Universidade Estadual de Campinas

Faculdade de Educação

Dissertação de Mestrado

Avaliação de políticas, programas e projetos: uma contribuição para a área educacional

Autora: **Paula Arcoverde Cavalcanti**

Orientador: **Prof. Dr. José Roberto Rus Perez**

Campinas/SP

Agosto de 2002

UNICAMP
BIBLIOTECA CENTRAL
SEÇÃO CIRCULANTE

UNIVERSIDADE ESTADUAL DE CAMPINAS

FACULDADE DE EDUCAÇÃO

DISSERTAÇÃO DE MESTRADO

Avaliação de políticas, programas e projetos: uma contribuição para a área educacional

tora: Paula Arcoverde Cavalcanti

ientador: Prof. Dr. José Roberto Rus Perez

Este exemplar corresponde à redação final da *DISSERTAÇÃO* defendida por Paula Arcoverde Cavalcanti e aprovada pela comissão julgadora.

Data: 12/08/02

Prof. Dr. José Roberto Rus Perez (orientador)

Comissão Julgadora:

Prof. Dr. José Roberto Rus Perez

Prof. Dr. José Dias Sobrinho

Prof. Dr. Romualdo Luiz Portela de Oliveira

© by Paula Arcoverde Cavalcanti, 2002.

UNIDADE BC
Nº CHAMADA T/UNICAMP C314a
V EX
TOMBO BC/ 52032
PROC. 16-124/03
C D X
PREÇO R$ 11,00
DATA 12/02/03
Nº CPD

CM00179202-2

BIB ID 278414

Catalogação na Publicação elaborada pela biblioteca da Faculdade de Educação/UNICAMP
Bibliotecária: Rosemary Passos - CRB-8ª/5751

C314a

Cavalcanti, Paula Arcoverde.
Avaliação de políticas, programas e projetos : uma contribuição para a área educacional / Paula Arcoverde Cavalcanti. -- Campinas, SP: [s.n.], 2002.

Orientador : José Roberto Rus Perez.
Dissertação (mestrado) – Universidade Estadual de Campinas, Faculdade de Educação.

1. Políticas publicas. 2. Política e educação. 3. Avaliação – Metodologia. 4. Avaliação de Pesquisa (Programa de ação social). 5. Projetos – Avaliação. I. Ruz Perez, José Roberto. II. Universidade Estadual de Campinas. Faculdade de Educação. III. Título.

02-140-BFE

Dissertação apresentada como exigência parcial para a obtenção do título de MESTRE em EDUCAÇÃO, na área de Política de Educação e Sistemas Educativos, da Faculdade de Educação/UNICAMP, junto à banca examinadora, em 12/08/2002, sob a orientação do Prof. Dr. José Roberto Rus Perez.

RESUMO

A avaliação de políticas, programas e projetos sociais e educacionais foi incorporada de fato, a agenda governamental Brasileira no início dos anos 90. Entre os fatores que contribuíram para isso destacam-se: o ajuste econômico e conseqüente redução dos recursos para a área social, as maiores exigências impostas pelos órgãos financiadores em relação ao controle de gastos e resultados. Uma dinâmica de racionalização, que incluiu a observância dos critérios de eficácia, efetividade e eficiência na utilização dos recursos financeiros e uma preocupação crescente com a *accountability*, passou a envolver a gestão pública brasileira.

Este trabalho tem como objetivo sistematizar a literatura acerca do tema, com vistas a proporcionar aos que se iniciam na avaliação das políticas, programas e projetos sociais e educacionais brasileiras um insumo metodológico para esta complexa atividade. Através da análise das contribuições de diferentes autores nacionais e estrangeiros e da comparação de seus enfoques, o trabalho discute aspectos tais como o processo de desenvolvimento do campo, os conceitos, modelos e tipos de avaliação, e os critérios de eficácia, efetividade e eficiência.

ABTRACT

Policy evaluation has been incorporated to the Brazilian government agenda in the early nineties. Among the factors contributing to this, the macroeconomic adjustment and the resultant decrease of resources oriented to social sectors, and the new requirements of financial sources on the control of their allocation, could be stressed. A dynamics towards rationalization, encompassing the observation of efficiency, efficacy and effectiveness criteria on resources utilization and an increasing concern with *accountability*, began to involve the Brazilian public decision-making.

This work systematizes the literature on the subject intending to provide the beginners on the complex activity of evaluating policies, programs and projects on the social and educational sectors with a methodological input. Through an analysis of foreign and Brazilian authors and a comparison of their approaches, the work discusses aspects as the field process of development, the concept, models and types of evaluation, and the efficiency, efficacy and effectiveness criteria.

A meu querido pai, Gilberto Cavalcanti e a minha querida mãe, Flavia Maria Faria Cavalcanti, pelo simples fato de terem me dado tudo.

A Gustavo Arcoverde de Vasconcelos, meu filho, pela força, pelo colorido, pela vida, pelo amor...

AGRADECIMENTOS

Mesmo fazendo um grande esforço, receio não conseguir agradecer a todas as pessoas que colaboraram comigo para a concretização deste trabalho.

Primeiramente a Renato Dagnino, pelo companheirismo, por compartilhar tristezas, ansiedades e alegrias durante o desenvolvimento deste trabalho e pelo profissional que é.

Ao Prof. Dr. José Roberto Rus Perez, pela orientação e por ter acreditado em mim.

Ao Departamento de Educação – Campus XIII da Universidade Estadual da Bahia, pela oportunidade que me foi concedida.

Ao meu Tio (irmão) Joaquim Faria e aos meus irmãos Gilberto e Fernanda por acreditarem no meu profissionalismo e sempre estarem ao meu lado.

A minha amiga de infância, Carla França, que sempre esteve presente em todos os momentos de minha vida, tenham sido estes tristes ou alegres, a amiga de Departamento, Silvia Maria Leite, que nos últimos anos muito ajudou em minha caminhada profissional e vida pessoal; á amiga Andréa Gatto, sempre disposta a me ouvir e ajudar.

A todos os meus alunos do Departamento de Educação-Campus XIII, da Universidade Estadual da Bahia, por serem minhas

"cobaias", meus incentivadores e por terem me ensinado que refletir , aprender e criar são verbos indispensáveis para um educador.

A Rudival Sampaio Queiroz, *in memorian*, um amigo que durante sete anos sempre "quebrou meus galhos" no Departamento de Educação-Campus XIII.

Aos amigos que aqui conquistei e que me deram força nessa caminhada: Adriana Momma, Cristiane Pavioti, Leila Maria Lopes Loureiro Saldanha, Prof. Dr. Luis Enrique Aguilar.

Ao pessoal da Secretaria de Pós graduação – Nadir, Rita e Gislene e a todos os funcionários desta Faculdade.

E especialmente ao LaPPlanE (Laboratório de Políticas Públicas e Planejamento Educacional), no qual convivi durante estes dois anos e meio e que tanto enriqueceu a minha vida profissional e pessoal.

ÍNDICE

Siglas ..... **xvii**

Lista de Quadros.....**xix**

INTRODUÇÃO..... **1**

CAPÍTULO I.....

O desenvolvimento da avaliação de políticas, programas e projetos sociais e educacionais.....**13**

1.1 O processo de construção das bases conceituais e metodológicas da avaliação nos Estados Unidos. 16

1.1.1 Primeiro momento: o enfoque na avaliação do rendimento do aluno.....22

.1.1.2 Segundo momento: o enfoque nos programas curriculares....27

1.2 A avaliação de programas educacionais no Brasil.....38

CAPÍTULO II.....

Conceito de avaliação de políticas, programas e projetos **47**

2.1. Conceitos de avaliação.....52

2.1.1 Conceito de avaliação: dimensão decisional, dimensão racional e dimensão integradora.....56

2.2. Avaliação como atividade de pesquisa.....67

CAPÍTULO III.....

Modelos de avaliação de políticas, programas e projetos.....**73**

3.1. Conjunto 1 - Modelos de avaliação: analítico, clássico ou quantitativo.....77

3.2. Conjunto 2 - Modelos de avaliação: global, alternativo ou qualitativos.....83

3.3. Conjunto 3 - Modelos de avaliação: analítico-global, clássico-alternativo ou misto.....89

3.4. Modelos de Avaliação: superação e críticas.....95

CAPÍTULO IV.....

Tipos de avaliação de políticas, programas e projetos.....**103**

4.1. Avaliação antes do início do programa.....107

4.2. Avaliação durante a execução do programa.....122

4.3. Avaliação após o término do programa.....147

CAPÍTULO V.....

Critérios de avaliação: eficácia, efetividade e eficiência.....**165**

5.1. Eficácia e efetividade.....169

5.2. Eficiência.....178

CONSIDERAÇÕES FINAIS.....**187**

REFERÊNCIAS BIBLIOGRÁFICAS.....**191**

BIBLIOGRAFIA CONSULTADA.....**197**

SIGLAS

CEA: Centro para o Estudo de Avaliação (Modelo de Avaliação).

BID: Banco Interamericano de Desenvolvimento.

CAD: Comitê de Assistência para o Desenvolvimento.

CEPAL: Comissão Econômica para a América Latina .

CIPP: Contexto, Insumo, Produto e Processo (Modelo de Avaliação).

ENC: Exame Nacional de Cursos (Provão)

ENEM: Exame Nacional do Ensino Médio

ESEA: Elementary end Secondary Education Act.

OEI: Organização dos Estados Ibero Americanos.

OCDE: Organização de Cooperação e Desenvolvimento Econômico.

OMS: Organização Mundial de Saúde.

ONU: Organização das Nações Unidas

ORELAC: Oficina Regional de Educação para a América Latina e o Caribe.

PAIUB: Programa de Avaliação Institucional das Universidades Brasileiras

SAEB: Sistema Nacional de Avaliação da Educação Básica

SIEMPRO: Sistema de Información, Monitoreo, y Evaluación de Programas

UNESCO: Organização das Nações Unidas Para a Educação, a Ciência e a Cultura.

UNICEF: Fundo das Nações Unidas para a Infância.

LISTA DE QUADROS

Quadro 1 - Relação dos autores analisados..... 12

Quadro 2 – Etapas de la concepción de la evaluación..... 54

Quadro 3 – Classificação dos modelos de avaliação por 3 conjuntos..... 76

Quadro 4 – Classificação do modelo de avaliação do Conjunto 1 por autor..... 82

Quadro 5 - Classificação do modelo de avaliação do Conjunto 2 por autor..... 87

Quadro 6 - Classificação do modelo de avaliação do Conjunto 3 por autor..... 92

Quadro 7 – Distribuição dos modelos de avaliação por autor segundo sua classificação..... 93

Quadro 8 – Nomenclatura dos tipos de avaliação por 3 Grupos segundo seus critérios..... 106

Quadro 9 – Nomenclatura utilizada para os tipos de avaliação do Grupo 1 por autor..... 121

Quadro 10 - Nomenclatura utilizada para os tipos de avaliação do Grupo 2 por autor..... 146

Quadro 11 - Nomenclatura utilizada para os tipos de avaliação do Grupo 3 por autor..... 164

INTRODUÇÃO

Questões sobre políticas educacionais fazem parte de minha vida profissional desde 1992, devido ao envolvimento com o sistema público de ensino sob duas óticas diferentes: enquanto implementadora de programas e projetos educacionais e como pesquisadora do cotidiano escolar onde estes se desenvolviam.

Com esse estrito contato, foi possível conhecer alguns dos esforços realizados pela Secretaria de Educação do Estado da Bahia em sistematizar avaliações do sistema escolar, a partir de algumas determinações do MEC (Ministério de Educação). Neste contexto, também foi percebida a crescente preocupação em avaliar as políticas, programas e projetos educacionais.

Nesse sentido, é possível apontar alguns dos fatores que contribuíram para a disseminação da necessidade em avaliar as políticas públicas de cunho social, no contexto brasileiro: o estabelecimento no país de padrões globalizantes de relacionamento internacional; levantamentos realizados pelos organismos financiadores internacionais sobre indicadores de pobreza; exigências dos órgãos nacionais e internacionais que financiam programas e projetos sociais, que passaram a ter como princípio básico o controle de gastos e avaliação de resultados; redução da intervenção estatal; ajuste econômico e conseqüente redução dos recursos financeiros para a área social.

Assim, este trabalho faz parte de um conjunto de preocupações a que faz alusão a literatura no que respeita à importância da avaliação de políticas públicas no atual contexto sócio-político e econômico brasileiro, uma vez que: é crescente a preocupação em "conhecer" se as políticas, programas e projetos sociais e educacionais estão tendo o efeito esperado; a avaliação vem sendo apresentada como um "instrumento democrático de controle das ações" governamentais; a avaliação é apontada como um instrumento que proporciona informações para planejamento e gestão; a avaliação é vista como um instrumento de feedback, para possíveis intervenções e correção das ações e atividades desempenhadas; e a avaliação é considerada como um instrumento que possibilita o controle dos recursos financeiros, cada vez mais escassos.

Em relação a essas questões controversas, pode-se afirmar que existe um consenso entre os estudiosos da área, no que diz respeito à necessidade de aprofundamento e reflexão acerca do tema, devido à sua complexidade, e por causa das proporções que a avaliação vem tomando dentro da agenda governamental, especificamente na área social e educacional.

Assim, acredita-se ser necessário destacar um aspecto sobre a literatura analisada, que trata da avaliação de políticas, programas e projetos, que consideramos ser importante para o entendimento deste trabalho: o ambiente "técnico" e de neutralidade, que cerca essa literatura. Salvo exceções, os autores encaram o tema da avaliação como se a política (programa, projeto) a avaliar estivesse simplesmente referida a (ou fosse o resultado de) critérios técnicos e que, portanto, sua avaliação poderia (ou deveria) também ser técnica. Como se a *policy* a ser avaliada não estivesse referida a interesses políticos (*politics*) dos atores que participam num processo que começa antes mesmo do momento da tomada de decisão, passa pela sua implementação e se encerra, dando origem a um novo ciclo, na sua avaliação. Como se o ambiente em que se desenvolvem esses momentos do ciclo da política não possuísse um intrínseco e inevitável duplo caráter técnico – *policy* – e político – *politics* – que a palavra denota (e encerra) nas línguas neo-latinas, a ser levado em conta por aqueles que querem avaliar seus resultados e impactos.

No entanto, mesmo tendo conhecimento de que a prática avaliativa está associada a questões políticas e ideológicas, a pesquisa, aqui apresentada, não entrará nessa discussão e deter-se-á sobre os aspectos teórico- metodológicos da avaliação de políticas, programas e projetos. Essa opção deve-se ao fato de que após a análise da literatura, foi possível encontrar dois aspectos que confirmam a necessidade de trazer à tona a discussão sobre o campo da avaliação.

O **primeiro** refere-se à diversidade analítico-conceitual da metodologia de avaliação de políticas, programas e projetos sociais e educacionais. O **segundo** aspecto diz respeito à escassez de trabalhos que busquem a sua sistematização.

A diversidade analítico-conceitual apontada parece ter origem na forma como se deu o surgimento da avaliação de política, programa e projeto, pois durante o período de amadurecimento e constituição do corpo teórico-metodológico específico dessa avaliação, houve uma contribuição de estudiosos de várias áreas do conhecimento, o que gerou uma variedade de correntes, conceitos, modelos e tipos de avaliação.

Ainda em relação a essa diversidade, encontram-se na literatura de avaliação de políticas sociais metodologias originárias da área social e educacional. Em relação a essas duas áreas é possível verificar a interseção dos pressupostos teórico-metodológicos entre ambas. Ou seja, a metodologia da área educacional é utilizada na área social, assim como também é possível encontrar o inverso.

Como exemplo disso, pode-se citar Briones (1996) que, em seu trabalho intitulado *Evaluación de Programas Sociales*, desenvolve um modelo de avaliação, tomando como base de análise modelos de avaliação que foram criados na área educacional. Além disso, o modelo de avaliação proposto por este autor é direcionado tanto para a área social como para a área educacional.

Isso talvez decorra porque a política educacional também é considerada de cunho social, o que de certa forma legitima a transferência e uso de metodologias entre as duas áreas, ainda que possuam especificidades.

Como exemplo dessa constatação pode-se apontar Weiss (1999, p. 13), ao afirmar que o estudo desenvolvido por ela tem como objetivo abordar os aspetos relacionados com a avaliação de políticas e programas sociais. Segundo ela, esses programas:

> *"...tienen como objetivo mejorar el sino de la gente. Estos programas son diversos: pueden tener como materia la educación, el bienestar social, la salud pública, la vivienda, la salud mental, los servicios jurídicos, las organizaciones correccionales, el desarrollo económico, y otros muchos campos. Pueden tener como fin cambiar el conocimiento, las actitudes, los valores y las conductas de las personas, o bien las instituciones con las que tratan o las comunidades en las que viven. La característica común de esos programas es la meta, que consiste en mejorar la vida y hacer que sea más satisfactoria para las personas a las que están destinados".*

Ainda em relação ao aspecto diversidade analítico-conceitual acerca da literatura sobre avaliação, é importante apontar

que os autores direcionam seus pressupostos teórico-metodológicos para três dimensões: política, programas e projetos[1].

Em relação à utilização desses pressupostos para avaliar objetos diferentes, Ballesteros (1996, p. 25) afirma:

> *"...política, programas, subprogramas y acciones suponen objetos de evaluación que difieren según un mayor / menor grado de complejidad."* [...] *Las políticas suponen él más alto nivel de complejidad y molaridad y conllevan, por tanto, mayor dificultad de evaluación; en definitivo las políticas pueden ser evaluadas a través de la evaluación de sus elementos constituyentes".*

Isto significa dizer que, ao desenvolverem seus pressupostos, os autores partem da perspectiva de que uma política é constituída por diversos programas e estes últimos englobam diferentes projetos. Portanto, para avaliar uma política é necessário que os programas e projetos que a integram sejam avaliados[2].

Neste sentido, pode-se apontar também como exemplo Cohen & Franco (1999, p. 16) que afirmam:

> *"...é necessário e factível avaliar as políticas sociais, no sentido de medir os resultados das ações realizadas. Mesmo os mais abstratos enunciados de política só podem ser postos em prática mediante sua operacionalização em programas e projetos que são modelados em ações concretas, sendo sempre possível (e necessário) avaliar se alcançaram os objetivos que queriam obter".*

---

[1] Os autores em sua grande maioria referem-se às três dimensões.
[2] Não está em discussão quem está avaliando.

Já em relação ao segundo aspecto — a escassez de sistematizações que tenham buscado aprofundar-se sobre questões metodológicas — pode-se afirmar que a análise minuciosa da produção científica, publicada a partir da década de 90, revelou um número bastante significativo de estudos que avaliam e analisam as ações governamentais ou detém-se às discussões políticas e ideológicas da avaliação no contexto brasileiro. No entanto, foi raro encontrar estudos e sistematizações que ofereçam um marco analítico-conceitual para a avaliação de políticas públicas.

Como possível explicação dessa escassez pode-se sinalizar que os estudiosos e pesquisadores de políticas públicas, em busca do "tempo perdido", devido ao período autoritário vigente no Brasil até meados da década de 80, preocuparam-se mais em avaliar do que promover sistematizações dessa natureza. O fato de existir uma vasta literatura de origem estrangeira que servia como base teórica às pesquisas realizadas é um fator adicional.

Essa realidade, ao nosso ver, deixou a preocupação com o aprofundamento teórico que levasse à conformação de um marco analítico-conceitual em "estado latente".

Já em relação ao uso da literatura estrangeira, acredito que ela deva servir como base para um aprofundamento e, a partir desse marco referencial, devam ser realizados estudos que promovam adequações ou até produzam um outro enfoque metodológico. Isso se justifica, considerando-se que a *policy making* e o contexto sócio-econômico e político-cultural dos países de onde se originam as metodologias de avaliação de políticas se diferenciam significativamente da realidade brasileira.

A constatação acerca do campo da avaliação de políticas, programas e projetos apontou para a necessidade de uma sistematização acerca do seu campo teórico-metodológico com o objetivo de realizar uma espécie de "crítica interna" à literatura, de modo a desfazer equívocos, devido à sua diversidade, e demonstrar as semelhanças ou diferenças entre os diversos autores e, assim, proporcionar um insumo metodológico útil a iniciantes na área de avaliação de políticas.

Neste sentido, e tendo conhecimento da dimensão do tema, tomou-se como ponto de partida para a construção do objeto de análise deste trabalho o texto *Avaliação política e avaliação de política: um quadro de referência teórica*, escrito por Figueiredo & Figueiredo (1986).

A escolha desse texto como estruturador do nosso trabalho deve-se ao fato de que este estudo foi um dos pioneiros no Brasil na área de avaliação de políticas públicas, com o objetivo de: estabelecer algumas conceituações sobre avaliação política e avaliação de políticas; abordar aspectos relacionados aos tipos de avaliação (processos e resultados); discutir sobre os critérios de eficácia eficiência e efetividade; fazer uma pequena retrospectiva histórica da avaliação de políticas sociais no Brasil.

Assim, foram estabelecidos como objeto de análise os seguintes aspectos relacionados com a avaliação de políticas, programas e projetos sociais e educacionais: o histórico da avaliação, o conceito, os modelos, a tipologia e os critérios utilizados para avaliar .

A partir da bibliografia citada por Figueiredo & Figueiredo (1986), foram selecionados os autores que abordassem em seus estudos os aspectos acima citados. Em relação a essa bibliografia, pode-se afirmar que a produção acerca de avaliação de política, programas e projetos, é distribuída em diferentes períodos históricos. O primeiro é a partir dos meados da década de 60 até a década de 80. Nesse período iniciou-se a construção do corpo de conhecimento da área nos países considerados centrais; onde foram desenvolvidos os estudos considerados pioneiros da área. Estes estudos em sua grande maioria são de língua inglesa, e serviram como base teórica para os estudiosos contemporâneos.

No segundo período, encontram-se os trabalhos na língua espanhola e portuguesa que, em sua grande maioria, foram desenvolvidos após a década de 80. Os autores desses trabalhos recorrem aos pressupostos teórico-metodológicos desenvolvidos pelos autores do período anterior.

Desta forma, não se teve a intenção de selecionar os autores utilizando como critério o período histórico em que as obras foram produzidas e sim os aspectos que elas abordavam. Então, para fins de análise, os estudos selecionados foram produzidos em inglês, espanhol e português e podem ser visualizados no **Quadro 1**. O motivo dessa escolha conforme foi inicialmente explicado nos parágrafos anteriores, pode será explicitado no decorrer do Capítulo I.

E, por fim, é também necessário enfatizar que a forma escolhida, para colaborar com os desafios acerca do campo teórico-metodológico da avaliação, foi através de um aprofundamento teórico, sob a forma de resenha bibliográfica.

Desta forma, procuramos apresentar no Capítulo I — *O desenvolvimento da avaliação de políticas, programas e projetos sociais e educacionais* — procuramos apresentar dentro de uma perspectiva histórica, a saber, o surgimento da avaliação de políticas, projetos e programas. O capítulo compreende duas seções, nas quais se abordam respectivamente, o processo de construção das bases conceituais e metodológicos nos Estados Unidos e o desenvolvimento da avaliação de programas educacionais no Brasil.

No Capítulo II — *Conceito de avaliação de políticas, programas e projetos sociais e educacionais* — são abordados os diversos conceitos produzidos pela literatura brasileira e estrangeira sobre avaliação. Nele procuramos demonstrar as diferenças e semelhanças entre avaliação e pesquisa avaliativa.

O Capítulo III — *Modelos de avaliação de políticas, programas e projetos* — discute como os autores classificam os diversos modelos de avaliação, procurando demonstrar as diferenças e semelhanças entre eles.

A tipologia de avaliação proposta pelos autores a partir de dois critérios de classificação, a saber, o momento em que se avalia e os aspectos que são objetos dessa avaliação, constituem o Capítulo IV, *Tipos de avaliação de políticas, programas e projetos.*

No Capítulo V — *Critérios de avaliação: eficácia, efetividade e eficiência* — além de abordarmos os aspectos relacionados a esses critérios, também procuramos apresentar as diferenças e semelhanças entre os autores que tratam do tema.

E, finalmente, as considerações finais, onde apresentamos um resumo geral dos capítulos e trazemos à tona alguns questionamentos a respeito da prática avaliativa no contexto brasileiro.

Quadro 1

Relação dos autores analisados

Autor	Nome do texto/obra
Alkin (1972)	Evaluation theory development
Aguilar & Ander - Egg (1995)	Avaliação de serviços e programas sociais
Ander – Egg (1990)	Evaluación de programas de trabajo social
Ballesteros (1996)	Evaluación de programas: una guía práctica en ámbitos sociales, educativos y de salud
BID (1997)	Evaluación: una herramienta de gestión para mejorar el desempeño de los proyectos
Briones (1998)	Evaluación de programas sociales
Cohen (1972)	Politics and research: evaluation of social action program in education
Cohen & Franco (1999)	Avaliação de projetos sociais
Faria (1998)	Avaliação de programas sociais: evoluções e tendências
Fernández (2000)	Evaluación y educación: modelos y propuestas
Figueiredo & Figueiredo (1986)	Avaliação política e avaliação de políticas: um quadro de referência teórica
Holister (1972)	The methodology of evaluation social action programs
Lara (1996)	Recogida y análisis de datos
Martinic (1997)	Diseño y evaluación de proyectos sociales: herramientas para el aprendizaje
Santacana & Benito (1996)	Evaluación de programas en educación
Saul (2000)	Avaliação emancipatória: desafio à teoria e à prática de avaliação e reformulação de currículo
Suchman (1977)	Evaluating educational program
Sulbrandt (1994)	A avaliação dos programas sociais: uma perspectiva crítica dos modelos usuais
Tenório (2000)	Avaliação de projetos comunitários: uma abordagem prática
SIEMPRO	Gestão Integral de programas sociais
Tripodi (1984)	Avaliação de programas sociais: diretrizes para administradores nos campos da educação, saúde e assistência social.
Viedma (1996)	Introducción a la evaluación
Weiss (1998) Weiss (1999)	Evaluation: methods for studying program and policies Investigación evaluativa: métodos para determinar la eficiência de los programas de acción
Worthen (1997)	Program evaluation: alternative approaches and practical guidelines

CAPÍTULO I

O desenvolvimento da avaliação de políticas, programas e projetos sociais e educacionais

Este capítulo tem por objetivo introduzir o leitor ao tema da avaliação de políticas, programas e projetos. A opção que presidiu sua elaboração foi de natureza histórica. Isto é, ao invés de iniciar pela apresentação dos conceitos e metodologias associadas ao tema (o que é feito a partir do Capítulo II), optou-se por situar o leitor ao longo do processo histórico que os originou. Esta apresentação, embora procure respeitar uma ordem cronológica, adota como critério expositivo o conceito de momento, tal como formulado por Matus (1996)[3].

