo chileno sobre o processo de avaliação *ex ante*, “*Instrucciones: proceso de evaluación ex ante de diseño de programas no sociales*”; o guia da Comissão Europeia “*Ex ante evaluation: a practical guide for preparing proposals for expenditure programmes*”; o volume 1 do “*Guia para avaliar políticas públicas*”, “A política é nova? Avaliação *ex ante*!”, desenvolvido pelo governo do Espírito Santo em parceria com o FGV EESP CLEAR; e também o “*Avaliação de políticas públicas: guia prático de análise ex ante*”, publicação do governo federal do Brasil.

Avaliação *ex post*

Neste tópico, vamos apresentar os tipos de avaliação que ocorrem durante, ou após, a implementação de uma política. Ao contrário das avaliações descritas no tópico anterior, aqui, as avaliações têm como foco políticas que estão em um estágio mais avançado do seu ciclo – após a etapa de formulação – e que podem, até mesmo, já não estarem mais sendo implementadas. Apesar de responderem a perguntas avaliativas diferentes, como vamos destacar ao longo do tópico, essas avaliações têm como objetivo analisar o desempenho da política implementada.

A seguir, apresentamos os seguintes tipos de avaliação *ex post*: avaliação de desenho, avaliação de processos, avaliação de impacto, avaliação de custo-benefício e avaliação de custo-efetividade. Também discutiremos a avaliação executiva, um processo simplificado de avaliação *ex post*, o qual permite que sejam feitas recomendações quanto a um projeto ou uma política com um baixo uso de recursos e de forma rápida.

Cabe destacar que os tipos aqui apresentados não incluem todas as possíveis avaliações *ex post*. Há outros tipos, como a avaliação da sustentabilidade de um projeto ou a avaliação participativa de um projeto.

Para saber mais sobre esses e outros tipos de avaliação, um recurso útil é a plataforma *Better Evaluation*,⁷ uma iniciativa de difusão de conhecimento sobre avaliação.

Também é importante destacar que não necessariamente é preciso realizar todos os tipos de avaliação aqui apresentados para que um processo de avaliação *ex post* de um programa, ou uma política, seja considerado completo. Muitas vezes, alguns desses tipos foram realizados como parte da avaliação *ex ante* da mesma política. Entretanto, todos esses tipos de avaliação trazem evidências importantes para compreender o ponto mais central do processo de avaliação *ex post*: o desempenho final da política. Deixar de realizar qualquer uma dessas avaliações pode prejudicar esse processo, levando a uma resposta incompleta quanto a esse desempenho. Isso ficará claro, ao longo do tópico, quando discutirmos os tipos de evidência que cada avaliação traz.

Avaliação de desenho

A avaliação de desenho envolve a análise da consistência do desenho de uma política, a partir da comparação com a sua Teoria do Programa, e permite responder a perguntas como:

- Faz sentido supor que uma política com esse desenho consiga potencialmente contribuir para resolver o problema social diagnosticado, ou atender à necessidade social identificada?
- A relação entre as atividades, que fazem parte dessa política, e os resultados esperados é lógica?

Esse tipo de avaliação é essencial para que uma boa avaliação *ex post* de uma política possa ocorrer, pois permite que sejam diagnosticadas falhas no desenho da política, as **falhas de teoria**, discutidas no módulo 1. Caso existam tais falhas, não será surpresa alguma se os resultados e os impactos previstos não tenham sido alcançados. Além disso, avaliações desse tipo também permitem reunir informações que poderão contribuir para o redesenho da própria política ou para o desenho de políticas similares.

Para realizar uma avaliação de desenho, seguindo Rossi, Lipsey e Freeman (2003), duas etapas são necessárias: (i) a explicitação da Teoria do Programa; (ii) a análise crítica dessa teoria. A primeira etapa pode já ter sido realizada, seja no processo de formulação da política ou do programa, seja no processo de avaliação *ex ante*. Nesse caso, cabe, ainda assim, verificar se essa explicitação – via, por exemplo, a elaboração da Teoria de Mudança, como discutido no módulo 1 – apresenta o nível de detalhamento necessário para a segunda etapa da avaliação e, caso necessário, ajustá-lo. Algumas vezes, entretanto, a política foi implementada sem que a Teoria do Programa fosse explicitada; nesse caso, será preciso explicitá-la.

Uma vez finalizada a primeira etapa, pode-se partir para a etapa de análise. Nesse momento, pretende-se checar quão adequadas são as ações da política ou programa para lidar com o problema – ou necessidade – e atingir os resultados esperados, ou seja, pretende-se examinar a consistência

⁷ Disponível em: <<https://www.betterevaluation.org/en>>.

lógica da própria Teoria do Programa. Aqui, é crucial que sejam contrastados os resultados esperados com o problema social – ou a necessidade – que se pretendia lidar. Novamente, caso a política tenha passado por um processo de avaliação *ex ante*, é possível que tal análise já tenha sido feita, como discutimos no tópico anterior.

Também é muito importante que sejam analisadas as relações causais entre as diferentes etapas da implementação, indicadas na Teoria do Programa. Aqui, como a política já está – ou já foi – implementada, é preciso considerar o contexto. Ou seja, é preciso entender se essas relações causais são válidas dado o contexto de implementação.

Avaliação de processos

A avaliação de processos, também conhecida como avaliação de implementação, responde a perguntas sobre o processo de implementação da política, por exemplo:

- Os insumos alocados foram suficientes para a realização das ações planejadas?
- As atividades planejadas foram realizadas no cronograma previsto?
- Os produtos planejados foram entregues aos beneficiários no cronograma previsto e com a qualidade esperada?

Com isso, a avaliação de processos permite verificar se há falhas na implementação do programa e identificar os fatores que comprometeram essa implementação, ao analisar os insumos utilizados, as atividades realizadas e os produtos entregues.

A avaliação de processos contribui para o aperfeiçoamento da política em análise ao reunir informações sobre possíveis problemas na operacionalização da sua implementação. Essas informações também podem servir de insumo para a implementação de outras políticas, em outros contextos. Além disso, entender se a implementação foi de fato bem-sucedida é importante para a análise dos resultados do processo de avaliação *ex post* como um todo, pois uma política mal implementada dificilmente alcançará todos os resultados esperados. Compreender se a implementação foi bem-sucedida – via uma avaliação de processo – é crucial, portanto, para definir se vale a pena realizar outros tipos de avaliação *ex post*, como uma avaliação de impacto ou uma análise de custo-benefício.

Para conduzir uma avaliação de processos, é preciso realizar uma comparação entre o que foi desenhado – e está sistematizado, por exemplo, na Teoria de Mudança quanto aos insumos, às atividades e aos produtos de uma política – e o que foi de fato executado na prática.

