

**Tribunal de Contas da União (TCU)
Senado Federal (SF)
Secretaria de Orçamento Federal (SOF)
Controladoria-Geral da União (CGU)**

ESPECIALIZAÇÃO EM ORÇAMENTO PÚBLICO

Instituto Serzedello Corrêa - TCU

**PERSPECTIVAS PARA AVALIAÇÃO DE
INDICADORES DE DESEMPENHO DE
PROGRAMAS DE GOVERNO**

Autor: Victor Raymond Steytler

Orientador: Me. Francisco Eduardo de Holanda Bessa

Brasília – DF
2010

**Tribunal de Contas da União (TCU)
Senado Federal (SF)
Secretaria de Orçamento Federal (SOF)
Controladoria-Geral da União (CGU)**

ESPECIALIZAÇÃO EM ORÇAMENTO PÚBLICO

**PERSPECTIVA PARA AVALIAÇÃO DE
INDICADORES DE DESEMPENHO DE
PROGRAMAS DE GOVERNO**

Victor Raymond Steytler

Orientador: Me. Francisco Eduardo de Holanda Bessa

Artigo apresentado ao Instituto Serzedello Corrêa
– ISC/TCU, como requisito parcial à obtenção do
grau de Especialista em Orçamento Público.

**BRASILIA - DF
2010**

DEDICATÓRIA

À minha família, pelo apoio incondicional, desde sempre.

AGRADECIMENTOS

Ao meu orientador, Eduardo Bessa, pela compreensão, pela empolgação e pela sintonia de ideias.

À Solange e ao Matheus, pelo apoio total.

À Martha, pelas opiniões e sugestões.

À Ane, pelas observações finais.

Aos meus pais, pela paciência e dedicação.

Ao Dilson e ao Celso, pelo empurrão inicial.

À Larissa, por todos os empurrões.

Aos colegas da DIE, pelas opiniões.

Aos colegas do curso, pela alegria e pelo companheirismo, que foram elementos essenciais para a minha motivação; e também pela seriedade e dedicação, que elevaram o nível do curso.

RESUMO

Diante da necessidade de avaliação de programas de governo por meio de indicadores de desempenho, conforme estabelecido no Manual de Elaboração do Plano Plurianual 2008-2011, surge a necessidade de verificar em que medida esse tipo de avaliação apresenta resultados que atendam às necessidades dos *stakeholders* (partes interessadas) e qual a qualidade dessas avaliações. Por essa trilha, este artigo apresenta propostas de diretrizes para uma meta-avaliação. Essa proposta é construída a partir de determinadas categorias de indicadores, submetendo-as a outras duas dimensões: as necessidades dos *stakeholders* e os critérios de qualidade dos indicadores. As categorias de indicadores e as assim chamadas “dimensões” foram estudadas por meio de uma revisão bibliográfica. Na sequência, foi estabelecido um parâmetro para a dimensão *stakeholders*, chamado de nível de importância de um conjunto de indicadores, e foi sugerido um índice de cobertura da necessidade de informação dos *stakeholders* para cada categoria de indicador delineada. Na fase seguinte, concernente à dimensão de “critérios de qualidade”, os referidos critérios foram compilados e separados em essenciais e desejáveis, para cada uma das categorias de indicadores existentes. Propõe-se que o não atendimento a um critério essencial resulta em classificação do indicador como insuficiente. Para os critérios desejáveis, foi considerada a relação entre a quantidade de critérios de qualidade atendidos pelo indicador e o total existente, resultando no índice de qualidade de cada categoria. Para um sistema de indicadores, esses índices de qualidade são ponderados e o resultado é multiplicado pelo índice de cobertura do indicador, proporcionando o índice de efetividade geral do sistema de indicadores analisado.

Palavras-chave: indicador, índice, desempenho, programa, meta-avaliação.

ABSTRACT

Considering the necessity to evaluate government programs through performance indicators, as established in the Manual for the Preparation of Multi-Year Plan 2008-2011, it is necessary to check how comprehensive this evaluation might be, so that it matches the needs of stakeholders and what should be the expected quality of these assessments. In such way, this article presents proposals for guidelines for a meta-evaluation. This proposal is built for certain categories of indicators which are submitted to two other dimensions: the needs of the stakeholders and the quality criteria. The categories of indicators and so-called "dimensions" were studied by means of a literature review. Following, it was established a parameter for the dimension stakeholders, which is the level of importance of a set of indicators, and it was suggested a coverage rate of the needs of stakeholders for each indicator category. In the next stage, the dimension of "quality criteria" was compiled, and the collected criteria were separated into essential and desirable for each category of existing indicators. It is proposed that the absence of the essential criteria should result in a classification of the indicator as insufficient. For the desirable criteria, it was considered the fraction between the quantity of existing quality criteria and the total, resulting in a quality index for each category. For a system of indicators, these quality scores are weighted and the result is multiplied by the coverage rate of the indicator, resulting in an index of overall effectiveness of the system of indicators analyzed.

Keywords: *indicator, index, performance, program, meta-evaluation*

SUMÁRIO

1. Introdução.....	1
1.1. Apresentação do Problema e necessidade de Pesquisa.....	1
1.2. Objetivos.....	3
1.2.1. Objetivo Geral.....	3
1.2.2. Objetivos Específicos.....	3
1.3. Desenvolvimento.....	3
2. Fundamentação teórica.....	4
2.1. O que são indicadores de desempenho.....	4
2.1.1. A utilidade dos indicadores de desempenho.....	5
2.1.2. Os componentes de um indicador de desempenho.....	7
2.1.3. Indicadores simples e complexos.....	9
2.1.4. As Categorias de indicadores de desempenho.....	10
2.1.5. As tipologias de indicadores de desempenho.....	13
2.1.6. Critérios de qualidade essenciais e desejáveis de um indicador de desempenho.....	14
2.1.6.1. Critérios de qualidade dos indicadores de desempenho.....	15
3. Metodologia.....	22
4. Proposta de abordagem para uma meta-avaliação.....	24
4.1. Considerações para a construção da abordagem proposta.....	24
4.2. A necessidade dos <i>stakeholders</i> em relação a cada categoria de indicador – primeiro confronto dimensional.....	25
4.2.1. O nível de importância das categorias de indicadores.....	27
4.2.2. O índice de cobertura de um sistema de indicadores de desempenho.....	29
4.3. Categorias de indicadores submetidas aos critérios de qualidade – segundo confronto dimensional.....	30
4.3.1. Critérios essenciais de qualidade.....	30
4.3.2. Critérios desejáveis de qualidade.....	30
4.3.3. O nível de efetividade geral do sistema de indicadores de avaliação de um programa.....	31
5. Conclusão.....	32
6. Referências bibliográficas.....	36

1. Introdução.

1.1. Apresentação do Problema e necessidade de Pesquisa.

O Manual de Elaboração do Plano Plurianual (PPA) 2008-2011 (BRASIL, 2007, p. 9) traz uma lista de princípios que devem orientar a elaboração e a gestão do referido Plano. Entre esses princípios destacam-se “a gestão estratégica dos projetos e programas considerados indutores do desenvolvimento para assegurar o alcance dos resultados pretendidos” e “o monitoramento e a avaliação dos projetos e programas de Governo, criando condições para a melhoria contínua e mensurável da qualidade e produtividade dos bens e serviços públicos”.

É clara a relação entre esses dois princípios apresentados. A qualidade dos bens e serviços públicos deve ser mensurável para que haja melhoria contínua. Esta melhoria se dá por meio da gestão estratégica, que se serve do monitoramento e da avaliação dos projetos e programas de Governo para realizar os ajustes necessários ao atingimento dos objetivos pretendidos.

O Manual do PPA estabelece também que “O Programa é o instrumento que articula um conjunto de ações (orçamentárias e não orçamentárias) suficientes para enfrentar um problema, devendo seu desempenho ser passível de aferição por indicadores coerentes com o objetivo estabelecido” (BRASIL, 2007, p. 41).

Por esse entendimento, fica evidente que uma das formas de aferição do desempenho dos programas de governo deverá ser realizada por meio do uso de indicadores. Ainda segundo o Manual, um indicador é um “instrumento capaz de medir o desempenho do programa. Deve ser passível de aferição e coerente com o objetivo estabelecido, ser sensível à contribuição das principais ações e apurável em tempo oportuno” (BRASIL, 2007, p. 45).

Nesse contexto, surge uma questão a respeito da qualidade do mecanismo de avaliação proposto para o PPA, ou seja, se essa avaliação realizada por meio de indicadores de desempenho é efetiva, no sentido de produzir informações que possam subsidiar as tomadas de decisão. Noutras palavras, impõe-se avaliar se os indicadores suportam adequadamente as decisões relativas à gestão dos programas de governo e subsidiam adequadamente as soluções propostas nos objetivos traçados para esses programas.

Na tentativa de responder a essa questão, surge a proposta da chamada meta-avaliação. O dicionário eletrônico Aurélio (FERREIRA, 2004) apresenta, entre as definições para o prefixo met(a)-, a de que se trata de uma “reflexão crítica sobre”. Desse modo, pode-se compreender a meta-avaliação como uma reflexão crítica a respeito de uma outra avaliação. Schwandt (apud HEDLER, 2009, p. 471) afirma que “a meta-avaliação é um método para checar a

qualidade da avaliação anterior”. A partir desses conceitos, foi estabelecida a abordagem deste trabalho, com a finalidade de propor uma abordagem para a avaliação de indicadores de desempenho de programas governamentais que, por sua vez, se propõem a avaliar os respectivos programas.

A avaliação da execução dos programas de governo e do cumprimento das metas do PPA é atividade finalística dos órgãos de controle, em especial da Controladoria-Geral da União, conforme o art. 20 da Lei nº 10.180, de 6 de fevereiro de 2001, segundo o qual :

“O Sistema de Controle Interno do Poder Executivo Federal tem as seguintes finalidades:
I - avaliar o cumprimento das metas previstas no plano plurianual, a execução dos programas de governo e dos orçamentos da União;
[...]” (BRASIL, 2001)

Essa avaliação pode ser, em grande parte, realizada por meio do acompanhamento das informações apresentadas pelos indicadores de desempenho, mesmo que em caráter complementar ou subsidiário às atividades atualmente desempenhadas pelos órgãos de controle.

A possibilidade de contribuir para a melhoria da atividade de controle consistiu no principal impulso para a realização deste trabalho, cuja proposta contempla uma possibilidade que, quando aperfeiçoada pela prática e pelo desenvolvimento de trabalhos futuros, permitirá aos órgãos de controle conhecer o nível de qualidade com que os indicadores refletem o desempenho dos programas de governo.

Mas os órgãos de controle não são os únicos atores potencialmente interessados em dispor de indicadores de desempenhos de qualidade. Os indicadores, enquanto mecanismos de aferição e divulgação de determinado desempenho, podem ter utilidade para diversos segmentos da sociedade: as chamadas partes interessadas, ou *stakeholders*¹.

Sem que exista uma parte interessada na obtenção de determinada informação, esta perde sua finalidade. Por outro lado, quanto maior o número de partes interessadas em determinada informação, maior sua importância. Desse modo, este trabalho pretende discutir a avaliação dos indicadores de desempenho sob o ponto de vista dos *stakeholders* ou partes interessadas na obtenção das informações produzidas pelos indicadores de desempenho de programas de governo, pela perspectiva de submetê-los ao crivo de determinados critérios de qualidade, selecionados a partir da literatura revisada.

1.2. Objetivos.

1.2.1. Objetivo Geral.

Propor uma abordagem para a avaliação de indicadores de desempenho de programas governamentais – meta-avaliação.

