

INVESTIGAÇÃO AVALIATIVA DE PROGRAMAS: RELEITURA DE UMA METODOLOGIA¹

Rosita Saupe *
Simone Beatriz Pedrozo Viana **
Ionice Maria Amaral ***

RESUMO

O trabalho apresenta um modelo para avaliação, considerada a etapa mais complexa no processo de desenvolvimento de programas. Reconhece a contribuição das metodologias existentes, mas procura recuperar o valor da expressão numérica para detectar regiões de fracasso, indefinição ou sucesso de projetos. Inicia com a construção de um diagrama de árvore, que representa a decomposição do programa em todas as suas variáveis; submete este diagrama ao método do júri, que atribui pesos conforme o valor relativo de cada componente; analisa a concordância ou não entre os pesos atribuídos pelos juizes; elabora os instrumentos para a coleta dos dados a partir do diagrama consolidado; aplica as equações para obtenção dos resultados. Conclui pela necessidade de se equalizar a importância de dados quantitativos e qualitativos para a efetivação de avaliações que produzam um conjunto completo de informações, o qual favorecerá o monitoramento de programas, tanto em relação aos seus objetivos quanto à satisfação dos envolvidos.

Palavras-chave: Métodos. Avaliação. Avaliação de programas.

INTRODUÇÃO

A avaliação é, sem dúvida, o desafio perene dos programas, sejam eles educacionais, assistenciais, gerenciais ou de pesquisa. A literatura sobre este tema é vasta e pode ser ilustrada por algumas propostas com as quais temos tido mais contato⁽¹⁻⁴⁾ e que, segundo nossa percepção, são amplas, mas frequentemente pouco mensuráveis. Elas resultam em boas informações para a tomada de decisão, mas, na maioria, não possibilitam uma expressão numérica da avaliação que permita acompanhar sua implementação e a comparação entre as várias etapas do desenvolvimento de um

mesmo programa ou com outros programas semelhantes. Na perspectiva de contribuir para o desenvolvimento de outras possibilidades mais sistematizadas, resolvemos resgatar uma metodologia que tem sua origem na *University of North Carolina*, proposta no início da década de setenta, introduzida no Brasil em meados da mesma década⁽⁵⁾ e utilizada na avaliação de alguns programas⁽⁶⁻¹⁰⁾.

Entendemos o momento atual como muito propício a este resgate, na medida em que tanto as políticas públicas voltadas para a educação quanto aquelas voltadas para a saúde, estão levando as instituições a elaborar, acompanhar e avaliar seus projetos. As

¹ Metodologia utilizada na pesquisa “Avaliação do processo de educação permanente em atenção básica”, projeto submetido ao Edital MCT-CNPq/MS-DAB/SAS nº 49/2005, aprovado e financiado conforme Processo 402044/2005-3, em desenvolvimento junto ao Programa de Mestrado Profissionalizante em Saúde e Gestão do Trabalho, da Universidade do Vale do Itajaí (UNIVALI).

* Enfermeira. Doutora em Enfermagem. Professora do Curso de Graduação em Enfermagem e do Programa de Mestrado Profissionalizante em Saúde e Gestão do Trabalho da UNIVALI.

** Fisioterapeuta. Mestre em Saúde da Família. Professora do Curso de Graduação em Fisioterapia da UNIVALI.

*** Enfermeira. Mestre em Saúde da Família. Professora do Curso de Graduação em Enfermagem da UNIVALI.

instituições formadoras, orientadas pela Lei de Diretrizes e Bases da Educação Nacional e pelas Diretrizes Curriculares para os Cursos na

Área da Saúde, estão construindo e implementando seus projetos pedagógicos; ao mesmo tempo, as instituições de saúde, na perspectiva de consolidação do Sistema Único de Saúde (SUS), centram esforços na implantação da Estratégia de Saúde da Família e em Programas de Educação Permanente. Todos estes projetos, bem como os subprojetos a eles vinculados, demandam de avaliações que subsidiem as mudanças e ajustes que se fizerem necessários⁽¹¹⁾.

A metodologia que estamos resgatando e ajustando às necessidades da dinâmica atual e às necessidades de avaliação das políticas públicas é proposta como capaz de detectar as possíveis falhas decorrentes do planejamento e da implantação de um programa (sentido amplo) e oferecer, ainda, parâmetros que reflitam o grau de desenvolvimento em seus múltiplos aspectos: cumprimento de metas; maximização em relação aos recursos; uniformidade da ação; otimização em confronto com outros programas⁽⁵⁾.