Rossi, Lipsey e Freeman (2003) apontam que uma avaliação de processos engloba tanto a análise da utilização dos serviços ou produtos por parte do público-alvo, como a comparação entre o planejamento inicial e o que foi de fato implementado. Importante destacar que esse tipo de avaliação inclui o que se denomina avaliação de resultados, que nada mais é que a comparação entre os produtos entregues e os planejados ou entre os resultados e impactos – utilizando a nomenclatura da Teoria de Mudança – alcançados e os planejados.

Para realizar uma boa avaliação de processos, é preciso contar, então, com uma Teoria do Programa sistematizada de forma detalhada, por meio de, por exemplo, uma Teoria da Mudança, apresentada no tópico 1.2.3 do módulo 1. Também é necessário que haja metas tenham sido bem definidas, tal como discutido no módulo 2. Para saber, por exemplo, se a política conseguiu entregar o pretendido, é preciso ter a referência do que era esperado, a meta de entrega.

Avaliação de impacto

A avaliação de impacto se refere a um desenho bastante específico de avaliação, com uma abordagem prioritariamente quantitativa, que pretende isolar e mensurar os impactos gerados por política. Assim, esse tipo de avaliação busca compreender o quanto a política conseguiu atacar o problema social.

Como os problemas sociais são compostos de diversas camadas, e tipicamente são afetados por diversos fatores, é preciso isolar, com a maior precisão possível, a contribuição específica daquela política para as mudanças ocorridas nos indicadores de interesse. Para tal, é preciso estabelecer uma relação de causalidade entre a presença de uma intervenção e mudanças no resultado de interesse, e, com isso, atribuir qual foi o impacto exclusivamente da intervenção. A pergunta avaliativa que guia essa avaliação é, então:

- Qual é o impacto do projeto sobre os resultados esperados?

No contexto de uma avaliação de impacto, o impacto é, conceitualmente, a diferença, nos resultados de interesse, entre dois cenários: (i) com presença da política; e (ii) na ausência da política.⁸ Pressupõe-se que, após a implementação da política, quaisquer diferenças observadas entre os cenários com relação aos resultados de interesse podem ser atribuídas à política. Para realizar tal análise, é preciso construir esse segundo cenário, denominado contrafactual. Para tal, deve-se selecionar um grupo de controle.⁹

Avaliação de custo-benefício e avaliação de custo-efetividade

A última etapa de um processo de avaliação *ex post* é a realização de uma avaliação de custo-benefício ou ainda de uma avaliação de custo-efetividade. Ambos os tipos de avaliação comparam os custos de uma política, ou programa, com os seus impactos, mas o fazem de formas diferentes. A avaliação de custo-benefício monetiza os impactos de uma política agregando diferentes tipos de impactos, em diferentes resultados, em uma mesma medida. Já a avaliação de custo-efetividade não realiza essa monetização e compara os custos com os impactos de uma política em diferentes resultados.

⁸ É um conceito distinto daquele de “impacto” no arcabouço da Teoria da Mudança. Naquele caso, os impactos se referem aos resultados de mais longo prazo da política; em avaliação de impacto, o que se entende por impacto é o efeito causal da política sobre a variável de interesse.

⁹ Discussões sobre as metodologias de definição desse grupo e de estimação do impacto podem ser encontradas em Gertler et al. (2011).

As perguntas avaliativas a que tais avaliações respondem são, por exemplo:

- Os benefícios de uma política compensam os seus custos?
- Quanto é preciso gastar para alcançar tal resultado?

Para responder a tais questões, esses tipos de avaliação utilizam evidências coletadas em outra avaliação *ex post* – a avaliação de impacto – para avaliar a eficácia de uma política. O primeiro passo para a realização desses dois tipos de avaliação é o cálculo do custo da política em análise. Uma boa referência para compreender o que deve ser considerado nesse cálculo encontra-se em JPAL (2016a). Ali, podem-se encontrar orientações precisas de como realizar o cálculo dos custos de uma política.

É importante destacar que, além dos custos operacionais associados à implementação de uma política, é preciso também considerar custos associados a investimentos em bens de capital e ao pagamento de todos os serviços básicos, necessários para o trabalho da equipe. Além disso, precisam ser inclusos os custos denominados de **custos de oportunidade**, aqueles que estão associados a insumos que não precisam ser pagos pelo uso, mas que, caso não fossem utilizados por essa política, seriam utilizados, e recompensados, de outra forma. Nesse caso, será preciso atribuir um valor a esses insumos.

No caso da avaliação de custo-benefício, torna-se necessário expressar os benefícios na mesma unidade de medida dos custos. Para isso, é preciso transformar os impactos da política em benefícios com valor monetário – processo que é denominado *monetização* – e depende de hipóteses quanto a possíveis ganhos, monetários, da sociedade diante do impacto da política. Após a transformação dos impactos em valores monetários, é feito um ajuste para considerar o efeito do tempo sobre o valor tanto dos custos quanto dos impactos – por meio do cálculo do *valor presente* – e, então, são calculados indicadores como o benefício líquido e a razão custo-benefício. Esses indicadores permitem comparar políticas que têm impactos em resultados diferentes de uma forma simples, além de permitir a agregação de diferentes impactos de uma mesma política. Vai além do escopo desse material discutir em detalhes essa metodologia e os processos de monetização e de cálculo do valor presente.¹⁰

No caso da avaliação de custo-efetividade, após o cálculo dos custos, esses são comparados, por meio de uma razão, com os impactos da política sobre os resultados de interesse. O indicador encontrado, a razão custo-efetividade, mede o investimento necessário para alcançar cada unidade de impacto. Com isso, torna-se possível comparar políticas que têm impactos nos mesmos resultados e identificar qual é capaz de gerar o maior impacto com o menor custo.¹¹

¹⁰ O capítulo 8 de Menezes Filho e Pinto (2017) é uma possível referência para quem desejar aprofundar-se nessa temática.

¹¹ Uma discussão em detalhes sobre esse tipo de análise pode ser encontrada em Dhaliwal et al. (2012).

Se por um lado a avaliação de custo-benefício permite comparar políticas que tenham objetivos similares, mas tenham impacto em resultados diferentes, por outro, a avaliação de custo-efetividade permite uma análise que não depende da monetização dos impactos, que nem sempre é possível ou recomendável. Ambas resumem os impactos de uma política de uma forma simples e sintética, com indicadores de fácil compreensão, o que permite comparações ao longo do tempo e em locais diferentes, mas dependem, por sua vez, de processos não triviais, tanto para a estimação do impacto, quanto para o cálculo do custo total, e, no caso da avaliação de custo-benefício, para a monetização dos impactos.

Avaliação executiva

A avaliação executiva também é um tipo de avaliação *ex post*, já que é realizada em políticas ou programas que já estejam em implementação. Este tipo de avaliação pretende ser uma ferramenta simples que possibilite ao gestor, por um lado, tomar decisões rápidas quanto à sua política e, por outro, definir se, e quais, outras avaliações *ex post* devem ser realizadas para responder, em maior profundidade, a questões que ficaram em aberto após esse primeiro processo avaliativo.