1.2.2. Objetivos Específicos.

- a) Por meio de revisão bibliográfica, relacionar as principais categorias de indicadores de desempenho que podem ser utilizados na avaliação de programas governamentais;
- b) Propor um mecanismo de avaliação para o nível de importância de cada uma das categorias de indicadores identificados, consideradas as necessidades de informação dos *stakeholders*;
- c) Propor o desenho de um índice de cobertura da necessidade de informação dos *stakeholders*, por meio da confrontação das categorias existentes para a avaliação de um programa com aquelas que são necessárias, conforme o nível de importância detectado a partir de “b”;
- d) Por meio de revisão bibliográfica, relacionar, como critérios de qualidade, os principais atributos essenciais e desejáveis para os indicadores de desempenho de programas governamentais;
- e) Propor mecanismo de avaliação para o índice de qualidade das categorias de indicadores utilizadas para avaliar um programa; e
- f) Propor uma possibilidade de cálculo do Nível de Efetividade Geral de um sistema de indicadores de avaliação de desempenho de programas governamentais.

1.3. Desenvolvimento.

Este artigo apresenta uma proposta de abordagem de meta-avaliação para indicadores de desempenho de programas de governo. No capítulo 2, fundamentação teórica, são apresentadas as definições de indicadores de desempenho, suas principais utilidades e componentes. Em seguida, apresentam-se as principais categorias de indicadores estudadas na bibliografia, as tipologias existentes e uma compilação bibliográfica de alguns dos principais critérios de qualidade que podem ser esperados de um indicador de desempenho.

No capítulo 3 é apresentada a metodologia que vai resultar em uma proposta de elaboração conceitual a partir da literatura estudada.

No capítulo 4 apresenta-se a abordagem de avaliação proposta. Nos itens e subitens que se seguem, após serem tecidas algumas considerações, discute-se a necessidade de informações por parte dos *stakeholders* e, a partir desses elementos, são apresentadas sugestões para as formas de avaliação e as fórmulas de cálculo para subsidiar uma meta-avaliação: o **nível de importância das categorias de indicadores**; o **índice de cobertura de um sistema de indicadores de desempenho**; a **insuficiência** ou **adequação** das categorias de indicadores existentes, em função do atendimento aos critérios **essenciais** de qualidade esperados daquelas categorias; os **índices de qualidade** de cada categoria de indicadores obtidos por meio do confronto dessas categorias com os critérios de qualidade desejáveis para os indicadores; e, finalmente, o **nível de efetividade geral de um sistema de indicadores de desempenho**.

Finalmente, o capítulo 5 conclui o trabalho, sintetizando os méritos da discussão, as necessidades de aprimoramentos a partir da abordagem proposta e as possibilidades de futuros estudos e encaminhamentos.

2. Fundamentação teórica.

2.1. O que são indicadores de desempenho.

Para o Tribunal de Contas da União - TCU (BRASIL, 2009b, p.4):

“Indicador de desempenho é um número, porcentagem ou razão que mede um aspecto do desempenho, com a finalidade de comparar esta medida com metas preestabelecidas. Alguns autores conceituam indicador de desempenho como um instrumento de mensuração quantitativa ou qualitativa de determinados aspectos do desempenho”.

O Guia Referencial para Medição de Desempenho e Manual para Construção de Indicadores da Secretaria de Gestão (SEGES) do Ministério do Planejamento acrescenta que:

“pode-se dizer que os indicadores possuem, minimamente, duas funções básicas: a primeira é descrever por meio da geração de informações o estado real dos acontecimentos e o seu comportamento; a segunda é de caráter valorativo que consiste em analisar as informações presentes com base nas anteriores de forma a realizar proposições valorativas” (BRASIL, 2009b, p. 12).

E complementa com a proposição de que, “de forma geral, os indicadores não são simplesmente números, ou seja, são atribuições de valor a objetivos, acontecimentos ou situações” (BRASIL, 2009, p. 13).

O TCU alerta que “o uso de indicadores de desempenho nas organizações deve ser um meio e não um fim em si” (BRASIL, 2009b, p. 7) e conclui apresentando informações a respeito da dinâmica inerente aos atos e aos fatos:

“a variação está presente em todos os aspectos da vida e nas características de produtos e processos, podendo ser descrita estatisticamente. Daí a origem do enfoque de tomadas de decisão por meio da utilização de indicadores de desempenho. A aplicação começa com o fato, o fato gera dados, os dados são processados e transformados em informações, as informações criam o conhecimento [...]” (BRASIL, 2009b, pp. 7-8).

2.1.1. A utilidade dos indicadores de desempenho.

Entre as várias possibilidades de informações que podem ser produzidas e selecionadas para a avaliação de um programa de governo, encontram-se os mecanismos de avaliação chamados indicadores de desempenho que, segundo Bittencourt (2006, p. 15), são “aproximações à realidade multidimensional da gestão” e fornecem “uma boa visão acerca do resultado que se deseja medir”. Bittencourt também sustenta que esses indicadores devem sempre receber dos seus usuários “uma interpretação compatível com o contexto em que estão inseridos” (BITTENCOURT, 2006, p. 5). Essa questão contextual está diretamente relacionada aos objetivos do programa a que os indicadores se propõem a mensurar.

Segundo o Guia Metodológico de Indicadores de Programas da Secretaria de Planejamento e Investimentos Estratégicos (SPI), “O gestor público deve ter em mente que o que se pretende com a utilização de indicadores de desempenho não é produzir números, mas aferir resultados” (BRASIL, 2010a, p. 50).

Os resultados que devem ser aferidos têm dois aspectos principais: o primeiro é a utilidade da informação, que está relacionada ao resultado que se quer conhecer e àquilo que se pode produzir com esse conhecimento. A segunda é o destinatário da informação, ou a parte interessada em conhecer os resultados do programa, os chamados *stakeholders*.

Quando os *stakeholders* são afetados por um determinado processo ou atividade, podem ter a necessidade ou o interesse de obter acesso a informações sistematizadas, que possam subsidiar suas decisões e seus comportamentos em relação ao referido processo ou atividade.

No Guia Metodológico de Indicadores de Programas, a SPI justifica que:

“é preciso quantificar os parâmetros que permitem aos gestores, políticos e cidadãos conhecer, opinar e decidir acerca das múltiplas e complexas atividades estatais, o que faz dos indicadores instrumentos basilares às melhores práticas de administração pública” (BRASIL, 2010a, p. 128).

E acrescenta que:

“os indicadores não servem apenas para o seu processo de planejamento, execução, monitoramento e controle gerencial, mas também outros *stakeholders* como os

demais órgãos da Administração Pública Federal (setoriais, centrais e de controle), os envolvidos em temas transversais, os administradores da agenda prioritária de governo, outros Poderes, as instituições privadas e do terceiro setor etc” (BRASIL, 2010a, p. 55).

Pela leitura do Manual de Elaboração do PPA 2008-2001 podem ser identificados outros *stakeholders*, inclusive os citados, contextualizados pelo aspecto da estrutura de um programa de governo. O primeiro stakeholder identificado é a Unidade Responsável que corresponde à “unidade administrativa responsável pelo gerenciamento do programa” (BRASIL, 2007, p. 41). Uma segunda parte potencialmente interessada que se identifica é o próprio público-alvo, ou seja:

“o(s) segmento(s) da sociedade ao(s) qual (is) o programa se destina e que se beneficia(m) direta e legitimamente com sua execução. São os grupos de pessoas, comunidades, instituições ou setores que serão atingidos diretamente pelos resultados do programa.” (BRASIL, 2007, p. 42).

A partir do mesmo documento, podem ser distinguidos o público-alvo dos beneficiários, que representam a “parcela do público-alvo atendida” (BRASIL, 2007, p. 113).

Segundo o Manual de Indicadores de desempenho para o Setor Público da Comissão Econômica para a América Latina e Caribe – CEPAL:

“Neste contexto também se assume que os próprios cidadãos interessados podem avaliar a qualidade, quantidade e oportunidade dos bens e serviços recebidos, de modo que, em alguns programas, incorpora-se a avaliação dos resultados por parte dos beneficiários ou das partes interessadas (stakeholders)” (BONNEFOY, 2005 p.10, tradução nossa)

Sobre o termo *stakeholder*, no contexto apresentado, o Manual da CEPAL traz nota esclarecendo que “abrange um grande grupo de pessoas físicas e jurídicas que são afetadas pelas decisões e ações públicas.” (BONNEFOY, 2005, p.10, tradução nossa) e acrescenta que:

“A demanda pela avaliação precisa ser gerada, especificada e articulada por atores internos e externos. É pouco provável que uma avaliação sem um ‘proprietário’, ou *stakeholder*, surta algum efeito. As barreiras institucionais à avaliação, como a resistência interna, podem ser reduzidas por meio da participação e da consulta aberta, criando confiança entre as partes.” (BONNEFOY, 2005, p.20, tradução nossa)

São exemplos, portanto, de *Stakeholders* em relação aos indicadores de programas de governo:

O gestor do programa: é um interessado intrínseco ao programa, pois tem a responsabilidade de prestar contas sobre os recursos aplicados e tem a atribuição de redirecionar o programa em função de seus resultados;

O público-alvo do programa: é interessado direto, tanto no resultado quanto na operacionalização do programa. Tem sua vida e suas necessidades afetadas pelo programa e depende dos seus benefícios;

O contribuinte: é o principal financiador do programa. Seu interesse está relacionado a conhecer os resultados da aplicação dos recursos que são gerados pelo pagamento dos tributos;

Os órgãos de controle: têm interesse e obrigação legal de acompanhar a boa e regular aplicação dos recursos públicos, a efetividade da execução dos programas, a adequação dos objetos, os custos de execução, a legalidade dos atos e propor melhorias para a gestão; e

A comunidade internacional: o interesse dos organismos internacionais está diretamente ligado à produção de estatísticas, rankings, avaliações comparativas entre os países. A obtenção dessas informações pelos diversos países pode ser revertida em aprendizado para possível implantação de programas similares (benchmarking). Para os investidores e bancos internacionais, essas informações podem gerar interesses para direcionar ou ampliar investimentos no país ou servir como referências de consolidação econômica, política, tecnológica, social, entre outras.

Outros possíveis *stakeholders* que podem ser citados são o governo central, a imprensa, os eleitores e alguns grupos de interesse específicos, conforme características do programa cujos indicadores sejam objeto de análise.

2.1.2. Os componentes de um indicador de desempenho.

Um indicador de desempenho é constituído por alguns componentes básicos. O primeiro é a **fórmula**, “padrão matemático que expressa a forma de realização do cálculo” (BRASIL, 2009a, p. 46), que necessita de tantos **dados** quantos sejam as **variáveis** que a compõem. Para o IBGE (apud BRASIL, 2010a, p. 21) “os indicadores são ferramentas constituídas de **variáveis** que, associadas a partir de diferentes configurações, expressam significados mais amplos sobre os fenômenos a que se referem”. O TCU (BRASIL, 2009b, p.14) cita, entre as **variáveis** que podem compor um indicador, o custo, a quantidade, a qualidade e o tempo. O referido Tribunal também registra a ideia de que “indicadores devem ser operacionalizados por informações fortemente baseadas em **dados** e fatos.” (BRASIL, 2009b, p.12). A SEGES acrescenta que “a geração de indicadores (medidas) relevantes requererá dados e informações específicos – cuja obtenção e tratamento demandarão mais ou menos tempo e representarão custos” (BRASIL, 2009a, p. 11) e também que “é necessário definir a origem de onde os dados serão extraídos, ou seja, a fonte de dados [...]”.

Além das **variáveis**, a fórmula pode conter operadores aritméticos e, em alguns casos, valores numéricos previamente estabelecidos.

Os **dados** podem ser produzidos, coletados, pesquisados e, quando estiverem todos disponíveis, alimentarão a **fórmula**. Esta, uma vez aplicada, produz um **resultado**, que será comparado com **referências** ou **metas**, indicando a situação em que se encontra o programa.

As **referências** são padrões de comparação estabelecidos, seja com base em programas similares, em desempenhos anteriores ou em estudos de mercado ou pesquisas. O padrão de comparação, segundo a SEGES, é um “índice arbitrário e aceitável para uma avaliação comparativa de padrão de cumprimento” (BRASIL, 2009, p. 46)

Já as **metas**, são os valores que o resultado do indicador procura igualar ou superar², conforme o objetivo do programa para um determinado período.