DESCRIÇÃO DA METODOLOGIA

A metodologia inclui as seguintes etapas: elaboração do “Diagrama de Árvore”; consulta a “*experts*”, denominado “Método do Júri”; verificação da concordância entre os juízes; construção e aplicação do(s) instrumento(s); avaliação do desempenho. A consolidação da avaliação ocorre através da identificação das aproximações e distanciamentos entre os resultados obtidos nas questões objetivas, em confronto com as respostas às perguntas abertas incluídas no mesmo instrumento e fornecidas pelos mesmos sujeitos, como veremos na seqüência do texto.

Primeira etapa: Diagrama de Árvore

O método inicia pela decomposição do programa a ser avaliado em todos os seus componentes, o mais detalhadamente possível. As fontes para a definição dos componentes se originam na experiência dos pesquisadores, no conteúdo do próprio projeto ou programa em avaliação e na literatura, podendo envolver a consulta a especialistas e peritos altamente respeitados na área de conhecimento específica. As contribuições dos peritos são obtidas através de entrevistas de profundidade. O conjunto obtido destas informações passa a orientar a definição de componentes ou variáveis a serem privilegiadas. Esta etapa pode ser considerada como a contextualização do fenômeno, capturando as várias perspectivas de perceber a realidade na qual está embutido o programa a ser avaliado.

Recomenda-se a organização de mapas a partir de descritores ou palavras-chave, evidenciados no processo de captura, os quais possibilitarão a identificação de aproximações e distanciamentos entre as várias perspectivas. O mapeamento ajuda a delimitar a realidade e estabelecer prioridades e subsidia a análise da viabilidade e possibilidades das proposições.

Passa-se então à construção do diagrama por parte dos pesquisadores ou técnicos envolvidos no projeto de avaliação, que, necessariamente, precisam ter uma perspectiva interdisciplinar e domínio da temática e estar capacitados a construir o primeiro exemplar do desenho da avaliação, que é denominado “Diagrama de Árvore” (figura 1), incluindo, de forma hierárquica, todos os aspectos que pretendem abordar. Esta matriz será então submetida, na etapa seguinte, a um grupo de profissionais, definidos entre especialistas na área.



Figura 1 – Diagrama de Árvore⁽⁶⁾.

Segunda etapa: método do júri

Estabelecida a matriz original representada pelo 'Diagrama de Árvore' (figura 1), considerada a primeira etapa do processo, passa-se à segunda fase. Nesta, considerando-se que cada componente do diagrama apresenta uma importância relativa, procede-se à ponderação de cada ramificação. Esta ponderação é efetivada através do *método do júri* ou consulta a *'experts'*, considerada uma forma científica de construir e avaliar programas e projetos. Parte do princípio de que a construção de modelos fica consolidada quando submetida à apreciação de especialistas na área. São considerados possíveis "juízes" os estudiosos e pesquisadores da área de conhecimento relacionada ao tema, convidados para atuarem como consultores, ficando o júri composto por número ímpar daqueles que aceitarem participar.

Para o estabelecimento dos pesos correspondentes a cada variável componente do Diagrama de Árvore pode-se obedecer a múltiplos critérios⁽⁵⁻⁸⁾: prioridade dos objetivos; custos ou benefícios; quantificação das atividades; "inputs" ou "outputs"; e importância da atividade/componente.

Neste texto optamos por utilizar a atribuição dos pesos através da **importância do componente**, conforme opinião dos juízes, ou seja, cada juiz atribuirá um valor, conforme importância atribuída a cada variável componente do diagrama, sempre em comparação com as demais em seu nível. Esta

opção deve-se ao fato de que este critério tem se mostrado o mais adequado em nossa experiência. A importância do componente será traduzida por um peso resultante da aplicação da seguinte escala:

Peso 1 = associa-se a um componente cuja importância relativa, em seu grupo e nível de desdobramento, é **muito baixa**;

Peso 2 = associa-se a um componente cuja importância relativa, em seu grupo e nível de desdobramento, é **baixa**;

Peso 3 = associa-se a um componente cuja importância relativa, em seu grupo e nível de desdobramento é **mediana**;

Peso 4 = associa-se a um componente cuja importância relativa, em seu grupo e nível de desdobramento, é **alta**;

Peso 5 = associa-se a um componente cuja importância relativa, em seu grupo e nível de desdobramento, é **muito alta**.