[3] *"Momento"* é uma instância repetitiva, pela qual passa um processo encadeado e contínuo, que não tem princípio nem fim bem demarcados (MATUS, 1996, p. 577). O conceito não tem uma característica meramente cronológica e indica instância, ocasião, circunstância ou conjuntura, pela qual passa um processo contínuo ou em cadeia, sem começo nem fim bem definidos.

Assim, cada momento do processo de construção de que trata este capítulo é abordado de modo a indicar os acontecimentos históricos e desenvolvimentos cognitivos que, ao lado da dimensão temporal, são responsáveis pela delimitação das bases conceituais e metodológicas da avaliação de programas educacionais. Isto sem que se pretenda evitar a natural superposição temporal que ocorre entre os momentos de um dado processo.

Historicamente, a trajetória da avaliação de programas educacionais percorre caminhos traçados por concepções e práticas diversas. Seu desenvolvimento emerge de um conjunto de práticas avaliativas na área educacional e social, agregando também princípios teórico-metodológicos da pesquisa social, da avaliação educacional e da avaliação de programas curriculares (primeiro enfoque adotado para a avaliação de programas), e avaliação de programas sociais.

Segundo autores como Cohen (1972), Weiss[4] (1998), Worthen[5] (1997) e Santacana & Benito (1996), durante muito tempo não se teve clareza do que realmente constituía uma avaliação, pois não existia um corpo de conhecimento específico que com ela pudesse ser identificado.

Na realidade, a avaliação de programas educacionais e sociais, tal como é conhecida na atualidade, teve um longo período de amadurecimento, através da colaboração de profissionais de diversas áreas do conhecimento. E foi justamente a diversidade desses profissionais que

---

[4] Esta obra foi publicada pela primeira vez em 1972.
[5] Esta obra foi publicada pela primeira vez em 1987.

gerou uma variedade significativa de tipos e modelos de avaliação de programas. O que, de certa forma, reflete um conhecimento, interesse e visão de avaliação distinta e, finalmente, aponta para o fato do corpo teórico-metodológico da avaliação de programas educacionais e sociais ter sido construído gradual e coletivamente.

Além dessa diversidade na construção teórica-metodológica da avaliação de programas educacionais, cabe apontar a ausência de registros sobre estudos avaliativos nesta época. A ausência de publicações sobre o tema, mesmo levando em conta que a avaliação de programas tal como é conhecida hoje, é ainda bastante recente, dificultou a abordagem histórica que aqui se apresenta sobre o processo de construção das bases conceituais e metodológicas da avaliação de programas educacionais.

Considerando esta realidade, e com o objetivo de facilitar o entendimento dessa primeira parte deste trabalho, é importante ressaltar dois aspectos. O primeiro é que, ao longo da apresentação do processo histórico[6] de construção da avaliação de programas, faremos referência a estudos avaliativos relacionados à área de saúde e à área educacional (avaliação educacional[7] e avaliação de programas). Esta referência a estudos com interesses distintos reafirma que o surgimento e evolução do corpo de conhecimento da avaliação de programas educacionais não foi linear e sofreu influência de diversas áreas do conhecimento. O segundo aspecto é que, no decorrer dessa apresentação, citaremos alguns dos principais acontecimentos sociais e políticos que contribuíram para que a avaliação de programas educacionais e sociais tivesse um desenvolvimento como o observado.

---

[6] Temos conhecimento de que foram realizadas várias avaliações (ou algo similar) em vários países da Europa e principalmente na França que influenciaram a história da avaliação. No entanto, não temos a intenção de relatar todas as avaliações que ocorreram durante o período hitórico aqui abordado, e nos deteremos em grande parte às avaliações que ocorreram nos Estados Unidos da América.

Assim, sem a intenção de compor uma exaustiva história da avaliação de programas, este capítulo compreende duas seções. A primeira retrata a construção das bases conceituais e metodológicas da avaliação da avaliação de programas sociais e educacionais nos Estados Unidos. A retrospectiva que se apresenta nessa seção é coerente com as características da literatura internacional pesquisada: o desenvolvimento da avaliação é centrado nos Estados Unidos e a maior parte da informação proporcionada pela literatura não se refere especificamente à avaliação de programas educacionais e sim à avaliação de programas sociais e mesmo à avaliação de políticas públicas em geral[8]. Esta realidade manifesta-se também tanto na seção seguinte como ao longo deste trabalho como um todo.

A segunda seção pode ser entendida como um complemento e um contraponto à primeira, abordando o processo ocorrido no Brasil. Deter-se-á no desenvolvimento da avaliação de programas educacionais.

1.1. O processo de construção das bases conceituais e metodológicas da avaliação nos Estados Unidos

Worthen (1997) apresenta um interessante registro sobre a história da avaliação. Segundo ele, processos formais de avaliação no setor público podem ser encontrados desde 2000 a.C., na China. De fato, as autoridades chinesas faziam exames de administração civil para medir a

---

[7] Também conhecida como avaliação de aprendizagem, avaliação de rendimento.
[8] A este respeito vale ressaltar que embora o foco deste trabalho seja a avaliação de programas educacionais e que, portanto, a quase totalidade das referências que nele se faz digam respeito a este tema, algumas vezes se omite o termo educacionais.

proficiência de seus funcionários públicos. Referindo-se à avaliação na educação, ele afirma que Sócrates já utilizava a avaliação da capacidade de expressão oral de seus discípulos como parte do processo de ensino-aprendizagem.

Após um longo período histórico sem muitos registros sobre a realização de avaliações de programas educacionais e sociais, ou pelo menos com algo parecido, Weiss (1998) aponta que, na Inglaterra, em 1660, foram desenvolvidos os primeiros estudos empíricos de problemas sociais. Porém, ressalta que esses estudos, não eram chamados de avaliações.

Referindo-se também às avaliações realizadas na Inglaterra, Worthen (1997) afirma que durante o século XIX o descontentamento com os programas educacionais e sociais gerou movimentos de reforma e que, para avaliar tais programas, foram designadas comissões constituídas por membros da realeza para ouvir e analisar testemunhos.

Ainda, em relação ao século XIX, Weiss (1998) destaca que o primeiro estudo 'etiquetado' como avaliação (estudo estatístico[9]) foi realizado em 1833. Porém, só em 1844 ocorreria o registro da primeira publicação de um estudo de avaliação (avaliação de estradas e canais)[10].

Referindo-se especificamente à avaliação na área educacional, Worthen (1997) afirma que, nesse mesmo período, entre 1840 e 1846, foram realizadas em Massachusetts avaliações educacionais, através de testes impressos concernentes a assuntos diversos. Estes testes serviram para medir objetivamente o estudante e avaliar a qualidade do sistema escolar, e podem ser considerados como a primeira avaliação realizada

---

[9] Estudo realizado por A. M. Guerry, que tentava mostrar que, na Inglaterra, a educação não estava reduzindo o crime.

sobre o estudante visando servir de base para comparar diferentes instituições de ensino.

No início do século XX, foram realizadas várias avaliações na área de saúde. Podemos citar três delas que colaboraram significativamente para o desenvolvimento da avaliação de programas. A primeira foi a avaliação para credenciamento das 155 'Escolas Médicas' que operavam nos Estados Unidos e Canadá. Esse estudo teve um grande impacto, pois apenas algumas das escolas avaliadas foram qualificadas. A segunda, realizada em 1912, avaliou a qualidade da diagnose médica. Esta avaliação, com enfoque semelhante ao da primeira, avaliava individualmente a qualidade do trabalho de profissionais. A terceira, em 1914, realizou a avaliação do desempenho dos cirurgiões.

Na década de 30, nos Estados Unidos, programas para salvar a economia americana foram implementados, e agências governamentais foram instituídas para "vigiar os programas nacionais", a fim de averiguar se os recursos públicos estavam sendo bem aplicados e surtindo os resultados esperados. Visando subsidiar essa avaliação, ocorreu uma gradativa inserção de cientistas sociais nas agências governamentais, com o objetivo de desenvolver pesquisa social aplicada às áreas de interesse público tais como: saúde, educação, transportes. Então, foram realizados estudos envolvendo um grande número de variáveis, relativas aos programas sociais e educacionais desenvolvidos pelo governo americano.

Para Worthen (1997), o esforço conjunto dos pesquisadores sociais com os profissionais das agências governamentais, além de proporcionar o desenvolvimento de pesquisas sociais nessas agências, foi

---

[10] Estudo realizado por Julies Depuit. que media o valor de um projeto de canal por meio de cálculos econômicos.

um dos fatores determinantes para o desenvolvimento da avaliação de programas.

O desenvolvimento da pesquisa social aplicada teve continuidade durante e depois da Segunda Guerra Mundial. O diferencial que se observou entre o durante e o depois da guerra é que, a partir da década de 50, os programas sociais aumentaram em extensão e escala, e os investigadores sociais começaram a observar os programas como um todo e com uma freqüência maior do que antes.

Em relação a essa mudança, Worthen (1997, p.29) afirma:

> *"Com este foco cada vez mais amplo e com uma freqüência cada vez maior, os cientistas sociais passaram a ver seus trabalhos como pesquisa de avaliação (métodos de pesquisa social aplicada para melhorar um programa particular). Se estendermos a definição de avaliação para cobrir a maioria dos tipos de coleta de dados em saúde e em programas de serviço social, podemos dizer que a avaliação teve um grande incremento nessas áreas nos anos 50 e início dos anos 60*[11]*."*

Indo um pouco mais além, e de certa forma indicando o surgimento da avaliação de programas sociais e educacionais, esse autor aponta a pesquisa social aplicada como a "precursora ancestral" da avaliação, tal como ela é conhecida hoje.

No que diz respeito ao desenvolvimento da avaliação de programas, Weiss (1998), concordando com Worthen (1997), aponta que, a partir do final dos anos 50 e início dos anos 60, a nação americana tomou consciência dos problemas sociais vivenciados pelos grupos menos favorecidos da população. Para atender às demandas sociais vigentes,

---

[11] Tradução do original em inglês.

foram implementados grandes programas de combate à pobreza. E, para proporcionar informação sobre a forma como se estava efetuando o gasto público e se os programas implantados estavam alcançando os resultados esperados, foram realizadas avaliações pelas agências governamentais que promoviam esses programas.

Referindo-se à avaliação de programas educacionais e sociais, Weiss (1998) é enfática em afirmar que a guerra contra a pobreza nos anos 60 marca o começo da avaliação dos programas sociais e educacionais nos Estados Unidos.

Santacana & Benito (1996) também se referem a esse período e, indo ao encontro de Weiss (1998), indicam mais dois fatos que marcaram o desenvolvimento da avaliação de programas. O primeiro, de ordem social, é a luta em defesa dos direitos humanos, a partir do início da década de 60:

> *"... el movimiento social en defensa de los derechos humanos (hacia 1965) impulsó también el desarrollo de programas educativos y sociales que aseguraran la igualdad de oportunidades para toda la población. De ello se derivaron programas de atención socioeducativa y sanitaria a las familias de niños de alto riesgo desde la perspectiva educativa (inmigrados, minorías sociales y étnicas, etc.) y programas de intervención educativa para niños con necesidades educativas especiales.* (SANTACANA & BENITO, 1996, p.242).

O segundo fato que, na visão dos autores, teria colaborado para o desenvolvimento da avaliação de programas, foi um movimento de ordem econômica, fruto da maturidade das sociedades democráticas, em que os gastos públicos devem ser justificados ideologica e economicamente, requerendo, portanto, dispositivos de controle.

Como foi observado, essas avaliações não se fundamentavam em um corpo teórico-metodológico específico para a avaliação de programas. Esses estudos baseavam-se nos conhecimentos de pesquisa social aplicada. No entanto, a idéia de utilizar a avaliação para obter informação e tomar decisões acerca do que estava sendo avaliado já estava presente.

De forma geral, até esse período, as avaliações de programas educacionais e sociais eram realizadas esporadicamente e muitas vezes resultavam do interesses individual de estudiosos, 'curiosos' em conhecer o que estava acontecendo nos programas desenvolvidos pelas instituições. Não obstante, é indiscutível que esses primeiros estudos foram importantes para o desenvolvimento da avaliação de programas, pois foi através dos acertos e erros acerca de como 'avaliar', que teve início a construção de um corpo teórico-metodológico para a avaliação de programas sociais e educacionais.

Na área educacional, a avaliação apresentou um percurso que pode ser dividido em dois momentos. No primeiro, a avaliação centrou-se no rendimento do aluno. No segundo, a avaliação preocupou-se com os programas curriculares e, gradativamente, foi direcionando seu foco para aspectos mais globais dos programas educativos de natureza administrativa, financeira e organizacional etc., embora a avaliação do rendimento do aluno continuasse sendo utilizada para avaliar esses programas. É importante salientar que estes momentos, que a seguir se aborda separadamente em detalhe, não são estanques mas se interpenetram no decorrer da construção do corpo teórico-metodológico da avaliação de programas.

1.1.1. Primeiro momento: o enfoque na avaliação do rendimento do aluno

Segundo a literatura, sobre a área educacional, a avaliação centrou-se primeiramente no rendimento do aluno. Nesse momento[12] em que o processo avaliativo concentrava-se na mensuração das habilidades individuais do aluno, começa a ser difundida a idéia de 'medida educacional[13]'. Como exemplo, podemos indicar que as provas educacionais como tecnologia de medida já haviam sido utilizadas por Thorndike e, mesmo antes de 1918, já teria nascido à prova objetiva, a qual penetrou rapidamente em todos os níveis de ensino.

Na década de 20 continuou o movimento científico de medição[14], "*que se centra en la invención de artefactos y técnicas para medir características de personalidad de los educandos*" (MÉNDEZ, 2001, p. 29), através dos 'testes com padrão de referência[15]' para medir os níveis de desempenho individual. Entretanto, poucos estudos foram publicados nessa época.

Ao longo do processo de consolidação dos "testes para medição", ocorre na década de 30, nos Estados Unidos, a padronização dos "testes de padrão de referência" e a sua transformação em um grande

---

[12]Não é nosso objetivo discutir os conceitos, os aspectos políticos-ideológicos da avaliação do rendimento do aluno ou mesmo algo similar. Apenas citamos esse enfoque para situar historicamente o processo evolutivo pelo qual passou a avaliação de programas educacionais.

[13] Não entraremos aqui na discussão acerca dos conceitos de medida, avaliação de rendimento ou avaliação de aprendizagem.

[14] Tradução do termo original em espanhol, *movimiento científico de medición* em MÉNDEZ (2001).

empreendimento comercial. Segundo Worthen (1997), os pedagogos deste período consideravam medida e avaliação quase como sinônimos.

Em relação aos testes de padrão de referência também conhecidos como testes padronizados, Dias Sobrinho (2002, p.20) afirma:

> *"Até 1930, caracterizou-se sobretudo como tecnologia dos testes padronizados e objetivos com a finalidade de medir a inteligência e o desempenho (testing period). No começo do século foram sistematizados os testes de inteligência e se tornaram conhecidas muitas mensurações pisicofísicas, tais como as capacidades sensoriais, os tempos de reação ou de realização de tarefas. A avaliação e medida eram termos com noções intercambiáveis ou equivalentes."*

Até aqui, a preocupação dos estudos avaliativos centrava-se no indivíduo enquanto receptor de conhecimentos através do ensino-aprendizagem. A preocupação com programas curriculares e seus resultados, emerge a partir da década de 30. Foi uma transição gradativa, demarcada por acontecimentos sociais, políticos e econômicos.

Dias Sobrinho (2002, p.20) destaca um acontecimento que contribuiu para o desenvolvimento da avaliação de programas na área educacional. Segundo esse autor,

> *"...depois de 30, os testes passaram a buscar mais propriamente medir aquisição a aquisição dos programas. A educação deveria participar mais vinculada e eficazmente dos esforços da recuperação econômica e da expansão da industria. Era necessário pois, formular mais claramente os programas educacionais e medir o seu cumprimento".*

---

[15] Tradução do termo original em inglês, "*norm-referenced test*".

Resumindo as idéias de outros autores, Santacana & Benito (1996) apontam que, no período que abrange as décadas de 30 e 40, dois acontecimentos contribuíram para o desenvolvimento da avaliação de programas na área educacional. O primeiro é o processo de "credenciamento[16]" através das avaliações de escolas e programas curriculares, realizadas pelas agências governamentais nos Estados Unidos durante os anos 30. Para tanto, foram criados critérios de avaliação que possibilitavam a comparação entre instituições educativas e que produziam informação objetiva para a tomada de decisão.

O segundo acontecimento ocorreu na mesma época e contribuiu para uma paulatina mudança do enfoque avaliativo na área educacional. Após a reivindicação de uma associação de pais por uma educação mais progressista do que a existente até então, foi desencadeado um estudo que durou oito anos (1932-1940). Este estudo foi realizado por Tyler e resultou em um modelo chamado de Avaliação Orientada por Objetivos[17].

Em relação a esse momento histórico, que delimita a primeira grande mudança no enfoque da avaliação na área educacional, Santacana & Benito (1996, p.241) são enfáticos ao afirmarem que:

> *"Aunque, hasta los años 30 no puede hablarse propiamente de evaluación de programas educativos, es obvio que en etapas anteriores se había realizado algún tipo de evaluación de acciones educativas. Todos los sistemas de control de los resultados educativos (exámenes de alumnos, exámenes de la calidad del profesorado, valoración general de los sistemas educativos, críticas a las estructuras educativas vigentes o "libros brancos", propuestas de cambios de planes educativos, etc.) realizados en busca de una mayor eficiencia educativa*

---

[16] Tradução do termo espanhol "*acreditación*", em SANTACANA & BENITO (1996) e do termo em inglês '*accreditation*', em WEISS (1998,1999) e WORTHEN (1997) e outros.
[17] Tradução de "*objectives-oriented evaluation approaches*", em WORTHEN (1997)

> *pueden ser considerados como aspectos de la evaluación de programas. Pero la evaluación de programas educativos, como conjunto organizado de acciones evaluativas que conducen a la formulación de un juicio de valor acerca de un programa y a la toma de decisión acerca de los cambios necesarios en el mismo a fin de incrementar su eficacia, no se produce hasta pasado el primer tercio de* siglo, y *tiene como pionero indiscutible a Tyler (1942)".*

Worthen (1997) e Weiss (1998), concordando com Santacana & Benito (1996), também destacam o estudo de Tyler[18], como responsável por uma significativa mudança na avaliação na área educacional. Para eles, o "enfoque de avaliação orientada por objetivos[19]", estabeleceu um novo viés. Ela, antes exclusivamente centrada nos processos individuais de aprendizagem, passou a interessar-se pelo o exame de programas curriculares, através da avaliação do nível em que os objetivos educacionais eram atingidos pelos alunos. O estudo de Tyler, portanto, parece ter estabelecido um novo padrão para avaliação na área educacional, pois, de certa forma, vinculou o processo educativo a objetivos preestabelecidos que deveriam ser atingidos pelos alunos.

Segundo Dias Sobrinho (2002), essa associação entre avaliação e objetivos, que aparentemente representa uma mudança de enfoque na área de avaliação, é a confirmação da educação como instrumento preponderante no processo de desenvolvimento econômico nos Estados Unidos. Associa-se a isso a ciência positivista vigente na época, que creditava à medição um caráter experimental e científico.

---

[18] Para maiores detalhes ver: WORTHEN (1997), SAUL (2000) e MÉNDEZ (2001).
[19] Tradução de "*objectives-oriented evaluation approaches*".

Desta forma, segundo Dias Sobrinho (2002, p.21) acreditava-se que a avaliação

> *"...deveria averiguar o quanto os estudantes, individual e coletivamente, conseguiam demonstrar, isto é, segundo a concepção de educação dominante, que as mudanças de comportamento se podiam observar ao final de um determinado processo, tendo em vista os objetivos estabelecidos. A avaliação deveria dedicar-se ao êxito na escolarização. Comprometia-se com a ideologia de eficiência social. Os objetivos educacionais eram estabelecidos em função do desenvolvimento industrial e com o esforço de recuperação da economia americana. Continua sendo fundamentalmente um processo de medida, porém, agora, mais preocupada com a adequação dos currículos às exigências sociais e econômicas, com a formulação de objetivos e a medida de seu cumprimento."*

Segundo Worthen (1997), além do trabalho desenvolvido por Tyler, outros acontecimentos devem ser considerados. Ele destaca, entre outros, a criação de fundações para avaliar a administração pública e o setor privado, o desenvolvimento de pesquisas internamente às agências responsáveis pelas políticas sociais nos Estados Unidos e a incorporação de cientistas sociais no 'staff' das agencias governamentais.

A importante contribuição de Tyler, que estabeleceu, definitivamente, a relação entre os resultados de aprendizagem desejável e aprendizagem real do aluno, não foi, contudo suficiente, para que a avaliação como processo de medida tivesse seu prestígio abalado. Ao contrário, a "medida de aprendizagem" que ele propunha seria utilizada mais tarde como um importante insumo para avaliar programas educacionais. Nesse período, segundo Dias Sobrinho (2002, p. 21), a avaliação

> *"...ganha um sentido mais operativo: ela passa a averiguar até que ponto os currículos e as práticas pedagógicas estão atingindo os objetivos de levar a escola a ser eficaz e alcançar as metas de eficiência que a economia exigia. Portanto, torna-se evidente a sua racionalidade instrumental. Ainda que permanecesse muito centrado nas escolas e sobretudo no eixo ensino aprendizagem, seu campo se ampliou na medida em que passou a articular os rendimentos dos alunos às questões de currículos e programas. Sua importância também cresceu como mecanismo de controle e seleção, segundo a lógica utilitarista que predominava, bem como aceitação de sua fidelidade e cientificidade, graças ao aperfeiçoamento dos testes e medidas".*

Durante esse momento histórico, pode-se perceber que, ao mesmo tempo em que a avaliação de programas educacionais ia se intensificando, também se "consolidavam as provas unificadas", o credenciamento das instituições de ensino, a avaliação por objetivos. Uma considerável mudança no enfoque da avaliação realmente só ocorreria mais tarde. É exatamente sobre essa mudança de enfoque que a seção seguinte se deterá.

1.1.2. Segundo momento: o enfoque nos programas curriculares

Apesar da mudança que desencadeou o estudo de Tyler, a prática da avaliação do rendimento do aluno, que se iniciou no começo do século XX, quando se desenvolveram instrumentos que possibilitavam medir o desempenho do aluno, e que se encontrava amplamente aceita na década de 40, continuou a ser utilizada na década de 50 e 60.

Um diferencial nesse período é que a preocupação, antes voltada basicamente para os instrumentos de medição, foi direcionada para o ensino aos pedagogos de taxionomias[20] que permitissem declarar objetivos educativos em condições explícitas e mensuráveis. A disseminação das taxionomias subsidiou a criação de "instrumentos de medição" e, de fato, tornou a "avaliação dos processos individuais de aprendizagem" ainda mais observáveis e mensuráveis objetivamente.

Santacana & Benito (1996) apontam um fato histórico que contribuiu para que fosse desencadeado o processo de mudança de enfoque da avaliação de programas educacionais, a partir do final da década de 50. Segundo eles a conquista do espaço pelos soviéticos desencadeou nos Estados Unidos uma série de mudanças no âmbito educacional. E, através do "*National Defense Education Act/1958*", as instituições educacionais incrementaram a formação técnico-científica nas escolas. Isto teria levado, segundo os autores, a um importante desdobramento:

> *"El movimiento del desarrollo curricular que se había iniciado en los años 40 con Tyler, se vio especialmente potenciado por el interés de programar y definir los curricula que pudieran desarrollar al máximo las capacidades cognitivas de los alumnos"* (SANTACANA & BENITO, 1996, p. 242).

Os vultosos recursos investidos pelos Estados Unidos, com o objetivo de realizar mudanças nos currículos, a fim de melhorar a capacidade técnico-científica dos alunos, fizeram com que os *policy makers*

---

[20] Como exemplo das taxionomias difundidas na época, Worthen (1997) aponta: a Taxionomia de domínio cognitivo (Bloom, 1956) e a Taxionomia domínio afetivo (Krathwohl, 1964).

passassem a se preocupar seriamente com a avaliação dos programas curriculares implementados pelas instituições educacionais.

Segundo Worthen (1997, p. 30), no entanto, os estudos realizados para avaliar essas mudanças curriculares

> *"...revelaram um empobrecimento conceitual e metodológico da avaliação. Planos inadequados e relatórios irrelevantes foram só alguns dos problemas observados. A maioria dos estudos dependeu e por isto importaram conceitos da pesquisa social e comportamental. E as técnicas, que eram apropriadas para tais pesquisas, não eram satisfatórias para a avaliação de programas escolares. Trabalhos teóricos que se relacionassem diretamente à avaliação (ao invés de pesquisa) não existiam e ficou evidente que mesmo o melhor método da pesquisa social comportamental não proporcionava orientação sobre como levar a cabo muitos dos aspectos envolvidos com a avaliação[21]".*

Isso, entretanto, não levou ao surgimento de um movimento forte o suficiente para proporcionar o desenvolvimento conceitual e metodológico da avaliação de programas educacionais. Tanto que, durante algum tempo, acreditou-se inclusive que os conceitos e práticas desenvolvidos no setor privado norte-americano, poderiam ser transportados e adaptados para avaliar os programas federais.

É nesse cenário que se configura o segundo momento da avaliação na área educacional, indicado anteriormente: aquele em que surge a preocupação com programas curriculares, sendo a avaliação do rendimento do aluno utilizada para avaliar estes programas, e tendo como enfoque avaliativo os aspectos mais globais de natureza administrativa, financeira e organizacional dos programas educacionais.

---

[21] Tradução do original em inglês.

Considerando esta realidade, podemos citar um acontecimento político que, de maneira muito particular, impulsionou o desenvolvimento e a realização da avaliação de programas, a partir do início dos anos 60. Trata-se da introdução, no setor público, *do Planning, Programming and Budgeting System (PPBS)*[22] utilizado pela Ford Motor Company. Esse sistema foi primeiramente introduzido no Departamento de Defesa dos Estados Unidos por Robert Mcnamara. E, gradativamente, até o final da década de 60, todas as agências governamentais já o tinham implantado; inclusive as universidades e sistemas escolares (pelo menos "no papel" [23]).

Segundo Worthen (1997, p.31), "*o emprego do PPBS visava melhorar a efetividade e eficiência das agências federais que administravam os programas de pobreza, definir objetivos organizacionais e ajudar nas decisões de alocação do orçamento*[24]".

Esta preocupação com controle e orçamento, proveniente do setor privado, alastra-se por todas as agências governamentais e, nesse período, prolifera a idéia de "custo-benefício" dos programas. Gradualmente, através de mudanças realizadas dentro das próprias agencias governamentais, o Sistema de Planejamento, Programação e Orçamento[25] foi transformando-se no que veio a se denominar 'administração por objetivos'.

---

[22] Tradução do termo Planning, Programming and Budgeting System - Sistema de Planejamento Programação e Orçamento. Para aprofundamento ver WORTHEN (1997) e WEISS (1998).