Relacionada ao resultado, deve existir uma orientação de **interpretação** que esclareça, por exemplo, qual o efeito para o programa, a depender de quanto maior ou menor seja o valor resultante do indicador apurado. Segundo a SEGES, a “interpretação do indicador informa se o bom desempenho é alcançado quando o resultado do indicador está aumentando, diminuindo ou permanecendo o mesmo” (BRASIL, 2009, p. 52) Como exemplo, no caso de um indicador de mortalidade infantil, a meta indica a busca da redução ao menor patamar possível. Assim, se o referido indicador apresentar um valor “acima” da meta, resultará, na verdade, em situação insatisfatória para o gestor governamental.

Outro componente importante é a **periodicidade** de produção dos dados que alimentam o indicador. Os dados devem estar disponíveis no momento da apuração do indicador, de forma que seja garantida a oportunidade da intervenção do gestor público para corrigir ou alterar uma situação. A SPI lembra que “uma proposta de elaboração de indicadores deverá permitir dispor de indicadores [...] com uma periodicidade que equilibre as necessidades de informação com os recursos técnicos e financeiros” (BRASIL, 2010a, p. 27). O Manual de Avaliação do PPA 2010 menciona que essa periodicidade é normalmente anual (BRASIL, 2010b, p. 16), mas o Manual de Elaboração (BRASIL, 2007, p. 46) e a SEGES (BRASIL, 2009a, p. 58) citam outros exemplos, como mensal, bienal, semanal e semestral. Essa informação é complementada pela SEGES ao acrescentar que “a periodicidade de monitoramento e revisão de cada indicador é definida conforme a natureza específica de cada um” (BRASIL, 2009c, p. 15). Esta última informação remete à noção de que diferentes programas podem ter ciclos de gestão ou de resultados mais curtos ou mais longos.

Finalmente, é preciso delimitar a **base geográfica** abrangida pelo indicador, que é o “menor nível de agregação geográfica de apuração do índice” (BRASIL, 2010a, p. 118) e a **unidade de medida** na qual o resultado será expresso, que consiste em um “padrão escolhido para

mensuração da relação adotada como indicador” (BRASIL, 2010a, p. 117). De acordo com a SEGES, uma unidade de medida “confere um significado ao resultado” (BRASIL, 2009a, p.50). A Secretaria de Gestão do Ministério do Planejamento, Orçamento e Gestão prossegue com exemplos de unidades de medida que podem ser expressas em um valor numérico, no caso dos indicadores simples, como o número de alunos matriculados no ensino médio. Já os indicadores compostos expressam a relação entre duas ou mais variáveis e as unidades de medida podem ser expressas na forma de uma proporção ou coeficiente, como o número de óbitos em relação à população total. Estes indicadores podem ainda ser expressos na forma de porcentagens, que são obtidas multiplicando-se o quociente de uma proporção por 100, ou na forma de uma razão ou índice, que se assemelham à proporção, dividindo-se um número A por um número B, com a diferença de que o termo razão é usado normalmente quando A e B representam categorias separadas e distintas. Quando este quociente é chamado de índice, significa que se trata de uma razão entre duas grandezas tais que uma não inclui a outra. São exemplos a densidade demográfica e a renda per capita. Finalmente, a taxa consiste em coeficientes multiplicados por uma potência de 10 e seus múltiplos, como a taxa de natalidade, que é calculada pela multiplicação do coeficiente de natalidade por 1.000 (BRASIL, 2009a, pp.50-52).

Cabe acrescentar a importância de que cada indicador seja associado a um **responsável** (BRASIL, 2009a, p. 58) por sua elaboração, controle e divulgação dos resultados obtidos, para facilitar os processos de comunicação, alteração e evolução do indicador. De modo geral, esse responsável é o gestor do programa, a unidade responsável pela sua execução ou, em última instância, o órgão responsável pelo programa. Também uma **descrição** pode facilitar a compreensão do indicador e de sua finalidade.

2.1.3. Indicadores simples e complexos.

Alguns indicadores de desempenho podem ter uma construção mais complexa, conforme duas situações específicas principais. A primeira, diz respeito a metas ou objetivos múltiplos. Numa situação como esta, caso não seja possível construir uma fórmula que reflita o atingimento de todos os objetivos e metas, pode-se desenvolver um indicador para cada objetivo ou grupos de objetivos. Uma vez obtidos os resultados correspondentes a cada um desses indicadores, os quais podem ser denominados **indicadores simples**, que “expressam valores absolutos” (BRASIL, 2009b, p. 6), será possível construir outro indicador, agora chamado de **indicador complexo**, cuja finalidade pode ser, entre outras, somar ou ponderar

os resultados obtidos nos indicadores anteriores. Os indicadores complexos, também chamados de agregados, são “o resultado da composição de diversos indicadores, cada um com o seu grau de importância ou de representatividade” (BRASIL, 2009b, p. 11).

Uma segunda situação diz respeito à dependência da análise dos resultados de vários indicadores para a verificação da realidade que se pretendeu modificar. Isso ocorre, por exemplo, quando se pretende avaliar uma política pública, composta por vários programas de governo, ou mesmo para avaliar um único programa, cuja execução dependa de várias ações que não possam ser verificadas em conjunto. Assim, os indicadores obtidos naqueles programas ou ações que vão compor a política pública ou o programa de governo, mesmo que sejam indicadores complexos ou agregados sob um ponto de vista inicial, podem agora, sob a perspectiva de uma análise macro, passar à condição de **indicador simples**, para a composição de um novo **indicador complexo**, que vai contemplar as diversas verificações anteriores, de forma mais ampla e abrangente.

Essa situação pode ser observada no cálculo do Índice de Desenvolvimento Humano (IDH), que é “uma medida composta com base nas metas das três dimensões básicas do desenvolvimento humano - uma vida longa e saudável, acesso à educação e um padrão de vida digno” (PNUD 2010a, p. 28). De acordo com a Nota Técnica de Apoio ao Lançamento do Relatório de Desenvolvimento Humano 2010 (PNUD, 2010b, p. 6), o IDH é composto por três índices: de esperança de vida; de rendimento; e de educação, “e estes indicadores são agregados segundo médias geométricas” (PNUD 2010a, p. 28). O indicador de educação, por sua vez, além de funcionar com indicador intermediário ou simples para o IDH, é um indicador complexo ou agregado resultante de outros dois índices: o de escolaridade esperado e o de média de anos de escolaridade (PNUD, 2010b, p. 6).

2.1.4. As Categorias de indicadores de desempenho.

Em relação aos aspectos medidos pelos indicadores, o TCU (BRASIL, 2000, p.14) entende que “o desempenho na obtenção de um determinado resultado pode ser medido segundo as seguintes dimensões de análise: **economicidade, eficiência, eficácia e efetividade.**”

O Guia Metodológico de Indicadores de Programas da Secretaria de Planejamento e Investimentos Estratégicos – SPI (BRASIL, 2010a, p. 128) apresenta, em relação à avaliação dos recursos alocados e dos resultados alcançados, essa mesma classificação para indicadores de avaliação de desempenho. O Guia explica a primeira dessas categorias, a dos indicadores de **economicidade**, como sendo aqueles que “expressam os gastos incorridos na alocação dos

insumos”. Já os indicadores de **eficiência** são aqueles que se propõem a verificar uma relação de produtividade, traçada entre os fins e os meios empregados, sejam estes expressos em termos de insumos, tempo, mão de obra ou outros recursos que se façam necessários. Para o Manual de Avaliação do PPA, a eficiência é “a medida da relação entre os recursos efetivamente utilizados para a realização de uma meta para um projeto, atividade ou programa frente a padrões estabelecidos” (BRASIL, 2010b, p. 14).

Para o Manual de Eaboração do PPA, os indicadores de **eficácia** são aqueles que demonstram “a medida do grau de atingimento das metas fixadas [...]” (BRASIL, 2007, p. 45). Finalmente, o Guia Metodológico informa que os indicadores de **efetividade** são, em geral, “multiprogramáticos e medem os efeitos positivos ou negativos na realidade social” (BRASIL, 2010a, p. 128). Este aspecto está diretamente relacionado aos objetivos de um programa e constitui-se em uma importante categoria de indicadores, uma vez que, além de poderem abranger mais de um programa, os indicadores de efetividade assumem características de análise de caráter múltiplo e podem vir a depender também de indicadores intermediários para sua construção. O Manual de Elaboração do PPA 2008-2011 propõe que a efetividade representa a própria medida de resultado de um programa, relacionada ao alcance do objetivo proposto, o que implica na “transformação ou mudança em uma realidade concreta a qual o programa se propôs modificar” (BRASIL, 2007, p. 37).

Dessas quatro categorias de indicadores de desempenho, três: eficiência, eficácia e efetividade, estão amplamente contempladas na literatura relacionada ao assunto. Além dessas, outras categorias também são contempladas na literatura. Bittencourt (2006, p. 19) acrescenta a dimensão da **equidade**, que seria, segundo ele, uma “medida da possibilidade de acesso aos benefícios de uma determinada ação por parte dos grupos sociais menos favorecidos em comparação com as mesmas possibilidades da média do país”.

Costa e Castanhar relacionam ainda outras categorias de indicadores, às quais denominam critérios. A primeira é a **sustentabilidade**, que “mede a capacidade de continuidade dos efeitos benéficos alcançados através do programa social, após o seu término” (MANUAL DO UNICEF apud COSTA e CASTANHAR, 2005, p.973). Em seguida, a **análise custo-efetividade**, que se relaciona à ideia de custo de oportunidade e cuja finalidade é informar a respeito de alternativas de ação para a obtenção de determinados resultados. Seguem dois exemplos possíveis para aplicação dessa categoria de indicador:

- a) o programa 1173 – Controle Interno, Prevenção e Combate à Corrupção, reduz a corrupção em X% em determinado período, enquanto que o programa 1164 – Prevenção e Combate à Lavagem de Dinheiro - reduz em Y%; ou
- b) para reduzir a corrupção em 1% no país, o programa 1173 gasta R\$ X,XX, enquanto que o programa 1164 gasta R\$ Y,YY para o mesmo percentual e no mesmo período.

Percebe-se que o primeiro exemplo apresentado possui características de eficácia e, o segundo, de eficiência. Por questões práticas que envolvem o termo efetividade, neste trabalho esta categoria será chamada de **custo de oportunidade**. Nos dois exemplos apresentados, nota-se que ocorreram comparações entre programas distintos.

Uma outra categoria é a **satisfação do beneficiário**, que “avalia a satisfação do usuário em relação à qualidade do atendimento que está obtendo do programa” (MANUAL DO UNICEF apud COSTA e CASTANHAR, 2005, p.973). Essa categoria de indicador busca obter a percepção, a impressão, ou a atitude do beneficiário em relação aos benefícios que está recebendo de um programa ou em relação ao produto obtido ou, ainda, em relação à qualidade do atendimento. Essa variação de possíveis percepções decorre do tipo de programa que se avalia e das características da informação que se pretende obter com o indicador utilizado. Uma dimensão que pode provocar fortes variações nas características das informações que são necessárias para esta categoria de indicador é a percepção dos *stakeholders*. Esta categoria, para uma maior abrangência e em consonância com a definição do Manual de Elaboração do PPA (BRASIL, 2007, p. 42), poderia também ser chamada de satisfação do público-alvo, já que este termo permite englobar as pessoas que deveriam ter sido atingidas pelo programa e não foram.

O Guia Referencial para Medição de Desempenho e Manual para Construção de Indicadores da SEGES apresenta seis dimensões para análise do desempenho, os 6Es. Esses seis tipos de indicadores são divididos em dois grupos: de resultado: **eficiência, eficácia e efetividade**; e de esforço: **economicidade, excelência e execução**. Este último indicador, segundo o Guia, “refere-se à realização dos processos, projetos e planos de ação conforme estabelecidos” (BRASIL, 2009a, p. 20). Pode ser utilizado como indicador de monitoramento. Seu conceito é associado ao acompanhamento das execuções física e financeira de determinada ação. O mesmo pode ocorrer com indicadores na categoria de **excelência** que, segundo o referido Guia, “é a conformidade a critérios e padrões de qualidade/excelência para a realização dos processos, atividades e projetos na busca da melhor execução e economicidade; sendo um elemento transversal” (BRASIL, 2009a, p. 20).