Para tal, o principal insumo são informações já existentes quanto à política que contribuam para uma análise desta e o diagnóstico de possíveis problemas: relatórios e indicadores de monitoramento, informações quanto à execução financeira e pesquisas realizadas com os beneficiários, por exemplo. Essas informações, acrescidas, quando possível, de outras coletadas de forma simples com os formuladores ou, até mesmo, com os beneficiários, são utilizadas para analisar a política desde o seu início, com o diagnóstico do problema social que motivou a criação de tal política, até os resultados alcançados. Esse processo culmina, então, em uma análise crítica que leva à elaboração de recomendações quanto à continuidade ou não da política, bem como quanto à necessidade de realização de outras avaliações em maior profundidade.

A avaliação executiva não permite responder a todos os tipos de perguntas avaliativas, mas é uma ferramenta poderosa para avaliar as dimensões de desenho, operação e resultados da política. A principal limitação talvez seja a sua dependência nos dados existentes, e na sua qualidade: políticas que contem com informações mais bem sistematizadas e sistemas de monitoramento sólidos terão avaliações executivas mais robustas. A principal vantagem, entretanto, é permitir que sejam geradas evidências para a tomada de decisão a partir de um processo simples e que aproveita todas as informações disponíveis de forma sistemática para gerar evidências para a tomada de decisão.

Vamos apresentar a seguir as principais etapas de uma avaliação executiva de forma resumida. Para se aprofundar no tema, boas referências incluem guias do governo chileno, “*Evaluación de programas gubernamentales*”, e do mexicano, “*Términos de referencia de la evaluación de consistencia y resultados*”; além do “*Magenta Book*”, material britânico sobre avaliações. Outras referências incluem o volume 3 do “*Guia para avaliar políticas públicas*”, “Avaliação ao alcance de todos: análise executiva”, desenvolvido pelo governo do Espírito Santo em parceria com o FGV EESP CLEAR; e também o “*Avaliação de políticas públicas: guia prático de análise ex post*”, publicação do governo federal.

Para iniciar o processo de avaliação executiva, é preciso caracterizar bem a política em análise, indicando o problema social que a motivou, descrevendo-a em detalhes. Em uma segunda etapa, é preciso avaliar o diagnóstico do problema social que motivou a formulação da política em análise. Aqui, é importante avaliar se o problema foi bem definido e caracterizado quando a política estava sendo formulada. Também é importante analisar se esse diagnóstico está atualizado; ou seja, é preciso entender se os indicadores que caracterizam esse problema estão sendo acompanhados, e caso não estejam, iniciar esse acompanhamento. O objetivo aqui deve ser compreender se o diagnóstico original, feito no momento da formulação da política, ainda é atual, se a dimensão do problema, assim como a sua distribuição no território ou na população, ainda é similar ao diagnosticado inicialmente. Mudanças no diagnóstico podem levar a recomendações de redesenho da política ou da sua implementação.

Após essa etapa, a avaliação executiva inclui uma análise do desenho da política. Novamente, a Teoria do Programa surge como componente central. Tal como na avaliação de desenho – discutida no início desta unidade –, durante a avaliação executiva, deve ser realizada uma análise crítica da Teoria do Programa, que deve ter sido sistematizada – por meio, por exemplo, da elaboração da Teoria da Mudança da política.

Na sequência, deve-se analisar a implementação. Essa etapa é similar à avaliação de processos, já discutida nesta unidade, mas, por ser parte de um processo simplificado, é realizada com base em informações já existentes e, caso seja necessário, em informações simples, coletadas com baixo custo e de forma rápida. O seu objetivo, porém, é o mesmo: compreender se a implementação foi feita de acordo com o planejado e identificar os aspectos responsáveis para o sucesso – ou não – dessa implementação. Nessa etapa, é importante realizar a análise dos resultados da política. Para tal, devem-se comparar os indicadores de resultado da política – com as suas metas – traçadas durante a formulação da política. Caso as metas não tenham sido cumpridas, é preciso investigar por que isso ocorreu, identificando possíveis problemas ou falhas. Caso metas não tenham sido traçadas para todos os resultados, podem-se utilizar políticas similares como comparação.

Pode ser interessante incluir na avaliação executiva uma etapa de avaliação da percepção dos beneficiários quanto à política, com base ou em dados já existentes ou em uma coleta simples com os beneficiários da política em análise. A ideia é compreender como os beneficiários percebem a política, os seus produtos e as suas atividades, o que contribui não apenas para entender a adequação da política às necessidades do seu público-alvo, mas também para detectar possíveis problemas de implementação ou entender por que certos resultados não são alcançados.

Assim, a avaliação executiva deve englobar uma avaliação de todas as etapas do ciclo de uma política: da identificação do problema à implementação da política. Para culminar esse processo, é preciso fazer uma análise crítica de todas as informações coletadas, de modo a embasar a última etapa desse processo, a de recomendações. Pode-se utilizar uma análise Swot para realizar essa análise.¹²

¹² Veja mais detalhes em “Avaliação de políticas públicas: guia prático de análise ex ante”.

Com base nessa análise crítica, e em todas as informações acumuladas ao longo do processo avaliativo, chega-se, então, às recomendações, a etapa final do processo de avaliação executiva. Essas recomendações podem englobar tanto indicações quanto à expansão – ou não – da política ou ao seu redesenho, como indicações quanto a quais avaliações devem ser desenvolvidas na sequência para aprofundar a avaliação daquela política em determinada área.

Por exemplo, caso a avaliação executiva indique um problema sério de implementação, pode ser recomendada uma avaliação de processo para que seja realizada uma análise mais aprofundada das falhas de implementação. Ainda nesse caso, a avaliação executiva não recomendaria a realização de uma avaliação de impacto, já que um programa mal implementado dificilmente alcançaria os resultados esperados, portanto, não seria recomendado realizar uma avaliação de impacto – e incorrer em custos altos – para, potencialmente, não encontrar nenhum resultado.

Outro exemplo envolve a etapa de análise de diagnóstico: um processo de avaliação executiva pode também recomendar que a implementação do projeto foque uma região diferente, após analisar o diagnóstico atualizado, na sua primeira etapa, e notar que o problema social identificado se tornou mais relevante nessa região. Por fim, uma avaliação executiva pode concluir que a política em análise consegue atingir as metas esperadas, com alto grau de satisfação da população-alvo e, com isso, recomendar a sua expansão.

Boxe 4: Elaborando um plano de avaliação

Como fazer, na prática, um plano de avaliação? Vamos retomar o exemplo do Programa de Infraestrutura Rodoviária, já discutido nas unidades anteriores, para ilustrar esse processo.

O ponto de partida deve ser a Teoria da Mudança do programa. É a partir da análise dessa teoria que vamos definir as perguntas avaliativas. Para avaliações *ex post*, vamos partir das diferentes etapas da Teoria da Mudança e, a depender da etapa da teoria que será base para a definição da pergunta, teremos avaliações de processo, que analisam os insumos, as atividades e os produtos; de resultado, que analisam o atingimento dos resultados; ou de impacto, que analisam o impacto do programa sobre os resultados e os impactos esperados.