[23] Mesmo tendo sido deliberado pelo presidente Johnson que a implementação do *PPBS* deveria ocorrer em todos os departamentos federais dos Estados Unidos, existia uma grande resistência por parte dos funcionários públicos. Devido a isso, em muitos casos o sistema ficou apenas "no papel".

[24] Tradução do texto original em inglês.

[25] Tradução do texto original em inglês "*Planning, Programming and Budgeting System*".

Referindo-se à área educacional, Worthen (1997) indica um outro acontecimento que colaborou para a difusão da avaliação de programas educacionais, durante a década de 60, nos Estados Unidos. Para ele, o projeto *Elementary end Secondary Education Act/1965 (ESEA/ 1965)*, do senador Kennedy, "*ativaria o evento que enviaria uma onda de choque pelo sistema de ensino norte-americano, despertando nos fazedores de política e profissionais a importância da avaliação sistemática*[26]." (WORTHEN, 1997, p.32).

Segundo Worthen (1997), o grande investimento de recursos no *ESEA*[27]*/1965* estabeleceu o nascimento da avaliação de programas contemporânea. Ou seja, a avaliação de programas contemporânea teria surgido da necessidade de realizar avaliações contínuas com o objetivo de controlar os gastos e informar se o dinheiro investido estava sendo aplicado adequadamente e se as alterações neste sistema estavam surtindo os efeitos esperados.

Indo ao encontro de Worthen (1997), Cohen (1972, p.137) afirma que:

> *"Antes de 1964 os objetos de avaliação em educação consistiam quase que exclusivamente em programas pequenos. Os avaliadores se interessavam com o desenvolvimento de currículo ou treinamento de professores. A avaliação geralmente acontecia em uma única escola ou distrito escolar e buscava produzir mudança educacional em local limitado, com orçamentos modestos e com muito pouco pessoal para pesquisa"*[28].

---

[26] Tradução do texto original em inglês.
[27] *Elementary end Secondary Education Act/ 1965.*
[28] Tradução do texto original em inglês.

Confirmando a posição de Worthen (1997) e Cohen (1972), em relação à mudança de enfoque da avaliação na área educacional, iniciada a partir do *ESEA / 1965*, Levy (1979, apud SAUL, 2000, p. 33) afirma:

> *"A expansão das atividades de desenvolvimento de currículo acentuou a necessidade de avaliar os programas educacionais. Tanto as agências financiadoras quanto os consumidores desejavam evidências de que os novos programas produziam resultados satisfatórios. As informações levantadas estavam relacionadas à relevância do novo programa para as necessidades da sociedade e do aluno, à significância e validade científica dos novos materiais de estudo, à capacidade de o programa eliciar certos comportamentos dos professores e alunos e aos resultados reais obtidos pela utilização de determinado conjunto de materiais instrucionais".*

Ristoff (2000, p. 115-116), concordando com Levy (1979), resume a análise da literatura realizada por Nevo (1979), para confirmar que a mudança de enfoque na área educacional teria ocorrido a partir de 1965. Para tanto, afirma que

> *"...professores e alunos sempre foram os objetos preferidos da avaliação educacional e quase toda a literatura sobre mensuração e avaliação nesta área até meados dos anos sessenta, tratava da avaliação da aprendizagem pelos alunos. Até aquela época era quase impossível encontrar alguma orientação substancial referente à avaliação de outros objetos, tais como projetos ou programas, materiais curriculares, ou de instituições educacionais como um todo. Só a partir de 1965 é que vários desenvolvimentos nos sistemas educacionais levaram a uma mudança significativa de enfoque. A partir daí, os objetos de avaliação educacional perderam a sua concentração sobre alunos e professores para contemplar projetos, programas e materiais instrucionais."*

Também, confirmando a importância do *ESEA/1965* para a mudança de enfoque da avaliação, Dias Sobrinho (2002, p. 23) afirma que:

> *"A lei sobre a educação primária e secundária promulgada pelo presidente Lyndon Johson em 1965 deu ênfase aos problemas dos estudantes que chegavam à escola com desvantagens sociais educacionais, em geral oriundos de famílias pobres e marginalizadas. Por proposta do senador Robert Kenedy, fez-se obrigatória uma avaliação dos programas especais destinados a esses alunos. Em seguida a avaliação se tornou obrigatória para todos os programas sociais e educativos norte-americanos".*

Além de indicar a necessidade de avaliar outros objetos do programa, o projeto ESEA/1965, exigiu que administradores escolares, psicólogos e pedagogos realizassem avaliações da prática de suas instituições escolares e autoavaliações. Isso revelou o despreparo desses profissionais para realizar avaliações e fez com que os administradores públicos incorporassem gradativamente ao 'staff' de avaliadores das agências governamentais profissionais de outras áreas do conhecimento.

Assim, a crescente demanda pela realização de avaliações nas áreas educacional e social, verificada a partir da década de 60, e a ausência de um corpo teórico-metodológico específico para avaliar os programas fizeram com que os avaliadores

> *"...utilizassem teorias provenientes de disciplinas afins, e até a compilar o que eles pudessem das metodologias melhor desenvolvidas como: desenho experimental e etnografia. O fato dos resultados estarem aquém do esperado, sublinhava a necessidade de desenvolver novas concepções e métodos mais adaptados às necessidades dos avaliadores"*[29] (WORTHEN, 1997, p. 38).

---

[29] Tradução do original em inglês.

A incorporação de profissionais de outras áreas e a utilização de outras técnicas e metodologias na prática avaliativa da época não foram suficientes para avaliar os projetos ou programas e apontaram para a necessidade de "formar" os especialistas em avaliação. Surgem, então, nesse período, cursos nos diversos níveis de ensino, com o objetivo de "treinar" avaliadores aptos para suprir as necessidades das agências governamentais. Esses cursos, segundo Worthen (1997, p.40)

> *"...já não eram apenas orientados á concepção de projeto de pesquisa, estatística, e sistema de medidas. E a psicologia já não era a principal disciplina que proporcionava as habilidades técnicas necessárias. Essas habilidades passaram a ser buscadas também em áreas como filosofia, sociologia, antropologia, ciência política, economia"*[30].

Esses cursos possibilitaram a interseção[31] disciplinar entre conhecimentos provenientes de diferentes áreas e passaram a influenciar as práticas avaliativas até a atualidade.

Nesse processo, e como resposta à necessidade emergente de produzir um corpo teórico-metodológico mais específico para a avaliação de programas educacionais, é que, a partir dos meados da década de 60, começam a surgir as primeiras contribuições para esse novo campo. Podemos citar alguns autores[32]: Cronbarch, Scriven, Stake, Stufflebean, Parlet & Hamilton, que se dedicaram a abordar múltiplos aspectos dos programas curriculares, utilizando-se de enfoques, instrumentos e técnicas variadas.

---

[30] Tradução do texto original em inglês.
[31] Trataremos dessa questão quando abordarmos o conceito de avaliação de programas sociais e educacionais.
[32] Os autores citados desenvolveram seus estudos na área de avaliação de programas curriculares. Com o passar dos anos, seus fundamentos foram incorporados pela

Segundo Dias Sobrinho (2002, p. 27), de forma geral, a avaliação desenvolvida nos Estados Unidos entre a década de 60 e a década de 80,

> *"...aproveitou-se da crença liberal segundo a qual os diversos problemas e dificuldades de grupos humanos, corretamente identificados e iluminados pelas ciências sociais, poderiam ser resolvidos através de adequadas políticas alimentadas por generosas inversões e dinheiro. A avaliação determinava a eficácia dessas políticas, preferencialmente quantificando os resultados".*

Os anos 70 são considerados por Weiss (1998) o ponto alto da história da avaliação de programas nos Estados Unidos. Nesse período as avaliações tornaram-se comum em todas as agências federais e, no final dessa década, quase todo o departamento da administração pública possuía um setor específico para realizar avaliações. Agências de financiamento e desenvolvimento supranacionais, como o Banco Interamericano de Desenvolvimento/BID e o Banco Mundial, seguiram a mesma tendência.

Conforme foi possível observar pela análise da literatura, durante as décadas de 60 e 70, ocorreram acontecimentos políticos, econômicos e sociais que contribuíram para a mudança de enfoque da avaliação, além de proporcionarem um gradativo amadurecimento do corpo de conhecimento da avaliação de políticas sociais e educacionais.

Esse processo, continuou na década de 80, quando a mudança ocorrida na forma de conduzir as políticas de cunho social

---

avaliação de programas sociais e educacionais, e é muito comum encontrá-los como referência na literatura especializada.

influenciou significativamente o processo avaliativo, estabelecendo uma outra perspectiva para essa prática. Em relação a essa realidade, em que se verifica na avaliação uma mudança de foco, Dias Sobrinho (2002, p. 28) afirma:

> *"A mudança de paradigma do Estado do Bem – Estar, para o neoliberalismo, engendrou um tipo novo de Estado Avaliador. Em outras palavras, o Estado deixou de ser o provedor de benefícios e serviços que a sociedade utilizava para superar seus problemas e passou a exercer severo controle e forte fiscalização. Observa-se aí uma transferência de ênfase. No primeiro caso, a avaliação tinha o propósito de analisar a eficácia dos programas com a finalidade de torná-los melhores e mais produtivos em termo sociais. No segundo prevalece a lógica do controle da racionalidade orçamentária, que efetivamente significa cortes de financiamento e rebaixamento da fé pública".*

Em relação à posição do Estado frente às políticas educativas e a mudança do processo avaliativo, Afonso (2000) e Dias Sobrinho (2002), afirmam que durante a década de 80, nasce um Estado Avaliador, em que a avaliação passa a ser vista como uma prestação de contas e controle, adquirindo um modelo de responsabilização (*accountability)*[33], baseado na lógica do mercado de trabalho. Nesse sentido, a avaliação recupera a epistemologia do positivismo e incorpora o modelo de gestão do setor privado, enfatizando produtos ou resultados do sistema educativo. Desta forma, a avaliação, segundo Dias Sobrinho (2002, p. 29),

---

[33] "*Accountability*: uma forma tecnocrática de valorar um procedimento burocrático de exigir o cumprimento de obrigações. É inevitável a conexão entre a *accountability* e a ideologia da eficiência. A responsabilidade, antes entendida nos âmbitos universitários como pertinência e eqüidade, ou em outras palavras, a prestação de contas à sociedade como um todo, referida como *accountability*, se transforma na exigência de demonstração da obtenção de determinados resultados através do emprego dos meios mais eficientes.É portanto a capacidade de prestar contas não à sociedade, mas aos governos e clientes." (DIAS SOBRINHO, 2002, p.29). Maior detalhes ver AFONSO (2000) e DIAS SOBRINHO (2002).

> *"...passou a ser um mecanismo fundamental dos governos nos seus esforços obsessivos de implantação de uma estrita cultura gerencialista e fiscalizadora (Henkel, 1991:121). Para o cumprimento dessa missão mais geral se tornou fundamental" colocar as escolas dentro de uma economia de mercado"(Beerstein,1990:142). Ganham grande proeminência os exames nacionais a revisar os padrões igualmente nacionais e estabelecer as classificações a servir informações fidedignas aos clientes e ao governo. Assim, a avaliação se transforma em "uma tecnologia de vigilância e de controle sobre os rendimentos das escolas e dos professores"(Eliott,1992:56), uma ferramenta para informar e legitimar as mediadas impopulares que os governos tiveram que tomar e que freqüentemente significam cortes de orçamento"(Hause, 1993:53), "uma estratégia estatal" que "forma parte da regulação do estado, controle e governo do estado (e) cumpre fins de política"(Pokewitz, 1992: 95-97)".*

De fato, vários especialistas do tema afirmam que durante o período das décadas de 60 e 80 foi bastante intenso o desenvolvimento do corpo teórico-metodológico da avaliação de programas sociais e educacionais. Muitos dos conceitos e metodologias desenvolvidos até então serviram como base para o desenvolvimento e a construção de outros pressupostos na América Latina e no Brasil.

É justamente o desenvolvimento da avaliação de políticas, programas e projetos no Brasil, que será abordado na próxima seção deste capítulo. Neste sentido, é necessário ressaltar também que trataremos alguns dos aspectos relacionados à avaliação na América Latina, por questões que parecem óbvias.

1.2. A avaliação de programas educacionais no Brasil

No Brasil, segundo alguns autores, a avaliação de políticas, programas e projetos educacionais é ainda muito recente. Os registros sobre essas avaliações no contexto brasileiro data a partir da década de 60. Como exemplo dessa constatação pode-se citar algumas práticas avaliativas: o Programa de Avaliação Somativa no Rio de Janeiro, desenvolvido pela Fundação Getúlio Vargas, na década de 60, com o objetivo de avaliar a capacitação de crianças ao término do 1º grau e a avaliação desenvolvida pela FUNBEC - Fundação Brasileira para o Ensino de Ciências de alguns de seus currículos.

Já na década de 70, segundo Saul (2000), foram desenvolvidas avaliações de programas curriculares. Nesta mesma década, foram realizadas, segundo Vianna (2002), avaliações do rendimento escolar, do Programa de Expansão e Melhoria do Ensino no meio Rural do Nordeste Brasileiro – EDURURAL.

Mas a literatura também sinaliza que essas avaliações ocorriam de forma assistemática e não faziam parte da agenda governamental como um sistema articulado e sistemático.

A mudança dessa realidade no contexto brasileiro iniciou-se com a crise dos anos 80 vivenciada pela América Latina e consolidou-se a partir dos meados da década de 80, com fim do período autoritário vigente até então.

Em relação à crise dos anos 80, pode-se afirmar que dentro do contexto sócio-político e econômico da época, em que as necessidades da população aumentavam e a quantidade de recursos financeiros diminuía, a avaliação de políticas passou a ser no âmbito governamental um instrumento para averiguar se os recursos públicos estavam sendo adequadamente utilizados, devido à redução de gastos no setor social.

Cohen & Franco (1999) afirmam que a redução de recursos para a área social forçou a racionalização, exigindo maior eficiência, eficácia e efetividade na aplicação dos recursos públicos. Segundo esses autores:

> *"A crise que afetou a região latino-americana nos anos oitenta reduziu os recursos disponíveis ao mesmo tempo em que agravou as necessidades da população. Então passou a ser necessário que as políticas sociais fossem transformadas em uma rede de segurança que impedisse a deterioração e, se possível ajudasse na recuperação dos níveis de vida da população"*. (COHEN & FRANCO, 1999, p.13-14).

Para Martinic (1997, p.11), o surgimento da avaliação sistemática de programas sociais e educacionais na América Latina está também associado à crise dos anos 80. Desde então, a

> *"América Latina vive importantes transformaciones sociales. Entre ellas, cabe destacar la reinserción de las economías en los mercados internacionales; un creciente aumento de relaciones entre los países a través de los acuerdos regionales; la consolidación de los sistemas democráticos y la transformación de los estados centralizados y de bienestar en estados más reducidos, descentralizados, que abren espacios para el funcionamiento del mercado y la autorregulación de los grupos en la vida social".*

Assim, segundo Martinic (1997), como conseqüência das mudanças que se verificavam, o estado, que desenvolvia suas atividades no âmbito social e educacional utilizando-se do 'enfoque universal' para atender as demandas, passa a utilizar-se do 'enfoque focalizado[34]'. Isto é, ao invés de tentar "*disminuir la desigualdad social, mejorar la distribución de los beneficios y reducir la pobreza, mediante una acción dirigida al conjunto de la sociedad*" o Estado latino-americano passa a buscar "*los mismos objetivos pero a través de la asignación de recursos a la población que, comparativamente, está más afectada*" (CIENES, 1995, p.1, apud MARTINIC, 1997, p.12).

Este novo enfoque, que surge no contexto de uma drástica redução do gasto social, procura fazer com que o Estado obtenha maior eficácia na gestão pública e aumente sua capacidade para atingir a sua população alvo. Como reflexo de sua adoção começam a verificar-se grandes reformas nos sistemas educativos, na distribuição de recursos e, em conseqüência, uma crescente necessidade de avaliar[35] de forma mais sistemática.

---

[34] Este enfoque surge en el contexto de una drástica disminución del gasto social donde el estado debe garantizar una mayor eficacia en cuanto a la gestión y capacidad para llegar a efectivamente a su población objetivo." (MARTINIC, 1997, p. 12).

[35] No entanto, cada país da América Latina, vivenciou essas mudanças de forma diferente e como conseqüência, desenvolveram seus sistemas de avaliação com concepções distintas e definidas por Lei. Entre os países da América Latina que implementaram algum tipo de sistema de avaliação da educação[35], podemos citar a Argentina, Colômbia, Chile, Equador, Paraguai, Peru, Uruguai e Venezuela. Nesses países a avaliação é considerada um instrumento de supervisão educativa e seus objetivos: "se pueden resumir de la manera seguinte: la oficina de supervisión coordina, evalúa y mantiene un seguimiento del sistema educativo promoviendo instrumentos e indicadores de evaluación y control. La evaluación de la efectividad y eficiência de los sistemas educativos tiene que tener en cuenta el uso de los recursos financieros para desarollar metodologías de análisis, instrumentos e indicadores de evaluación". (PÉREZ, 1998, p.70)

Ainda em relação ao surgimento da demanda de avaliação, que ocorreu na América Latina a partir da década de 80, podendo também ser atribuída ao caso brasileiro, pode-se apontar três situações que Wiesner (2000) utilizando-se dos pressupostos de Guerrero (1999), afirma terem contribuído para que isso tenha ocorrido e que de certa forma confirma os fatores políticos e econômicos, apontados anteriormente.

> *"...la demanda de evaluación surge de tres situaciones: i) al enfrentar correcciones fiscales, los países tienen que evaluar bien dónde hacen recortes y donde mantienen recursos; ii) frente al fracasso de algumas políticas o algunos proyectos, los gobiernos advierten que no puedem permanecer passivos si no quieren pagar un alto precio político, y iii) la sociedad civil sé há vuelto en una mayor demanda real de evaluaciones y del desarrollo da la capacidad de exigirles rendición de cuentas a los gobiernos"*(WIESNER, 2000, p.16).

Especificamente, em relação ao caso brasileiro, verifica-se que mesmo diante das exigências imposta pela crise, só a partir do fim do período autoritário vigente até os meados da década de 80, a avaliação de políticas, programas e projetos sociais e educacionais foi, de fato, incorporada à agenda governamental brasileira.

Em relação à ausência da avaliação de políticas, programas e projetos, como atividade regular no âmbito do setor público no período autoritário, Lobo (1998) destaca dois fatores que teriam contribuído para que isso ocorresse: a inexistência de uma cultura de avaliação, devido ao constrangimento de que a administração pública autoritária poderia ser alvo ao ter suas políticas avaliadas; as dificuldades burocráticas que são colocadas, sob a forma de falta de apoio financeiro, político e institucional, para a realização de avaliações.

Já no que diz respeito à opressão exercida pelo governo na sociedade civil (instituições, organizações, etc.), a literatura indica que o regime autoritário, que se manteve até meados da década de 80, além de não ter interesse em avaliar suas ações, também obstaculizava a realização de avaliações por parte da sociedade civil, o que teria levado a uma virtual ausência da avaliação de políticas públicas até o início dos anos 90 e de certa forma explicaria o fato de esta área ser recentemente implantada no País. Ou seja, só após a abertura política no Brasil, foi "permitido" à sociedade acompanhar as ações governamentais.

O processo de abertura econômica e política que se seguiu ao período autoritário fez emergir as instituições acadêmicas como atores importantes no campo da avaliação. Foram criados núcleos de estudos de políticas públicas nas áreas de saúde, educação, etc., que, através de suas pesquisas, começaram a realizar avaliações das ações governamentais. O desenvolvimento desses núcleos parece ter sido permeado pela idéia de que a avaliação é um "*instrumento democrático de controle e monitoramento da ação estatal*" (ARRETCHE, 1998).

Ainda em relação ao fim do período autoritário, e ao maior desenvolvimento da avaliação no contexto brasileiro, mas referindo-se a área social, Canesqui (2000, p.91) afirma:

> *"O crescente interesse na avaliação de políticas e programas sociais no Brasil incrementou-se com a restauração da democracia e os processos de reforma das políticas sociais, desde a segunda metade da década de 80, introduzindo-se paulatinamente na agenda pública o necessário aperfeiçoamento das políticas e programas sociais nas suas operações e resultados, para torná-las mais efetivas no atendimento às necessidades de seus beneficiários".*

Referindo-se especificamente à área educacional, Pestana (1998, p. 55) aponta três razões "*que explicam a eclosão de investimentos e propostas na área de avaliação*", no contexto brasileiro. São elas: ênfase na qualidade do ensino, políticas de descentralização e pressão política e social (conhecido nos países anglo-saxãos como *accountability*) [36].

Ainda, sobre a incorporação da avaliação na agenda governamental brasileira, a partir da década de 90, Brown (2001, p. 10) afirma:

> *"El interese en la evaluación entre los países de desarrollo fue fomentado en gran parte por el Departamento de Evaluación de Operaciones del Banco Mundial, que intentó en los primeros años del decenio de 1990 crear la capacidad de evaluación ampliado su Sistema de evaluación de resultados de la gestión. En virtud de este programa, el Banco Mundial asistió al comienzo a seis países (entre ellos Brasil, Colombia Chile e India) y a un banco regional de desarrollo para ampliar su capacidad de evaluación".*

Portanto, é possível aceitar a afirmação de que, no contexto brasileiro, a disseminação da necessidade de avaliar as políticas, programas e projetos, além dos fatores como redução do gasto na área social e educacional e o estabelecimento de padrões globalizantes de relacionamento internacional está associado às exigências de organismos supranacionais que financiam tais atividades.

Esses fatores, portanto, além de proporcionarem ao País uma "flexibilização" econômica, também serviram para "enquadrá-lo" em uma dinâmica de racionalização, o que levou à utilização de critérios de eficácia, eficiência e efetividade na alocação e no emprego dos recursos

---

[36] Para maiores detalhes dessas razões, ver PESTANA (1998).

financeiros e passou a ter como princípio básico, o controle de gastos e avaliação de resultados.

Desta forma, afirma-se que foi a partir da década de 90, que a avaliação das políticas, programas e projetos educacionais foi de fato incorporada à agenda governamental brasileira, sob a forma de política, através dos programas nacionais de avaliação da educação em todos os seus âmbitos.

Como exemplo dessa constatação, pode-se citar algumas avaliações[37] (de políticas educacionais, sistemas educativos, programas educacionais etc.), desenvolvidos pelo Ministério da Educação e Cultura (MEC): o Sistema Nacional de Educação Básica – SAEB, Exame Nacional de Cursos – ENC (Provão - Modelo de Avaliação da Graduação), Exame Nacional do Ensino Médio - ENEM, o Programa de Avaliação Institucional das Universidades Brasileiras - PAUB.

O fato de a avaliação, como foi dito, ser ainda recente no Brasil, parece explicar a dificuldade encontrada no levantamento de informações a respeito do desenvolvimento da avaliação de políticas educacionais (sistemas educativos, programas, projetos, etc.). Apesar disso, é possível fazer algumas considerações que acreditamos serem importantes, através da comparação do desenvolvimento da avaliação no Brasil e nos Estados Unidos.

A primeira é que a construção das bases conceituais e metodológicas da avaliação de políticas, programas e projetos sociais e educacionais ocorreu nos Estados Unidos através de um processo de amadurecimento iniciado, a partir da década de 40 até os anos 80.

---

[37] Não temos intenção de discutir os aspectos políticos, ideológicos e teórico-metodológicos desses exames.

No Brasil, no entanto, nesse mesmo período (a partir da década 60), verifica-se a apropriação e utilização desses pressupostos, para o desenvolvimento das avaliações no contexto brasileiro.

A segunda consideração é que, em relação à mudança de enfoque da avaliação, na área educacional, pode-se observar que ela ocorreu no Brasil de forma semelhante à observada nos Estados Unidos. Primeiramente houve ênfase na avaliação de rendimento[38] e, posteriormente, ocorreu grande preocupação com a avaliação de programas curriculares[39]. No entanto, existe uma diferença entre os dois países no que diz respeito ao contexto político-econômico no qual essas mudanças ocorreram.

A terceira é que enquanto nos Estados Unidos a preocupação em avaliar e controlar a educação (políticas educacionais, sistemas educativos, programas, etc.) de forma sistemática ocorreu a partir dos meados da década de 60, no Brasil isso se inicia a partir dos meados da década de 80.

A quarta consideração é que na América Latina, e no Brasil, tal como nos Estados Unidos, o desenvolvimento da avaliação está ligado à grandes reformas na educação, à tentativa de solução de grandes problemas sociais, à redução de gastos públicos, a altos investimentos financeiros. E, adicionalmente, no nosso caso, a avaliação foi colocada como uma exigência pelos organismos supranacionais que financiam as políticas educacionais.

---

[38] Não é nossa intenção detalhar esse tema pois não é nosso objeto de estudo.
[39] Para maior detalhamento ver SAUL (2000).

Por último fica a consideração de que, embora a maioria da literatura brasileira utilize as bases conceituais e metodológicas provenientes de outros países, existem especificamente trabalhos relacionados à área educacional, como o realizado por Ana Maria Saul, orientados a desenvolver metodologias de avaliação mais adequadas ao contexto brasileiro.

Fatos como esses talvez estejam indicando que estejamos passando para um novo momento em que a literatura estrangeira venha a servir como uma referência para a realização de estudos que promovam as adaptações necessárias e que, ao servirem de base para um aprofundamento, venham a gerar um novo enfoque metodológico significativamente distinto. O esforço associado a esse novo momento parece justificar-se uma vez que tanto o ambiente da *policy making* como os contextos sócio-econômico e político-cultural dos países onde se originam as metodologias de avaliação de políticas se diferenciam significativamente, da realidade brasileira e latino-americana.

É com base na retrospectiva histórica do processo de construção das bases conceituais e metodológicas da avaliação de políticas, programas e projetos sociais, que se infere a diversidade conceitual que o tema apresenta.

Assim, o Capítulo II deter-se-á sobre os principais elementos conceituais e metodológicos que hoje conformam o instrumental da avaliação de políticas, programas e projetos.

CAPÍTULO II

Conceito de avaliação de políticas, programas e projetos

O termo avaliação é derivado do termo latim "valere" que significa valorar. E que implica na ação de testar, medir ou atribuir valor a alguma coisa. Na linguagem ordinária, portanto, o termo avaliação é uma palavra bastante elástica e que tem empregos muito diferentes, mas que de forma geral é utilizada para emitir juízo, fazer uma valoração ou medir algo.

Devido à elasticidade do termo avaliação, os autores que tratam do tema têm-se preocupado em distinguir entre duas maneiras de avaliar. A avaliação informal, também conhecida como avaliação no sentido 'lato', considerada inerente à atividade humana presente em nossa vida cotidiana, em nossa vida profissional quando avaliamos os outros ou nós mesmos. E a avaliação formal, também chamada de avaliação em sentido estrito ou avaliação sistemática. Esta denotaria uma ação mais complexa e requer procedimentos científicos e pressupostos teórico-metodológicos que proporcionem maior sistematização, credibilidade e fidedignidade da informação colhida e produzida.