Terceira etapa: concordância dos juízes

Após atribuição dos pesos pelos diversos juízes, verifica-se a concordância entre estes. Esta concordância ocorre quando TODOS os avaliadores atribuem o mesmo valor ao componente em questão. Quando a concordância não se expressa por esta unanimidade, o método recomenda que sejam aplicados testes estatísticos, utilizando-se de duas provas^(12, 13): o Teste de *Wilcoxon*: quando desdobramento apresenta duas dimensões; e o Coeficiente de Concordância de *Kendall*, quando o desdobramento contém mais de duas dimensões.

Outras possibilidades referem-se ao uso da média ou da mediana dos pesos atribuídos pelos juízes. Em nossos estudos temos dado preferência a esta opção, que tem respondido adequadamente às nossas necessidades. Resulta daí um mapa detalhado em todos os seus componentes, com seu valor relativizado frente aos demais componentes, possibilitando a análise pretendida. Os valores obtidos pela consulta aos juízes, após serem submetidos aos testes escolhidos, são inseridos no Diagrama de Árvore, passando a representar o PESO do componente.

Quarta etapa: construção e aplicação do instrumento e coleta dos dados.

Completado o Diagrama de Árvore com todos os seus componentes discriminados e pesos atribuídos, passamos à construção do ou dos instrumentos, focalizando os componentes de último nível e ajustando-os a uma escala Likert de 1 a 10. Além disto, é importante incluir questões abertas, que possibilitem expressões individuais dos informantes, evidenciando aspectos comuns ou singulares de suas percepções quanto ao programa em avaliação. Com isto obteremos um conjunto de dados que possibilitará inúmeras análises. O instrumento deve ser testado num grupo de sujeitos que compõem a população selecionada para avaliar o programa e sofrer os ajustes daí resultantes. Recomendam-se estudos censitários, para a obtenção de informações do maior número possível de sujeitos que são diretamente influenciados pelo projeto ou programa. A seguir apresentamos um modelo de questão vinculada a um componente:

QUESTÃO 1 - Avalie seu nível de satisfação com os programas de educação permanente que tem freqüentado, assinalando o número correspondente, considerando o 1 como "avaliação **TOTALMENTE NEGATIVA**" e o 10 como "avaliação **TOTALMENTE POSITIVA**". No intervalo de 2 a 9 você encontra as opções intermediárias entre estes extremos.

ESCALA LIKERT									
1	2	3	4	5	6	7	8	9	10
Avaliação Totalmente Negativa									Avaliação Totalmente Positiva

Por **FAVOR**, justifique ou complemente sua opção:

A forma de aplicação do(s) instrumento(s) para a coleta dos dados será decidida pelos avaliadores, admitindo-se a entrevista, a consulta por correio ou internet, os grupos focais ou outras técnicas de domínio dos pesquisadores, lembrando-se que, quanto menor a interação entre pesquisador e sujeito, mais detalhadas devem ser as orientações para o devido preenchimento do instrumento. Seja qual for a técnica selecionada, deve-se dar especial atenção aos cuidados éticos envolvidos e ao processo de sensibilização, aproximação e interação com os sujeitos respondentes.

Quinta etapa: avaliação do desempenho.

Todas as medidas se iniciam pelas atividades de último nível e vão prosseguindo pelas ramificações do Diagrama de Árvore até chegarem ao programa como um todo. No caso do desempenho, distinguem-se dois processos⁽⁵⁻⁸⁾.

O primeiro processo é a medida das atividades de último nível, cujos critérios podem ser vários. Recomendamos a utilização do critério subjetivo, que consiste em estabelecer a correspondência entre os valores de uma escala ordinal – através da qual foram medidas as variáveis de avaliação no instrumento - e a escala-padrão de medida do desempenho, apresentada na figura 2. Recomendamos ainda que a escala ordinal corresponda diretamente à escala-padrão, ou seja, de 1 (um) a 10 (dez) para ambas. Assim, a equação a ser aplicada é a seguinte:

$$(VEO1 \times Nr1) + (VE2 \times Nr2) + \dots$$

Até incluir todas as questões.
 $Nr1 + Nr2 + Nr3 + \dots$
 Até incluir todas as questões.

Legenda:
 VEO1 = Valor da escala ordinal na questão 1, e assim sucessivamente.
 Nr1 = Número de respostas obtidas na questão 1, e assim sucessivamente.

O segundo processo é a medida dos componentes dos demais níveis, que consiste na aplicação da equação:

$$D = \frac{\sum Da Wa}{\sum Wa}$$

Legenda: D = Desempenho
 $\sum Da Wa$ = somatória dos índices resultantes da avaliação das atividades de último nível e multiplicados pelo seu peso.
 $\sum Wa$ = somatória dos pesos

Ou seja, o desempenho de cada componente é igual à somatória dos índices resultantes da avaliação das atividades de último nível e multiplicados pelo seu peso, conforme obtido na ‘terceira etapa’; este resultado é dividido pela somatória dos pesos.