Exemplos de perguntas avaliativas, baseadas na Teoria da Mudança do Programa de Infraestrutura Rodoviária, apresentada em uma videoaula do módulo 1, poderiam incluir:

- Para uma avaliação de processo – A malha rodoviária foi pavimentada, ampliada e recuperada?
- Para uma avaliação de resultados – Após o programa, houve um aumento na segurança viária?
- Para uma avaliação de impacto – Qual é o impacto do Programa sobre o desenvolvimento local?

Também podemos pensar em avaliações *ex ante* a partir da análise da Teoria da Mudança. A mais direta seria a avaliação da própria Teoria do Programa, ou seja, a avaliação da teoria subjacente à própria Teoria da Mudança. Um exemplo de pergunta avaliativa poderia ser:

- Há evidências de que pavimentar, ampliar e recuperar a malha rodoviária levaria a uma redução no congestionamento rodoviário e uma facilitação do deslocamento intermunicipal?

Já para pensar em avaliações de necessidade, outro tipo de avaliação *ex ante*, deve-se partir do problema em análise e fazer um diagnóstico completo, incluindo a caracterização completa desse problema e a identificação das suas causas e consequências. No caso do Programa de Infraestrutura Rodoviária, foram dois os problemas diagnosticados que levaram à formulação da política: alto número de acidentes de trânsito e dificuldade de acesso à malha rodoviária pavimentada. Uma avaliação de necessidades desses problemas responderia, por exemplo, às seguintes perguntas:

- Quantos acidentes de trânsito ocorreram no ano passado nas cidades em análise?
- Como o congestionamento rodoviário prejudica a população local?

Assim, a partir da análise da Teoria de Mudança, e também do problema escolhido, podem-se elaborar perguntas avaliativas. Essas perguntas podem, então, ser reunidas em um plano de avaliação, ilustrado na tabela 8, a seguir, e já discutido no módulo 1.

Tabela 7 – Plano de avaliação

pergunta a ser respondida	subperguntas	tipo de avaliação	metodologia	fonte de dados	órgão responsável pela análise de dados

Fonte: Adaptada de World Bank Institute (2009).

Além das perguntas avaliativas, o plano de avaliação reúne algumas outras informações importantes para o planejamento das avaliações. Na segunda coluna, as subperguntas devem ser indicadas. Essas devem ser entendidas como detalhamentos da pergunta principal, que foquem em algum subgrupo específico ou algum indicador específico. Para o caso da pergunta avaliativa de avaliação de impacto detalhada anteriormente, uma subpergunta poderia ser: “O impacto foi superior em cidades que já contavam com certo nível de desenvolvimento?”. Não é necessário desenvolver subperguntas para todas as perguntas avaliativas.

Metodologias para avaliação de políticas

A metodologia escolhida para uma avaliação depende tanto do tipo de avaliação que está sendo realizada, como também das perguntas avaliativas. Diferentes metodologias permitem coletar informações diferentes e, por isso, a escolha de metodologia deve ser feita após a definição do tipo de avaliação e das perguntas avaliativas. A seguir, vamos discutir brevemente métodos quantitativos, qualitativos e mistos.

Cabe, antes de prosseguir, reforçar que não existem escolhas metodológicas melhores ou piores. Existem escolhas metodológicas mais, ou menos, adequadas aos tipos de avaliação e às perguntas avaliativas, por isso é tão importante, primeiro, definir bem as perguntas avaliativas – no plano de avaliação, como discutimos no módulo 1, no tópico 3.2 – e entender o tipo de avaliação que será feito – como discutimos na unidade 2 deste módulo – para então, fazer as escolhas metodológicas.

Nos próximos tópicos, ao discutirmos cada um dos métodos, vamos focar, especialmente, as perguntas avaliativas a que cada um consegue responder. Não é escopo deste curso o aprofundamento em discussões metodológicas.

Métodos quantitativos

Métodos quantitativos são utilizados para analisar aquilo que pode ser mensurado ou quantificado de modo a gerar dados numéricos (REINO UNIDO, 2018; 2020). Essas análises podem incluir desde o uso de simples estatísticas descritivas e a análise de indicadores – como os discutidos no módulo 2 – até a realização de análises econométricas mais complexas. São exemplos de métodos quantitativos: a análise de regressão, a avaliação de impacto aleatorizada e a comparação da evolução de indicadores no tempo.

Tais métodos podem utilizar dados quantitativos coletados como parte do processo avaliativo, por meio de entrevistas, questionários e testes pré-programa e pós-programa. Métodos quantitativos podem também usar dados já existentes previamente, como aqueles encontrados em bases de dados secundárias, como o Censo Escolar do Inep, a Pesquisa Nacional por Amostra de Domicílios (PNAD) do IBGE, o Recenseamentos da População e Habitação do Instituto Nacional de Cabo Verde e o Inquérito sobre Orçamento Familiar (IOF) do Instituto Nacional de Moçambique.

Métodos quantitativos permitem a comparação e a agregação estatística de dados, além de possibilitarem a descrição de tendências gerais e o teste de hipóteses quanto a relações, incluindo causais, entre diferentes fenômenos (REINO UNIDO, 2018; 2020; PRIME, 2020).

Assim, esses métodos permitem responder a perguntas como: “Qual a evolução ao longo do tempo dos indicadores de resultado?”, “Qual foi o impacto desse projeto?”, “Quanto do público-alvo foi atingido pela política?” e “Quanto custou?”. Além disso, esses métodos permitem medir a profundidade e a amplitude da implementação de um programa ou uma política, como o número de pessoas que participaram, o número de pessoas que concluíram o programa, as mudanças ocorridas na renda dos beneficiários e a redução da pobreza e desigualdade.

Por outro lado, análises quantitativas também permitem mensurar tanto os resultados quanto os impactos de um programa ou uma política, por exemplo, a redução do desemprego após a implementação de um programa de treinamento e a contribuição deste para essa redução. Esses métodos permitem ainda comparar o custo de um programa em relação ao seu benefício. Considerando o tipo de perguntas a que tais métodos conseguem responder, estes são os métodos prioritariamente utilizados em avaliações de impacto, de custo-benefício e custo-efetividade.

As vantagens dos métodos quantitativos são a sua capacidade de trazerem resultados representativos – de certa população – além de consistentes e precisos. Entretanto, tais métodos requerem dados que nem sempre são de simples obtenção. Muitas vezes, por exemplo, é preciso coletar de dados, o que pode implicar altos custos. Além disso, esses métodos dependem muito da qualidade dos dados coletados; dados de má qualidade geram análises que também serão de má qualidade. Por fim, apesar de permitirem testar hipóteses sobre relações, incluindo causais, tais métodos nem sempre permitem compreender em profundidade os mecanismos que levam a tais relações.