Pelo exposto, verifica-se que na literatura os tipos de indicadores podem ser chamados de dimensões, aspectos, medidas ou critérios. Por questões de conveniência e de alinhamento aos objetivos deste trabalho, que compreendem a tarefa de relacionar as principais categorias de indicadores, ou seja, de especificá-los conforme as diferentes avaliações a que se propõem, optou-se por denominá-los, simplesmente, de categorias.

Podem existir situações nas quais seja possível avaliar o desempenho de um programa de governo por meio de uma única categoria de indicadores. Mas o usual é que se tenha um conjunto de categorias de indicadores, desde que se deseje promover uma avaliação completa de um programa. A esse conjunto de categorias será atribuído o nome **sistema de indicadores**. Para Jannuzzi (2005, p. 154) “a lógica de acompanhamento de programas requer a estruturação de um sistema de indicadores”. Para a SPI (Brasil, 2010a, p. 22) um sistema de indicadores “constitui um conjunto de indicadores que se referem a um determinado tema ou finalidade programática”. Assim, um sistema de indicadores pode ser compreendido como um conjunto de categorias de indicadores de desempenho empregado para realizar a avaliação de um programa.

2.1.5. As tipologias de indicadores de desempenho.

Várias são as classificações possíveis para os indicadores de desempenho, tratadas de formas distintas pelos autores citados neste artigo. Classificam-se os indicadores quanto à sua natureza, objetividade, objetivo, complexidade, fluxo de implantação, modalidade, entre outras.

Essas tipologias, entretanto, são mais úteis ao processo de desenvolvimento do que ao esforço de avaliação dos indicadores de desempenho, uma vez que se constituem em características que, estando compreendidas, podem orientar o desenvolvimento de indicadores tendo em vista o objetivo de cada programa.

A título de exemplo, uma das tipologias propostas por Jannuzzi é “a diferenciação dos indicadores entre indicador-insumo, indicador-processo, indicador-resultado e indicador-impacto” (2005, p. 144). E explica que:

“Os indicadores-processo ou fluxo são indicadores intermediários, que traduzem, em medidas quantitativas, o esforço operacional de alocação de recursos humanos, físicos ou financeiros (indicadores-insumo) para a obtenção de melhorias efetivas de bem-estar (indicadores- resultado e indicadores-impacto)” (2005, p. 144).

Neste trabalho, por outro lado, optou-se por estudar os critérios de qualidade dos indicadores em cada categoria. Assim, as categorias de indicadores serão confrontadas com requisitos

essenciais e desejáveis de qualidade, a fim de avaliar se cumprem satisfatoriamente sua finalidade de avaliar um programa de governo.

Para as finalidades deste trabalho, preferiu-se não exaurir a análise sobre todas as classificações encontradas na literatura, em que pese a sua importância, a fim de evitar a sobreposição de conceitos e informações, e tendo em vista o objetivo de propor um caminho de avaliação de indicadores, muito mais do que efetuar ampla e completa revisão da literatura que trata das diversas possibilidades de classificação.

Ainda assim, quando necessário, algumas tipologias são mencionadas, como ocorre na situação em que é exposta a classificação feita pela SEGES, que divide as categorias de indicadores em dois grupos: de esforço e de impacto.

2.1.6. Critérios de qualidade essenciais e desejáveis de um indicador de desempenho.

Segundo o TCU (BRASIL, 2009b, p. 12) “os indicadores devem apresentar alguns critérios ou requisitos de qualidade”. Em seguida (BRASIL, 2009b, pp. 12-13), o TCU apresenta aqueles que considera serem os principais atributos para um indicador.

Além do TCU, outros órgãos e autores citados neste artigo propõem critérios, que variam na sua definição e nomenclatura, ocorrendo casos nos quais um critério pode ser fruto de uma combinação ou de uma dissociação de outros, ou ainda pode ser derivado de um agrupamento ou uma divisão de outros critérios. Enfim, a literatura existente não sistematizou um padrão para os critérios de qualidade que devem compor um indicador.

Um ensaio para a compilação desses critérios será apresentado no próximo tópico.

Os critérios de qualidade podem ser divididos em dois grupos: essenciais e desejáveis. Segundo a SPI, as propriedades essenciais “são aquelas que qualquer indicador de Programa deve apresentar” (BRASIL, 2010a, p.25). A condição de essencialidade de um critério é atribuída conforme se identifica que o seu não atendimento implicaria em distorções ou deficiências no indicador que poderiam acarretar erro de avaliação do programa.

Já os desejáveis, aos quais a SPI (BRASIL, 2010a, p. 26) chama de complementares, podem agregar valor à qualidade da informação produzida.

Em síntese, pode-se entender como essencial um critério de qualidade cuja ausência venha a comprometer a qualidade da informação gerada pelo indicador ou o funcionamento do indicador. Os critérios serão considerados desejáveis quando sua ausência diminuir a qualidade esperada do indicador ou da informação produzida, sem, no entanto, comprometê-lo intrinsecamente.

Em consonância com a abordagem de avaliação que será proposta neste trabalho, é necessário que um avaliador possa atribuir a condição de essencial ou desejável a cada um dos critérios que pretende avaliar, conforme as peculiaridades das categorias de indicadores que está avaliando e sempre considerando as características e as condições do programa que é objeto de avaliação.

A seguir, apresentam-se os principais critérios de qualidade encontrados na literatura.

2.1.6.1. Critérios de qualidade dos indicadores de desempenho.

Praticidade: O indicador deve subsidiar a tomada de decisão. A SEGES chama os indicadores que detêm este atributo de investigativos: “os dados devem ser fáceis de analisar, sejam estes para registro ou para reter informações e permitir juízos de valor” (BRASIL, 2009a, p. 47).

No mesmo sentido, segundo o TCU, “o indicador deve realmente funcionar na prática e permitir a tomada de decisões gerenciais” (BRASIL, 2000, p. 12).

Definição: o mesmo que Bittencourt chama de “critérios definidos”, e para quem “o indicador deve ter a ele associada uma clara definição dos critérios aplicáveis para determinação de seu valor” (2006, p. 17).

Factibilidade: para a SPI:

“os dados necessários para as medições se constituem em informações que fazem parte dos processos de gestão da instituição e, portanto, obtidas através de instrumentos de coleta, seja por amostra ou censo, estatísticas, aplicação de questionários, observação etc., dependendo do aspecto a ser medido. Uma proposta de elaboração de indicadores deverá permitir dispor de indicadores de medição factível, em momentos adequados e com uma periodicidade que equilibre as necessidades de informação com os recursos técnicos e financeiros” (BRASIL, 2010a, pp. 26-27).

Este é um critério múltiplo, que abarca a factibilidade e a periodicidade. A periodicidade é aspecto tratado em critério distinto. Assim, trata-se aqui exclusivamente da factibilidade do indicador. Como as questões referentes ao tempo e aos custos também são tratados em outros critérios, este pode ser restringido à factibilidade técnica.

Para se afirmar que existe factibilidade técnica em um indicador é necessário que se esteja seguro de que os métodos, o conhecimento técnico e os recursos tecnológicos existentes na unidade responsável pela operacionalização de um indicador são plenamente compatíveis com os métodos, o conhecimento técnico e os recursos tecnológicos necessários à operacionalização desse indicador. Isso significa que deve haver capacidade técnica disponível na unidade que gerencia o indicador, ou acessível a esta unidade, compatível com

as necessidades técnicas para o desempenho das atividades de: coleta de dados, operacionalização do indicador, manutenção do indicador (atualização; ajuste de fórmula, de amostra, de região sócio demográfica; etc.), análise das informações obtidas e utilização e divulgação dessas informações.

Acessibilidade: “deve haver possibilidade e facilidade de acesso às informações primárias, bem como de registro e manutenção para o cálculo dos indicadores” (BITTENCOURT 2006, p. 17).

Há comprometimento do indicador na hipótese de dificuldade de acesso aos dados primários.

Auditabilidade: para a SPI, “qualquer pessoa deve sentir-se apta a verificar a boa aplicação das regras de uso dos indicadores (obtenção, tratamento, formatação, difusão, interpretação)” (BRASIL, 2010a, p. 26).

Confiabilidade: Bittencourt entende que:

“A fonte de dados utilizada para o cálculo do indicador deve ser confiável, de tal forma que diferentes avaliadores possam chegar aos mesmos resultados empíricos. (É possível, naturalmente, que existam diferentes interpretações para um ID, mas não se deve escolher um ID cujos valores medidos diferem significativamente entre avaliadores diferentes.)” (BITTENCOURT 2006, p. 16).

Cobertura: Para Jannuzzi (2000, p. 13), “os indicadores devem representar adequadamente a amplitude e a diversidade de características do fenômeno monitorado, resguardado o princípio da seletividade e da simplicidade”. O autor acrescenta que (2005, p. 140) “deve-se procurar empregar indicadores de boa cobertura territorial ou populacional, que sejam representativos da realidade empírica em análise”.

É preciso compreender que a cobertura máxima consiste, muitas vezes, em trabalho censitário, o que não é possível na maioria dos casos.

Seletividade: por este critério: “deve-se estabelecer um número equilibrado de indicadores que enfoquem os aspectos essenciais do que se quer monitorar” (BRASIL, 2000, p. 13). Assim, é importante verificar se o indicador é indispensável para a avaliação do programa.

Em complemento, para aquilo que a SEGES chama também de importância, o indicador deve fornecer “informações sobre as principais variáveis estratégicas e prioridades definidas de ações, produtos ou impactos esperados” (BRASIL, 2009a, p. 46).

A ausência deste critério diminui a qualidade do sistema de indicadores, pois este será composto por uma quantidade de indicadores acima da necessária.

Simplicidade: segundo o TCU, por este critério, “o indicador deve ser de fácil entendimento, qualquer pessoa deve ser capaz de tirar conclusões a partir da análise do indicador. Os nomes

e expressões devem ser conhecidos e entendidos por todos os envolvidos de forma homogênea, garantindo transparência e validade” (BRASIL, 2009b, p. 13).

A SPI complementa que “devem ser de fácil obtenção, construção, manutenção, comunicação e entendimento pelo público em geral, interno ou externo” (BRASIL, 2010a, p. 25).

Este critério está relacionado à compreensão das informações prestadas pelo indicador.

Periodicidade: diferente da periodização que, como se verá adiante, pressupõe a associação de informações a um período específico, este critério tem relação estreita com o conceito de tempestividade. Segundo Jannuzzi, “para que se possa acompanhar a mudança social, avaliar o efeito de programas sociais implementados, corrigir eventuais distorções de implementação, é necessário que se disponha de indicadores levantados com certa regularidade” (2005, pp. 141-142). Esta regularidade deve estar diretamente relacionada a períodos definidos na própria concepção do indicador, uma vez que cada tipo de informação a ser produzida demanda uma periodicidade distinta.

O TCU também chama este critério de atualização periódica, no sentido de que “o indicador deve permitir atualização de forma a representar a situação mais atual possível” (BRASIL, 2009b, p. 13).

Se o indicador não define a frequência com que as informações precisam estar disponíveis, ou se a define, mas as informações não são produzidas com a tempestividade necessária, ocorre um impacto negativo na possível correção de desvios na execução do programa, pelos gestores. Em alguns casos, a informação que não for atualizada periodicamente perde a sua utilidade.

Periodização: diferente do critério da periodicidade, para Bittencourt:

“o indicador deve ser capaz de correlacionar-se, sem ambiguidades, com os resultados de um período de tempo predefinido; caso contrário, não conterá informação capaz de associar um determinado ato ou política (que são executados sempre ao longo de um período de tempo também determinado) aos resultados observados” (2006, p. 17).

Este critério diz respeito a uma delimitação temporal do período a que se refere, valorizando a questão do período assim como o critério da cobertura valoriza a questão geográfica.