No Capítulo I, mostrou-se como a construção do corpo teórico-metodológico da avaliação foi gradativa e como, em função da influência de diversas áreas do conhecimento, foi-se aos poucos alterando o enfoque dos trabalhos sobre o tema. Como resultado, conta-se hoje com um corpo teórico-metodológico que passou por um processo de amadurecimento significativo e que é capaz de proporcionar uma considerável diversidade conceitual[40]. Essa realidade sugere que se destaque três aspectos observados na literatura que consideramos importantes para o entendimento deste capítulo.

O primeiro é que o conceito de avaliação de políticas programas ou projetos é encontrado tanto na literatura específica da área educacional, produzida por educadores, quanto na que trata da área social que a abarca, produzida por sociólogos, economistas e outros cientistas sociais.

O segundo aspecto é que se pode identificar com freqüência na literatura uma interseção de referências teórico-metodológicas. Isto é, muitos autores utilizam conceitos, modelos e tipos de avaliação provenientes da área educacional[41] nos seus estudos sobre a área social (pelas razões que se aponta em seguida parece ser menos freqüente o movimento inverso).

---

[40] Neste capítulo nosso objetivo é demonstrar a divesidade conceitual apontada anteriormente e apresentar as dimensões encontradas nos conceitos de avaliação, conforme observado na literatura especializada. No entanto, é imprtante indicar que não se tem a intenção de discutir em que condições políticas, sociais, econômicas e ideológicas os conceitos foram construídos e são utilizados para avaliar as políticas, programas e projetos sociais e educacionais. Também acreditamos ser necessário destacar que não nos deteremos na aplicabilidade desses conceitos de avaliação no contexto internacional ou brasileiro, mesmo porque o objetivo do nosso trabalho não é discutir sob que condições políticas e ideológicas ocorrem as avaliações.

[41] Nesse sentido, encontramos, inclusive, autores que desenvolveram seus estudos na área de avaliação de rendimento do aluno ou avaliação de aprendizagem. Como exemplo podemos citar o trabalho realizado por Tyler.

Além disso, e dentro do que poderia ser considerado a área educacional, é comum encontrar autores que, tendo desenvolvido trabalhos na área de avaliação de aprendizagem, direcionaram posteriormente seu arsenal teórico-metodológico para avaliar programas curriculares. Ocorre, nesses casos, uma utilização nem sempre adequadamente justificada do mesmo instrumental empregado para a avaliação do processo de ensino-aprendizagem para a avaliação de políticas, programas ou projetos sociais e educacionais[42]. De qualquer forma, e sem levar em conta o juízo que se possa fazer deste procedimento, é importante reconhecer que esses conceitos, independentemente do foco para qual foram inicialmente direcionados, têm servido como base para o estudo de processos avaliativos de atividades ou fenômenos na área social e educacional de natureza diversa.

O terceiro aspecto a destacar é que são encontrados na literatura alguns termos que se referem ou que são utilizados como sinônimos de avaliação de políticas, programas e projetos sociais e educacionais. São eles: pesquisa avaliativa, pesquisa de avaliação, estudos avaliativos, investigação avaliativa e, na área educacional, avaliação educacional.

Em relação ao termo avaliação educacional deve-se considerar que ele é utilizado para referir a avaliações de vários aspectos do âmbito educacional, inclusive à avaliação de políticas, programas ou projetos educacionais.

---

[42] Algo que se poderia comparar com a utilização de ferramentas de análise concebidas para tratar a formulação de estratégias empresariais, a fim de analisar o processo de decisão que tem lugar em agências governamentais.

Como exemplo dessa constatação, pode-se citar Franco (1990) e Méndez (2001), que utilizam o termo avaliação educacional como sinônimo de avaliação de aprendizagem ou do rendimento do aluno. Isso se deve, provavelmente, ao fato de que, durante muitas décadas, a avaliação de aprendizagem foi o único tipo de avaliação utilizado na educação, ficando conhecida como avaliação educacional. Esta é, portanto, a referência mais comum e também a mais antiga que encontramos na literatura.

Já Vianna (1989, 1992, 2000) e Afonso (2000) utilizam o termo avaliação educacional para se referirem à avaliação de aprendizagem, avaliação de programas educacionais, programas curriculares, avaliação de materiais instrucionais, sistemas de ensino, avaliação do rendimento escolar etc.

Para esses autores, essas avaliações são dimensões da avaliação educacional. A avaliação educacional é, portanto, um termo que abrange a avaliação de vários fenômenos e atividades no campo da educação, mas isso não implica que as avaliações que se realiza tenham sempre o mesmo enfoque avaliativo.

Nessa perspectiva, Vianna (1989, p. 41) afirma que:

> *"Avaliar é determinar o valor de alguma coisa para um determinado fim. A avaliação educacional visa, pois, à coleta de informações para julgar o valor de um programa, produto, procedimento ou objetivo" (Worthen & Sanders, 1973) ou, ainda, "apreciar a utilidade potencial de abordagens alternativas para atingir determinados propósitos. A avaliação refere-se, assim, a atividades sistemáticas ou formais para o estabelecimento do valor de fenômenos educacionais quaisquer que seja" (Pophan, 1975)".*

Também, numa perspectiva parecida, encontramos Saul (2000), que utiliza o termo avaliação educacional para se referir à avaliação de currículos, avaliação de aprendizagem e avaliação de programas educacionais. No entanto, mesmo usando o termo avaliação educacional de forma abrangente, a autora, ao desenvolver sua proposta de Avaliação Emancipatória[43], especifica claramente que esta serve para avaliar programas tanto na área educacional como na social.

Considerando, desta forma, que a metodologia de avaliação pode ser proveniente da área social e educacional e que existem diversos termos que podem fazer referências à avaliação, optamos por utilizar o termo avaliação de políticas, programas e projetos sociais e educacionais durante todo o texto deste trabalho.

Após essas delimitações, entramos propriamente no tema deste capítulo, que será dividido em duas seções. Na primeira serão apresentados os diversos conceitos produzidos pela literatura estrangeira e brasileira sobre avaliação de políticas, programas ou projetos. E, na segunda seção, procuraremos mostrar as diferenças e semelhanças que existem entre avaliação e pesquisa avaliativa.

---

[43] Segundo Saul (2000), a Avaliação Emancipatória caracteriza-se como um processo de descrição, análise e crítica de uma dada realidade visando explicitamente a sua transformação.

2.1. Conceitos de avaliação

Como dito anteriormente, encontramos na literatura diversos conceitos para fazer referência à avaliação de políticas, programas e projetos sociais e educacionais. Essa diversidade conceitual deve-se, possivelmente, a dois fatores. O primeiro diz respeito às alterações que o conceito foi sofrendo devido à influência de diversas áreas do conhecimento. Devido a isso encontramos conceitos e modelos de avaliação provenientes da interseção de algumas áreas do conhecimento tais como: administração, economia, ciências sociais, educação etc. O segundo fator refere-se à necessidade de avaliar aspectos diferentes de uma mesma política, programa e projeto social e educacional, estabelecendo uma estreita relação entre as necessidades vigentes e os interesses de estudiosos, possibilitando assim que as conceituações reflitam focos de interesses considerados relevantes.

Partindo dessa constatação, podemos inferir que o conceito de avaliação de políticas, programas e projetos sociais e educacionais passou por um processo de amadurecimento gradativo durante o período em que o seu corpo teórico-metodológico[44] foi sendo construído e consolidado, e possibilitou um número significativo de conceitos para esse tipo de avaliação.

---

[44] Para maior detalhes ver Capítulo I.

Alguns desses conceitos foram inicialmente produzidos na área de avaliação de aprendizagem. Essa atividade avaliativa foi concebida apenas como um processo de medição e quantificação do desempenho do aluno através de testes de objetivos, e logo permeou outros campos na área educacional.

Segundo Briones (1998), a concepção de avaliação proveniente da área educacional, entendida como processo de medição e quantificação para possível julgamento, foi sendo mecanicamente utilizada para avaliar atividades muito diferentes daquelas para as quais ela tinha sido projetada.

A avaliação foi, portanto, durante um longo período, considerada um instrumento para medir e julgar objetos ou pessoas através de um processo de quantificação que permite estabelecer números ou pontuações sobre o que está sendo avaliado.

À medida que o processo avaliativo passou a ter uma função importante na administração e gestão das agências governamentais, a avaliação, entendida apenas como julgamento e medição de programas e projetos, já não atendia às exigências sociais, econômicas, políticas. O conceito de avaliação foi agregando dimensões e, gradativamente, foi sendo substituído por concepções que abrangiam novos enfoques.

Essas alterações na concepção da avaliação foram sintetizadas pelo Banco Interamericano de Desenvolvimento (BID, 1997) em quatro "gerações" conforme **Quadro 2.**

Com respeito às concepções de avaliação presentes nesse quadro, podemos inferir que a partir da segunda geração do conceito de avaliação, ela passa a ser um mecanismo que proporciona informações sobre os resultados ou objetivos de um determinado objeto ou atividade que está sendo avaliada para uma possível tomada de decisão.

Quadro 2

Etapas de la concepción de la evaluación

Etapa/Período	Interés
Primera generación-medición Fines del siglo XIX	**Identificación del nivel de logro de ciertos objetivos específicos, mediante la medición cuantitativa con instrumentos uní modales**
Segunda generación-descripción / comparación **Décadas de 1920 a 1950**	**Identificación de conjuntos de aspectos optimizadores y limitantes con respecto a ciertos objetivos establecidos; comparación de enfoques usando métodos experimentales, o de fenómenos, en grupo y situaciones de ocurrencia natural.**
Tercera generación-servicios de valor **Décadas de 1950 a 1980**	**Comparación de resultados múltiples con metas y normas establecidas a priori; evaluación normativa relacionada con hitos e indicadores de referencia.**
Cuarta generación-Transparencia e responsabilidad ejecutiva y desempeño. **Década de 1980 a la fecha**	**Transparencia, responsabilidad ejecutiva y desempeño coordinados por medio del análisis de todos los dados disponibles, usando una variedad de métodos y mediciones múltiples, e incorporando el conocimiento, las perspectivas y los valores de todos los involucrados.**

Reproduzido de: BID,1997, p. 3.

Partindo do **Quadro 2** e da análise preliminar, pode-se apontar a existência de três dimensões[45] complementares e historicamente cumulativas (no sentido de que cada uma se adiciona à anterior), que identificamos no conceito de avaliação de políticas, programas e projetos que utilizamos para organizar a próxima seção.

A primeira é a dimensão decisional (prática) da avaliação e relaciona-se com a tomada de decisão e, está virtualmente em todos os conceitos analisados.

A segunda é a dimensão racional. Refere-se à importância que pouco a pouco passou a ser conferida à racionalização da decisão e dos recursos disponíveis através da utilização de critérios[46] de eficiência, eficácia e efetividade.

A terceira é a dimensão que denominamos integradora, que busca incorporar a tendência observada de que a atividade avaliativa vem sendo crescentemente entendida como um processo e como parte integrante de um processo abrangente, que inclui o planejamento, a formulação, a implementação de uma política, programa ou projeto social e educacional.

---

[45] As dimensões apresentadas na próxima seção foram 'criadas' com o objetivo de facilitar o entendimento dos conceitos propostos pelos autores. Também é necessário destacar que a prática avaliativa no contexto brasileiro, pode não seguir os preceitos desses conceitos. Isto significa dizer que, as avaliações realizadas no Brasil podem seguir outros conceitos ou incorporar outros pressupostos analítico-conceituais, que não os contemplados na seção.

[46] Esses critérios serão abordados no Capítulo V.

2.1.1. Conceito de avaliação: dimensão decisional, dimensão racional e dimensão integradora

A dimensão decisional é encontrada virtualmente em todos os autores. Podemos observá-la desde os conceitos mais remotos até os mais recentes. Como exemplo, citaremos dois autores que propuseram, na década de 70, as seguintes definições de avaliação de políticas, programas ou projetos sociais e educacionais[47].

Segundo Houston (1972, p. 51),

> *"...avaliar um programa de ação social é colecionar evidências acerca de sua efetividade*[48]*.(...)Estas evidências proporcionam ao avaliador a capacidade de persuadir os outros a modificar, continuar ou cancelar o programa avaliado* [49]*".*

E sendo mais explícito quanto à dimensão decisional (prática) da avaliação, Alkin (1972, p.107) afirma que:

---

[47] A ênfase que se dá nesta seção aos aspectos particulares da obra dos autores que denotam cada uma das dimensões analisadas tem por objetivo melhor caracteriza-las. Não significa, no entanto, que os mesmos não tenham considerado as demais dimensões identificadas.

[48] Para esse autor, a efetividade é definida como sendo a "capacidade do programa para causar mudanças naqueles que estão expostos a ele". (HOUSTON, 1972, p 51). Tradução do original em inglês.

[48] Tradução do texto original em inglês.

> *"A avaliação é o processo de averiguar as áreas de 'preocupação' que necessitam de decisão, colecionando, selecionando e analisando informações apropriadas para informar dados sumariamente úteis para a tomada de decisão e indicar alternativas possíveis*[50]*".*

A preocupação com a tomada de decisão, que caracteriza o que denominamos dimensão decisional, é bastante difundida na avaliação de políticas, programas ou projetos sociais e educacionais e sociais.

Assim, Weiss (1998) define a avaliação de programas como uma investigação sistematizada que utiliza os métodos das ciências sociais para medir os resultados de um programa e compará-los às metas planejadas e, desta forma, concluir acerca da eficácia, do valor e do êxito daquilo que está sendo avaliado. A avaliação descreve e compreende as relações entre variáveis, estabelecendo uma seqüência causal entre causa e efeito, e proporciona informação precisa que possibilite a tomada de decisão.

Como se pode observar, e tal como apontamos acima, essa autora também aponta para a dimensão racional da avaliação e enfatiza a necessidade de estabelecer critérios que possibilitem a racionalização da tomada de decisão.

Nessa perspectiva, para Weiss[51] (1998) a avaliação (investigação ou pesquisa avaliativa) é procedimento concebido para ajudar aos responsáveis por programas ou políticas a trabalhar e alocar seus

---

[51] A primeira publicação desta obra foi em 1972.

recursos da melhor maneira possível. A avaliação é uma maneira para aumentar a racionalidade das decisões ou, segundo a autora,

> *"...é uma pesquisa sistemática da operação e / ou resultados de um programa ou política comparando-os a um jogo de padrões explícitos como meio para contribuir para melhoria do que esta sendo avaliado*[52]*".* (WEISS, 1998, p.4).

Ainda, ressaltando a dimensão decisional (prática) e racional do conceito da avaliação, Briones (1998, p.13) afirma que o processo avaliativo é:

> *"Un tipo de investigación que analiza la estructura, el funcionamiento y los resultados de un programa con el fin de proporcionar información de la cual se deriven criterios útiles para la toma de decisiones en relación con su administración y desarrollo".*

Esse autor destaca, ainda, que a avaliação não é uma "ciência", é pesquisa interdisciplinar e por isso utiliza instrumentos da pesquisa social e de diversas outras áreas do conhecimento que permitam coletar, analisar e sistematizar de maneira mais precisa as informações para uma tomada de decisão. Assim, para Briones (1998, p.15),

> *"...en el proceso de la investigación evaluativa se establece una especial relación entre teoría y práctica, de tal modo que más que la aplicación en abstracto de leyes y principios generales se utilizan conceptos y experiencias específicas y apropiadas al problema o situación que se analiza".*

---

[52] Tradução do original em inglês.

Tripodi[53] (1984, p.17) propõe uma Avaliação Diferencial[54] e define avaliação de programas sociais como sendo:

> *"...a acumulação sistemática de fatos para fornecer informações sobre a realização dos requisitos e objetivos do programas em relação aos seus esforços, eficácia e rendimento em qualquer dos estágios de seu desenvolvimento. Os fatos da avaliação podem ser obtidos através de várias técnicas relativamente sistemáticas e são incorporados em determinado sistema de valores para a tomada de decisão sobre os programas sociais":*

Esse autor, ao conceituar a avaliação, ainda que de forma não explícita, a entende como um processo contínuo para planejamento e desenvolvimento das políticas, programas e projetos sociais e educacionais (dimensão integradora).

Viedma (1996, p.1) nessa mesma perspectiva, afirma que a avaliação é um instrumento que deve estar presente em todas as fases do programa, desde o planejamento até a conclusão das atividades. Assim, para a autora,

> *"En realidad, la evaluación es un instrumento que sirve para medir, monitorizar y seguir las actividades de un programa determinado". [...] La evaluación no es un caos sino un medio para saber de dónde arranca, por dónde pasa y dónde termina una actividad. Esto se hace por medio de tecnologías validadas de medición, arrojando resultados de gran valor especialmente para quienes tienen la responsabilidad de planificar y adoptar decisiones que afectan a colectivos".*

---

[53] A primeira edição desta obra foi em 1975.

Também destacando a dimensão integradora da avaliação no processo de planejamento e desenvolvimento de programas e projetos sociais e educacionais, Martinic (1997, p.7-8) conceitua a avaliação como

> *"...una actividad permanente e integrada al ciclo de vida de los proyectos* [...] En un sentido amplio la evaluación es el juicio o apreciación que emite una persona o un equipo sobre las actividades y resultados de un proyecto [...] *La evaluación contribuye no solo para la toma de decisiones sino que, fundamentalmente, al aprendizaje y los conocimientos que los propios ejecutores tienen del problema que abordan".*

Já Cohen & Franco (1999) para comporem sua definição recorreram aos conceitos propostos por duas organizações supranacionais: a Organização das Nações Unidas (ONU) e a Organização Mundial de Saúde (OMS). Ambas enfatizam a avaliação como uma atividade intrínseca ao processo de planejamento e fundamental, para melhorar as atividades que estão sendo implementadas.

Esses autores enfatizam, também, a dimensão racional ao indicar que a avaliação deve possibilitar a racionalização de recursos para o desenvolvimento do programa, através de critérios associados à eficiência e eficácia[55] dos programas que estão sendo avaliados.

Assim, segundo Cohen & Franco (1999, p.77), a avaliação é entendida como

> *"...uma atividade que tem como objetivo maximizar a eficácia dos programas na obtenção de seus fins e a eficiência na alocação de recursos para a consecução dos mesmos".*

---

[55] Esses termos serão detalhados posteriormente no Capítulo V.

Introduzindo o tema da avaliação *ex-ante*, que trataremos em detalhe posteriormente, mas que é importante para entender a contribuição desses autores, é conveniente chamar a atenção para o fato de que eles consideram a avaliação como um instrumento para decidir se um programa ou projeto é o modo mais eficiente e eficaz para viabilizar a intervenção numa determinada área ou problema. Também destacam que a avaliação desempenha um papel importante no momento de escolha entre dois programas ou projetos alternativos que podem ser implementados.

Nessa perspectiva, a avaliação possibilitaria informações para uma tomada de decisão racional (acerca de qual é o melhor, eficaz, eficiente e efetivo), dentre um rol de programas que foram elaborados. Assim, segundo Conhen & Franco (1999, p.73),

> *"A avaliação não deve ser concebida como uma atividade isolada e auto-suficiente. Ela faz parte do processo de planejamento da política social, gerando uma retroalimentação que permite escolher entre diversos projetos de acordo com sua eficácia e eficiência. Também analisa os resultados obtidos por esses projetos, criando a possibilidade de retificar ações e reorientá-las em direção ao fim desejado".*

Na mesma perspectiva e enfatizando o conjunto das três dimensões, encontramos Ander-Egg (1990, p.19) que afirma:

> *"...la evaluación consiste en utilizar una serie de procedimientos destinados a comprobar si se han conseguido o no las metas y objetivos propuestos, identificar los factores o razones que han influido en los resultados (éxito o fracaso) y formular las recomendaciones pertinentes que permitan tomar decisiones con el fin de introducir correcciones o los reajustes que sean necesarios. En este último sentido, la evaluación es un instrumento de retroalimentación".*

Considerando as dimensões decisional, racional e integradora em seu conjunto, o autor faz uma relação entre o ato de planejar e o ato de avaliar. Para Ander-Egg (1990, p.20), então

> *"Si planificar es introducir organización y racionalidad en la acción para el logro de determinadas metas y objetivos, la evaluación es una manera de verificar esa racionalidad, midiendo el cumplimiento - o perspectivas de cumplimiento de los objetivos y metas previamente establecidos y la capacidad para alcanzarlos".*

Ballesteros & Benito (1996, p.23) também enfatiza a avaliação como um instrumento que proporciona informações para a tomada de decisão. Para eles, a *"evaluación de programas es la sistemática investigación a través de métodos científicos de los efectos, resultados y objetivos de un programa con el fin de tomar decisiones sobre él."* Para chegar a essa conceituação de avaliação de programas, eles recorrem às definições estabelecidas por diversos autores. Entre eles, podemos encontrar Tyler, Scriven, Cronbah[56], que são considerados por muitos como os 'pais da avaliação'.

É interessante perceber que apesar de se referirem à avaliação sem especificar área ou tipo de programa, eles recorreram aos estudiosos que desenvolveram seus fundamentos conceituais para a área de avaliação de programas educativos e, mais especificamente, de programas curriculares.

---

[56] Esses autores desenvolveram seus pressupostos na área de programas educacionais. Esses pressupostos foram utilizados e serviram como base para diversas atividades avaliativas, inclusive na área de programas sociais.

Essa recorrente utilização de conceitos provenientes de diversas áreas do conhecimento para compor aqueles relativos a avaliação de programas sociais e educacionais também pode ser observado em outros autores.

Em Aguilar & Ander-Egg (1995) sugere-se um percurso bastante complexo para conceituar a avaliação. Nesta obra, os autores selecionaram aproximadamente 20 definições de avaliação, provenientes de diversas áreas do conhecimento e realizadas em diferentes períodos históricos. Dentre os autores dessas definições podemos encontrar alguns que direcionaram seus estudos para a área de saúde, educação ou do desenvolvimento comunitário. Também incluem conceitos propostos por organismos supranacionais, como a UNESCO (Organização das Nações Unidas para a Educação, Ciência e Cultura) e a OMS (Organização Mundial da Saúde)

A partir dos conceitos que analisaram, Aguilar & Ander-Egg (1995, p. 31-32) conceituam a avaliação como

> *"...uma forma de pesquisa social aplicada, sistemática, planejada e dirigida; destinada a identificar, obter e proporcionar de maneira válida e confiável dados e informação suficientes e relevantes para apoiar um juízo sobre o mérito e o valor dos diferentes componentes de um programa (tanto na fase de diagnóstico, programação ou execução), ou de um conjunto de atividades específicas que se realizam, foram realizadas ou se realizarão, com o propósito de produzir efeitos e resultados concretos; comprovando a extensão e o grau em que se deram essas conquistas, de forma tal que sirvam de base ou guia para uma tomada de decisão racional e inteligente entre cursos de ação, ou para solucionar problemas e promover o conhecimento e a compreensão dos fatores associados ao êxito ou fracasso de seus resultados."*

Também, ressaltando as três dimensões da avaliação e apontando a avaliação como um instrumento sistêmico de valoração ou julgamento, que proporciona informações para planejamento de programas futuros, Lara (1996, p. 91) afirma que:

> *"La evaluación consiste en un examen crítico del desarrollo y de los efectos actuales o potenciales de un programa, y se orienta a aumentar la eficacia del programa en curso y a proporcionar una base orientadora para el diseño e implementación de programas futuros. Por ello, la evaluación es, fundamentalmente, un proceso diseñado para ser usado por la dirección operativa de un programa y por los responsables de tomar decisiones sobre la financiación, continuidad, expansión y acreditación del programa".*

Avançando na sua conceituação de avaliação de políticas, programas ou projetos sociais e educacionais, esse autor conclui que, "*evaluar es emitir juicios sobre las actividades y los objetivos de un programa, a partir de información empírica, recogida, rigurosa y sistemáticamente*" (LARA, 1996, p.105).

Brown (2001) utiliza-se da definição proposta pela OCDE (Organização de Cooperação e Desenvolvimento Econômico) ao definir avaliação. Para ele,

> *"La evaluación se ha definido como la valoración, en la forma más sistemática y objetiva posible, de un proyecto, programa o política en marcha o terminado, incluyendo su diseño, ejecución y resultados".* (OCDE, 1986, apud BROWN, 2001, p.9).

Esse autor, completa a conceito de avaliação afirmando que:

> *"...el propósito principal de una evaluación podría ser el de determinar la importancia y cumplimiento de objetivos, la eficiencia, el impacto o la capacidad de sustentación. La evaluación puede realizarse antes, durante o después de la ejecución"(BROWN, 2001, p. 9)*

Desta forma, salienta a dimensão racional, através da utilização de critérios de eficiência e eficácia para a racionalização de recursos e desenvolvimento do que está sendo avaliado.

Dois organismos supranacionais – o BID (Banco Interamericano de Desenvolvimento) e o SIEMPRO (*Sistema de Información, Monitoreo y Evaluación de Programas Sociales*) – contribuíram de forma significativa ao proporem conceitos de avaliação de políticas, programas ou projetos sociais e educacionais. Para o primeiro, que incorpora o conceito criado pelos especialistas da OCDE/CAD (Organização e Cooperação de Desenvolvimento Econômico/Comitê de Assistência ao Desenvolvimento)[57], a avaliação é:

> *"...un escrutinio — lo más sistemático y objetivo posible — de un proyecto, programa o política en ejecución o terminado, y sus dimensiones de diseño, ejecución y resultados. El propósito es determinar la pertinencia y logro de los objetivos y la eficiencia, efectividad, impacto y sustentabilidad del desarrollo". Una evaluación debe proporcionar información que sea creíble y útil, para permitir la incorporación de la experiencia adquirida en el proceso de adopción de decisiones tanto del prestatario como de los donantes".* (OCDE/CAD 1991, apud BID, 1997. p. 2)

---

[57]OCDE/CAD: Comitê de Assistência ao Desenvolvimento da Organização de Cooperação para o Desenvolvimento Econômico.

O segundo organismo define a avaliação como *"un proceso permanente y continuo de indagación y valoración de la planificación, la ejecución y la finalización del programa social"*. (SIEMPRO, 1999, p.55).

Essas duas conceituações consideram a avaliação como ferramenta para gerar informação para a tomada de decisão (dimensão decisional) e acentuam a dimensão integradora (processo inerente ao planejamento). Porém, o BID (1997) utiliza o Modelo de Marco Lógico[58], como fundamento para suas atividades avaliativas e o SIEMPRO (1999) desenvolve a avaliação dentro da perspectiva de Gestão Integral de Programas Sociais Orientada para Resultados[59].