Para se aprofundar em métodos quantitativos, duas boas referências são “*Quantitative social science: an introduction*”, para análises mais simples; e “*Introdução à econometria: uma abordagem moderna*”, para análises mais complexas. Se o interesse for aprofundar-se em métodos quantitativos no contexto de avaliações, uma boa referência é o livro “*Avaliação de impacto na prática*”.

Métodos qualitativos

Métodos qualitativos são, em geral, definidos como aqueles que usam palavras e textos tanto na coleta quanto na análise de dados (REINO UNIDO, 2018; 2020). Tais métodos envolvem a análise, a comparação, o contraste e a interpretação de padrões a partir de dados coletados, além da descrição de comportamentos e contextos. Nesse caso, a coleta de dados se dá por meio da observação direta do programa – em qualquer etapa do ciclo – além da realização de entrevistas, grupos focais e estudos de caso – com beneficiários, com os gestores do programa, com diferentes partes interessadas, em diferentes contextos de implementação do programa – e também da análise de documentos escritos. São métodos qualitativos: entrevistas em profundidade, grupos focais, análise de documentos, etnografia e observação.

Métodos qualitativos permitem entender, de forma mais aprofundada, resultados e impactos, e também explorar novas questões, além de possibilitarem o aprofundamento maior na análise de processos, relações, valores, comportamentos e atitudes (PRIME, 2020). Assim, esses métodos permitem responder a perguntas como: “Quais são os problemas relevantes?”, “Existem entraves na implementação? Quais?”, “Qual é a percepção dos agentes envolvidos?”, “Como as pessoas atuam?” e “Quais possíveis canais ou mecanismos explicam as mudanças observadas?”.

Métodos qualitativos são capazes de responder a questões relacionadas às barreiras de implementação e também à percepção, dos diferentes atores envolvidos na política, em relação à sua adequação ao problema social ou à qualidade dos produtos entregues. Também permitem

compreender como se dão as relações entre os diferentes responsáveis pela implementação. Dado o tipo de pergunta avaliativa respondido por métodos qualitativos, esses são os mais utilizados em avaliações de processo e de desenho. Entretanto, mesmo em avaliações de impacto, predominantemente quantitativas, métodos qualitativos podem contribuir para entender o objeto de estudo e para compreender os impactos encontrados.

As vantagens desses métodos provêm da capacidade de gerar dados contextuais e da contribuição na análise de assuntos complexos. Além disso, métodos qualitativos são essenciais para a interpretação de resultados provenientes dos dados quantitativos. Entretanto, esses métodos também possuem limitações. Em geral, os seus resultados não podem ser generalizáveis ou representativos – para toda a população beneficiada pelo programa, por exemplo.

Para se aprofundar em métodos qualitativos, uma sugestão é o livro *“Investigação qualitativa e projeto de pesquisa: escolhendo entre cinco abordagens”*.

Métodos mistos

Métodos mistos são aqueles que combinam as duas outras metodologias, quantitativas e qualitativas, tanto no desenho da avaliação, quanto na coleta e na análise de dados e interpretação dos resultados. Quando implementados corretamente, métodos mistos permitiriam uma sinergia entre os métodos apresentados anteriormente, de forma que ambos contribuam para a validação e qualidade das evidências encontradas.

A ideia, ao utilizar métodos mistos, não é apenas realizar parte das análises com métodos quantitativos e parte com métodos qualitativos, tampouco usar um método apenas para completar ou explicar os resultados do outro método. Ao contrário, métodos mistos devem envolver a utilização de análises quantitativas e qualitativas de forma combinada ao longo de todo o processo de pesquisa. O desenho da pesquisa deve ser guiado pelo objetivo desta, que levará à elaboração das perguntas que precisam ser respondidas e, então, aos métodos escolhidos. O objetivo é que o uso de métodos mistos leve a uma análise mais completa e de maior qualidade, que dê informações mais ricas para o processo de tomada de decisão.

Um bom exemplo seria o uso de métodos mistos para realizar uma avaliação de processo: os métodos quantitativos poderiam ser utilizados para analisar os indicadores de monitoramento da implementação, e os qualitativos para compreender as barreiras à implementação e a percepção dos beneficiários quanto à qualidade dessa implementação. Ambos os tipos de informação são úteis para o gestor da política e enriquecem o processo avaliativo.

Outro exemplo seria a realização de uma avaliação de impacto com métodos mistos: os métodos quantitativos poderiam ser utilizados para isolar o impacto da política, por meio, por exemplo, da realização de uma avaliação de impacto aleatorizada, e os métodos qualitativos poderiam ser utilizados para compreender os mecanismos que levaram ao impacto encontrado ou, então, entender as heterogeneidades encontradas na análise quantitativa. Novamente, o uso dos diferentes métodos enriquece o resultado da análise de impacto e dá ao gestor da política informações mais ricas para a sua tomada de decisão.

Na prática, esses ainda são muito pouco utilizados, apesar da utilidade e das vantagens de combinar ambos os métodos anteriores. Talvez uma importante razão para essa falta de uso seja o fato de que pesquisadores, tanto qualitativos quanto quantitativos, muitas vezes, são treinados em uma só metodologia e, geralmente, não valorizam, tampouco estão acostumados com a utilização da outra metodologia, tendo dificuldade de implementar um desenho de pesquisa que seja de fato baseado em métodos mistos. Além disso, para realizar uma pesquisa baseada em métodos mistos, é preciso contar com pesquisadores especialistas em ambos os métodos, e tal configuração da equipe nem sempre é possível.

Para se aprofundar em métodos mistos, uma sugestão são os livros “*Q-squared: combining qualitative and quantitative approaches in poverty analysis*” e “*Designing and conducting mixed methods research: second edition*”.

Boxe 5: Elaborando um plano de avaliação (continuação)

Para finalizar o plano de avaliação, apresentado inicialmente no boxe 4, para o exemplo do Programa de Infraestrutura Rodoviária, é preciso definir a metodologia, preenchida na terceira coluna; as fontes de dados, detalhadas na quarta coluna; e, caso seja possível, também os responsáveis pela análise dos dados, apontados na quinta coluna.

Como discutido na unidade anterior, a metodologia será definida a partir da pergunta avaliativa e do tipo de avaliação. Recuperando o boxe 4, podemos, por exemplo, definir que a avaliação de resultado com a pergunta avaliativa “Após o programa, houve um aumento na segurança viária?” será respondida via comparação dos indicadores de segurança viária antes e depois da implementação do programa, ou seja, por meio de métodos quantitativos.

Já para a pergunta relacionada à avaliação de necessidades, “Como o alto número de acidentes de trânsito prejudica a população local?”, podemos optar por realizar grupos focais com a população local, ou seja, respondê-la por meio de um método qualitativo. Por fim, para a pergunta de avaliação de impacto, “Qual o impacto do Programa sobre o desenvolvimento local?”, podemos optar por utilizar métodos mistos e primeiro utilizar métodos qualitativos para compreender quais seriam os possíveis impactos e depois, aplicar uma pesquisa quantitativa com a população local para mensurar tais impactos.