Validade: indicadores devem ter “capacidade de representar, com a maior proximidade possível, a realidade que se deseja medir e modificar” (BRASIL, 2010a, p. 25); devem ser “pertinentes e adequados” (BRASIL, 2007, p. 120) e devem “refletir o fenômeno a ser monitorado” (BRASIL, 2000, p. 13).

O TCU ainda dá a esta qualidade fundamental o nome de relação causa-e-efeito, explicando que:

“deve-se verificar se as relações de causa-e-efeito traçadas para os objetivos estratégicos também são válidas para os indicadores de desempenho (nova validação da hipótese estratégica). Não sendo válidas, as relações de causa-e-efeito, os objetivos estratégicos e/ou os indicadores de desempenho deverão ser revistos” (BRASIL, 2009b, p. 8).

Não sendo válidas as relações de causa e efeito entre os objetivos previstos no programa e as informações medidas pelo indicador, as informações obtidas não podem ser utilizadas para a tomada de decisão.

Para compreender melhor essa relação de causa e efeito, é possível admitir que, para cada verbo existente em um objetivo geral proposto, corresponda um objetivo específico a ser alcançado, ou seja, para cada ação (verbo) proposta, esperar-se-á um resultado. Também, para tantos quantos forem os objetos associados a cada um dos verbos existentes no objetivo, corresponderá uma meta a ser atingida, ainda que tal meta não tenha sido estabelecida e respeitadas as condições impostas, quando estas estiverem presentes nos objetivos. Por este raciocínio, o objetivo estabelecido para o programa 1173, por exemplo, pode ser desdobrado como no quadro que segue:

Quadro 1- possíveis relações de causa-e-efeito para o programa 1173.

AÇÃO	OBJETO	CONDIÇÃO
fortalecer as atividades de	investigação	das irregularidades no Poder Executivo
	apuração	
	repressão	
prevenir	a corrupção	
combater	a impunidade	
ampliar	a transparência	da gestão pública
desenvolver	atividades do sistema de controle interno do Poder Executivo Federal	em cumprimento ao disposto nos artigos 70 e 74 da Constituição Federal de 1988

Neste exemplo, cada uma das seis primeiras linhas da coluna central, correspondente a um objeto, deverá estar relacionada a uma meta específica.

Na última linha, as ações apresentadas ainda devem ser desdobradas, conforme cada uma das atividades (objetos) correspondentes ao Sistema de Controle Interno do Poder Executivo Federal (SCI) e os artigos 70 e 74 da CF/88, cujos desdobramentos implicarão em um conjunto ainda maior de ações, objetos e condições, para os quais podem ser verificadas as relações de causa-e-efeito.

Cabe lembrar que, para cada uma dessas relações, pode existir um indicador específico. Então, deve ser sempre verificado o propósito do indicador, tendo em vista o objeto específico de medição a que este se propõe.

Independência: De acordo com o TCU, “o indicador deve medir os resultados atribuíveis às ações que se quer monitorar, devendo ser evitados indicadores que possam ser influenciados por fatores externos” (BRASIL, 2000, p. 13). Isso significa dizer que abrangências geográficas, temporais ou de outras naturezas, que podem eventualmente serem medidas além do necessário ou além dos limites de abrangência do programa, prejudicarão a qualidade da informação prestada pelo indicador. Este critério é complementar ao da cobertura. Naquele critério, a preocupação se dá com a suficiência dos dados; neste, com o excesso.

Desagregabilidade: complementar à cobertura, trata-se de uma evolução daquele critério. É a “capacidade de representação regionalizada de grupos sociodemográficos, considerando que a dimensão territorial se apresenta como um componente essencial na implementação de políticas públicas” (BRASIL, 2010a, p. 26).

É um critério que valoriza a informação obtida com o indicador, permitindo que seja desagregada em componentes territoriais menores, o que possibilita acesso a informações parciais que, por sua vez, proporcionam a identificação daqueles grupos sociodemográficos que são mais ou menos afetados pelo programa.

Estabilidade: para o TCU, “O indicador deve permanecer estável ao longo de um determinado período, permitindo a formação de uma série histórica” (BRASIL, 2009b, p. 13).

Homogeneidade: Para Bittencourt:

“Na construção de indicadores e na sua interpretação devem ser consideradas apenas variáveis homogêneas. Assim, não se pode comparar historicamente custos unitários quando as composições respectivas foram alteradas ao longo do tempo; não se pode também comparar dados de natureza ou composição heterogêneas. A homogeneidade é requisito essencial da comparabilidade, que, como vimos, é indissociável da utilização dos ID” (2006, p. 16).

Quando Bittencourt afirma que a homogeneidade é requisito essencial da comparabilidade, isso não quer dizer que o critério seja essencial para a existência do indicador, mas especificamente para possibilitar a comparabilidade.

Comparabilidade: no entendimento da SEGES, “os indicadores devem ser facilmente comparáveis com as referências internas ou externas, bem como séries históricas de acontecimentos” (BRASIL, 2009a, p. 47).

É desejável que os indicadores sejam comparáveis com referências (indicadores similares; programas similares em outros países). No caso das séries históricas, existe critério específico, o da estabilidade. Deve-se ter ciência de que nem sempre isso é possível, como nos casos em que o programa ou o indicador utilizado são inéditos.

Rastreabilidade: no entendimento do TCU, dever haver “facilidade de identificação da origem dos dados, seu registro e manutenção”.

Transparência: segundo Jannuzzi:

“A boa prática da pesquisa social recomenda que os procedimentos de construção dos indicadores sejam claros e transparentes, que as decisões metodológicas sejam justificadas, que as escolhas subjetivas – invariavelmente freqüentes – sejam explicitadas de forma objetiva.

Transparência metodológica é certamente um atributo fundamental para que o indicador goze de legitimidade nos meios técnicos e científicos, ingrediente indispensável para sua legitimidade política e social” (JANNUZZI, 2005, p. 141).

Denominação clara: segundo o Manual de Elaboração do PPA 2008-2011, a denominação do indicador deve ser clara, precisa e autoexplicativa (BRASIL, 2007, p. 120).

Economicidade: este foi o critério mais listado na literatura, presente em toda a bibliografia consultada a respeito de critérios de qualidade. Foram verificadas algumas variações de nomenclatura, como custo-efetividade, econômicos, e de custo razoável, atributo dado por Bittencourt, de quem foi emprestada a definição de que “as informações necessárias ao cálculo do indicador devem ser coletadas e atualizadas a um custo razoável; em outras palavras, a manutenção da base de dados não pode ser dispendiosa frente aos benefícios esperados da utilização do indicador” (2006, p. 17).

Uma questão pendente, e que, pela sua abrangência, não será tratada neste artigo, corresponde à definição do que são custos razoáveis, critério ainda discricionário e subjetivo.

Em função disso, entende-se que o critério, quando aplicado, deve verificar também a existência de um estudo sobre custos relacionados à apuração e manutenção dos indicadores.

Existem casos, como em situações de risco para a população, nos quais a obtenção de informações a respeito da evolução do problema pode justificar a elevação dos custos de obtenção da informação.

Devido às variações nas definições, optou-se por adicionar outro critério, denominado operacionalidade, comentado a seguir.

Operacionalidade: neste caso, será utilizada parte da definição do TCU para economicidade, segundo a qual “não deve ser gasto tempo demais procurando dados, muito menos pesquisando ou aguardando novos métodos de coleta. Os benefícios decorrentes do indicador devem ser maiores que os custos incorridos na medição” (BRASIL, 2009b, p. 13).

Assim, diferentemente da economicidade, que se preocupa com os custos financeiros, este critério se relaciona aos demais insumos (tempo; pessoal; equipamentos) utilizados na operacionalização do indicador.

Publicidade: condizente com o princípio da transparência, a SPI entende que “os indicadores devem ser públicos, isto é, conhecidos e acessíveis a todos os níveis da instituição, bem como à sociedade e aos demais entes da administração pública” (BRASIL, 2010a, p. 26).

Este critério difere do critério denominado transparência, que diz respeito a justificativas pertinentes para as decisões metodológicas e explicitação de forma objetiva das escolhas subjetivas. Aquele critério trata da clareza de métodos, critérios e escolhas que envolvem a produção da informação: coleta, tratamento, armazenamento, processamento e utilização de dados e informações. Este, por sua vez, diz respeito à divulgação, disseminação e publicação dos resultados obtidos por meio do emprego de um indicador.

Para Jannuzzi, que dá o nome de comunicabilidade a este critério:

“Comunicabilidade é outra propriedade importante, com a finalidade de garantir a transparência das decisões técnicas tomadas pelos administradores públicos e a compreensão delas por parte da população, dos jornalistas, dos representantes comunitários e dos demais agentes públicos” (JANNUZZI, 2005, p. 141).

Sensibilidade: se um indicador possui capacidade de “refletir tempestivamente as mudanças decorrentes das intervenções realizadas” (BRASIL, 2010a, p. 26), então é um indicador sensível. Jannuzzi também chama este critério de especificidade (2005, p. 141).

Adaptabilidade: para o TCU, é a “capacidade de resposta às mudanças de comportamento e exigências dos clientes. Os indicadores podem tornar-se desnecessários ao longo do tempo e devem ser eliminados ou substituídos por outros de maior utilidade” (BRASIL, 2009b, p. 12).

Disponibilidade: de acordo com o TCU, indica a “facilidade de acesso para coleta, estando disponível a tempo, para as pessoas certas sem distorções, servindo de base para que decisões sejam tomadas” (BRASIL, 2009b, p. 13).

Entende-se ser este um critério muito amplo, que agrega outros critérios já listados, como: acessibilidade, periodicidade e praticidade.

Investigativos: para a SEGES, “os dados devem ser fáceis de analisar, sejam estes para registro ou para reter informações e permitir juízos de valor” (BRASIL, 2009a, p. 47). Da mesma forma, entende-se que este critério corresponde ao da praticidade.

Mensurabilidade: entende-se que este critério, apesar de ser demandado pelo Manual de Elaboração do PPA 2008-2011 (BRASIL, 2010a, p. 26), segundo o qual é desejável, já está coberto por outros critérios, quais sejam: periodicidade, confiabilidade e factibilidade.

Representatividade: o indicador “deve expressar bem a realidade que representa ou mede” (BRASIL, 2009b, p. 13); deve também “ser a expressão dos produtos essenciais de uma atividade ou função” (BITTENCOURT 2006, pp. 15-16).

A essência que se pode extrair dessas proposições diz respeito à capacidade de o indicador realmente produzir informações de acordo com os seus objetivos, e de conseguir atender a todos os objetivos a que se propõe.

Entende-se que representatividade corresponde ao mesmo critério já listado, denominado validade.

Temporalidade: critério utilizado pela SPI que, por ter características agregadoras, foi coberto por outros critérios já anteriormente listados.

Além desses, o Manual de Elaboração do PPA 2008-2011 (BRASIL, 2007, p. 120) afirma que os indicadores devem ser verificáveis, passíveis de aferição, coerentes com o objetivo estabelecido, sensíveis à contribuição das principais ações e apuráveis em tempo oportuno.

Pelo estudo realizado, concluiu-se que essas exigências são atendidas pelos critérios de auditabilidade, validade, cobertura e periodicidade.

Relevância: Jannuzzi defende que “a relevância para a agenda político-social é a primeira e uma das propriedades fundamentais de que devem gozar os indicadores escolhidos em um sistema de formulação e avaliação de programas sociais específicos.” (2005, p. 139).

Pelo entendimento citado, pode-se considerar que esta característica é de fundamental importância para os programas sociais, uma vez que sempre estarão em uma condição de maior relevância os indicadores dos programas de governo que estiverem sob o foco da agenda política ou da mídia.

Além disso, por mais que um determinado indicador refira-se a um programa de relevância mediana, ainda assim poderá ser um indicador de extrema qualidade. Dessa forma, perde o sentido fazer uma análise da qualidade do indicador baseada na sua relevância. A relevância pode, sim, ser utilizada como critério para a escolha de indicadores que serão analisados.

Vale lembrar que, para efeitos deste trabalho, é considerada a relevância de cada categoria de indicadores para os *stakeholders*, sendo esta uma das dimensões de avaliação da abordagem que será descrita a seguir.