Para concluir esta abordagem acerca das três dimensões contidas nos conceitos propostos pelos autores que consideramos relevantes, apresentaremos algumas considerações a respeito do tema.

A primeira é que os conceitos de avaliação de políticas, programas e projetos sociais e educacionais aqui apresentados caracterizam -na como um tipo de pesquisa eminentemente prática.

A segunda é que as dimensões aqui propostas — decisional, racional e integradora —, embora nem sempre explicitamente abordadas por todos os autores analisados, estão presentes nos conceitos de avaliação de políticas, programas e projetos formulados tanto na área educacional como da área social.

---

[58] Ver BID (1997).
[59] Ver SIEMPRO (1999).

A terceira consideração é que os diversos conceitos aqui apresentados parecem corroborar o que apontamos no primeiro capítulo deste trabalho acerca da pluralidade das contribuições disciplinares, observadas no processo de construção do corpo teórico-metodológico da avaliação de políticas, programas e projetos.

2.2. Avaliação como atividade de pesquisa

Além das diferenças estabelecidas entre avaliação informal e avaliação formal (avaliação sistematizada), encontramos na literatura autores que apontam diferenças entre avaliação e pesquisa avaliativa[60] e autores que utilizam esses dois termos como sinônimos.

Suchman (1977) afirma a esse respeito que avaliar é apenas um método que determina o grau em que os objetivos anteriormente planejados foram alcançados, enquanto que a pesquisa avaliativa é muito mais do que isso.

Segundo ele, a pesquisa avaliativa envolve mais do que o simples ato de "reprovar", o que está sendo avaliado ou até mesmo, determinar o grau de alcance dos objetivos planejados de uma política, programa ou projeto.

---

[60] Este termo pode também ser encontrado na literatura como: *evaluation research*, pesquisa de avaliação, pesquisa avaliativa, *investigación evaluativa*, estudos avaliativos ou ainda o termo avaliação.

A pesquisa avaliativa procura as condições que afetam o sucesso ou fracasso do programa através de hipóteses que estabelecem uma relação causal entre as atividades desenvolvidas e os resultados atingidos pelo programa. Pode-se dizer que a pesquisa avaliativa preocupa-se com: 'o quê?', 'como' e 'por quê?' foram alcançados os resultados do programa avaliado.

Portanto, esse tipo de pesquisa pergunta sobre o tipo de mudança desejada, os meios pelos quais esta mudança vai ser provocada, e os sinais através dos quais tais mudanças podem ser reconhecidas, de modo a subsidiar a tomada de decisão acerca do que está sendo avaliado.

Nessa mesma perspectiva, Weiss (1998, 1999) afirma que avaliar é julgar o mérito de algo utilizando critérios, enquanto que a pesquisa avaliativa[61] é um método de avaliação que, para emitir um juízo, utiliza as ferramentas da pesquisa social para fazer de forma mais precisa e objetiva o processo de julgamento. É, portanto, uma forma de pesquisa que

> *"...trata de describir, de comprender, las relaciones entre variables y establecer la secuencia causal. Porque está estudiando un programa que interviene en las vidas de las personas con la intención de producir un cambio, la evaluación puede hacer, a menudo, interferencias directas acerca de los vínculos causales que conducen desde el programa hasta el efecto".* (WEISS, 1999, p.21).

[61] Tradução do termo investigación evaluativa do original em espanhol.

Sem utilizar o termo pesquisa avaliativa, porém na perspectiva de que avaliação é uma forma de pesquisa, Lara (1996) indica que avaliar é emitir juízo de valor sobre uma política, programa ou projeto, a partir de informações coletadas empiricamente, rigorosamente e sistematicamente. Para a coleta dessas informações, a avaliação utiliza procedimentos da pesquisa social. Considerando essa relação, o autor afirma que toda avaliação supõe uma pesquisa e, portanto, toda avaliação deve ser entendida como uma pesquisa avaliativa.

Além disso, como toda pesquisa, a avaliação partiria de hipóteses formuladas em consonância com um modelo de relação causal e deveria ter sempre presente os aspectos que influenciam o desenvolvimento do programa. Isso significa dizer que a pesquisa avaliativa deve identificar e analisar as causas devido às quais o programa atinge (ou não) seus objetivos.

Partindo então desses pressupostos, Lara (1996, p.106) afirma que:

> *"La investigación evaluativa, en tanto que proceso científico, requiere de la aplicación de procedimientos sistemáticos y rigurosos para recoger y avalizar la información que ha de proporcionar las evidencias empíricas capaces de contrastar las hipótesis planteadas. Pero la evaluación no acaba con la presentación de lo resultados de esa contrastación en la forma de proposiciones de conocimiento: su objetivo específico es emitir juicios de valor sobre el objeto que se está evaluando (un determinado programa o servicio social, o una intervención puntual) que permitan aumentar la eficacia del programa en curso y tomar decisiones acerca de su continuidad o expansión, o proporcionen una base orientadora para el diseño e implementación de programas futuros".*

Nessa mesma perspectiva, Martinic (1997) também afirma que a pesquisa avaliativa parte do pressuposto de que as relações entre os diversos elementos de um projeto originam as hipóteses no âmbito de um modelo causal. Ela estuda a validez das hipóteses que imputam as relações entre elementos do projeto e destes com o problema que sofrerá a intervenção e com as condições para o seu desenvolvimento. Assim, segundo este autor,

> *"...la evaluación de proyectos es el estudio de las relaciones que proponen los proyectos a través de sus hipótesis y lógicas de intervención. La evaluación, por lo tanto, supone que existen relaciones causales o de probabilidad entre los elementos del proyecto y entre estos y los factores del entorno que inciden en su desarrollo. Una relación causal es diferente a la correlación que tienen las variables* [...] *Para que una relación sea causal debe existir correlación pero, además, la variable independiente debe ocurrir antes que el efecto e incidir, claramente, en ella* (Hernández, R., et al, 1994:86). *De este modo la relación causal no solo supone correlación entre dos o más variables sino que establece una dirección en dicha relación".* (MARTINIC, 1997, p. 123).

Weiss (1999), concordando com Martinic (1997), Suchman (1977) e também com Lara (1996), no que diz respeito à pesquisa avaliativa, enfatiza o fato de que este tipo de pesquisa é também uma maneira de aumentar a racionalidade das decisões acerca das políticas, programas ou projetos sociais e educacionais. Assim, para Weiss (1999, p.15):

> *"...la aplicación de los conocimientos y la metodología de las ciencias sociales tendrá efectos beneficiosos: mejorará la toma de decisiones, dará lugar a que se planeen mejores programas y, de esta manera, rendirá un servicio a los participantes en los programas de maneras más beneficiosas, eficientes".*

Procurando estabelecer diferenças entre avaliação e pesquisa avaliativa, Aguilar & Ander-Egg (1994, p.22) afirmam que avaliar é medir a consecução dos objetivos e emtas previamente estabelecidos, enquanto que a pesquisa avaliativa utiliza-se dos métodos e técnicas da pesquisa social

Estes autores afirmam que a pesquisa avaliativa é conceituada como

> *"...uma forma de valoração sistemática que se baseia no emprego de procedimentos que, apoiados no uso do método científico, servem para identificar, obter e proporcionar a informação pertinente, julgar o mérito e o valor de algo, de maneira justificável"* (AGUILAR & ANDER-EGG,1994, p. 23).

Os autores acima mencionados, ao estabelecerem as diferenças entre avaliação e pesquisa avaliativa enfatizam aspectos distintos. Os primeiros — Suchman (1977), Weiss (1999) e Lara (1996) — apontam que os propósitos da avaliação e da pesquisa avaliativa são diferentes. Segundo eles, a avaliação preocupa-se com processo de julgamento do mérito e do grau de alcance dos objetivos, etc. A pesquisa avaliativa, de outro modo, além de realizar esse julgamento, se preocupa com a maneira pela qual os resultados foram alcançados, estabelecendo uma relação causal entre, por um lado, os diversos elementos, atividades, etc., e, por outro, destes com os resultados do programa que está sendo avaliado.

Já Aguilar & Ander-Egg (1995), apontam que a diferença entre avaliação e pesquisa avaliativa não está relacionada com a questão dos propósitos (ou objetivos) que cada uma delas contempla. E, sim, uma questão de técnicas métodos utilizados para a obtenção, sistematização e análise da informação.

A partir das diferenças apontadas por esses autores, podemos considerar que tanto a avaliação quanto a pesquisa avaliativa, são formas de emitir algum juízo de valor utilizando critérios previamente estabelecidos para, assim, determinar se o que está sendo avaliado atingiu os objetivos estabelecidos.

A característica apontada pelos autores como diferenciadora entre avaliação e pesquisa avaliativa é a utilização de métodos e técnicas da pesquisa social e o estabelecimento, através de hipóteses, de uma relação causal entre atividades ou objetivos e resultados alcançados. Na perspectiva desses autores, a avaliação supõe um processo de pesquisa e, portanto, toda avaliação de uma política, programas ou projeto é uma pesquisa avaliativa.

Partindo desse pressuposto, para que a pesquisa avaliativa seja operacionalizada é necessário escolher o modelo avaliativo. No entanto, assim como existe a diversidade dos conceitos de avaliação, também existe uma grande variedade de modelos, desenvolvidos ao longo do processo de amadurecimento e conformação do marco analítico-conceitual da avaliação de políticas, programas e projetos sociais e educacionais.

Assim, o Capítulo III, deter-se-á sobre as classificações de alguns dos diversos modelos de avaliação, provenientes tanto da área social quanto da área educacional, bem como apresentará modelos que não foram classificados pelos autores analisados.

CAPÍTULO III

Modelos de avaliação de políticas, programas e projetos

Conforme foi apresentado nos capítulos anteriores, o corpo teórico-metodológico da avaliação de políticas, programas e projetos sociais e educacionais passou por um processo de amadurecimento e sofreu influências de diversas áreas do conhecimento. Processo este que teve como resultado a grande diversidade conceitual e metodológica que observamos na literatura existente e um significativo número de modelos[62] de avaliação que são empregados nas atividades avaliativas .

Para um melhor entendimento do conteúdo abordado por este capítulo, é conveniente explicitar a definição de modelo de avaliação de políticas, programas ou projetos sociais e educacionais encontrada na literatura.

---

[62] Não temos a intenção de detalhar todos os aspectos teóricos e metodológicos dos modelos de avaliação apresentados neste capítulo. Faremos algumas referências aos aspectos conceituais, metodológicos e técnicos desses modelos, para situar o leitor em relação às classificações propostas pelos autores. Ver mais sobre os modelos em ALKIN (1972), SAUL (2000), BRIONES (1998), MARTINIC (1997), VIANNA (2000), BALLESTEROS (1996), FERNANDEZ (2000), WORTHEN (1997) e outros.

Segundo Briones (1998, p.31) um modelo de avaliação é:

> *"...un esquema o diseño general que caracteriza la forma de la investigación evaluativa que ha de realizarse, las técnicas o procedimientos para la recolección y análisis de la información, el conocimiento final que se desea obtener y los usuarios principales de los resultados del estudio".*

Martinic (1997, p.123), quando estabelece a definição de modelo de avaliação, enfatiza a relação de causalidade que, como foi visto no Capítulo II, deve estar presente na pesquisa avaliativa. Para este autor,

> *"Un modelo é entonces un modo conceptual de organizar las relaciones que el proyecto establece y que representan los resultados esperados del proyecto y los factores que parecen incidir en los mismos. Existen distintos tipos de modelos de acuerdo al tipo de complejidad de las relaciones que establecen".*

Completando sua definição de modelo, Martinic (1997, p.126) afirma que:

> *"Un modelo evaluativo es un conjunto de supuestos, relaciones hipotéticas y preferencias que representan o describen la ejecución y los resultados de los proyectos. Estos se construyen desde una perspectiva conceptual que relaciona la teoría o marco de referencia del propio proyecto con el conocimiento especializado del evaluador o de investigación acumulada sobre el tema".*

A partir dessas definições, pode-se inferir que modelo de avaliação é o conjunto de aspectos que englobam desde a concepção de avaliação até os instrumentos que serão utilizados na coleta, análise e divulgação das informações.

Na literatura observada, encontram-se diversos modelos de avaliação que possuem características metodológicas semelhantes ou distintas. Essas características, que englobam desde o planejamento do processo avaliativo passando pela de coleta de informação e finalizando com a utilização dos dados da avaliação, foram utilizadas como categorias e possibilitaram aos estudiosos agrupar alguns dos vários modelos de avaliação em três conjuntos distintos.

Em relação a esse agrupamento, pode-se observar três aspectos: o primeiro é que cada conjunto reúne um grupo de modelos com características semelhantes e possui um termo para classifica-lo. No entanto, parece não existir entre esses autores consenso no que diz respeito aos termos utilizados para classificar e referir-se a esses modelos. Isto pode ser visualizado no **Quadro 3.**

A partir desse quadro, pode-se inferir que um mesmo conjunto de modelos possui três termos diferentes para classificá-lo. Ou seja, autores referem-se a um conjunto de modelos utilizando uma nomenclatura diferente. Isto significa dizer que nas próximas seções serão apresentadas as classificações dos autores que, mesmo utilizando termos diferentes, referem-se a um mesmo conjunto de modelos.

Quadro 3

Classificação dos modelos de avaliação por 3 conjuntos

Conjunto 1	Conjunto 2	Conjunto 3
modelos: • **analítico** • **clássico** • **quantitativo**	**modelos:** • **global** • **alternativo** • **qualitativo**	**modelos:** • **analítico-global** • **clássico-alternativo** • **misto**

O segundo aspecto observado é que além da diversidade de termos para um só conjunto de modelos, também existem divergências entre alguns autores no que diz respeito à classificação de determinados modelos. Ou seja, um mesmo modelo de avaliação pode ser classificado por autores de maneiras diferentes (para um autor o modelo pode fazer parte do **Conjunto 2** e para um outro autor, o mesmo modelo é classificado como pertencente ao **Conjunto 3**). E o terceiro aspecto é que existem modelos citados que não foram classificados em nenhum dos conjuntos apresentados e que segundo os autores representam a tentativa de superar os modelos tradicionais de avaliação.

Considerando os três aspectos destacados acerca da classificação dos modelos de avaliação, o presente capítulo compreende quatro seções, sendo que as três primeiras seções foram organizadas a partir da configuração da linha vertical do **Quadro 3**. Assim, a primeira seção apresenta o conjunto de modelos classificados como clássico, analítico ou quantitativo. A segunda trata do conjunto de modelos denominados de global, alternativo ou qualitativo. E a terceira seção detém-se sobre o conjunto de modelos que são chamados de analítico-global, clássico – alternativo ou misto.

E, por fim, a quarta seção apresenta os modelos que não foram incorporados pelos autores estudados em nenhuma das classificações anteriores e as críticas aos modelos considerados por alguns autores como "tradicionais".

3.1. Conjunto 1 - modelos de avaliação: analítico, clássico ou quantitativo

Briones (1998) em sua obra direcionada à avaliação de programas sociais classifica os modelos de avaliação em três conjuntos: modelos analíticos, modelos globais e modelos analíticos-globais. Esta seção deter-se-á nos primeiros.

Assim, para esse autor, os modelos analíticos:

> *"...están relacionados con paradigmas epistemológicos explicativos y, a través de ellos, con posiciones realistas u objetivistas". En la evaluación, al igual que en otro tipo de investigaciones la finalidad del conocimiento consiste en proporcionar explicaciones de modalidades de funcionamiento del programa y de sus resultados (positivos o negativos) en término de alguna teoría o mediante la apelación a factores causales de diversa naturaleza (estructurales, demográficos, motivacionales, etc.). De ahí la necesidad de considerar el programa como integración de diversos componentes y dimensiones y estás, a su vez, conformadas por características o variables con las cuales sea posible determinar relaciones específicas (variables independientes y dependientes, en muchos casos* (BRIONES, 1998, p.31).

Dentre os modelos analíticos, encontramos o modelo CIPP (Contexto-Insumo-Processo-Produto), criado por Stufflebean para o campo educacional (mais especificamente para avaliar programas curriculares).

Esse modelo, segundo Briones (1998, p.32), define a avaliação como "*un proceso destinado a delinear, obtener y proporcionar informaciones útiles para el juzgamiento o ponderación de decisiones alternativas*".

O outro modelo analítico de avaliação apontado por Briones (1998, p.35), foi criado por ele e é chamado de modelo de Referentes Específicos. Este modelo é

> "*...un tipo de investigación que analiza el contexto, los objetivos, los recursos, el funcionamiento, la población y los resultados de un programa con el fin de proporcionar información de la cual se puedan derivar criterios útiles para la toma de decisiones en diversos niveles del programa*".

Ainda, segundo este autor, este modelo "*proporciona diversas orientaciones para la evaluación de componentes específicos de un programa, sea en el campo de la educación, de la salud, de la agricultura*" (BRIONES, 1998, p.35).

Na perspectiva adotada por esse modelo, a avaliação pode referir-se a todos ou a alguns componentes do programa que está sendo avaliado. O modelo de Referentes Específicos[63] proporciona diversas informações através da: avaliação de contexto, de recursos, do

---

[63] Para maiores detalhes, ver BRIONES (1998).

funcionamento, da população, de resultados, da efetividade, de eficiência e de eficácia[64].

Martinic (1997) em seu trabalho sobre a avaliação de projetos sociais, ao classificar os modelos de avaliação, utiliza a mesma definição proposta por Briones (1998) na sua obra publicada pela primeira vez em 1991. Porém, ao detalhar as distinções de cada um deles, enfatiza os aspectos relacionados ao tamanho e a cobertura dos projetos e programas. Para ele, os modelos analíticos de avaliação, além de possuírem as características apontadas por Briones (1998), são direcionados para avaliar grandes projetos e com ampla cobertura.

Assim, para Martinic (1997, p.127), os modelos analíticos:

> *"...tienen por objeto proyectos de amplia cobertura los que son abordados a través de distinciones analíticas y desagregación de sus componentes. Buscan relaciones de tipo causal y proponen explicaciones de los hechos observados".*

Como exemplo de um modelo analítico, esse autor cita o modelo CIPPP criado por Stufflebean, que através do estudo das relações causais entre os componentes de um programa, proporciona informações para auxiliar a tomada de decisão.

Santacana & Benito (1996), referindo-se especificamente à avaliação de programas educativos, usam um termo diferente do utilizado por Briones (1998) e Martinic (1997) para classificar modelos de avaliação. Para esses autores, os modelos chamados de analíticos são modelos clássicos de avaliação.

---

[64] Estes critérios serão tratados no Capítulo V.

Os modelos clássicos podem ser divididos em dois subconjuntos: os modelos que se preocupam em avaliar em que medida ocorreu à consecução de metas e objetivos e os que, independentemente do que se avalia e do foco de interesse dos *decision makers*, têm como principal objetivo auxiliar a tomada de decisão.

Esses autores citam como exemplo de modelo clássico o proposto por Tyler, na avaliação que realizou com o objetivo de averiguar a consecução de objetivos educacionais atingidos pelos alunos. Segundo ela,

> *"El primer modelo de evaluación de programas denominado de consecución de objetivos o metas, fue propuesto por Tyler* (1942), *para quien el objetivo de la evaluación consiste en la determinación de la medida en que se han logrado los objetivos pretendidos. Para evaluar un programa se precisa definir operativamente y organizar los objetivos, seleccionar las experiencias educativas, los instrumentos y elaborar un diseño pre-post de intervención"*. (SANTACANA & BENITO, 1996, p. 243)

Em relação ao segundo subconjunto dos modelos clássicos, Santacana & Benito (1996, p.243) afirmam que "*los modelos de toma o facilitación de decisiones enfatizan la provisión de información útil para controlar de forma adecuada el programa, racionalizando todas las tomas de decisión*".

Como exemplo desse segundo subconjunto os autores apontam o modelo CIPP[65] (avaliação de contexto, avaliação que se refere ao

---

[65] "Las decisiones educativas deben realizarse en los cuatro elementos que define el modelo. La evaluación del Contexto (lugar de aplicación, población, necesidades, problemas subyacentes a las necesidades, etc.) va dirigida a planificar el cambio y a fijar los elementos en relación a los cuales se van a juzgar los resultados. La evaluación referida al Input (análisis de recursos, diseños, presupuestos, estrategias alternativas, etc. Permite estructurar las actividades conducentes al cambio, que serán juzgadas bajo la perspectiva de Implantación. La evaluación del Proceso (identificación de barreras, inspección del proceso, etc.) permite ejercer el control de la planificación de un programa

Input (insumo), avaliação de processo e avaliação de produto), elaborado por Stufflebean.

Em relação a esse modelo clássico de avaliação, Santacana & Benito (1996, p.243) afirmam que "*en el modelo CIPP la evaluación de un programa es el proceso por el cual se delinea, obtiene y proporciona información útil para enjuiciar las alternativas de decisión*".

Esses autores apontam mais dois modelos que, segundo eles, pertencem ao conjunto de modelos clássicos que auxiliam a tomada de decisão. O primeiro é o modelo de Análise das Discrepâncias[66] proposto por Provus[67], que ressalta o fato de que a avaliação, ao abranger diferentes atividades, destaca a preocupação dos responsáveis pela tomada de decisão.

O segundo modelo apontado por Santacana & Benito (1996) como sendo também um modelo clássico — tomada de decisão — é o Modelo de *CSE - Center for the Study of Evaluation* (Centro para Estudos de Avaliação)[68], desenvolvido por Alkin.

---

y, finalmente la evaluación del producto (resultados) permite la toma de decisiones acerca de su continuidad, o de su finalización". (SANTACANA & BENITO, 1996, p. 243-244)

[66] Tradução do termo original em inglês, *Discrepancy Evaluation* encontrado em ALKIN (1972) e do termo em espanhol, *Análisis de las Discrepancias* em SANTACANA & BENITO (1996).

[67] A idéia básica dessa abordagem centra-se no fato de que, ao avaliarmos, fazemos comparações que somente são possíveis se padrões forem fixados, entendendo-se por padrão a descrição de alguma coisa - objeto, pessoa, programas, etc. - deva ser. Fixados os padrões, é necessário obter medidas do desempenho (performance); Portanto, nesta abordagem, avaliar é comparar desempenhos a padrões estabelecidos e essas comparações nos vão fornecer possíveis discrepâncias entre ambos – desempenhos e padrões. Avaliação seria, assim, o julgamento do valor de algo, estabelecido a partir da discrepância entre os padrões e a performance." (VIANNA 2000, p. 61)

[68] O modelo proposto por Alkin foi desenvolvido durante a construção dos pressupostos de sua teoria sobre avaliação. Na prática, os pressupostos desenvolvidos por Alkin, orientaram as atividades de pesquisa do Centro para o Estudo de Avaliação da *University of California at Los Angeles*, ficando conhecido como um modelo de avaliação conhecido como Modelo CSE (*Center for the Study of Evaluation*).

Segundo Alkin (1972), esse modelo de avaliação desenvolve-se em cinco áreas de necessidades de avaliação assim distribuídas: avaliação de sistemas, do planejamento do programa, da implementação do programa, para melhoria do programa e para certificação do programa.

Já Fernández (2000), em sua obra *"Evaluación y Educación: modelos e propuestas*, utiliza a classificação feita por Tejada (1989). Nela o autor classifica os modelos a partir sua característica principal que é a ênfase nas informações de natureza qualitativa. Como exemplo, o autor cita dois modelos, que também são apontados por Santacana & Benito (1996). No entanto, usa uma outra nomenclatura para referir-se aos modelos. O que para Santacana & Benito (1996) são modelos clássicos, para Fernández (2000) são modelos quantitativos.

No **Quadro 4**, pode-se visualizar melhor o conjunto dos modelos de avaliação propostas pelos autores e o primeiro aspecto destacado acerca dessa classificação, em que um único modelo de avaliação possui três nomenclaturas distintas.

Quadro 4

Classificação dos modelos de avaliação do Conjunto 1 por autor

Autor	Classificação	Modelo
Briones (1998)	**analítico**	• **CIPP (contexto, *input* (insumo), processo e produto)** • **Referentes específicos**
Martinic (1997)	**analítico**	• **CIPP (contexto, *input* (insumo), processo e produto)**
Santacana & Benito (1996)	**clássico**	• **CIPP (contexto, *input* (insumo), processo e produto);** • **Avaliação por objetivos;** • **Análise das discrepâncias;** • **ACE (avaliação dos centros de estudos)**
Fernández (2000)	**quantitativo**	• **CIPP (contexto, *input* (insumo), processo e produto);** • **Avaliação por objetivos**

De forma geral, a partir das classificações propostas pelos autores, pode-se inferir que independentemente dos termos utilizados (analítico, clássico ou quantitativo), o conjunto de modelos apresentado nesta seção possui em comum a dimensão decisional. Isto é, servem como apoio para a tomada de decisão. Além desta, pode-se destacar duas características que denotam a concepção de avaliação que o **Conjunto 1** representa. A primeira refere-se à hipótese de que existe uma relação de causalidade entre os componentes daquilo que está sendo avaliado. E a segunda, de que estes modelos privilegiam informações de natureza quantitativa

Além desses modelos, os autores classificam um outro conjunto, que mesmo tendo como objetivo auxiliar a tomada de decisão, possui características metodológicas distintas do **Conjunto 1** de modelos (analítico, clássico ou quantitativo). Esse segundo conjunto de modelos pode ser chamado de global, alternativo ou qualitativo e será objeto da próxima seção .

3.2. Conjunto 2 - modelos de avaliação: global, alternativo ou qualitativo

Briones (1998), além do conjunto de modelos apontado na seção anterior, sugere uma classificação para outro conjunto que possui características distintas dos modelos analíticos.

Esse segundo conjunto de modelos é chamado por ele de modelos globais. Na visão desse autor, esses modelos foram desenvolvidos com o intuito de compreender, interpretar e dar significado as ações e atividades que se desenvolvem no programa, com a perspectiva de enfatizar as informações de natureza qualitativa.

Sendo assim, para Briones (1998, p.31-32), nos modelos globais:

> *"Su enfoque es global, holístico, sin refere ncia a factores causales ni a variable particulares que pudieran ser sometidos a tratamientos estadísticos. Por el lado de sus paradigmas epistemológicos de base, muchos de estos modelos globales se relacionan con posiciones subjetivistas e idealistas."*

Como exemplo de modelos global, Briones (1998, p.54), aponta a Avaliação Iluminativa[69], criada por Parlett & Hamilton. Essa avaliação, foi "*modelada sobre las bases de los programas educacionales de pequeña escala*", com o objetivo de aumentar o conhecimento e compreensão dos

---

[69] "La Evaluación Iluminativa, como los enfoques similares que han crecido junto a ésta, es un intento de articular y practicar un auténtico enfoque alternativo. Por lo tanto, es *holístico*; los evaluadores presentan los variados contextos del programa evaluado y buscan retratarlo como proceso total, como una construcción organizativa individual que necesita ser simultáneamente examinado desde diferentes perspectivas, es sensible, porque los investigadores trabajan unidos para proporcionar todo lo relacionado con un programa en forma útil, que puede tomar diferentes formas y acercarse a diversas fuentes y métodos, pero que se diseña para interesar, para informar y agregar algo más a su comprensión (...) El enfoque que es heurístico porque el diseño de investigación sufre continuas redefiniciones a medida que aumentan los conocimientos básicos (...) Es *interpretativo* porque desarrolla descripciones estrechamente entrelazadas con ejemplos, con hechos y cifras significativas, se realiza una descripción general del programa que realza la inherente complejidad del mismo. También arroja luz (en consecuencia iluminativo) sobre facetas poco conocidas o que fueran olvidadas previamente y que resultan para su vida y carácter" (PARLETT, 1985, apud, BRIONES, 1998, p.54- 55).

programas, através de informações de natureza qualitativa e, assim, contribuir para a tomada de decisão.