Os tipos de fonte de dados serão consequência direta da metodologia definida e dos tipos de dados existentes. Cabe destacar que é possível utilizar tanto dados primários, coletados para fins de uma pesquisa específica, quanto dados secundários, de outras instituições. As diferenças entre esses tipos de dados foram discutidas na unidade sobre Fonte de Dados no módulo 2.

Por fim, para definir os responsáveis pela análise, é importante mapear as organizações envolvidas na implementação do projeto, de modo a identificar quais delas teriam a capacidade de realizar diferentes avaliações. Para realizar uma avaliação de impacto, por exemplo, seria importante contar com um especialista em metodologias de avaliação de impacto na equipe.

Aqui também vale destacar outra discussão: a depender do tipo de projeto, pode ser interessante contar com avaliadores externos ao projeto. Por exemplo, caso o projeto seja financiado por uma organização internacional, é muito comum que seja exigida uma avaliação de impacto externa. Assim, essa questão também deve ser levada em consideração na definição dos responsáveis por cada tipo de avaliação.

Com isso, completamos o preenchimento do plano de avaliação. Ao fim desse processo, o gestor contará com um planejamento inicial das avaliações que serão realizadas ao longo do ciclo de uma política. Essa ferramenta será muito útil no processo de gestão com base em evidências!



CONCLUSÃO

Ferramentas de Monitoramento & Avaliação no campo de políticas públicas são cruciais para gerar evidências ao longo do ciclo da política, possibilitando ao gestor identificar o que funciona e o que não funciona. Monitoramento e avaliação permitem a identificação de boas práticas, por um lado, e de pontos a serem aperfeiçoados no caminho da implementação da política, por outro. Assim, M&A são fundamentais para uma gestão com base em resultados, já que permitem gerar informações que possibilitam o aperfeiçoamento da execução das políticas públicas, podendo auxiliar assim na melhoria da qualidade do gasto público.

Esta apostila apresentou conceitos introdutórios sobre monitoramento e avaliação de políticas. Para o aluno que deseja aprofundar-se nos diferentes tópicos mencionados ao longo deste curso, sugerimos seguir a literatura indicada nas diferentes unidades, e acompanhar o trabalho executado pelo FGV EESP CLEAR, cuja missão é promover políticas públicas baseadas em evidências e apoiar a capacitação de profissionais em ferramentas de M&A.

BIBLIOGRAFIA

ALKIRE, S.; SANTOS, M. E. *Acute multidimensional poverty: a new index for developing countries*. Technical Report 38. Oxford: University of Oxford, 2010.

ALKIRE, S.; FOSTER, J. Counting and multidimensional poverty measurement. *Journal of Public Economics*, v. 5, n. 7, p. 476-487, 2011.

ANAND, S.; SEGAL, P.; STIGLITZ, J. (Eds.) *Debates on the measurement of global poverty*. Oxford: Oxford University Press, 2010.

ATKINSON, A.; BOURGUIGNON, F. The comparison of multi-dimensioned distributions of economic status. *Review of Economic Studies*, v. 49, n. 2, p. 183-201, 1982.

BAKEWELL, O.; GARBUTT, A. *The use and abuse of the logical framework approach*. Stockholm: Sida, 2005.

BRASIL. Casa Civil da Presidência da República, Instituto de Pesquisa Econômica Aplicada. *Avaliação de políticas públicas: guia prático de análise ex ante*. Brasília: Ipea, 2018. v. 1. Disponível em: <http://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=32688>.

BRASIL. Casa Civil da Presidência da República, Instituto de Pesquisa Econômica Aplicada. *Avaliação de políticas públicas: guia prático de análise ex post*. Brasília: Ipea, 2018. v. 2. Disponível em: <http://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=34504>.

BUSSAB, W. O.; MORETTIN, P. *Estatística básica*. São Paulo: Saraiva, 2017.

CANADÁ. Centre of Excellence for Evaluation. *Guide for Rapid Evaluation*. Disponível em: <<https://www.canada.ca/en/treasury-board-secretariat/services/audit-evaluation/centre-excellence-evaluation/guide-rapid-impact-evaluation.html>>. Acesso em: maio 2020.

CHAKRAVARTY, S. R.; SILBER, J. Measuring multidimensional poverty: the axiomatic approach. In: *Quantitative approaches to multidimensional poverty measurement*, p. 192-209. Springer, 2008.

CHILE. Dirección de Presupuestos do Ministerio de Hacienda. *Instrucciones*: proceso de evaluación ex ante de diseño de programas no sociales. Santiago, 2016. Disponível em: <http://www.dipres.gob.cl/598/articles-73342_Instructivo_2015_VF.pdf>.

CHILE. Dirección de Presupuestos do Ministerio de Hacienda. *Evaluación de Programas Gubernamentales* (EPG). Disponível em: <<http://www.dipres.gob.cl/598/w3-article-111762.html>>. Acesso: maio 2020.

COMIM, F.; BORSI, M. T.; VALERIO MENDOZA, O. *The multi-dimensions of aporophobia*. [S.l.: s.n.], 2020.

COMISSÃO EUROPEIA. *Ex ante evaluation*: a practical guide for preparing proposals for expenditure programmes. Disponível em: <https://www.mfcr.cz/assets/cs/media/EU-MFCR_Pr-007_2001-12_Hodnoceni-ex-ante-Prakticke-pokyny-pro-zpracovani-navrhu-vydajovych-programu.pdf>. Acesso: jun. 2020.

CRESWELL, J. *Investigação qualitativa e projeto de pesquisa*: escolhendo entre cinco abordagens. Porto Alegre: Penso, 2014.

CRESWELL, J; CLARK, V. L, P. *Designing and conducting mixed methods research*. 3. ed. Thousand Oaks, CA: Sage, 2018.

DHALIWAL, I.; DUFLO, E.; GLENNERSTER, R.; TULLOCH, C. *Comparative cost-effectiveness analysis to inform policy in developing countries*: a general framework with applications for education. Abdul Latif Jameel Poverty Action Lab (J-PAL), MIT, 2012. Disponível em: <https://www.povertyactionlab.org/sites/default/files/resources/CEA%20in%20Education%202013.01.29_0.pdf>.

FIELDS, G. S. *Distribution and development*: a new look at the developing world. Cambridge: The MIT Press, 2002.

FOSTER, J.; GREER, J.; THORBECKE, E. A class of decomposable poverty measures. *Econometrica*: journal of the econometric society, p. 761-766, 1984.

FOSTER, V.; BRICEÑO-GARMENDIA, C. M. (Eds.). *Africa's infrastructure*: a time for transformation. Washington, DC: The World Bank, 2010. Disponível em: <<https://openknowledge.worldbank.org/handle/10986/2692>>.

GERTLER, P. J.; MARTINEZ, S.; PREMAM, P.; RAWLINGS, L. B.; VERMEERSCH, C. M. J. *Avaliação de impacto na prática*. Washington, DC: Banco Mundial, 2011. Disponível em: <<https://openknowledge.worldbank.org/bitstream/handle/10986/2550/9781464800887.pdf?sequence=19>>.