Os valores obtidos em cada variável são convertidos para uma Escala de Medida do Desempenho (figura 2). Conforme esta escala, os valores situados entre 1 (um) e 4 (quatro) correspondem à “região de fracasso”, indicando a extrema fragilidade dos componentes que ficarem neste patamar; entre 4,1 (quatro e um décimo) e 7 (sete) encontra-se a “região de indefinição”, que atribui ao componente um desempenho intermediário. O sucesso é alcançado quando o componente atingir um patamar acima de 7,1 (sete e um décimo), evidenciando o pleno alcance do objetivo relativo ao mesmo.



Figura 2 – Escala-padrão de medida do desempenho⁽⁵⁾.

Estes resultados são inseridos no Diagrama de Árvore, completando o ciclo quantitativo da avaliação; ou seja, cada componente ou variável avaliada terá sua representação, incluindo sua descrição, o peso relativo de seu valor em confronto com os demais componentes de seu nível de desdobramento, a explicitação sobre se houve ou não concordância entre os juízes quanto ao valor relativo do componente e uma medida de seu desempenho, sempre em relação ao conjunto dos aspectos avaliados.

As respostas às perguntas abertas são analisadas em seu conteúdo conforme método adaptado de um conjunto de orientações⁽¹⁴⁻¹⁶⁾, iniciando pela organização do “corpus”, entendido como o conjunto de informações coletadas através das entrevistas. Nesta etapa as falas ou depoimentos escritos dos sujeitos são transcritos, impressos e organizados em painéis. Cada linha horizontal corresponde a um sujeito e cada coluna, a uma questão ou tema abordado. Este modelo possibilita a visualização concomitante de vários depoimentos. A fase da leitura exaustiva procura impregnar o(s) pesquisador(es) das expressões verbais dos depoentes, facilitando a etapa seguinte de mapeamento das categorias, conforme agrupamento dos pólos de significado denominados:

- pólos gerais ou universais, quando presentes com frequência, caracterizando a unidade, homogeneidade, igualdade e uniformidade dos dados;
- pólos particulares ou singulares, quando indicarem a diversidade, a singularidade ou expressão particular do informante;
- pólos omissos, aspectos que deveriam ou poderiam estar presentes, mas não foram evidenciados nos dados.

Este processo permite apreender a perspectiva coletiva, não nega a contribuição dos sujeitos autônomos, que apresentem opiniões pessoais, diferenciadas e originais e inclui expectativas do pesquisador e indicações da literatura que não foram contempladas.

CONCLUSÕES

A perspectiva desta metodologia incorpora a contribuição de Lord Kelvin, que afirma que a expressão numérica de um fenômeno ajuda a compreendê-lo e a definir suas possibilidades⁽¹⁷⁾, e para isto este modelo contribui, sem sombra de dúvida. Não obstante, nossa posição procura ultrapassar a dicotomia quantidade *versus* qualidade e defende a necessidade de integrar dados quantitativos e qualitativos nos processos de avaliação. Entendemos que, ao romperem com um destes paradigmas ou radicalizarem em posições polarizadas, os profissionais empobrecem suas pesquisas, no caso, a investigação com objetivos de avaliação de programas.

A metodologia resgatada neste artigo procura salientar a contribuição de uma avaliação que torna visível, através da atribuição de um valor numérico objetivo, o desempenho de cada componente ou variável de um programa. Assim, possibilita sua comparação com os demais componentes no mesmo nível de importância ou em nível diferente, o

estabelecimento de estratégias para a conservação ou melhoria deste desempenho e a reconsideração das metas originais. Quando complementada com dados qualitativos, produzirá um conjunto completo de informações, que favorecerá o monitoramento do programa tanto em relação aos seus objetivos quanto à satisfação dos envolvidos.

EVALUATIVE STUDY OF PROGRAMS: RE-READING A METHODOLOGY

ABSTRACT

This work presents a model for evaluation, which is considered the most complex stage in the process of program development. It recognizes the contribution of the existing methodologies, but seeks to recover the value of numerical expression for detecting areas of failure, lack of definition, or success of the programs. It begins by constructing of a tree diagram, which breaks down the program into all its variables; this diagram is then submitted to the jury method, which attributes weights according to the relative value of each component; the study analyzes the level of agreement between the weights attributed by the judges; it elaborates the data collection instruments based on the consolidated diagram; and applies the equations for obtaining the results. It concludes with a need to balance the importance of the quantitative and qualitative data for carrying out evaluations that produce a complete set of information, which will benefit the monitoring of programs, both in relation to their objectives and the satisfaction of those involved.