Martinic (1997) segue a mesma classificação proposta por Briones (1998) no que diz respeito aos modelos globais de avaliação. No entanto, ressalta que esses modelos são mais indicados para a avaliação de projetos pequenos e que possuem uma cobertura mais focalizada.

Segundo esse autor, os modelos globais

> *"...tienen por objeto proyectos más pequeños o focalizados y se preocupan de realizar una interpretación amplia e profunda del conjunto de dimensiones y de interacciones que inciden en la marcha y resultados de la intervención. Analizan las probabilidades de os cambios y buscan comprender las relaciones que tienen las variables con dichos cambios."* (MARTINIC, 1997, p. 127)

Martinic (1997) cita três modelos globais de avaliação. O primeiro é a Avaliação Iluminativa, conforme já foi apontada por Briones (1998). O segundo é o modelo de Avaliação Sensitiva[70] (responsiva), proposta por Stake[71], que tem como objetivo descrever e julgar os pontos fortes e fracos de um programa educacional. E o terceiro modelo global , é a Avaliação Focalizada[72], desenvolvida por Patton para ser aplicada na área

---

[70] Tradução do termo original em espanhol, *Evaluación Sensitiva* em MARTINIC, (1997). Também encontrado na língua espanhola, *Evaluación respondente* em FERNANDÉZ (2000). Como tradução do termo original em inglês, *Evaluation Responsive* em STAKE (1975). Encontramos na língua portuguesa o termo Avaliação Responsiva.

[71] "Uma avaliação é responsiva na proposta de Stake, se sua orientação está voltada para as atividades do programa e menos para os seus objetivos; se ela é correspondente à necessidade de informação propostas por diferentes públicos e se diferentes perspectivas de valor são apresentadas no relatório sobre o sucesso ou fracasso do programa. O ponto central, numa avaliação, é o plano de observação para a coleta e o registro dos dados, que devem ser de qualidade para que em um retrato fiel das ações e reações possa ser apresentado e discutido". (VIANNA, 2000, p. 37)

[72] Tradução do termo *Evaluación Focalizada* em MARTINIC (1997). Também encontrada em Briones (1998) como *Evaluación de Utilización Focalizada*. Como tradução do termo original em inglês, *Utilization – Focused Evaluation*, utilizaremos o termo Avaliação

educacional, enfatiza o uso de dados de natureza qualitativa, que auxiliam na tomada de decisão.

Já Santacana & Benito (1996) afirmam que existem modelos alternativos de avaliação. Esses são considerados como modelos qualitativos ou naturalistas e têm como característica principal o fato de enfatizarem o papel dos beneficiários ou participantes como agentes fundamentais para a valorização do que está sendo avaliado.

Como exemplo de modelos alternativos, Santacana & Benito (1996) citam a Avaliação Responsiva (ou participativa)[73] proposta por Stake, direcionada para avaliar programas educacionais e, à Avaliação Iluminativa proposta por Parlet & Hamilton. Esse modelo, segundo os autores

> *"Pretende alejarse del uso de la metodología experimental a fin de poder captar el enfoque sistemático de la realidad o conjunto de hechos interrelacionados que intervienen a lo largo de la implementación de un programa"*. (SANTACANA & BENITO 1996, p.244)

Fernández (2000) também possui uma classificação para o conjunto apresentado anteriormente. Para ele, o modelo de Avaliação Iluminativa proposto por Parlett, anteriormente classificado por Briones

---

Focalizada. Esta avaliação será um pouco mais detalhada na próxima seção deste capítulo.

[73] Tradução do termo original em espanhol, *Evaluación responsable o participativa*. "El modelo de evaluación responsable o participativa de Stake (1983) se basa en el principio de que cada usuario de un programa (staff, ejecutores, usuarios, grupos, de sujeto afectados, etc.) lo percibe desde su propia perspectiva. El evaluador debe reflejar las múltiples realidades implicadas en un programa y su análisis se lleva a cabo mediante un continuo intercambio informativo entre los usuarios, con el objetivo de descubrir, investigar y proponer soluciones de forma participativa" (SANTACANA & BENITO, 1996 p.244).

(1998) e Martinic (1997) de modelo global e por Santacana & Benito (1996), de modelo alternativo, é um tipo de modelo qualitativo[74].

Também como um tipo de modelo qualitativo, Fernández (2000, p.32) cita a avaliação Crítica Artística desenvolvida por Eisner[75], que visa "*ofrecer una representación y valorización de la práctica educativa y sus consecuencias dentro de su contexto*".

No **Quadro 5,** o **Conjunto 2** — modelos de avaliação: global, alternativo ou qualitativo —, pode ser visualizado com sua respectiva classificação.

Quadro 5

Classificação dos modelos de avaliação do Conjunto 2 por autor

Autor	Classificação	Modelo
Briones (1998)	**global**	• **Avaliação iluminativa**
Martinic (1997)	**global**	• **Avaliação iluminativa;** • **Avaliação responsiva;** • **Avaliação focalizada**
Santacana & Benito (1996)	**alternativo**	• **Avaliação iluminativa;** • **Avaliação responsiva**
Fernández (2000)	**qualitativo**	• **Avaliação iluminativa;** • **Avaliação crítica artística**

---

[74] Como foi dito na seção anterior, FERNANÉZ (2000) utiliza a classificação feita por TEJADA (1989).

[75]"A preocupação de Eisner é especificamente, com os problemas relacionados à arte e à criação artística, e, devemos ressaltar, daí sua posição anticomportamental na linha tyleriana. A posição de Eisner é coerente com as suas idéias, pois diz, claramente, que ensinar exige *artistry* (ensinar para ele é uma arte) e que a educação é um processo que difere de um contexto para outro. Eisner na sua crítica a avaliação, baseia-se em dois conceitos: competência educacional (*educational connoiseurship*) e criticismo educacional" (VIANNA, 2000, p. 64).

No **Quadro 5**, observa-se também que no **Conjunto 2** um mesmo modelo de avaliação pode ser classificado pelos autores com termos diferentes.

Partindo da classificação e das características apontadas pelos autores, pode-se inferir três aspectos relacionados aos dois conjuntos de modelos de avaliação apresentados nas seções anteriores. O primeiro é que tanto o **Conjunto 1** (analítico, clássico ou quantitativo) quanto o **Conjunto 2** (global, alternativo ou qualitativo) foram desenvolvidos para auxiliarem a tomada de decisão e, portanto, possuem em sua concepção de avaliação a dimensão decisional.

O segundo refere-se ao tipo de informação que ambos privilegiam e utilizam para a tomada de decisão. O **Conjunto 1** representa os modelos que enfatizam informações de natureza quantitativa e o **Conjunto 2** de natureza qualitativa. Isso porém não implica que os modelos utilizem só um ou outro tipo de informação.

E o terceiro aspecto, relaciona-se especificamente ao **Conjunto 2** (global, alternativo ou quantitativo). Esses parecem que foram criados na perspectiva de superar os modelos do **Conjunto 1** (analítico, clássico ou quantitativo), vigentes na época em que cada foi desenvolvido.

Além do **Conjunto 1** e **Conjunto 2** apresentados anteriormente, os autores estabelecem um terceiro conjunto de modelos de avaliação chamados de analítico-global, clássico-alternativo ou misto. Na próxima seção serão apresentadas as características e a classificação desses modelos.

3.3. Conjunto 3 - modelos de avaliação: analítico-global, clássico-alternativo ou misto

Briones (1998, p.32), além dos conjuntos de modelos classificados de analíticos ou globais, aponta "*un modelo que se ubica entre los paradigmas explicativos e interpretativos*". Os modelos que possuem características dos dois tipos de modelos formam um terceiro conjunto que é chamado de analítico-global.

Como exemplo desse tipo de modelo, Briones (1998) cita a Avaliação de Utilização Focalizada[76]. Segundo ele, a avaliação criada por Patton e aplicada à área educacional, propõe o uso de explicações indutivas e qualitativas. Para esse autor,

> "*...el esquema de evaluación de utilización focalizada representa un intento de superar la oposición abierta entre los paradigmas que destacan la explicación basada en sistemas hipotético-deductivos, diseños rígidos de control de variables; el uso preferente o exclusivo de dados cuantitativos y los paradigmas holísticos que privilegian la descripción de tipo etnográfica, la interpretación comprensiva y el uso de dados cualitativos*" (BRIONES, 1998, p. 46).

---

[76] "La evaluación de utilización focalizada se inicia con la identificación y organización de las personas que toman las decisiones relevantes frente a un programa, como también de las personas que emplearán la información producida por la evaluación (no audiencias vagas y pasivas). El evaluador trabaja con esas personas con el fin de localizar (focalizar) las preguntas de evaluación relevantes". (BRIONES, 1998, p.47)

Para melhor compreensão do **Conjunto 3** de modelos e das variações de classificações encontradas, foi construído o **Quadro 6**.

Quadro 6

Classificação dos modelos de avaliação do Conjunto 3 por autor

Autor	Classificação	Modelo
Briones (1998)	analítico-global	• Avaliação focalizada
Martinic (1997)	N.C*	• N.C*
Santacana & Benito (1996)	clássico - alternativo	• avaliação "goal-free" (avaliação sem referência a objetivos)
Fernández (2000)	misto	• Avaliação "goal-free" (avaliação sem referência a objetivos); • Avaliação responsiva

***N.C: não classifica.**

No **Quadro 6**, pode-se perceber que Martinic (1997), mesmo utilizando a classificação e nomenclatura estabelecida por Briones (1998) para os **Conjuntos 1 e 2** (ver **Quadros 4 e 5**) não compactua com este ou qualquer um dos outros autores, no que diz respeito ao **Conjunto 3** de modelos de avaliação.

Após a classificação dos três conjuntos, foi elaborado o **Quadro 7**, que constitui a junção dos **Quadros 4, 5 e 6,** possibilitando, assim, a visualização das diferentes classificações para um mesmo modelo de avaliação, conforme foi destacado anteriormente.

O **Quadro 7** foi assim organizado: na primeira coluna são apresentados os modelos de avaliação que foram classificados. Na segunda coluna estão as classificações (representadas por conjunto) e na última coluna encontram-se os autores.

No **Quadro 7,** é possível observar as variações de classificação de um mesmo modelo de avaliação. Como exemplo, pode-se apontar o modelo de Avaliação Responsiva (Stake) que é classificado em dois conjuntos distintos. Para Santacana & Benito (1996) e Martinic (1997) esse modelo pertence ao **Conjunto 2** e para Fernández (2000), ao **Conjunto 3**.

Quadro 7

Distribuição dos modelos de avaliação por autor segundo sua classificação

Modelos	Classificação			Autor			
	Conj.1	Conj.2	Conj.3	Briones	Martinic	Santacana & Benito	Fernandez
CIPP	XX			XX	XX	XX	XX
Avaliação de referentes específicos	XX						
Avaliação por objetivos	XX			XX		XX	XX
Análise das discrepâncias	XX					XX	
ACE (avaliação de centros de estudos)	XX					XX	
Avaliação Iluminativa		XX		XX	XX	XX	XX
Avaliação responsiva		XX	XX	XX	XX	XX	XX
Avaliação Focalizada		XX	XX	XX	XX		
Avaliação crítica artística		XX					XX
Avaliação " goal-free" (sem referência a objetivos)			XX			XX	XX

Após as classificações propostas nesta seção, pode-se inferir que os modelos analítico-global, clássico-alternativo ou misto representam um conjunto de modelos que, além de serem criados com o objetivo de auxiliarem a tomada de decisão, mesclam as características dos modelos dos conjuntos apresentados nas seções anteriores.

Após as explanações pode-se fazer algumas considerações a respeito das classificações propostas pelos autores. A primeira é que cada autor ao classificar os modelos de avaliação destaca características que consideram mais importantes. A segunda é que os modelos, independentemente das características metodológicas e da classificação realizada pelos autores, têm como objetivo principal proporcionarem informações para auxiliarem na tomada de decisão.

A terceira consideração é que os estudiosos durante alguns anos propuseram modelos na perspectiva de superar algum tipo de modelo proposto anteriormente.

E quarta consideração é que além de existir uma variedade de modelos, esses podem ser classificados de diversas formas. Essa constatação nos remete a um pensamento de Scriven, sintetizado por Vianna (2000), que acreditamos ser muito atual e que de certa forma demonstra o porque da diversidade conceitual, metodológica e classificatória encontrada no campo da avaliação de políticas, programas e projetos sociais e educacionais.

> *"As várias áreas do conhecimento, especialmente as de estruturação mais recente, como é o caso da avaliação, estão sujeitas a disputas de natureza ideológica. A avaliação faz uso de diferentes modelos, frutos de concepções as mais diversas, e enfrenta complexos problemas metodológicos dificeis de consenso. Tudo isso traduz a intensa fermentação que existe em seu campo de atuação, o que é um sintoma bastante positivo, conforme Scriven* (1973)" (VIANNA, 2000, p.99).

3.4. Modelos de avaliação: superação e críticas

Além dos modelos[81] propostos nas seções anteriores e que foram classificados em três conjuntos distintos, os autores apontam outros modelos sem a preocupação de estabelecer uma classificação.

Martinic (1997) aponta alguns modelos que representam a tendência cada vez mais acentuada, de se criarem novas concepções que mesclem características dos modelos existentes ou que, a partir de um modelo, sejam feitas alterações e adaptações para se tornarem mais adequados a determinados programas sociais e educacionais. Em relação a esta possibilidade, o autor afirma que:

> *"... la tendencia de la investigación evaluativa es hacia la integración de los modelos más que hacia la separación y dicotomización de los mismos. En efecto, en la práctica misma el evaluador o los equipos de evaluación combinan enfoques metodológicos y modelos evaluativos en función del objeto o programas a abordar. Un evaluador 'analítico' puede utilizar enfoque cualitativo y global para ilustrar con profundidad aspectos particulares da realidad observada. A la inversa, quien ha realizado un estudio comprensivo e interpretativo puede formular hipótesis y buscar relaciones explicativas a comprobar estadísticamente y de modo extensivo".* (MARTINIC, 1997, p.127-128)

---

[81] Não temos a intenção de detalhar todos os modelos de avaliação apresentados nesta seção. Apenas serão destacados alguns aspectos considerados importantes para o entendimento do leitor. Para maiores detalhes ver BRIONES (1998), BID (1997), CARVALHO (1998), COHEN & FRANCO (1999), MARTINIC (1997), SAUL (1998, 2000), SIEMPRO (1999), SULBRANDT (1994), TRIPODI (1984).

Como exemplo dessa tendência, Martinic (1997) cita a avaliação criada pela OECD, que adotou os componentes centrais do modelo CIPP e desenvolveu um modelo de relações sistêmicas entre componentes do projeto para ser utilizado na área educacional.

Além desse modelo, o autor afirma que a partir das concepções e relações entre contexto, insumos, processos e produtos, proposto inicialmente por Stufflebean por meio do modelo CIPP, vários autores desenvolveram novos enfoques e modelos de avaliação. Como exemplo cita o modelo de Custo-Impacto desenvolvido por Cohen & Franco[82], e o modelo de Referentes Específicos[83], criado por Briones.

Um outro modelo encontrado é a Avaliação Participativa. Segundo Briones (1998, p.61), essa avaliação não possui formas preestabelecidas, constituindo uma modalidade que qualquer um dos modelos de avaliação pode vir a assumir. Ou seja, a avaliação participativa é uma opção metodológica que pode ser desenvolvida por qualquer investigação avaliativa. Para esse autor,

> *"...en la evaluación participativa la idea es que las personas que pertenecen a los programas tomen en sus manos las tareas que implica un estudio de ese tipo, con claras ventajas para ellos en lo que toca a lograr los propósitos buscados: aprendizaje grupal e individual, autorrealización, etc.; y para el programa, mejor determinación de los problemas, propuesta de soluciones realistas, mayor compromiso de los usuarios tanto con los objetivos de la evaluación como las soluciones, etc."* (BRIONES, 1998, p.59-60).

---

[82] Para maiores detalhes ver COHEN & FRANCO (1999).

[83] Este modelo de Avaliação já foi abordado neste capítulo e classificado pelo seu criador como sendo um modelo analítico de mvaliação (Conjunto 2). Para maiores detalhes ver BRIONES (1998).

Também direcionando seu enfoque para a Avaliação Participativa, pode-se encontrar Carvalho (1998) que, ao referir-se a esse Modelo de avaliação como o próprio nome diz, destaca o papel dos atores envolvidos em todo o desenvolvimento do programa.

Ela afirma que a avaliação participativa possui dois objetivos. O primeiro é a inclusão de todos os atores envolvidos na política, programas ou ação institucional no processo avaliativo. Dessa forma, "*a avaliação participativa permite aflorar as diversas verdades sobre os propósitos e resultados do programa ou ação institucional*" (CARVALHO, 1998, p.91).

E o segundo objetivo é desencadear um processo de aprendizagem social, através da democratização das informações e conhecimentos produzidos na investigação avaliativa, gerando uma maior transparência da ação pública, objeto da avaliação.

Um outro modelo pode ser apontado: a Avaliação Diferencial[84]. Esse modelo, segundo Tripodi (1984), foi criado na perspectiva de suprir as novas necessidades de avaliação, surgidas a partir das inovações e relevância no planejamento e desenvolvimento de programas.

Para tanto, esse modelo tem como objetivo avaliar diferentes estágios do programa[85], através de diferentes questões sobre os esforços, a eficácia e rendimento do programa, utilizando técnicas apropriadas aos objetivos da avaliação.

Ao sugerir o modelo de Avaliação Diferencial, Tripodi (1984, p.11) afirma que :

---

[84] Primeira edição em 1975.
[85] Segundo Tripodi (1984) os estágios que compõem um programa são: iniciação do programa, contato do programa e implementação.

> *"Fontes de financiamento, grupos profissionais, clientela e uma opinião pública mais sofisticada têm solicitado aos diretores de programas que demonstrem não apenas as necessidades às quais seus programas se dirigem, mas também sua contribuição para resolver ou aliviar problemas sociais. Além disso, têm sido constantemente levantadas questões acerca da administração dos recursos financeiros e da eficiência e custos relativos de programas alternativos. Para o administrador, a viabilidade, conveniência e adequação da avaliação de programas podem determinar o sucesso e/ou sobrevivência de um programa."*

Um outro modelo de avaliação, que não foi classificado nas seções anteriores, é a Avaliação Emancipatória , proposta por Saul (2000). Esse modelo nasceu no contexto brasileiro, e como diz a própria autora, destina-se à avaliação de programas sociais e educacionais e tem como compromisso principal *"fazer com que as pessoas direta ou indiretamente envolvidas em uma ação educacional escrevam sua própria história e gerem as suas próprias alternativas de ação"* (SAUL, 2000, p. 61).

A Avaliação Emancipatória, segundo Saul (2000, p.53), foi criada na perspectiva de superar

> *"A marca autoritária que caracteriza a avaliação educacional nos paradigmas[86] clássicos em uso, predominantemente expressa na relação pesquisador-pesquisado, associada à rejeição de seus pressupostos teóricos-metodológicos e acrescida do baixo poder de impacto dos resultados avaliativos no contexto educacional..."*

---

[86] "O termo paradigma está sendo utilizado na mesma acepção a ele atribuído por T. S. Kuhn (1970), um conceito abrangente com significado semelhante a visão de mundo, filosofia ou mesmo ortodoxia intelectual. Um paradigma prescreve áreas de problemas, métodos de pesquisa e padrões de solução e explicação aceitáveis pela comunidade acadêmica que o adota" (SAUL, 2000, p. 53).

Também na perspectiva de superar os modelos de avaliação que considera 'tradicional', Sulbrandt (1994) propõe a Avaliação Global ou Integral. Para esse autor, existem três formas de avaliações tradicionais ou clássicas[87]. A primeira é a avaliação de metas, que tem como objetivo medir os efeitos de um programa com relação às metas propostas. Segundo o autor

> *"De acordo com este esquema de avaliação, são atribuídos valores a um conjunto de metas e trata-se de estabelecer o êxito relativo do programa conforme o grau em que essas metas tenham sido cumpridas. As perguntas de pesquisa fazem referência direta às metas como uma maneira de assegurar que os resultados do estudo tenham influência prática, introduzindo mudanças nas políticas"* (SULBRANDT, 1994, p.370).

A segunda forma é a avaliação de impactos, que tem como "objetivo verificar o grau "*em que os grupos-alvos ou a comunidade foram ajudados e experimentaram uma mudança em sua situação devido a estes programas*" (SULBRANDT, 1994, p. 373).

E a terceira forma de avaliação tradicional apontada por esse autor, é a avaliação de processos que é realizada durante o desenvolvimento do programa e:

> *"Seu objetivo é destacar possíveis defeitos na elaboração dos procedimentos, identificar barreiras e impedimentos para sua implementação e, por meios dos registros de eventos e atividades fornecer a informação necessária para decisões sobre sua reprogramação. O uso adequado da informação produzida no programa permite introduzir mudanças durante o decorrer da sua execução"* (SULBRANDT, 1994, p.377).

---

[87] Não é intenção detalharmos essas avaliações. Apenas serão abordados alguns aspectos conceituais e metodológicos para melhor entendimento dessa seção. Essas avaliações serão tratadas no Capítulo IV.

Então, na perspectiva de superar essas avaliações, Sulbrandt (1994) sugere a Avaliação Global e Integral que, segundo o autor, não é propriamente um modelo, mas sim uma nova perspectiva de avaliação, com intenção de complementar o enfoque dos modelos tradicionais existentes. Segundo o autor,

> *"A avaliação global pretende complementar os enfoques tradicionais com outros renovados, de tal forma que, em conjunto, mostrem seus logros e junto com eles os diferentes fatores significativos do programa, sua implementação em um determinado meio institucional e social, bem como a ação dos diferentes grupos significativos vinculados a um programa a fim de complementar o que falta nas elaborações básicas, porém mantendo seu nível de análise e seu alto grau de rigor."* (SULBRANDT, 1994, p. 391).

Além dos modelos propostos pelos estudiosos das áreas social e educacional, pode-se encontrar modelos de avaliações desenvolvidos pelas agências supranacionais como o BID (1997) e a SIEMPRO (1999).

A primeira, desenvolve o Modelo Lógico de Avaliação, por meio do Método de Análise do Marco Lógico[88]. Este modelo segundo BID (1997, p.17) é *"un instrumento que ayuda a los diseñadores de proyectos a*

---

[88] "El marco Lógico es uno de los principales instrumentos utilizado hoy por las agencias de desarrollo para el diseño y planificación de proyectos. Concebido por la Agencia para el Desarrollo Internacional de los Estados Unidos (USAID) a fines de los años setenta, el marco lógico es una herramienta de trabajo com la cual un evaluador puede examinar el desempeño de un proyecto en todas sus etapas [...] La matriz del marco lógico está basado en dos principios básicos; primero, las relaciones lógicas verticales de causa-efecto entre las diferentes partes de un problema, que corresponden a los cuatro niveles horizontales que se relacionan las actividades (o insumos), los componentes (o productos), el propósito y el fin como el conjunto de objetivos jerarquizados del proyecto; segundo, el principio de la correspondencia (lógica horizontal), que vincula cada nivel de objetivos a la medición del logro (indicadores y medios de verificación) y las condiciones que pueden afectar su ejecución y posterior desempeño (o supuestos principales)" (BID, 1997, p. 15).

entender mejor la naturaleza de los problemas que están tratando de resolver".

A segunda agência, desenvolve um modelo de avaliação dentro do Marco de Gestão Integral[89] de Programas Sociais Orientada para Resultados. Nessa perspectiva, a avaliação é concebida como

> *"...un proceso permanente e continuo de indagación y valoración de la planificación, la ejecución y la finalización del programa social".*
> *Su finalidad es generar información, conocimiento y aprendizaje dirigidos a alimentar la toma de decisiones oportunas y pertinentes para garantizar, la eficacia y la calidad de los procesos, los resultados y los impactos de los programas, todo ello en función del mejoramento de las condiciones de vida de sus poblaciones beneficiarias"* (SIEMPRO, 1999, p.55).

Considerando a diversidade de modelos e classificações apontadas neste capítulo, pode-se fazer algumas considerações.

A primeira diz respeito às constantes referências que os autores contemporâneos fazem aos modelos de avaliação criados há mais de três décadas. Isto nos leva a aceitar que esses modelos foram tão importantes para o desenvolvimento do campo da avaliação de políticas, programas e projetos, que se tornaram modelos seminais.

A segunda consideração é que de forma geral, os modelos apresentados, ao serem desenvolvidos pelos seus criadores, surgem na perspectiva de superar o modelo de avaliação vigente em cada momento histórico e social.

---

[89] "Por gestión integral se entiende la articulación permanente de los procesos de planeación, ejecución y evaluación que se retroalimentan mutuamente y permiten el rediseño continuo de las acciones públicas" (WIESNER, 2000, p.5).

A terceira é que existe na atualidade uma tendência de se criar ou utilizar modelos heterogêneos de avaliação. Essa possibilidade pode ser pautada em duas vertentes: a de se tentar compactar em um modelo o que cada um dos modelos já produzidos têm de 'bom' ou a de se atender às necessidades sócio-educacionais e políticas-econômicas vigentes do momento. Isso porém não significa dizer que não existam práticas avaliativas pautadas exclusivamente em um dos modelos anteriormente citados neste capítulo.

A quarta consideração refere-se ao constante movimento de interseção entre os modelos de avaliação, provenientes especificamente da área educacional com os da área social. Ou seja, muitos dos modelos que foram inicialmente produzidos na área educacional de certa forma foram absorvidos e gradativamente influenciaram os estudiosos da área social e vice-versa.

E, por fim, a quinta consideração é que, além de existir uma variedade de modelos com suas respectivas características, também pode-se observar uma diversidade de classificações e nomenclaturas que são utilizadas pelos autores, para se referir a esses modelos.

Essa diversidade, no entanto, também pode ser encontrada ao nos depararmos com os tipos de avaliação encontrados na literatura.