GORGENS, M.; KUSEK, J. Z. *Making monitoring and evaluation systems work: a capacity development toolkit*. Washington, DC: The World Bank, 2009.

IMAI, K. *Quantitative social science: an introduction*. Princeton: Princeton University Press, 2017.

IMAS, L. G.; RIST, R. *The road to results: designing and conducting effective development evaluations*. Washington, DC: The World Bank, 2009.

INSTITUTO JONES DOS SANTOS NEVES. *Guia para avaliar políticas públicas: a política é nova? Avaliação ex ante!* Vitória, ES: IJSN, 2018a. v. 1. Disponível em: <<http://www.ijsn.es.gov.br/component/attachments/download/6418>>.

INSTITUTO JONES DOS SANTOS NEVES. *Guia para avaliar políticas públicas: avaliação ao alcance de todos – análise executiva*. Vitória, ES: IJSN, 2018b. v. 3. Disponível em: <<http://www.ijsn.es.gov.br/component/attachments/download/6420>>.

INSTITUTO JONES DOS SANTOS NEVES. *Guia para avaliar políticas públicas: e quando a política está em andamento? Avaliação ex post*. Vitória, ES: IJSN, 2018c. v. 4. Disponível em: <<http://www.ijsn.es.gov.br/component/attachments/download/6421>>.

INTERNATIONAL FEDERATION OF RED CROSS AND RED CRESCENT SOCIETIES. *Project/Programme Monitoring and Evaluation (M&E) Guide*. Geneva: IFRC, 2011.

J-PAL – ABDUL LATIF JAMEEL POVERTY ACTION LAB. *J-PAL Costing Guidelines*. Disponível em: <<https://www.povertyactionlab.org/sites/default/files/resources/costing-guidelines.pdf>>. Acesso em: maio 2020.

KRAFT, M. E.; FURLONG, S. R. *Public policy: politics, analysis, and alternatives*. [S.l.]: CQ Press, 2019.

KUSEK, J. Z.; RIST, R. *Ten steps to a results-based monitoring and evaluation system: a handbook for development practitioners*. Washington, DC: The World Bank, 2004.

MENEZES FILHO, N.; PINTO, C. C. X. *Avaliação econômica de projetos sociais*. São Paulo: Fundação Itaú Social, 2017.

MÉXICO. Consejo Nacional de Evaluación de la Política de Desarrollo Social (Coneval). *Términos de Referencia de la Evaluación de Consistencia y Resultados*, 2017. Disponível em: <https://www.coneval.org.mx/Evaluacion/MDE/Documents/TDR_ECR.pdf>. Acesso em: maio 2020.

MOSSE, Roberto; SONTHEIMER, Leigh Ellen. *Performance monitoring indicators handbook*. Washington, DC: The World Bank, 1996.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. Development Assistance Committee. Working Party on Aid Evaluation. *Glossary of key terms in evaluation and results-based management*. Paris: OCDE, 2002. Disponível em: <<https://www.oecd.org/dac/evaluation/2754804.pdf>>.

ORTEGÓN, E.; PACHECO, J. F.; PRIETO, A. *Metodología del marco lógico para la planificación, el seguimiento y la evaluación de proyectos y programas*. Santiago: Cepal, 2005, p. 13-28.

OSBORNE, D.; GAEBLER, T. *Reinventing government: how the entrepreneurial spirit is transforming the public sector*. Reading, MA: Addison-Wesley, 1992.

PEW-MACARTHUR RESULTS FIRST INITIATIVE. *Evidence-based policymaking: a guide for effective government*. Washington, DC: The Pew Charitable Trusts, John D. and Catherine T. MacArthur Foundation, 2014.

PNUD. *Indicators for policy management: a practical guide for enhancing the statistical capacity of policy-makers for effective monitoring of the MDGs at the country level*, 2005. Disponível em: <https://www.undp.org/content/dam/aplaws/publication/en/publications/poverty-reduction/poverty-website/indicators-for-policy-management/Indicators_for_Policy_Management.pdf>.

PNUD. *Human Development Report 2019*. Beyond income, beyond averages, beyond today: inequalities in human development in the 21st century, 2019.

PNUD. Independent Evaluation Office. *UNDP Evaluation Guidelines*. Jan. 2019. Acesso em: maio 2020. Disponível em: <http://web.undp.org/evaluation/guideline/documents/PDF/UNDP_Evaluation_Guidelines.pdf>.

PNUD. United Nations Development Programme and Oxford Poverty and Human Development Initiative. *Charting pathways out of multidimensional poverty: achieving the SDGs*, 2020. Disponível em: <http://hdr.undp.org/sites/default/files/2020_mpi_report_en.pdf>. Acesso em: out. 2020.

PROGRAM IN RURAL MONITORING AND EVALUATION. *Fundamentals of M&E 1. 2. ed.* Presentation. [S.l.]: PRiME, 2018. Disponível em: <<https://www.primetraining.global/materials/fundamentals>>.

RAVALLION, M. *The economics of poverty: history, measurement and policy*. Oxford: Oxford University Press, 2016.

REINO UNIDO. Public Health. *Guidance: evaluation methods*, 2018. Disponível em: <<https://www.gov.uk/government/publications/evaluation-in-health-and-well-being-overview/evaluation-methods#combining-qualitative-and-quantitative-research>>. Acesso em: jun. 2020.

REINO UNIDO. HM Treasury. *The Magenta Book*, 2020. Disponível em: <<https://www.gov.uk/government/publications/the-magenta-book>>. Acesso em: maio 2020.

ROBERTS, D.; KHATTTRI, N. *Designing a results framework for achieving results: a how-to guide*. Washington, DC: Independent Evaluation Group, World Bank, 2012.

ROSSI, P. H.; LIPSEY, M. W.; FREEMAN, H. E. *Evaluation: a systematic approach*. Thousand Oaks, CA: Sage, 2003.

RUBIO, G. M. The design and implementation of a menu of evaluations. The Nuts & Bolts of M&E Systems, n. 6. Washington, DC: World Bank, 2011.

SAUNDERS, R. P.; EVANS, M. H.; JOSHI, P. Developing a process-evaluation plan for assessing health promotion program implementation: a how-to guide. *Health Promotion Practice*, v. 6, n. 2, p. 134-147, 2005.

SECCHI, L. *Políticas públicas: conceitos, esquemas de análise, casos práticos*. São Paulo: Cengage Learning, 2014.

SEN, A. Poor, relatively speaking. *Oxford Economic Papers*, p. 153-169, 1983.

SEN, A. A decade of human development. *Journal of Human Development*, v. 1, p. 17-23, 2000.

WATKINS, R.; WEST MEIERS, M.; VISSER, Y. L. *A guide to assessing needs*: essential tools for collecting information, making decisions, and achieving development results. Washington, DC: World Bank, 2012. Disponível em: <<https://openknowledge.worldbank.org/handle/10986/2231>>.