Key words: Methods. Evaluation. Program evaluation.

INVESTIGACIÓN EVALUATIVA DE PROGRAMAS: RELECTURA DE UNA METODOLOGÍA

RESUMEN

El trabajo presenta un modelo para evaluación, considerada la etapa más compleja en el proceso de desarrollo de programas. Reconoce la contribución de las metodologías existentes, mas trata de recuperar el valor de la expresión numérica para detectar regiones de fracaso, indefinición o éxito en los mismos. Comienza con la construcción de un diagrama de árbol que representa la descomposición del programa en todas sus variables; somete este diagrama al método del jurado que atribuye pesos según el valor relativo de cada componente; analiza la concordancia o no entre los pesos atribuidos por los jueces; elabora los instrumentos para la colecta de datos a partir del diagrama consolidado; aplica las ecuaciones para la obtención de los resultados. Concluye con la necesidad de equalizar la importancia de datos cuantitativos y cualitativos para la efectución de evaluaciones que produzcan un conjunto de informaciones completas que favorecerán el monitoreo de programas, tanto en relación a sus objetivos cuanto a la satisfacción de los involucrados.

Palabras Clave: Métodos. Evaluación. Programa evaluación.

REFERÊNCIAS

1. Starfield B. Atenção primária: equilíbrio entre necessidades de saúde, serviços e tecnologia. Brasília, DF: UNESCO: Ministério da Saúde; 2002.
2. Popham WJ. Manual de avaliação: regras práticas para o avaliador educacional. Petrópolis: Vozes; 1977.
3. Saul AM. Avaliação emancipatória: desafio à teoria e à prática de avaliação e reformulação de currículo. 3. ed. São Paulo: Cortez; 1995.
4. Luckesi C. Avaliação da aprendizagem escolar. São Paulo: Cortez; 1995.
5. Spínola MRP, Pereira EM. Avaliação de Programas - uma metodologia. *Projeção*. 1976;7(1):26-32.
6. Spínola MRP, Pereira EM. Avaliação do Programa Imposto de Renda - 77. *Projeção*. 1977;19 Supl Esp:1-11.
7. Saupe R. Proposição de uma metodologia para avaliação de cursos de graduação em enfermagem. 1979 [Dissertação de Mestrado]. Florianópolis (SC): Programa de Pós-Graduação em Enfermagem - UFSC; 1979.
8. Saupe R. Avaliação do curso de enfermagem da UFSC conforme opinião de egressos e empregadores. *Rev Cienc Saude*. 1983 jun.;2(3):55-70.
9. Viana SBP. Competências dos fisioterapeutas para a Atenção Básica em Saúde da Família. 2005. [Dissertação de Mestrado]. Itajaí: Universidade do Vale do Itajaí; 2005.
10. Amaral IM. Competência do enfermeiro para promoção e educação em saúde da família. 2005. [Dissertação de Mestrado]. Itajaí: Universidade do Vale do Itajaí; 2005.

11. Scochi MJ, Silver L. Municipalização e avaliação de qualidade de serviços de saúde: uma análise localizada. *Cienc Cuid Saude* 2002 jan/jun; 1(1): 55-9.
12. Kendall MG. Rank correlation methods. London: Griffin; 1948.
13. Siegal S. Estatística não-paramétrica para as ciências do comportamento. São Paulo: McGraw-Hill; 1975.
14. Bogdan R, Birten SK. Qualitative research for education: an introduction for theory and methods. Boston: Allyn and Bacon; 1982.
15. Minayo MCS. O desafio do conhecimento: pesquisa qualitativa em saúde. São Paulo: Hucitec; Rio de Janeiro: Abrasco; 1992.
16. Triviños ANS. Introdução à pesquisa em ciências sociais: a pesquisa qualitativa em educação. São Paulo: Atlas; 1987.
17. Paim L. Quantitativos e qualitativos do cuidado de enfermagem. [Tese Livre Docência]. João Pessoa: Universidade Federal da Paraíba; 1979.

Endereço para correspondência: Rosita Saupe. Rua Mediterrâneo, 172, apto 401. CEP: 88037-610 Florianópolis (SC) saupe@univali.br

Recebido em: 25/09/2006

Aprovado em: 26/02/2007