Partindo dessa premissa, o Capítulo IV abordará os diversos conceitos, classificações e nomenclaturas referentes aos tipos de avaliação de políticas, programas e projetos sociais e educacionais.

CAPÍTULO IV

Tipos de avaliação de políticas, programas e projetos

Como foi visto nos capítulos anteriores, no campo da avaliação de políticas, programas e projetos[90] sociais e educacionais, existe uma diversidade teórico-metodológica proveniente talvez da maneira como foi construído o seu campo de conhecimento, conforme foi demonstrado no Capítulo I.

A diversidade apontada acerca dos conceitos e das classificações dos modelos de avaliação, apresentados respectivamente nos Capítulos II e III, também pode ser observada no que diz respeito aos tipos[91] de avaliação. Em relação a essa diversidade podemos destacar três aspectos. O primeiro refere-se ao número significativo de tipos de avaliação. O segundo diz respeito à vasta nomenclatura para cada tipo de avaliação. Ou seja, um mesmo tipo de avaliação pode ter duas ou mais

---

[90] Para que o texto não se torne cansativo, durante o seu desenvolvimento faremos a utilização dos termos avaliação de programas e projetos, avaliação de programas ou avaliação de projetos.

[91] Não temos a intenção de detalharmos todos os aspectos teórico-metodológicos dos tipos de avaliação. Nesta seção serão abordados alguns aspectos conceituais e metodológicos, com o objetivo de servirem como ponto de análise para demonstração das semelhanças e diferenças entre os diversos autores.

nomenclaturas para significar a mesma coisa. E o terceiro aspecto relaciona-se com as semelhanças ou diferenças entre os autores em relação aos aspectos teórico–metodológicos dos tipos de avaliação.

Mesmo assim, é possível afirmar que os autores constróem suas tipologias a partir de quatro critérios. O primeiro serve para classificar as avaliações a partir de quem a realiza[92]. O segundo critério é utilizado pelos autores para estabelecer as avaliações segundo sua função[93]. O terceiro critério serve para classificar as avaliações a partir do momento em que o programa ou projeto é avaliado. E o quarto critério é utilizado pelos autores para estabelecer as avaliações a partir dos aspectos ou atividades que serão avaliados. É importante destacar que esses tipos de avaliações, segundo os autores, não são excludentes e sim complementares.

Este capítulo, no entanto, deter-se-á sobre as avaliações que foram classificadas a partir dos dois últimos critérios: quanto ao momento e quanto ao aspecto/atividade que é objeto da avaliação. Esta escolha deve-se ao fato de que nesses tipos de avaliação foi encontrada uma maior diversidade, conforme os aspectos destacados anteriormente.

É importante, porém, destacar que alguns dos autores analisados, ao estabelecerem os tipos de avaliação, nem sempre utilizam os critérios apontados anteriormente ou, ainda, recorrem a outros critérios. Mesmo assim, é possível, diante das definições de avaliação

---

[92] Utilizando este critério, as avaliações podem ser classificadas como: interna, externa ou mista. Para maiores detalhes ver BRIONES (1998) e AGUILAR & ANDER-EGG (1995), dentre outros.

[93] As avaliações classificadas a partir desse critério são chamadas de somativa e formativa. No entanto, os autores também utilizam essas avaliações para se referir a avaliações que ocorrem durante ou ao final de um programa. A avaliação somativa e formativa serão abordadas nas próximas seções, quando trataremos dos tipos de avaliação classificados a partir do momento em que se avalia.

propostas por eles, classificar as avaliações dentro de um dos três grupos estabelecidos anteriormente.

Partindo dos dois critérios e das definições dos tipos de avaliação, foi possível organizar três grupos conforme pode ser observado no **Quadro 8.**

Considerando que os autores se referem às diversas atividades que podem ser avaliadas, utilizando muitos termos e conceitos que ora se assemelham ora se diferenciam, os três grupos que representam os tipos de avaliação foram assim denominados: **Grupo 1**— tipos de avaliação que podem ser realizados antes do início do programa, **Grupo 2** — tipos de avaliação que podem ser realizados durante a execução do programa e o **Grupo 3** — tipos de avaliação que podem ocorrer após o término do programa.

No **Quadro 8,** encontram-se na primeira coluna os critérios utilizados para classificar os tipos de avaliação. A segunda coluna agrupa para cada critério (momento em que se avalia e o aspecto/atividade que é avaliado) as avaliações que pertencem ao **Grupo 1**. Esta mesma distribuição é valida para o **Grupo 2 e Grupo 3** das colunas posteriores. Para melhor entendimento da organização desse quadro, é necessário destacar que:

- os tipos de avaliação foram agrupados sem a preocupação de indicar os autores, pois isso será realizado nas próximas seções;

- de fato existem avaliações que possuem o mesmo termo ou são classificadas da mesma maneira por autores diferentes. Tendo em vista essa realidade, essas avaliações só aparecem uma única vez no **Quadro 8**, evitando-se assim uma repetição desnecessária.

Quadro 8

Nomenclatura dos tipos de avaliação por 3 Grupos segundo seus critérios

Grupos critérios	Grupo 1	Grupo 2	Grupo 3
Critério A: **momento em que se avalia**	avaliação *ex-ante;* avaliação inicial; avaliação de pré - decisão; análise; avaliação formativa; avaliação antes.	avaliação de processo; acompanhamento; avaliação formativa; evaluación intermédia; avaliação parcial; avaliação da gestão; avaliação contínua; avaliação concomitante; evaluación de la ejecución; monitoramento; avaliação de processo ou avaliação formativa ou avaliação de eficácia.	avaliação para a certificação do programa; avaliação de resultados (custo-benefício); avaliação somativa; avaliação (resultados ou impactos); evaluación final; avaliação *ex-post;* evaluación de la finalización; evaluación de pos-terminación o *ex-post;* evaluación terminal o *ex-post;* avaliação de resultados (custo-benefício).
Critério B: **aspecto ou atividade avaliada**	avaliação do planejamento; estudo diagnóstico; avaliação da programação; análisis de objetivos; pre-evaluación y especificación del programa; avaliação do plano e conceitualização; avaliação do estudo e pesquisa, avaliação do diagóstico; evaluación del diseño y conceptualización.	acompanhamento e monitoramento (avaliação da eficiência); avaliação formativa (processo ou eficácia); evaluación concurrente; avaliação concorrente; monitoramneto ou avaliação cocomitante; monitoreo y seguimiento; evaluación de la cobertura; evaluación de la productividade del programa; avaliação diagnóstica; avaliação da implementação; avaliação de processo; evaluación de objetivos internos; evaluación de la ejecución del proyecto; evaluación del perfil estructural y funcional; evaluación de los procedimientos; evaluación de la perspectiva de los beneficiários; avaliação do esforço; avaliação da produtividade; avaliação do ambiente organizacional; avaliação do rendimento pessoal; evaluación operativa (intermedia o ejecución).	evaluación de impacto o somativa; evaluación de resultados; evaluación de la coordinación; evaluación de los efectos; evaluación del producto(evaluación del a eficacia, efectividad, evaluación del impacto); continuidad y justificación del programa; evaluación del impacto; avaliação da eficácia; avaliação da eficiência (rentabilidade económica); avaliação do processo de gestão; avaliação de objetivos e metas; avaliação comparativa; avaliação custo-benefício; opinión de los beneficiarios; avaliação do fim do programa.

Este capítulo compreende três seções. A primeira tratará dos tipos de avaliação que são realizadas antes do início do programa ou projeto. Aqui se encontra a avaliação *ex-ante*, seus sinônimos e os aspectos relacionados ao que os autores afirmam poder avaliar durante este período.

A segunda seção refere-se aos tipos de avaliação que podem ser realizados durante a execução do programa. Nessa seção será abordada a avaliação de processo, a avaliação parcial, etc., e o que se pode avaliar nesse período.

E, por fim, a terceira seção abordará as avaliações que podem ocorrer ao término do programa. Nesta seção, portanto, encontra-se a avaliação chamada de *ex-post*, avaliação final, etc., e quais aspectos podem ser objetos de avaliação .

4.1. Avaliação antes do início do programa

Como foi dito anteriormente, em relação aos tipos de avaliação, foram encontradas algumas diferenças e semelhanças entre os autores. Essa realidade pode ser observada no que diz respeito às avaliações que foram reunidas no **Grupo 1** (ver **Quadro 8**).

Esta seção, tratará das avaliações que ocorrem antes do início do programa, partindo dos critérios avaliação quanto ao momento em que se avalia e quanto aos aspectos e atividades que são objetos de avaliação. Tem como objetivo apresentar os diversos conceitos e termos que são

utilizados, apontando em que medida os autores se assemelham ou divergem.

Dando início à nossa análise, podemos apontar Alkin (1972). Este autor, dentro do seu Modelo CEA[94] (Centro para o Estudo da Avaliação), afirma que diferentes tipos de decisões podem requerer tipos diferentes de avaliação. Mesmo sem fazer referência explicita a avaliação que pode ocorrer antes do início do programa, esse autor aponta para a necessidade de avaliar alguns dos seus aspectos. Um desses aspectos é o planejamento do programa.

Para Alkin (1972), a avaliação do planejamento do programa[95] tem como objetivo promover informações sobre o planejamento de diversos programas para, posteriormente, compará-los e, desta forma, obter dados suficientes para escolher o programa que possa satisfazer às necessidades educacionais específicas, anteriormente detectadas.

Ander-Egg (1990), da mesma forma que Alkin (1972), não utiliza um termo específico para referir-se à avaliação que ocorre no início do programa. Porém, aponta para a necessidade de realizar avaliações sobreos aspectos do programa antes de executá-lo e que fazem parte da área de coerência interna do programa. São elas: avaliação do estudo (diagnóstico) e avaliação da fase de programação.

Ander – Egg (1990) afirma que essas duas avaliações dentro da área de coerência interna coletam e avaliam as informações acerca da sustentabilidade e factibilidade do programa ou projeto em si mesmo, tendo em vista a solução do problema que o suscita.

---

[94] Tradução do termo original em inglês, *Model CSE (Center for the Study of Evaluation)*.
[95] Tradução do termo original em inglês *'program planning'* em, ALKIN (1972).

Ballesteros (1996), assim como Alkin (1972) e Ander-Egg (1990), não utiliza um termo específico para referir-se à avaliação que ocorre antes do início do programa, mas aponta para aspectos do programa que devem ser focos de avaliação.

Para apontar as avaliações desses aspectos a autora parte da lógica implícita no *ciclo de intervención social*[96], proposto por Palumbo (1987), e afirma que a avaliação de alguns aspectos do programa na fase de programação é fundamental para o seu desempenho, pois ao realizá-la está se avaliando a suficiência e adequação do programa de intervenção social.

Como exemplo, Ballesteros (1996) cita as avaliações chamadas de *analisis de objetivos, pré-evaluacion y especificación del programa*, que são feitas para saber se os objetivos e os meios estabelecidos são pertinentes, para solucionar os problemas que serão alvos da intervenção através do programa.

Considerando, portanto, os conceitos propostos por Ander-Egg (1990) e Ballesteros (1996), pode-se inferir que as avaliações denominadas de estudo (diagnóstico), avaliação da programação, *análisis de objetivos* e a *pré-evaluacion y especificación del programa*, tem como objetivo, de modo geral, avaliar a pertinência e viabilidade do programa, tendo em vista a solução dos problemas para o qual foi elaborado.

Já Aguilar & Ander- Egg (1995), diferentemente dos autores até aqui mencionados, utilizam os critérios — momento em que se avalia e aspecto e atividade a ser avaliada — para construir sua tipologia de avaliação.

---

[96] Na lógica do *ciclo de intervención social*, é necessário que os planejadores tenham informações a cerca das diversas fases desse ciclo. BALLESTEROS (1996).

Para esses autores a avaliação que ocorre antes do início do programa é chamada de avaliação *ex-ante*, tendo como sinônimos os termos: avaliação antes, avaliação inicial ou pré-decisão.

Para Aguilar & Ander–Egg (1995, p.41) a avaliação que ocorre antes da execução do programa "*é feita antes de tomar a decisão de empreender um programa ou projeto, ou de aprovar a sua realização*". Esses autores afirmam que a avaliação *ex-ante*

> "*Consiste em avaliar o programa ou projeto em si mesmo, mediante a estimação crítica de sua pertinência, viabilidade e eficácia potencial, com a finalidade de proporcionar critérios racionais sobre a conveniência ou não de levar a cabo um programa ou projeto. Geralmente, trata de "comparar a eficácia ou rentabilidade econômica das diferentes ações para alcançar os objetivos desejados".*(Pineault & Daveluy, 1987)" (AGUILAR & ANDER- EGG, 1995, p. 41).

Essa avaliação tem como aspecto a ser avaliado o plano e a conceitualização do programa durante sua elaboração, "*pelo qual se julga a pertinência formal e potencial de um programa levando em conta a coerência interna entre seus diversos componentes*" (AGUILAR & ANDER-EGG, 1995, p.51).

A avaliação do plano e conceitualização do programa abordam três aspectos: avaliação do estudo e pesquisa, avaliação do diagnóstico e, propriamente, a avaliação do plano e concepção do programa[97]. Com isso o autor afirma que, para o programa atingir seus objetivos, deve-se partir

---

[97] "Embora tenhamos denominado, de modo geral, estes três aspectos que podem ser avaliados do "plano e conceituação do programa", em sentido estrito, avaliar o plano e concepção do programa é avaliar o modelo de intervenção social ao qual se recorre para resolver determinada situação problemática. Isto compreende em seus aspectos mais formais as estratégias de ação, as atividades concretas que foram propostas e, sobretudo, os métodos, técnicas e procedimentos que se pretende utilizar". (AGUILAR & ANDER–EGG, 1994, p.52).

primeiramente do estudo e diagnóstico da situação–problema e só após essas análises, partir para elaboração de um programa que possa intervir em tal realidade.

Faria (1998), sem se preocupar com critérios para estabelecer a sua tipologia de avaliação, traz uma conceituação da avaliação que ocorre antes do início do programa que é bem semelhante a dos autores mencionados anteriormente.

Para a autora, essa avaliação é chamada de avaliação *ex-ante* pois ocorre antes e durante a elaboração do programa e

> *"...consiste no levantamento das necessidades e estudos de factibilidade que irão orientar a formulação e o desenvolvimento do programa".Inclui a definição de seus objetivos, âmbito de aplicação, caracterização dos beneficiários e suas necessidades"* (FARIA, 1998, p.4).

Já Cohen & Franco (1999) utilizam como critério para estabelecer sua tipologia, o momento em que é realizada a avaliação. Dentro dessa especificação, os autores afirmam que a avaliação que ocorre antes do início do programa é chamada de avaliação *ex-ante.*

No entanto, esses autores, mesmo utilizando o termo *ex-ante* para se referir à avaliação que ocorre antes do início do programa e da mesma forma que Aguilar & Ander-Egg (1995) e Faria (1998), percorrem um caminho diferenciado, pois se detêm sobre os aspectos mais econômicos do programa.

Eles afirmam que a avaliação *ex-ant,* tem suas raízes na economia e que as metodologias mais adequadas para realizar esse tipo avaliação são: a análise de custo–benefício e análise-custo efetividade. Esta última é mais apropriada para avaliar os projetos sociais.

Para Cohen & Franco (1999, p.108) a avaliação *ex-ante*:

> *"Tem por finalidade proporcionar critérios racionais para uma decisão qualitativa crucial: se o projeto deve ou não ser implementado. Também permite ordenar os projetos segundo sua eficiência para alcançar os objetivos perseguidos".*

Partindo da citação acima, infere-se que para Cohen & Franco (1999) a avaliação *ex-ante* pode ser utilizada tanto para a averiguação das reais condições do projeto em atingir os seus objetivos quanto para auxiliar na escolha entre dois ou mais projetos que foram elaborados para solucionar um problema.

Martinic (1997), assim como Cohen & Franco (1999) também afirma que a avaliação *ex-ante* pode ser utilizada para avaliar a viabilidade do projeto durante sua fase de concepção e assim proporcionar critérios técnicos para a tomada de decisão a respeito de qual projeto é mais adequado para ser implementado. Esse tipo de avaliação tem como objeto a ser avaliado o plano do projeto no qual se encontram todas as suas diretrizes.

Martinic (1997), além de destacar as mesmas metodologias citadas por Cohen & Franco (1997), também sinaliza que a análise de custo–efetividade é a mais adequada para a avaliação de projetos sociais.

Comparando Martinic (1997) e Cohen & Franco (1999), no que diz respeito à utilização da avaliação *ex-ante*, é possível afirmar que eles partem do pressuposto de que existe uma variedade de projetos que foram elaborados para solucionar um problema e que, portanto, torna-se necessário a realização de avaliações para averiguar qual dos programas é o mais viável e coerente para ser implementado.

Tenório (2000) não utiliza critérios para estabelecer os tipos de avaliação, mas se refere a uma atividade avaliativa que ocorre antes do programa ser posto em prática. Ao referir-se a esse tipo de avaliação, o autor utiliza o termo análise. Para ele, a análise é a primeira fase do processo de avaliação de projetos e tem como objetivo identificar sua viabilidade, além de auxiliar na escolha entre vários projetos que foram elaborados.

A análise, segundo Tenório (2000, p. 22-23),

> *"...é um conjunto de atividades que tem por objetivo principal avaliar previamente o projeto que se pretende implantar,permitindo que a comunidade verifique antes mesmo de começar a desenvolver um determinado projeto, se será possível realiza-lo totalmente, tendo como ponto de partida o que foi inicialmente planejado".*

Para realizar essa análise, Tenório (2000, p. 26-27) afirma que a técnica mais apropriada é a análise de custo–benefício. Para ele, esse tipo de análise:

> *"Serve para a realização da análise tanto sob a ótica do privado, em que se espera obter algum lucro com a realização de determinado projeto, como sob a ótica social, em que o aspecto de lucro não é privilegiado, sendo necessário, portanto, algumas adaptações no método para que se possa utilizá-lo para avaliar projetos".*

Em relação a esse autor, pode-se observar que no que diz respeito ao termo que utiliza para referir-se à avaliação que ocorre antes do início do programa, ele se diferencia dos demais. Ou seja, enquanto Cohen & Franco (1999), Aguilar & Ander-Egg (1995), Martinic (1997) e Faria (1998), utilizam o termo avaliação *ex-ante*, ele usa o termo análise.

Porém, ao definir o que para ele é análise, Tenório (2000) vai ao encontro das definições de avaliações aqui apresentadas, inclusive no que se refere às metodologias ou técnicas utilizadas para realizar essa análise.

Em se tratando do tipo de avaliação que se realiza antes do início do programa, a literatura desenvolvida pelos organismos supranacionais, como o BID (1997) e o SIEMPRO (1999) trazem também algumas diferenças e semelhanças entre si.

Assim, segundo BID (1997) dentro da perspectiva do Marco Lógico, a avaliação que ocorre antes do início do projeto, mais especificamente durante sua preparação, é denominada de avaliação *ex-ante*.

Segundo o BID (1997, p. iv)[98] essa avaliação é:

> *"Una determinación o estimación de necesidades de avaluación, aplicada en la fase ex-ante del ciclo de evaluación, que incluye estudios de factibilidad, la identificación de los objetivos del proyectos y todas aquellas otras funciones realizadas antes de comenzarlo".*

[98] Esta citação foi retirada do glossário do BID (1997).

Para o BID (1997, p. 11), a avaliação ex-ante tem uma função formativa[99]. Ou seja, tem como função (papel) "melhorar o diseño y la planificación del proyecto estableciendo así las condiciones para as actividades de evaluación que se deben llevar a cabo a lo largo del ciclo del proyecto, o sea, la "evaluabilidad" del proyecto" e sugerir melhores enfoques para futuros projetos.

Para desenvolver a avaliação *ex-ante* o BID (1997) aponta quatro ferramentas de avaliação assim denominadas: *Marco Lógico, instrumentos de análisis económico, financeiro e institucional, diagnóstico de evaluabilidad e listado de datos de referência*[100].

Já o SIEMPRO (1999) na perspectiva da *Gestión Integral de Programas Sociales*, afirma que a avaliação se inicia desde o momento em que se detecta os problemas que darão origem ao projeto social até após sua execução. Partindo desse pressuposto, afirma que durante a planificação do programa deve se realizar uma avaliação que tem como objetivo:

> *"...la indagación y la valoración de todos os aspectos relacionados con el proceso de definición, formulación y establecimiento de viabilidad inicial de un programa. Esta evaluación en términos generales, sirve para corregir la planificación y la formulación del programa, para replanificarlo y para conocer si es viable la ejecución del programa planificad."* (SIEMPRO, 1999, p. 59).

---

[99] Este conceito foi proposto por Scriven em 1967. Este autor estabeleceu as diferenças entre dois tipos de avaliação a partir das funções que podem desempenhar no processo decisório. A primeira é a avaliação formativa, cuja função está diretamente relacionada com a decisão de desenvolver um programa, sua modificação ou ainda sua revisão. E a segunda é a avaliação somativa, cuja função está diretamente relacionada com a decisão de acabar, continuar, aplicar novamente o projeto que está sendo avaliado. Esses dois tipos de avaliação foram muito difundidos no meio acadêmico e têm sido utilizados tanto na área educacional como na área social. Essa realidade será vista durante o desenvolvimento deste capítulo.

Para tanto é utilizada uma ferramenta intitulada de avaliação *ex-ante* que

> *"...permite realizar una revisión integral del programa propuesto antes de iniciar su ejecución.*
> *La evaluación ex-ante consiste en un análisis de la formulación del programa para conocer si los elementos operativos del mismo (actividades, tareas, recursos, presupuestos, etcétera) resultan adecuadamente consistentes en /los documento/s del programa propuesto"* (SIEMPRO, 1999, p.74).

Em relação à avaliação *ex-ante*, SIEMPRO (1999, p.74) afirma que:

> *"Esta herramienta también indaga la viabilidad de los componentes y de las actividades planificadas, al mismo tiempo que propone alternativas para aquellas que tengan menor nivel de viabilidad y selecciona la opción de inversión más eficiente. De esta manera se busca racionalizar la intervención social".*

Em relação à avaliação da viabilidade do programa, SIEMPRO (1999) afirma que isto só deve acontecer quando for detectado algum problema relacionado com seus componentes e atividades. Essa avaliação é realizada através da: análise da viabilidade econômico-financeira (custo-efetividade), análise da viabilidade política, análise da viabilidade institucional e análise da sustentabilidade[101].

Ao fazermos uma comparação entre os pressupostos do BID (1997) e o SIEMPRO (1999), pode-se destacar que mesmo o primeiro denominando a avaliação *ex-ante* como um tipo de avaliação e o segundo como uma ferramenta, notar-se-á que, de forma geral, ambos partem do

---

[100] Para maiores detalhes ver BID (1997).
[101] Para maiores detalhes ver SIEMPRO (1999).

pressuposto de que a avaliação realizada antes da execução do programa tem como objetivo principal a melhoria do planejamento e, conseqüentemente, do desenvolvimento do que está sendo avaliado.

Viedma (1996), utilizando o critério momento em que se avalia para estabelecer sua tipologia de avaliação, aponta para um tipo de avaliação que ocorre antes do início do programa. Essa avaliação é chamada de avaliação formativa e tem como objetivo melhorar a planificação do programa.

Em relação a esse tipo de avaliação, o autor, afirma que:

> *"La evaluación formativa consiste en proporcionar información sobre cómo mejorar un programa ya en proceso, o la gestión del mismo. También consiste en producir datos que ayudarán a diseñar y desarrollar un programa [...] la evaluación formativa tiene, por tanto, mucho que ver con la planificación del proyecto. En este sentido, cabe preguntar-se si el proyecto debería o no ponerse en marcha; si los materiales de intervención son los adecuados..."* (VIEDMA, 1996, p.3).

A utilização do termo avaliação formativa[102], no entanto, aponta-nos para a necessidade de rever alguns dos aspectos propostos por Scriven em 1967[103]. Esse autor estabeleceu as diferenças entre metas e papéis (usos ou funções) da avaliação. E a partir dos papéis (funções) que a avaliação pode desempenhar no processo decisório, foram estabelecidos dois tipos de avaliação: a formativa e somativa[104].

---

[102] A avaliação formativa é comumente utilizada para referir-se à avaliação que acontece durante a execução do programa, que será tratada na próxima seção deste Capítulo.
[103] O texto utilizado em nossa análise foi publicado em 1972, mas os pressupostos de Scriven foram desenvolvidos no ano de 1967.
[104] Este tipo de avaliação será abordado mais detalhadamente na última seção deste capítulo, quando forem abordadas as avaliações que ocorrem ao término de um programa ou projeto.

Ambas têm como meta fazer o julgamento, estimar o mérito e valor de um programa. Porém a avaliação formativa tem como função promover a melhoria do programa e está relacionada com a decisão de modificar ou revisar (aprimorar o que está sendo implementado) e a avaliação somativa tem como função promover informações que afetam as decisões de continuar ou acabar um programa, projeto, curso, atividades, etc.

As diferenças entre essas avaliações, como pode ser visto, foram apontadas a partir das funções que podem assumir e não a partir do momento em que acontecem, mesmo se tendo conhecimento de que em cada fase de um programa, avalia-se com um propósito.

Partindo desse pressuposto, pode-se inferir que Viedma (1996) aponta a avaliação formativa como um tipo de avaliação que ocorre antes do início do programa, relacionando-a não só com o critério do momento em que se avalia para construir sua tipologia, como também recorrendo ao sentido estrito (função ou papel) que esse tipo de avaliação pode desempenhar, dentro do processo decisório de uma determinada fase do programa.

Retomando a questão das avaliações que ocorrem antes do início do programa, pode-se apontar a tipologia proposta por Lara 1996). Este autor estabelece "*una topología*[105] *de evaluaciones según centren sus objetivos*".

No entanto, em nosso entendimento, o critério utilizado por Lara (1996), é equivalente ao critério aspecto/atividade que pode ser

---

[105] As avaliações classificadas por LARA (1996) e que serão abordadas em outras seções deste capítulo são: *evaluación del proceso (aspectos de la ejecución y funcionamento, implementación, monitorización y seguimento), evaluación del producto (impacto, eficiencia, eficácia y efectividad) e evaluación de la pertinencia..*

avaliado, que já foi apontado no inicio deste capítulo e que serviu inclusive para organizar esta seção.

Considerando, portanto essa correlação, Lara (1996) aponta para o aspecto *diseño y conceptualización* que deve ser avaliado, antes do programa ser executado. A avaliação do *diseño y a conceptualización del programa*, segundo essa autora,

> *"...el análisis evaluativo se centra en determinar si existen razones para esperar que el programa, tal y como ha sido diseñado, será capaz de alcanzar a la población objetivo, sin que existan problemas importantes de cobertura; si existen razones para esperar que el modelo y las estrategias de intervención en que se basa el programa serán los adecuados en el contexto concreto en el que el programa se aplica; si la secuencia de actividades del porgrama es correcta y se han previsto los posibles desajustes, retrasos y obstáculos que pueden llegar a comprometer el logro de los objetivos del programa, y si los recursos asignados al programa son los adecuados"* (LARA, 1996, p.110).