3. Metodologia.

Este artigo constitui-se em uma elaboração conceitual, a partir da literatura estudada, com o objetivo de apresentar uma proposta de abordagem para a avaliação de indicadores de desempenho de programas de governo, ou, em outras palavras, caminhos e diretrizes para uma meta-avaliação.

Inicialmente, foi realizada uma revisão bibliográfica, por meio da qual foram identificados os aspectos relativos à utilidade dos indicadores de desempenho. O principal fato inerente a essa utilidade que foi identificado na literatura consiste na necessidade de informação por parte dos *stakeholders*. Em seguida foi possível identificar que os indicadores de desempenho de programas governamentais podem ser divididos em categorias, conforme o tipo de informação que disponibilizam. Finalmente, foram estudados os critérios de qualidade dos indicadores, para que as informações por eles produzidas sejam mais precisas e melhor aproveitadas. A partir deste levantamento foi realizada, com base na literatura estudada, uma compilação dos principais critérios de qualidade que podem ser demandados de um indicador de desempenho.

As categorias de indicadores que foram listadas foram então submetidas ao confronto com outras duas dimensões que de certa forma definem e qualificam os indicadores de desempenho: *stakeholders* e critérios de qualidade.

O primeiro confronto parte do princípio de que uma informação destina-se a um ou mais interessados. Assim, a análise sobre a utilidade de determinada categoria deriva da reflexão sobre as necessidades dos *stakeholders*, considerando a natureza e tipologia de informação produzida em cada categoria de indicadores. Esse confronto, ou simplesmente essa verificação, permitiu estabelecer um parâmetro inicial, que foi chamado de nível de importância de um conjunto de indicadores, ou de categorias de indicadores, referentes a determinado programa de governo, sob a perspectiva dos *stakeholders*. As categorias de indicadores puderam, assim, ser pontuadas conforme sua importância para o programa e conforme a quantidade de grupos (*stakeholders*) interessados nas informações em cada categoria.

Estabelecido o nível de importância, percebeu-se ainda que é possível verificar em que medida esse parâmetro é atendido, no que diz respeito à existência total ou parcial das categorias de indicadores que sejam demandadas pelos *stakeholders*. Uma vez que seja identificado se um programa possui as categorias de indicadores que são demandadas pelos *stakeholders*, é possível estabelecer um índice de cobertura da necessidade de informação.

A fase seguinte consistiu em propor um segundo confronto, agora entre os critérios de qualidade compilados e cada uma das categorias de indicadores existentes. Para este exercício, considerou-se que os critérios podem ser divididos em dois grupos: essenciais e desejáveis. A inexistência de qualquer um dos critérios incluídos no grupo de “essenciais” compromete o indicador, o que leva à sua classificação como insuficiente, neste caso. Quanto

ao grupo de critérios “desejáveis”, é considerada a relação entre a quantidade de critérios de qualidade desejáveis atendidos e a quantidade total existente, com a evidenciação do percentual de critérios desejáveis identificados. Obteve-se, assim, o índice de qualidade para cada uma das categorias de indicadores existentes. Esses índices foram enfim ponderados e, o resultado, multiplicado pelo índice de cobertura do indicador, o que resultou, finalmente, no índice de efetividade geral do sistema de indicadores analisado.

4. Proposta de abordagem para uma meta-avaliação

4.1. Considerações para a construção da abordagem proposta.

O exercício de construção de uma proposta de avaliação multidimensional de indicadores consiste em um processo sequencial de análise que contempla três dimensões passíveis de análise: as **categorias** de indicadores de desempenho; os interesses ou necessidades dos **stakeholders**; e os **critérios** de qualidade.

Esta construção foi dividida em cinco etapas:

- a) identificação das necessidades dos *stakeholders* para cada categoria listada;
- b) verificação da medida em que as categorias existentes em um determinado programa atendem às necessidades dos *stakeholders* (medida de cobertura)
- c) verificação da adequação ou insuficiência de cada categoria de indicadores, pela identificação da existência dos critérios de qualidade considerados essenciais;
- d) verificação da qualidade de cada categoria de indicadores, pela identificação da quantidade de critérios desejáveis existentes
- e) a apuração da efetividade geral dos indicadores, pela conjunção das etapas anteriores.

Cabe destacar que esta abordagem se presta a realizar duas análises distintas. Uma consiste em um exame mais simples, na análise específica de um único indicador, que pode ser categorizado, avaliado e qualificado. Isso permite o exame, por exemplo, de um indicador intermediário, que será utilizado para compor um indicador complexo. Esta análise pode ser realizada tantas vezes quantas forem necessárias, para cada uma das categorias de indicadores que se desejar. Como poderá ser percebido adiante, esse tipo de análise não abrange todas as etapas. Considerando as etapas: necessidade – cobertura – adequação – qualidade – efetividade, a análise de um único critério, separadamente, não pode verificar a cobertura total do sistema de indicadores de um programa. Também, não é possível verificar a efetividade

geral do sistema, a não ser que essa única categoria de indicador seja suficiente para avaliar todo o programa.

A abordagem completa (com 5 etapas), por sua vez, permite avaliar o conjunto de indicadores complexos, de distintas categorias, portanto, para um dado programa. Trata-se da análise de um sistema de indicadores. Neste caso, além da análise de cada uma das categorias de indicadores de desempenho necessárias e disponíveis para o programa, que contemplam sua categorização, avaliação e qualificação, também será considerado se o conjunto de indicadores complexos disponível satisfaz as necessidades de informação dos *stakeholders*, e com que nível de qualidade.

Observa-se que essa segunda análise contempla a possibilidade de avaliar todo o conjunto de indicadores de desempenho de um determinado programa de governo. Da mesma forma poderá ser utilizada para, a partir de indicadores complexos que agreguem os indicadores de desempenho de vários programas, verificar a qualidade do indicador ou conjunto de indicadores propostos para a avaliação de uma determinada política pública.

A seguir, são detalhadas as cinco etapas sequenciais da abordagem de avaliação proposta.

4.2. A necessidade dos *stakeholders* em relação a cada categoria de indicador – primeiro confronto dimensional.

A perspectiva de análise proposta inicia-se pela verificação, sob a ótica de cada grupo de *stakeholders*, a respeito da necessidade ou não de receber informações relacionadas a cada uma das categorias de indicadores propostas para a composição de um sistema de indicadores. Nesta primeira etapa, confrontam-se as dimensões das **categorias** de indicadores com as necessidades dos *stakeholders*, identificando-se qual a relevância de cada uma das categorias propostas, na perspectiva de cada *stakeholder*.

É certo que existe uma dificuldade objetiva para aferir a percepção dos *stakeholders*. A expectativa não é de que sejam conduzidas pesquisas específicas com os *stakeholders*, o que elevaria sobremaneira o custo de uma meta-avaliação. Assim, a abordagem desta etapa do presente trabalho propõe, para fins de uma aplicação imediata, que o próprio usuário examinador dos indicadores realize esse exercício, assumindo a perspectiva de cada um dos *stakeholders* e estabelecendo, sob o provável ponto de vista de cada tipologia de *stakeholder*, qual sua necessidade informacional, em relação às categorias de indicadores listadas.

É possível, e mesmo necessário, que esse cruzamento (categoria x *stakeholders*) possa ser utilizado em um momento prévio ao da execução do programa, ainda na fase de seu

planejamento. Nessa fase de planejamento, com todas as ferramentas que lhe são intrínsecas, será possível definir com melhor precisão, para cada programa, quais são os *stakeholders* interessados naquele programa e quais as suas necessidades de informação. Assim, para eliminar a subjetividade das avaliações posteriores e as conseqüentes diferenças que podem surgir em função de diferentes avaliadores, este cruzamento pode ser alimentado no momento do desenho do próprio programa e dos seus indicadores.

Segue, a título de exemplo, uma lista de possíveis perguntas que pode ser utilizada como ponto de partida para verificar a necessidade de determinada categoria de indicadores de desempenho, para cada *stakeholder*:

Quadro 2 - sugestões de perguntas dirigidas à dimensão dos stakeholders em função de sua necessidade de informação relacionada às categorias de indicadores.

CATEGORIA	SUGESTÃO DE PERGUNTA
	Este <i>stakeholder</i> tem a necessidade de ser informado a respeito...
Eficácia	do cumprimento das metas que foram estabelecidas para este programa?
Eficiência	do nível de utilização de insumos para este programa?
Efetividade (impacto)	do nível de sucesso dos objetivos deste programa?
Economicidade	dos custos incorridos com os insumos alocados para a execução deste programa?
Custo de oportunidade	das alternativas existentes para a consecução dos objetivos deste programa?
Sustentabilidade	da continuidade dos efeitos futuros produzidos pelas ações atuais deste programa?
Satisfação do público-alvo	do nível de satisfação por parte do público-alvo do programa?

Para a verificação da necessidade de cada uma das categorias de indicadores selecionadas para a composição do sistema de indicadores sob análise, sugere-se uma medição com sistema binário, atribuindo-se um ponto para cada *stakeholder* que necessite da informação e zero, quando a categoria não interessar ao *stakeholder*, como exemplificado no Quadro 3.

Quadro 3 – Exemplo de verificação da necessidade de informações a respeito de cada categoria de indicadores selecionadas para compor um sistema de indicadores de um programa, em relação a grupos de *stakeholders* selecionados como potenciais interessados nessas informações.

STAKEHOLDERS	CATEGORIAS						
	Eficácia	Eficiência	Efetividade (impacto)	Economicidade	Custo de Oportunidade	Sustentabilidade	Satisfação do público-alvo
Beneficiário do Programa	1	1	1	1	0	0	1
Gestor do Programa	0	1	1	1	1	0	1
Contribuinte	1	0	1	0	0	0	0
Órgãos de controle	1	0	1	0	1	0	1
Comunidade internacional	0	1	1	1	0	0	1

TOTAL	3	3	5	3	2	0	4
--------------	----------	----------	----------	----------	----------	----------	----------

4.2.1. O nível de importância das categorias de indicadores.

Em vez de se verificar apenas quais categorias de indicadores são necessárias para compor um sistema de indicadores para determinado programa, podem ser estabelecidos níveis de necessidade, ou de importância, para cada categoria de indicadores em relação a cada um dos *stakeholders*, conforme necessidades específicas da análise que se está realizando.

Conforme a quantidade de *stakeholders* que necessite de cada categoria de indicadores ou, em um exame mais profundo, atribuindo-se pesos para esses *stakeholders*, podem ser ponderados os resultados obtidos e obtido o primeiro índice desta perspectiva de meta-avaliação: o **nível de importância de cada uma dessas categorias de indicadores**, sob a perspectiva dos *stakeholders*. No quadro 3, a título de exemplo, foram somadas a quantidade de *stakeholders* que apresentaram necessidades em relação a cada categoria, o que traz uma aproximação da importância relativa. Neste caso, não foram relativizadas/ponderadas as necessidades dos *stakeholders*, o que também pode ser feito, se for o caso.

É possível que, neste ponto da análise, em função de determinada categoria não se mostrar necessária para nenhum dos *stakeholders*, que não haja a necessidade da existência desta categoria de indicadores para a avaliação do programa em exame, o que sugeriria o seu descarte. É o que ocorre com a categoria da **sustentabilidade** que, no exemplo hipotético apresentado no quadro 3, mostrou-se desnecessária para o sistema de indicadores analisado.

Mas este **nível de importância** pode ser ainda mais sofisticado. No ato do desenvolvimento de um sistema de indicadores para um programa, ou no momento da avaliação desse sistema, alguns indicadores podem ter sua importância considerada superior à dos demais. Eficiência, eficácia e efetividade, são as categorias mais trabalhadas na literatura pesquisada, uma vez que estão contempladas em Bittencourt (2006), Brasil, SEGES (2009a); Brasil, SPI (2010); Brasil, SPI (2007); Brasil, TCU (2009b); Brasil, TCU (2000); e Jannuzzi (2005). Considerando estas mesmas referências, a categoria da economicidade não foi contemplada em Jannuzzi (2005) e, especialmente, também não foi contemplada no Manual de Elaboração: plano plurianual 2008-2011 (Brasil, 2007). Quanto às demais categorias, a frequência de suas considerações na literatura revisada é menor.