WOOLDRIDGE, J. M. *Introdução à econometria*: uma abordagem moderna. São Paulo: Cengage CTP, 2010.

WORLD BANK INSTITUTE. *Capacity development and results*. Monitoring and evaluation for results: training for trainers, 2009.

WORLD BANK. *Operations policy and quality department*. Results framework and M&E: Guidance Note, 2013. Disponível em: <http://siteresources.worldbank.org/PROJECTS/Resources/409401365611011935/Guidance_Note_Results_and_M&E.pdf>.

FONTES DE DADOS

Banco Africano de Desenvolvimento.

Disponível em: <<https://www.afdb.org/en/knowledge/statistics>>.

Biblioteca de Microdados do Banco Mundial.

Disponível em: <<https://microdata.worldbank.org/index.php/catalog>>.

Demographic and Health Surveys (DHS).

Disponível em: <<https://www.dhsprogram.com/What-We-Do/Survey-Types/DHS.cfm>>.

Departamento de Informática do Sistema Único de Saúde (Datasus).

Disponível em: <www.datasus.saude.gov.br>.

Global SDG Indicators Database.

Disponível em: <<https://unstats.un.org/sdgs/indicators/database>>.

Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep).

Disponível em: <www.portal.inep.gov.br>.

Instituto Nacional de Estatística (Angola).

Disponível em: <<https://www.ine.gov.ao/indicadores-estatisticos>>.

Instituto Nacional de Estatística (Cabo Verde).

Disponível em: <<http://ine.cv/estatisticas-por-tema>>.

Instituto Nacional de Estatística (Guiné-Bissau).

Disponível em: <<http://www.stat-guinebissau.com>>.

Instituto Nacional de Estatística (Moçambique).

Disponível em: <<http://www.ine.gov.mz/estatisticas>>.

Instituto Nacional de Estatística (São Tomé e Príncipe).

Disponível em: <<https://www.ine.st>>.

Ipeadata.

Disponível em: <www.ipeadata.gov.br>.

Sistema IBGE de Recuperação Automática (Sidra).

Disponível em: <www.sidra.ibge.gov.br>.

UNdata.

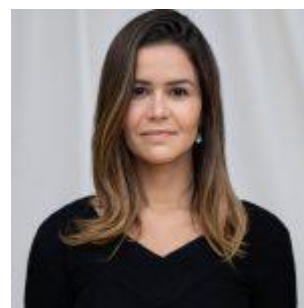
Disponível em: <<http://data.un.org>>.

PROFESSORES-AUTORES

André Portela de Souza, diretor do FGV EESP CLEAR, é graduado em Economia, pela Universidade Federal da Bahia (UFBA), mestre em Economia, pela Universidade de São Paulo (USP), e doutor em Economia, pela Universidade Cornell. Foi professor doutor da Universidade de São Paulo, professor assistente adjunto da Universidade Vanderbilt e professor visitante da Universidade Cornell. Atualmente, é professor da Escola de Economia de São Paulo, da Fundação Getulio Vargas (FGV EESP), coordenador do Centro de Microeconomia Aplicada (C-Micro) e diretor do FGV EESP CLEAR.



Lycia Lima, vice-diretora do FGV EESP CLEAR, é doutora em Administração Pública e Governo, com ênfase em Política e Economia do Setor Público, pela Escola de Administração de Empresas de São Paulo, da Fundação Getulio Vargas (FGV EAESP), com período sanduíche no Massachusetts Institute of Technology (MIT), mestra em Economia do Desenvolvimento, pela Universidade de Londres (Soas), e graduada em Economia, pela Universidade Federal de Minas Gerais (UFMG). Atua como professora do curso de graduação da FGV EAESP, pesquisadora associada do Centro de Microeconomia Aplicada (C-Micro – EESP FGV) e do Centro de Economia e Política do Setor Público (FGVCEPESP). Previamente, ocupou posições no Banco Mundial, em Washington; na ONU, em Genebra; na Secretaria de Assuntos Estratégicos da Presidência da República; e no Governo de Minas Gerais. A sua principal área de pesquisa é Microeconomia Aplicada, principalmente na avaliação de políticas públicas que focam investimento em capital humano: primeira infância, educação e mercado de trabalho. Também desenvolve pesquisas nas áreas de Economia Urbana e Finanças Públicas.



EQUIPE TÉCNICA

Camila Soares é economista formada pela Faculdade de Ciências Econômicas, da Universidade Federal de Minas Gerais (Face-UFMG), com mestrado, também em Economia, pelo Programa de Pós-Graduação em Economia, da Universidade Federal do Rio de Janeiro (UFRJ), e doutoranda do Programa de Administração Pública e Governo, da FGV EAESP. Como parte do doutorado, foi *visiting fellow* do Departamento de Sociologia, da Universidade de Harvard, em 2019. Anteriormente, trabalhou como analista na área de Política Pública do Insper, em projetos da Cátedra Instituto Ayrton Senna. Tem larga experiência em políticas públicas, com especial ênfase em avaliação de políticas.



Gabriel Weber Costa é bacharel em Ciências Econômicas, pela Universidade Federal de Pelotas (UFPel), mestre em Economia, pela FGV EESP, e doutorando em Economia, pela mesma instituição. Atualmente, pesquisador do FGV EESP CLEAR. Atuou como consultor do Banco Mundial. Possui experiência em metodologias de monitoramento e avaliação, em particular desenhando e conduzindo avaliações de impacto, incluindo projetos nas áreas de Educação, Saúde e Políticas Sociais.



Juliana Camargo é doutora em Economia, pela FGV EESP, mestra e graduada em Economia, pela Universidade Federal do Rio Grande do Sul (UFRGS). Possui experiência na área de Avaliação de Impacto de Políticas Públicas e em programas sociais com foco em educação, primeira infância, habilidades socioemocionais, desigualdade, desenvolvimento local e pobreza. Possui experiência coordenando equipe de coleta de dados primários para a avaliação de impacto e desenvolvimento de Teoria da Mudança de programas. Atualmente, é pesquisadora do Centro FGV EESP CLEAR e do C-Micro – Centro em Microeconomia Aplicada.



Marina Ribeiro é graduada em Jornalismo, pela Escola de Comunicações e Artes da Universidade de São Paulo (ECA-USP). Atuou como repórter nas principais empresas de comunicação do País, até fazer uma transição de carreira para atuar com políticas públicas no terceiro setor. É mestranda em Políticas Públicas, no Insper, onde também atuou como assistente de pesquisa no Núcleo de Medição para Investimentos de Impacto Socioambiental (Insper Metricis). Atua com difusão de evidências e conhecimento e relações institucionais no FGV EESP CLEAR.





*CENTRO DE APRENDIZAGEM EM
AVALIAÇÃO E RESULTADOS PARA
O BRASIL E A ÁFRICA LUSÓFONA*