A respeito da possibilidade de se relativizar a importância dos indicadores, existe uma proposta da SEGES, em seu Guia Referencial para Medição de Desempenho e Manual para Construção de Indicadores, para que estas três categorias (eficiência, eficácia e efetividade,

chamadas de dimensões de resultado) sejam ponderadas com 60% do peso em relação às demais, que ficariam com 40%, com o pertinente argumento de que vale mais medir o resultado do que o esforço. A SEGES propõe ponderação de 25% para efetividade, 20% para eficácia e 15% para eficiência.

De qualquer forma, devem ser consideradas as particularidades de cada programa avaliado para que se estabeleça uma ponderação entre as diferentes categorias de indicadores para a avaliação do programa.

A partir das totalizações do quadro 3, pode-se aplicar a proposta da SEGES, com a finalidade de ponderar as categorias necessárias, em função de suas importâncias relativas. Vale lembrar que se trata da análise de um sistema de indicadores de um programa hipotético. Assim, descartando-se a categoria da **sustentabilidade**, que fora hipoteticamente considerada desnecessária para o caso em análise, tem-se que:

Quadro 4 – Exemplo de verificação da necessidade de informações a respeito de cada categoria de indicadores selecionadas para compor um sistema de indicadores de um programa, em relação a grupos de *stakeholders* selecionados.

	CATEGORIAS						
	Eficácia	Eficiência	Efetividade (impacto)	Economicidade	Custo de Oportunidade	Satisfação do público-alvo	TOTAL
PONTUAÇÃO TOTAL	3	3	5	3	2	4	20
PONDERAÇÕES (%)	20	15	25	15	10	15	100
NÍVEL DE IMPORTÂNCIA (%)	17	13	35	13	5	17	100

Vale lembrar que esta etapa consiste em um levantamento de necessidades de informação dos *stakeholders* e que independe da existência ou não dos indicadores, ou seja, pode se dar tanto na fase de planejamento quanto em uma fase avaliativa, na qual pode ser detectada a carência ou o excesso de determinada categoria de indicadores.

Definido o **nível de importância** para cada categoria de indicadores, esse nível deve ser traduzido em termos numéricos, seja em números absolutos ou relativos, de forma que a soma desses valores seja um número inteiro a partir do qual se possa calcular a próxima etapa, o índice de cobertura de um sistema de indicadores. No exemplo apresentado no quadro 4, a pontuação total de cada categoria foi multiplicada pela respectiva ponderação e, em seguida, esses valores foram transformados em números percentuais³.

4.2.2. O índice de cobertura de um sistema de indicadores de desempenho.

Identificadas e pontuadas as categorias de indicadores necessárias à avaliação de um programa, o passo seguinte corresponde à primeira fase da avaliação propriamente dita e pretende verificar o quanto daquela necessidade dos *stakeholders* em relação às categorias de indicadores está atendida pelas categorias de indicadores existentes em um determinado caso. Neste momento verifica-se quais são as categorias de indicadores que, tendo sido consideradas necessárias na etapa anterior, foram contempladas entre as categorias efetivamente existentes na avaliação do programa.

Atribui-se, então, para cada categoria necessária e existente a mesma pontuação do seu nível de importância que fora calculada na primeira etapa, conforme quadro 4. Para as categorias inexistentes, atribui-se o valor zero. Em seguida, somam-se esses valores e divide-se o valor obtido por 100. O resultado é uma relação que, colocada em termos percentuais, traduz o **índice de cobertura do sistema de indicadores** sob exame, como segue:

$$\frac{\sum \text{pontos obtidos pelos indicadores que são necessários e existentes}}{100} = \text{índice de cobertura do sistema (em \%)}$$

Essa relação demonstra em que medida as necessidades máximas de obtenção de informação por parte dos *stakeholders* estão sendo cobertas pelas categorias de indicadores existentes no sistema de indicadores que é utilizado pela gestão do programa.

Quadro 5 – Cálculo do índice de cobertura de um sistema de indicadores.

	CATEGORIAS						
	Eficácia	Eficiência	Efetividade (impacto)	Economicidade	Custo de Oportunidade	Satisfação do público-alvo	TOTAL (%)
NÍVEL DE IMPORTÂNCIA (%)	17	13	35	13	5	17	100
CATEGORIAS EXISTENTES	17	0	35	0	0	17	69

Pela observação do quadro 5, verifica-se que o índice de cobertura desse conjunto (sistema) de indicadores de desempenho hipotético é de sessenta e nove por cento. Neste exemplo, o sistema de indicadores do programa contempla as categorias de eficácia, efetividade e satisfação do público-alvo, passando ao largo das categorias de eficiência, economicidade e custo de oportunidade.

4.3. Categorias de indicadores submetidas aos critérios de qualidade – segundo confronto dimensional.

Este segundo confronto dimensional, corresponde à segunda fase da avaliação dos indicadores, de natureza qualitativa, e inicia-se pela terceira etapa da perspectiva de avaliação proposta.

4.3.1. Critérios essenciais de qualidade.

Nesta terceira etapa, as categorias de indicadores são submetidas ao critérios de qualidade essenciais para cada uma das categorias.

Para os indicadores considerados essenciais, o critério de avaliação da resposta é eliminatório, isto é, se um mesmo que apenas um único critério de qualidade essencial não seja atendido por determinada categoria de indicador, aquele indicador será considerado **insuficiente**. Assim, não há “pontuação” extra pelo atendimento a mais de um critério “essencial”, já que todos, nesta abordagem, são considerados imprescindíveis.

Essa forma de escolha de critérios tem similaridade com aquela sugerida pela SPI, em seu Guia Metodológico de Indicadores de programas, em sua Matriz de Priorização de Indicadores, a ser utilizada na fase de planejamento, com a finalidade de selecionar, entre diversos indicadores possíveis, quais são os mais adequados para a utilização em determinado programa (BRASIL, 2010a, pp. 61-62).

4.3.2. Critérios desejáveis de qualidade.

A quarta etapa da abordagem proposta consiste na aplicação dos critérios desejáveis de qualidade às categorias de indicadores que compõem o sistema analisado.

Para esses critérios desejáveis, a existência do atributo de qualidade por cada uma das categorias de indicadores avaliada deverá ser pontuada. Várias possibilidades podem ser testadas para a aferição dessa pontuação. No exemplo que se apresenta no quadro 6, optou-se, mais uma vez, pela pontuação binária: um = critério atendido e zero = critério não atendido. Mas outras formas de avaliação podem ser utilizadas, inclusive relativizando/ponderando a importância de cada um dos critérios para cada uma das categorias avaliadas.

Quadro 6 – Verificação dos critérios desejáveis de qualidade de uma categoria de indicadores de desempenho.

Cobertura	1
-----------	---

Independência	0
Desagregabilidade	1
Estabilidade	1
Homogeneidade	1
Comparabilidade	1
Rastreabilidade	1
Transparência	0
Denominação clara	1
Economicidade	0
Operacionalidade	1
Periodização	1
Publicidade	0
Seletividade	1
Sensibilidade	1
Simplicidade	1
	75%

No exemplo hipotético apresentado, a categoria verificada atende a 75% dos critérios de qualidade desejáveis. Este pode ser considerado o índice de qualidade desta categoria de indicador. Caso os critérios selecionados tenham importâncias variáveis para o programa ou para o sistema de indicadores analisados, então esse exame deverá passar por uma ponderação de importância entre os critérios de qualidade, a exemplo do que foi feito na verificação do nível de importância das categorias de indicadores.

Pela observação do quadro 6, verifica-se que o resultado obtido pode ser transcrito por meio da seguinte fórmula:

$$\frac{\sum \text{pontos atribuídos para os critérios desejáveis}}{\text{pontuação máxima possível para os critérios desejáveis}} \times 100 = \text{índice de qualidade do indicador (em \%)}$$

O percentual resultante da aplicação da fórmula reflete o **índice de qualidade** do indicador, para cada categoria. Neste caso, a avaliação é qualitativa e classificatória. Cada uma das categorias de indicadores selecionadas para compor o sistema de indicadores em análise, apresentará um valor percentual, depois de submetidas aos critérios desejáveis de qualidade.

Obtidos esses percentuais está finalizada a quarta etapa e, finalmente, passa-se ao último cálculo proposto (quinta etapa), que é o nível de efetividade geral do sistema de indicadores de um programa.

4.3.3. O nível de efetividade geral do sistema de indicadores de avaliação de um programa.

O valor aqui proposto, também expresso em termos percentuais, é dado pela seguinte fórmula:

Nível de efetividade geral = índice de cobertura do indicador x média aritmética dos índices de qualidade das categorias de indicadores que compõem o sistema.

É importante ressaltar que a nomenclatura aqui atribuída ao índice – **nível de efetividade geral** pressupõe que a análise está se desenvolvendo sob um sistema de indicadores finais de avaliação de um programa. Isso significa que, sejam os indicadores avaliados compostos ou simples, o nível de efetividade geral presta-se a avaliar o conjunto de indicadores que representa a maior agregação possível para os indicadores de um programa. Caso contrário, quando da análise de um conjunto parcial, a nomenclatura do índice deve ser adaptada, como nível de efetividade parcial ou nível de efetividade intermediário.

Com o cálculo do Nível de Efetividade Geral, finaliza-se a avaliação de um sistema de indicadores de determinado programa ou, de outro modo, realiza-se a meta-avaliação desse sistema de indicadores.

Para completar o exemplo simulado apresenta-se resultados hipotéticos para as três categorias de indicadores existentes (a partir de seis categorias necessárias) avaliados nos moldes do quadro 6 e, admitindo que nenhuma dessas categorias tivesse sido considerada insuficiente quando submetida aos critérios essenciais, é possível obter um cálculo parecido com o que segue:

$$\frac{(75\% + 100\% + 40\%)}{3} \times 69\% = 49\%$$

Neste exemplo, é obtida a média dos percentuais dos **índices de qualidade** das três categorias de indicadores analisadas. Em seguida, o valor obtido é multiplicado pelo índice de cobertura desse sistema de indicadores, obtido anteriormente (quadro 5) e resultando, enfim, no **nível de efetividade geral** do sistema de indicadores analisado.

Observa-se que, caso alguma dessas categorias tivesse obtido um **índice de qualidade** de 0% ou tivesse sido considerada insuficiente na verificação de atendimento aos critérios essenciais, este valor também deveria compor o cálculo, reduzindo a média do índice de qualidade.

5. Resultados e discussão

Este artigo possibilitou a apresentação de idéias para a avaliação dos indicadores de desempenho de programas de governo. Não apenas porque os indicadores devem existir, por força normativa, mas pela identificação de vários públicos que, por interesse ou necessidade, com certeza por direito, podem se fazer “proprietários” das informações que são ou que deveriam ser geradas pelos indicadores de desempenho. A abordagem aqui apresentada não esgota o rol desses *stakeholders*. Também não é definitivo o rol dos critérios de qualidade

aplicáveis aos indicadores. Nesse aspecto, há necessidade de permanecer a perspectiva de uma evolução constante desses critérios.

Ainda sobre esses critérios de qualidade, a compilação aqui apresentada resultou de uma tentativa de síntese. O que foi encontrado em termos de definições, de nomenclatura, de unidade, está ainda longe de permitir uma sistematização desses critérios. Sobre suas definições, não são claras o suficiente para que possam ser precisamente delimitados. Para uma sistematização, inicia-se um longo caminho de discussão e estudo. Em especial porque as classificações existentes são, quase sempre, voltadas para o desenvolvimento dos indicadores e não com o foco na sua avaliação (ou meta-avaliação).

De qualquer forma, foi possível estabelecer uma proposta de caminho que contempla uma avaliação completa de um sistema de indicadores de desempenho para programas de governo. Partiu-se da necessidade de informações por parte dos *stakeholders*, para diagnosticar o **nível de importância** das categorias de indicadores. Com esse dado, é possível verificar o **índice de cobertura** de um sistema de indicadores de desempenho e, então, compreender em que medida um indicador ou um sistema de indicadores cumpre sua função de produzir as informações esperadas pelos interessados nas informações.

Em seguida, foi possível propor a verificação qualitativa dos indicadores, contemplando as especificidades de cada categoria e a importância relativa dos critérios de qualidade existentes para cada uma delas. Verificou-se a possibilidade de dividir os critérios de qualidade em dois grupos: essenciais e desejáveis. Isso permite avaliar os indicadores, primeiro, sob o aspecto eliminatório, para aqueles critérios cuja carência comprometeria o resultado da informação e, em seguida, sob o aspecto classificatório, compreendendo que a presença de uma quantidade maior de critérios desejáveis tem o poder de aumentar a qualidade da informação. Resta a discussão sobre a importância deste ou daquele critério e ainda a avaliação quanto a ponderação dos critérios de qualidade para cada caso específico de sistema de indicadores.

Finalmente, ponderando-se os diversos índices de qualidade obtidos, tem-se a média da qualidade de um sistema de indicadores avaliado. Essa média, multiplicada pelo índice de cobertura, traz o nível de efetividade geral de um sistema de indicadores. Os dois índices que compõem a efetividade geral do sistema de indicadores sintetizam a suficiência, a necessidade e a qualidade da informação.

Esta proposta de uma meta-avaliação de indicadores de desempenho é, sem dúvida, uma proposição dinâmica, aberta, e receptiva a aperfeiçoamentos e avanços que poderão ser propostos a partir de estudos futuros.

Cabe, principalmente, definir uma escala de qualificação para os Níveis de Efetividade Gerais esperados para os sistemas de indicadores, com a finalidade de estabelecer metas de qualidade para os indicadores. A partir dessa definição, podem ser selecionados os programas de governo que poderão ser avaliados, com certo nível de credibilidade, por meio de indicadores, conforme sua posição nessa escala de avaliação a ser criada.

O principal resultado obtido por meio deste artigo reflete a possibilidade de se desenvolver um método objetivo para a avaliação dos indicadores de desempenho de programas de governo.

Uma vez adaptadas às sistemáticas de avaliação da CGU, as perspectivas aqui propostas podem ser imediatamente aplicadas, como suplemento ao acompanhamento sistemático da execução de programas de governo, para avaliar a qualidade dos indicadores dos programas de governo envolvidos nesse acompanhamento. Essa aplicação, além de permitir a proposição de melhorias na estrutura dos indicadores avaliados, pode permitir a utilização daqueles sistemas de indicadores que se mostrarem mais efetivos, com o intuito de auxiliar no processo de avaliação dos programas de governo.

O mesmo pode ser feito para os programas não hierarquizados no acompanhamento realizado pela CGU, a título de complemento ao acompanhamento sistemático. Do mesmo modo, é possível escalonar os sistemas de indicadores de desempenho dos diversos programas do governo federal e, isso feito, servir-se dos melhores indicadores e recomendar melhorias àqueles indicadores que apresentarem um menor nível de efetividade.

O resultado que foi apresentado por este trabalho remete à essência do acompanhamento gerencial dos programas governamentais. A avaliação de indicadores de desempenho pode ajudar a reduzir o tempo que hoje é destinado à realização de auditorias ou, por outro lado, ampliar a capacidade de acompanhamento gerencial de programas de governo. Não porque suas propostas estejam completas, mas justamente pelas possibilidades de evolução dessas propostas. Novas dimensões podem ser acrescentadas às análises. Os critérios podem ser revistos e relativizados/ponderados. Pode-se ampliar o estudo para uma análise transversal de sistemas de indicadores, por meio da qual seria possível verificar as interferências de um programa em outro. Enfim, a proposta de uma meta-avaliação dos indicadores de programas pode ser ampliada e aperfeiçoada, tendo em vista os caminhos e as possibilidades que se pretendeu provocar a partir deste trabalho.

NOTAS

¹ Partes interessadas. No caso, os *stakeholders* são referidos como sendo cada um dos grupos que são potencialmente interessados na obtenção de algum tipo de informação a respeito do desempenho dos programas de governo.

² Dependendo dos objetivos de um programa, pode não ser interessante superar uma meta, seja pela economia de recursos, seja pelo excesso de serviço gerado, que pode conduzir à ociosidade ou ao desperdício.

³ No caso do custo de oportunidade, o arredondamento do valor, que deveria se dar para o próximo número inteiro, foi forçosamente conduzido ao número inteiro anterior, para permitir o uso de números inteiros e, assim, facilitar a compreensão dos cálculos realizados.

6. Referências bibliográficas.

BONNEFOY, Juan Cristóbal; Armijo, Marianela. CEPAL. ILPES. Naciones Unidas. Serie Manuales, n° 45. **Indicadores de desempenho en el sector público**. Santiago, novembro de 2005. 106p. Disponível em: <<http://www.eclac.org/publicaciones/xml/2/23572/manual45.pdf>>. Acessado em 07 fev. 2011.

BITTENCOURT, F.M.R. Indicadores de desempenho como instrumentos de gestão, auditoria e análise econômica. **Revista de Conjuntura**: Editada pelo Conselho Regional de Economia do Distrito Federal, Brasília, Ano VI, n° 23, jul/set de 2006. 2000 - . Trimestral. 38 p. Disponível em: <<http://www.corecondf.org.br/discovirtual/revdownload/revista23.pdf>>. Acesso em: 20 out. 2010.

BRASIL. Lei n° 10.180, de 6 de fevereiro de 2001. Organiza e disciplina os Sistemas de Planejamento e de Orçamento Federal, de Administração Financeira Federal, de Contabilidade Federal e de Controle Interno do Poder Executivo Federal, e dá outras providências. Disponível em: http://www.planalto.gov.br/ccivil_03/Leis/LEIS_2001/L10180. Acesso em: 29 dez. 2010.

BRASIL. Ministério do Planejamento, Orçamento e Gestão – Secretaria de Gestão. **Guia Referencial para Medição de Desempenho e Manual para Construção de Indicadores** (Melhoria da gestão pública por meio da definição de um Guia Referencial para Medição do Desempenho da Gestão, Controle para o Gerenciamento dos Indicadores de Eficiência, Eficácia e de Resultados do Programa Nacional de Gestão Pública e Desburocratização, Produto 4), Brasília: MP, 2009a. 112 p. Disponível em: <[http://www.gespublica.gov.br/biblioteca/pasta.2010-04-26.8678882873/Guia%20-%20Indicadores%20\(versao%20preliminar%20Dez%202009\).pdf/view](http://www.gespublica.gov.br/biblioteca/pasta.2010-04-26.8678882873/Guia%20-%20Indicadores%20(versao%20preliminar%20Dez%202009).pdf/view)>. Acesso em: 20 out. 2010.

BRASIL. Ministério do Planejamento, Orçamento e Gestão – Secretaria de Gestão. **Mapeamento Bibliográfico e do Estado da Arte sobre Indicadores de Gestão** (Melhoria da gestão pública por meio da definição de um Guia Referencial para Medição do Desempenho da Gestão, Controle para o Gerenciamento dos Indicadores de Eficiência, Eficácia e de Resultados do Programa Nacional de Gestão Pública e Desburocratização, Produto 1), Brasília: MP, 2009c. 316 p. Disponível em: <https://conteudo.gespublica.gov.br/biblioteca/pasta.2010-12-08.2954571235/produto_1_indicadores_versao_final_110809.pdf>. Acesso em: 4 jan. 2011.

BRASIL. Ministério do Planejamento, Orçamento e Gestão. Secretaria de Planejamento e Investimentos Estratégicos – SPI. **Indicadores de Programas: Guia Metodológico**, Brasília: MP, 2010a. 128 p. Disponível em: <http://www.planejamento.gov.br/secretarias/upload/Arquivos/spi/publicacoes/100324_indicadores_programas-guia_metodologico.pdf>. Acesso em: 20 out. 2010.

BRASIL. Ministério do Planejamento, Orçamento e Gestão. Secretaria de Planejamento e Investimentos Estratégicos – SPI. **Manual de avaliação: plano plurianual 2008/2011: exercício 2010: ano base 2009**, Brasília: MP, 2010b. 47 p. Disponível em:

<http://www.sigplan.gov.br/download/avaliacao/manualdeavaliacao_2010.pdf>. Acesso em: 12 jul. 2010.

BRASIL. Ministério do Planejamento, Orçamento e Gestão. Secretaria de Planejamento e Investimentos Estratégicos – SPI. **Manual de Elaboração: plano plurianual 2008-2011**, Brasília: MP, 2007. 126p. Disponível em: <<http://www.sigplan.gov.br/download/manuais/ManualdeElabora%C3%A7%C3%A3o2008-2011.pdf>>. Acesso em: 20 out. 2010.

BRASIL. TRIBUNAL DE CONTAS DA UNIÃO. **Indicadores de Desempenho**– Brasília: TCU, Secretaria-Geral da Presidência - Secretaria de Planejamento e Gestão, 2009b. 23 p. Disponível em: <http://www.controlepublico.org.br/evento/upload/28_111-indicadores%20de%20desempenho.pdf>. Acesso em: 21 out. 2010.

BRASIL. TRIBUNAL DE CONTAS DA UNIÃO. **Técnicas de Auditoria: Indicadores de Desempenho e Mapa de Produtos** – Brasília: TCU, Coordenadoria de Fiscalização e Controle, 2000. 32p. Disponível em: <<http://portal2.tcu.gov.br/portal/pls/portal/docs/670038.PDF>>. Acesso em: 21 out. 2010.

COSTA, Frederico Lustosa da; CASTANHAR, José Cezar. **Avaliação de programas públicos: desafios conceituais e metodológicos**. Revista de Administração pública, Rio de Janeiro, v.37, n.5, p. 969-992, Set./Out. 2003. Disponível em: <http://www.fgv.br/ebape/nova-ebape/comum/arq/Costa_castanha.pdf>. Acesso em: 14 nov. 2010.

HEDLER, Helga Cristina; TORRES, Cláudio Vaz. **Meta-avaliação de Auditorias de Natureza Operacional do Tribunal de Contas da União**. RAC, Curitiba, v. 13, n. 3, art. 7, p. 468-486, Jul./Ago. 2009. Disponível em: <http://www.anpad.org.br/periodicos/arq_pdf/a_878.pdf>. Acesso em: 21 out. 2010.

FERREIRA, Aurélio Buarque de Holanda. **Novo Dicionário Eletrônico Aurélio versão 5.11a**. Novo Dicionário Aurélio da Língua Portuguesa corresponde à 3ª. edição, 1ª. impressão da Editora Positivo, revista e atualizada do Aurélio Século XXI, O Dicionário da Língua Portuguesa. Editora Positivo. Edição eletrônica autorizada à POSITIVO INFORMÁTICA LTDA. 2004.

JANNUZZI, Paulo de Martino. Indicadores para diagnóstico, monitoramento e avaliação de programas sociais no Brasil. **Revista do Serviço Público**, Brasília, nº 56 (2): 137-160, abr/jun de 2005. 1937 - . Trimestral. 254 p. Disponível em: < <http://www.conei.sp.gov.br/ind/ind-sociais-revista-serv-publico.pdf>>. Acesso em: 24 out. 2010.

PNUD. **Relatório de Desenvolvimento Humano 2010**. Edição do 20º aniversário. A verdadeira Riqueza das Nações: Vias para o Desenvolvimento Humano - Nova Iorque: Communications Development Incorporated, Washington, D.C. 2010a. 253p. Disponível em: <http://hdr.undp.org/en/media/HDR_2010_PT_Complete.pdf>. Acesso em: 07 nov. 2010.

PNUD. Nota Técnica de apoio ao Lançamento do Relatório de Desenvolvimento Humano 2010 “A verdadeira Riqueza das Nações”, Brasil, 2010b. 11p. Disponível em: <<http://www.pnud.org.br/arquivos/arqui1288884144.zip>>. Acesso em: 07 nov. 2010.