Lobo (1998) em relação à avaliação *ex-ante*, não a conceitua como os autores até aqui apresentados, mas faz uma crítica, destacando a importância da adaptação desse tipo de avaliação para os programas de cunho social. Segundo ela, a avaliação *ex-ante*

> *"...sempre foi muito estimulada e induzida nos programas financiados pelos organismos multilaterais de financiamento, principalmente aqueles voltados à infra-estrutura econômica e ao desenvolvimento urbano. Trata-se de análises custo-benefício, de custo- efetividade, das taxas de retorno econômico dos investimentos previstos* [...] *As fórmulas e metodologias existentes para uma avaliação ex-ante, não condizem com as questões tratadas em áreas sociais. Problemas de educação e saúde, por exemplo dificilmente se submetem a categorias de análise e previsão de resultados tais como sugeridos por estas metodologias"* (LOBO, 1998, p.76)".

A autora afirma que as metodologias e técnicas utilizadas na avaliação *ex-ante*, provenientes da área econômica, não são as mais adequadas para avaliar os programas e projetos sociais e educacionais, devido as suas especificidades.

Como proposta para superar esta realidade, Lobo (1998) afirma que é necessário utilizar uma "análise de viabilidade política, econômico-financeira e institucional dos programas ou projetos a serem implantados".

Após a explanação acerca dos tipos de avaliação que podem ser realizadas antes do início do programa, foi elaborado o **Quadro 9** que reúne a tipologia dos autores aqui analisados.

Após a apresentação das conceituações e termos referentes, à avaliação que pode ser realizada antes de se iniciar a execução do programa, é possível ressaltar alguns aspectos relacionados com o tema.

O primeiro é que, de forma geral, as avaliações pertencentes ao **Grupo 1** (que podem ser realizadas durante a elaboração do programa ou logo após, mas sempre antes de sua implementação), é conhecida virtualmente conhecida como avaliação *ex-ante*. Alguns autores, porém referem-se a esse tipo de avaliação com termos similares (ver **Quadro 9**).

O segundo aspecto é que, independentemente dos termos utilizados, parece existir um consenso entre os autores no que diz respeito à importância da avaliação para o bom desenvolvimento do programa.

E o terceiro é que, a partir das conceituações e características gerais apontadas pelos autores, e independentemente dos termos utilizados, pode-se inferir que a avaliação que ocorre antes do início do programa, tem como objetivo principal averiguar as reais possibilidades de sua execução durante a fase de elaboração e planejamento de suas

atividades e, assim, minimizar os possíveis desvios na sua fase de execução (implementação).

Após apresentarmos os tipos de avaliação que podemocorrer antes da execução da política, programa e projetos, a próxima seção deter-se-á sobre o **Grupo 2** — avaliações que ocorrem durante a execução do programa (ver **Quadro 8**).

Quadro 9

Nomenclatura utilizada para os tipos de avaliação do grupo 1 por autor

	Nomenclatura utilizada para as avaliações	
Autor	**Critério A**	**Critério B**
Alkin (1972)	**N.C**	**avaliação do planejamento**
Ander-Egg (1990)	**N.C**	**estudo diagnóstico; avaliação da programação**
Ballesteros (1996)	**N.C**	**analisis de objetivos; pre – evaluación y especificación del programa**
Aguilar & Ander-Egg (1995)	***ex-ante;* avaliação antes; avaliação inicial ou pré decisão**	**avaliação do plano e conceitualização**
Faria (1998)	**avaliação *ex-ante***	**N.C**
Cohen & Franco (1999)	**avaliação *ex-ante***	**N.C**
Martinic (1997)	**avaliação *ex-ante***	**avaliação do plano**
Tenório (2000)	**análise**	**N.C**
BID (1997)	**avaliação *ex-ante***	**N.C**
SIEMPRO (1999)	**avaliação *ex-ante***	**evaluación de la planificación**
Viedma (1996)	**evaluación formativa**	**N.C**
Lara (1996)	**N.C**	**evaluación de diseño y conceptualización del programa**
Lobo (1998)	**avaliação *ex-ante***	**N.C**

4.2. Avaliação durante a execução do programa

Conforme foi apresentado na introdução deste capítulo, dois dos critérios que são utilizados pelos autores para classificar os tipos de avaliação são: o momento em que se avalia e o aspecto/atividade que é avaliado.

Partindo desses critérios, na seção anterior foram abordadas as semelhanças e divergências entre os autores que tratam da avaliação que ocorre antes do início do programa. A partir dessa análise, foi possível identificar que existe uma diversidade conceitual e diferentes termos para nomear um mesmo tipo de avaliação.

Essa realidade, porém, parece não ser exclusividade desse tipo de avaliação e pode ser observada nas avaliações que ocorrem durante o processo de execução do programa.

Considerando, esse pressuposto, esta seção tem como objetivo analisar e apresentar as possíveis semelhanças e ou divergências entre os autores analisados, abordando os diversos termos e os aspectos conceituais e metodológicos da avaliação que pode ser realizada durante a execução do programa.

Iniciaremos nossa análise por Alkin (1972). Este autor não utiliza os critérios anteriormente mencionados para realizar sua tipologia de avaliação. Porém, dentro do seu modelo de avaliação CEA[106] (Centro para o Estudo da Avaliação), ele afirma que diferentes decisões exigem

---

[106] Tradução do termo original em inglês *Model CSE (Center of the Study Evaluation).*

tipos diferentes de avaliações, e o *decision making* é quem determina a natureza do que será avaliado.

Mesmo sem fazer alusão explicita ao momento em que se deve realizar a avaliação, Alkin (1972) afirma que dentro de um projeto existem duas áreas de necessidades de avaliação: a área de implementação do programa e a área de melhoria do programa[107].

Em relação à primeira área, Alkin (1972, p.11) afirma que "*a tarefa da avaliação nesta fase é determinar o grau em que o planejado no programa coincide com o que foi implementado*[108]". Já a segunda área — melhoria do programa — tem como objetivo obter informações sobre o funcionamento do programa, para aperfeiçoar o seu andamento.

Considerando os conceitos propostos por Alkin (1972), infere-se que essas duas avaliações podem ser consideradas como avaliações que ocorrem durante a execução do programa e se detêm sobre os aspectos do seu processo de implementação e funcionamento, com o objetivo de obter informações que auxiliem na tomada de decisão.

Holister (1972), de maneira mais direta que Alkin (1972), afirma que a avaliação que ocorre durante a execução do programa é chamada de avaliação de processo. Essa avaliação consiste no monitoramento das funções e atividades administrativas, da contabilidade e das transações financeiras de um programa.

Segundo Holister (1972), o objetivo principal da avaliação de processo é saber se o programa está sendo administrado honestamente e eficazmente e desta forma proporcionar informações que possibilitem as

---

[107] Tradução do termo original em Inglês.
[108] Tradução do texto original em inglês, ALKIN (1972).

mudanças necessárias na sua execução. Para ele, qualquer programa deve ser monitorado ou avaliado.

Worthen (1997), diferentemente de Alkin (1972) e Holister (1972), afirma que a avaliação que ocorre durante a execução do programa é chamada de avaliação formativa.

Utilizando as definições propostas por Scriven[109], Worthen (1997) afirma que esse tipo de avaliação possui um papel específico dentro do processo de execução e desenvolvimento do programa, que é o de melhorar o seu funcionamento.

Segundo esse autor, a avaliação formativa "*conduz (ou deveria conduzir) às decisões sobre o desenvolvimento do programa (inclusive modificar, revisar ou deixar como está)*[110]". E destaca sua importância ao afirmar que: "*o processo de desenvolvimento do programa se torna incompleto e ineficiente caso não ocorra à avaliação formativa*"[111] (WORTHEN, 1997, p. 15).

Baseado, em Scriven, Worthen (1997) afirma que a avaliação formativa tem como meta atribuir mérito ou valor, e como função aperfeiçoar o que for necessário durante a execução do programa.

Assim, é possível afirmar que o autor, ao sugerir a avaliação formativa, a relaciona não só com o momento em que essa avaliação ocorre, mas também, com a função que desempenha dentro do processo de desenvolvimento do programa.

---

[109] Segundo Worthen (1997, p. 8), Scriven "notou que a avaliação formal pode ter muitos papéis porém uma única meta: determinar o valor ou mérito de tudo que é avaliado. Scriven, fez a distinção entre meta e papéis da avaliação: a meta da avaliação é prover respostas às perguntas significantes, e os papéis da avaliação referem-se ao modo pelo qual essas respostas são utilizadas." Tradução do texto original em inglês.
[110] Tradução do texto original em inglês.
[111] Tradução do texto original em inglês.

Continuando a análise, pode-se apontar Weiss[112] (1998, 1999). Mesmo não explicitando os critérios que serviram para nortear a sua tipologia, essa autora aponta para dois tipos de avaliação que, no nosso entendimento, podem ser consideradas como avaliações que ocorrem durante a execução de um programa: avaliação de processo e avaliação formativa.

No que diz respeito ao primeiro tipo, Weiss (1998) afirma que os pesquisadores (avaliadores) ao concentrar seus estudos nos aspectos do processo de desenvolvimento do programa estão realizando uma avaliação de processo. Esse tipo de avaliação tem como objetivo conhecer até que ponto o programa está seguindo as práticas prescritas (fidelidade do programa para com o seu planejamento). Ao tratar desse tipo, Weiss[113] (1998) o está relacionando, mesmo que implicitamente, com o critério aspecto/atividade do programa que é objeto da avaliação.

Já no que se refere ao segundo tipo de avaliação — a formativa — Weiss (1998, 1999), defende as mesmas idéias que Worthen (1997) e nos remete mais uma vez para as idéias propostas por Scriven[114].

Segundo a autora, a avaliação formativa foi inicialmente utilizada para avaliar os aspectos relacionados ao desenvolvimento de programas curriculares, tendo como função[115] proporcionar informações para melhoria do programa durante sua execução. Para Weiss (1999, p.30), a avaliação formativa:

> *"...produce la información que se retroalimenta durante o desarrollo de un currículo para ayudar a mejorarlo. Presta*

---

[112] A primeira publicação desta obra foi em 1972.
[113] A primeira publicação desta obra foi em 1972.
[114] Scriven distinguiu duas funções (ou propósitos) que a avaliação pode assumir e estabeleceu a partir daí dois tipos de avaliação: a formativa e somativa.
[115] Também encontrado na literatura como: uso, propósito, papel, fim.

servicios a las necesidades de quienes lo están desarrollando".

Ainda, segundo Weiss (1998, 1999), mesmo tendo sido criada para avaliar programas curriculares, a avaliação formativa, pode ser utilizada para outros tipos de programas, uma vez que a função desse tipo de avaliação pode ser aproveitada em qualquer processo avaliativo, de qualquer programa ou projeto social.

Após o esclarecimento dos dois tipos de avaliação — avaliação de processo e avaliação formativa —, Weiss (1998) faz uma correlação entre elas.

Segundo a autora, a avaliação de processo recebe este nome por se relacionar com os aspectos a serem avaliados. Já a avaliação formativa, é assim chamada por se relacionar com a intenção, propósito ou uso das informações dentro do processo de execução e desenvolvimento do programa.

Tendo em vista que ambas ocorrem durante a execução do programa, Weiss (1998) afirma que o avaliador, ao realizar a avaliação do processo de execução de um programa, pode utilizar-se da função da avaliação formativa (que é melhorar o programa em desenvolvimento). Portanto, a associação entre o momento em que se avalia e o aspecto do programa, que é objeto da avaliação e função da avaliação, parece ser a causa para que esses dois tipos de avaliação sejam comumente utilizados como sinônimos, neste caso, a utilização procede.

Enquanto os autores anteriormente citados não fazem referência aos critérios que foram utilizados para estabelecer os tipos de avaliação, Viedma (1996) afirma que as avaliações são classificadas a partir do momento em que são realizadas.

Esse autor, seguindo esse critério para estabelecer a sua tipologia, afirma que a avaliação que ocorre durante a execução do projeto é denominada de avaliação de processo. Essa avaliação "*consiste en monitorizar todo cuanto se hace o no se hace durante o desarrollo del proyecto. También se monitorizan barreras que este encuentra en su desarrollo, así como las acciones tomadas para supéralas*" (VIEDMA, 1996, p.4). Ou seja, o autor se preocupa com os aspectos e elementos internos, relacionados ao funcionamento do programa.

Sem levar em conta os termos que a avaliação realizada durante a execução do programa possa receber, mas, considerando, as definições desse tipo de avaliação e dos aspectos/atividades que são avaliados, pode-se inferir que as avaliações propostas por Alkin (1972), Holister (1972), Worthen (1997), Weiss (1998, 1999) e Viedma (1996), são equivalentes, pois têm como foco de interesse o processo de execução e funcionamento do programa na perspectiva de poder melhorá-lo. O que nos parece é que as diferenças entre os autores até aqui analisados estão muitos mais relacionadas à forma como eles tratam do tema do que propriamente a questões conceituais.

Continuando nossa análise, pode-se apontar Briones (1996, p.23) que mesmo afirmando que "*hay diversos tipos de evaluación, diferenciados de acuerdo con otros tantos critérios*", ao realizar sua tipologia não especifica que critérios foram utilizados. Mesmo assim, a partir dos conceitos propostos por Briones (1996), é possível afirmar que os quatro

tipos de avaliação apontados por ele, relacionam-se no nosso entendimeto com o critério do momento que se avalia e o critério do aspecto/atividade que é avaliado durante o processo de execução do programa.

Assim, para efeito de análise, a *evaluación intermedia o contínua, evaluación formativa, evaluación de proceso e evaluación de objetivos internos*[116]" (BRIONES, 1998) são avaliações que ocorrem durante o processo de execução do programa.

Em relação as *evaluación intermedia e evaluación formativa*, Briones (1998, p.24) afirma que ambas podem ser vistas como sinônimos. Segundo ele, isso se deve ao fato de que "*el termino la evaluación formativa, originado en el campo de la educación, se há transferido a otros campos de actividad donde se la utiliza como sinónimo de evaluación intermédia*".

Então se a avaliação formativa procedente da área educacional,

> "*...es el studio que se realiza durante el proceso de elaboración de un currículo y tiene como propósito, según lo indica su nombre, proporcionar información de retroalimentación que permita mejorar (formar) ese instrumento o procedimiento de instrucción*" (BRIONES 1998, p.24)

e as *evaluaciones intermedias* são provenientes da área social e

> "*...buscan información durante el desarrollo del programa con el fin de detectar problemas de funcionamiento y buscarles soluciones adecuadas y oportunas. En todo programa lo deseable es que estas evaluaciones intermedias sean continuas, de modo, que sirvan de apoyo permanente a numerosas actividades que se desarrollan durante su realización*". (BRIONES, 1998, p.23).

---

[116] Os objetivos internos, "se refieren a situaciones que han de alcanzarse en la población em la cual acúa directamente el programa" (BRIONES, 1998, p.25).

Ao considerar o sentido estrito das avaliações formativa e intermédia pode-se inferir que: ambas ocorrem durante a fase de execução do programa e possuem como propósito melhorar o seu desenvolvimento, podendo assim ser consideradas como sinônimos (vimos algo similar ao analisar Weiss (1998, 1999)).

Já no que diz respeito à avaliação de processos, Briones (1998, p.24) afirma que essa avaliação tem como objetivo avaliar o "*funcionamiento y otros aspectos dinámicos de un programa.*" Mais adiante porém, esse autor diferencia dois tipos de processo: *proceso de mantedimiento*[117] *e proceso de cambios y innovación*[118].

Em relação à avaliação de objetivos internos Briones (1998, p. 25-26) afirma que esse tipo de avaliação é destinado a "*verificar el logro de tales objetivos y la relación com la organización del programa, la calidad de los instructores, los recursos utilizados, etecetera*" e pode ser considerada como similar da *evaluación intermedia o continua.*

Ao analisar os quatro tipos de avaliação — avaliação intermédia, avaliação de processos, avaliação de objetivos internos e avaliação formativa — propostos por Briones (1998), pode-se inferir que todos têm como foco de interesse o processo de funcionamento do programa e tem como objetivo melhorar o seu desenvolvimento.

---

[117] "Por los procesos de mantedimiento, entendemos las acciones y el manejo de los recursos necesarios para poner en práctica la estratégia y metodologia del programa; vale decir; los procedimientos instrumentales directamente relacionados com el logro de los objetivos" (BRIONES, 1998, p.145).

[118] "Los procesos de cambio e innovación están contituidos en su mayoría por las interacciones que se producen entre el personal del equipo técnico y los usuários del programa, o solo entre estos últimos, com el fin de lograr los objetivos propuestos" (BRIONES, 1998, p.151).

Considerando as definições realizadas por Briones (1998), e relacionando-as com os critérios apontados anteriormente, pode-se dizer que a avaliação intermédia ou formativa é assim chamada por relacionar-se com o momento em que se avalia e conseqüentemente, com seu propósito. Enquanto que a avaliação de processos e avaliação de objetivos além de se relacionarem com o momento em que se avalia, também são assim chamadas devido aos aspectos que são objetos de avaliação.

Ballesteros (1996), diferentemente de Briones (1998), faz sua tipologia, utilizando como critério o momento em que ocorre a avaliação. Para essa autora, a avaliação que ocorre durante a execução do programa é chamada de avaliação formativa e "*tiene por objetivo esencial la mejora y perfeccionamento del programa*" (BALLESTEROS, 1996, p. 33).

Ballesteros (1996), assim como Worthen (1997), Weiss (1998, 1999) e Briones (1998), remete-nos aos pressupostos de Scriven, ao se referir à avaliação formativa. É importante destacar, mais uma vez, que essa avaliação foi assim chamada devido à função que desempenha, mas é comumente relacionada ao momento em que se avalia[119]. Esta dicotomia se deve ao fato de que a avaliação que ocorre durante a execução de um programa tem uma função específica que pode ser considerada como uma função formativa.

No entanto, Ballesteros (1996), em um segundo momento, cita outro tipo de avaliação. Partindo da lógica implícita no *ciclo de intervención social*[120], proposto por Palumbo (1987), aponta para a necessidade de avaliar os aspectos relacionados à implementação durante a fase de

---

[119] Esta relação foi abordada na seção anterior, quanto se tratou da tipologia apresentada por WEISS (1998, 1999).

[120] Na lógica do *ciclo de intervención social*, é necessário que os planejadores tenham informações a cerca das diversas fases desse ciclo.

execução de um programa. Esse tipo de avaliação é denominado de *evaluación de la implementacion del programa.*

Após a indicação da avaliação formativa e avaliação da implementação do programa, Ballesteros (1996, p.68) relaciona esses dois tipos de avaliação e afirma que:

> *"Cuando se realiza una evaluación de proceso, al tratar de averiguar si los elementos que la constituyen están surtiendo micro-efectos y/o si está desarrollando en el sentido previsto, estamos juzgando sobre el progreso del programa. Ello es importante no solo porque nos va permitir mejorar el programa sobre la marcha (como así es uno dos objetivos de la evaluación formativa), sino porque nos permitirá establecer basta qué punto los distintos elementos constituyentes del programa se están dispensando convenientemente".*

Isto significa que, segundo Ballesteros (1996), o avaliador ao realizar a avaliação da implementação de um programa (processo), a faz na perspectiva da função da avaliação formativa (melhorar o desenvolvimento).

Ballesteros (1996) também afirma que não se pode pensar em avaliar ao final do programa sem conhecer o que realmente ocorreu durante a sua execução, mais propriamente, sobre os aspectos e ou atividades que foram implementados ou não durante o seu processo de execução.

Ander-Egg (1990) também utiliza o critério momento em que se avalia para estabelecer sua tipologia de avaliação. Esse autor afirma que a avaliação que ocorre durante a execução de um programa é chamada de *evaluación parcial* e tem como objetivo introduzir correções durante o desenvolvimento do programa ou projeto. Este tipo de avaliação

deve estar presente em todas as fases de execução de um programa, na perspectiva de revisar, analisar e orientar o que está sendo realizado.

Segundo Ander-Egg (1990), dentro do processo de execução de um programa deve-se escolher quais aspectos serão objetos de avaliação. Para o autor, esses aspectos estão divididos em duas áreas: *área de coherencia interna e area de coherencia externa.*

Dentro da área de coerência interna pode se realizar quatro subtipos[121] de avaliação, são elas: *evaluación de la ejecución del proyecto, evaluación de la implementación, evaluación del perfil estructural y funcional de la institución e evaluación de los procedimientos utilizados.*

Para Ander-Egg (1990), os subtipos apresentados anteriormente são assim definidos:

- a avaliação da execução do projeto verifica até que ponto as metas e objetivos propostos foram alcançados;

- a avaliação da implementação avalia os meios, recursos, atividades que são necessários para o desenvolvimento do projeto. Neste momento avalia se ocorreram tarefas de preparação e motivação para as pessoas envolvidas no projeto; também avalia a disponibilidade de recursos (humanos, materiais, financeiros técnicos);

- avaliação do perfil estrutural e funcional da instituição *que* consiste em uma análise da estrutura organizacional;

---

[121] Para maiores detalhes ver ANDER-EGG (1990).

- e, por fim, *a* avaliação dos procedimentos utilizados, que tem como objetivo verificar de que forma e quais foram os procedimentos utilizados durante o desenvolvimento do programa (forma operativa).

Lara (1996), diferentemente de Ander-Egg (1990) e Ballesteros (1996), realiza sua tipologia de avaliação utilizando o critério "*según centren sus objetivos*". No entanto, para nosso entendimento, esse critério pode ser considerado equivalente ao critério aspectos/atividades do programa que são avaliados.

Considerando, essa equivalência, Lara (1996) afirma que os aspectos relacionados à execução e funcionamento do programa podem ser avaliados durante sua fase de execução. Essa avaliação é conhecida como avaliação de processo.

Com a intenção de detalhar esse tipo de avaliação, Lara (1996) afirma que dentro da avaliação de processo podem-se distinguir cinco subtipos[122] conforme seu alcance, o momento em que são realizadas (dentro do processo) e especificidade de seus objetivos. São elas: *evaluación de la implementación, monitoreo e seguimiento, evaluación del esforzo, evaluación de la cobertura e evaluación de la productividad.*

Segundo Lara (1996, p.112) a avaliação da implementação tem como objetivo *"conocer como se ha puesto en marcha un programa, y determinar si existen diferencias significativas entre el funcionamiento del programa y lo previsto en su diseño".*

---

[122] Para maiores detalhes ver LARA (1996).

Já o monitoramento e seguimento, segundo a autora,

> *"...contempla las operaciones de evaluación del funcionamiento del programa que normalmente se llevan a cabo por los propios gestores, utilizando procedimientos sistematizados y continuados de recogida de información, con el fin de detectar posibles desviaciones y mejorar la gestión del programa"* (LARA, 1996, p.113).

Dando continuidade as diferenciações entre os subtipos de avaliação de processo, Lara (1996, p.113-114), afirma que

> *"...la evaluación del esfuerzo se orienta a conocer si realmente el programa esta aplicando los recursos de que dispone con la intensidad adecuada y en los plazos previstos"* [...] *La evaluación de la cobertura orienta a determinar hasta qué punto el programa llega a incidir sobre la población objetivo".*

Por fim, Lara (1996, p.114) afirma que a avaliação de produtividade do programa:

> *"...se orienta a determinar qué productos se han obtenido como consecuencia de la realización de las actividades del programa, y que relación existe entre el volumen y la calidad de los productos obtenidos y los insumos utilizados para ello".*

Já Aguilar & Ander-Egg (1995), ao realizarem sua tipologia recorrem aos dois critérios apontados no início deste capítulo: momento em que se avalia e aspectos do programa que são objetos da avaliação.

Em relação ao primeiro critério, os autores afirmam que a avaliação que ocorre durante a execução do programa, além de ser chamada de avaliação de processo, pode ser conhecida também como:

"avaliação da gestão, avaliação contínua, avaliação concorrente, monitoramento ou avaliação concomitante".

Para Aguilar & Ander-Egg (1995, p.41), a avaliação de processo, independentemente do nome que possa receber, tem como objetivo "*avaliar as mudanças situacionais, isto é, estabelecer até que ponto se está cumprindo e realizando o programa ou prestando um serviço de acordo com a proposta inicial*".

Em relação ao segundo critério, os autores afirmam que diferentes aspectos relacionados com o funcionamento do programa podem ser avaliados durante sua execução. Isto significa que, dentro da avaliação de processo, cada um desses aspectos estabelece um subtipo de avaliação. Neste sentido, Aguilar & Ander-Egg (1995) seguem a mesma linha de raciocínio apontada por Ander-Egg (1990) e Lara (1996).

No entanto, Aguilar & Ander (1995) apontam para quatro subtipos de avaliação dentro da avaliação de processo, enquanto Ander-Egg (1990) e Lara (1996) apontam para cinco subtipos.

Segundo Aguilar & Ander-Egg (1995), os quatro subtipos[123] da avaliação de processo são assim definidos:

- a avaliação da cobertura verifica até que ponto o programa está alcançando a população objeto;
- a avaliação da implementação avalia os aspectos técnicos da implementação, ou seja, se o programa está sendo operacionalizado (verificação dos meios e instrumentos que foram programados);

- a avaliação do ambiente organizacional, verifica até que ponto os aspectos estruturais e funcionais do organismo responsável pelo programa favorece ou dificulta o seu desenvolvimento;

- e, por fim, a avaliação do rendimento pessoal, que tem como objetivo avaliar a capacidade, competência e habilidade de um indivíduo para realizar as atividades que lhe foram atribuídas dentro do programa.

Ao comparar os subtipos da avaliação dentro da avaliação de processos propostos por Lara (1996) e Aguilar & Ander-Egg (1995), pode-se sinalizar algumas semelhanças e diferenças.

Em relação aos subtipos de avaliação de processo — avaliação de cobertura e avaliação de implementação — esses autores compartilham da mesma perspectiva. Porém, no que diz respeito ao monitoramento e seguimento, Lara (1996) e Aguilar & Ander-Egg (1995) assemelham-se quanto ao objetivo mas se divergem no que se refere classificação. Ou seja, para Lara (1996), monitoramento e seguimento é um subtipo da avaliação de processo enquanto que para Aguilar & Ander-Egg (1995), a própria avaliação de processo é, por si só, um monitoramento e seguimento.

Também se referindo a avaliação que ocorre durante o processo de execução do programa, pode-se citar dois organismos supranacionais: BID (1997) e SIEMPRO (1999). Como foi visto anteriormente, nos Capítulos II e III, esses organismos desenvolvem a avaliação sob duas perspectivas diferentes.

---

[123] Para maiores detalhes ver AGUILAR & ANDER-EGG (1995).

Assim, para o BID (1997) a partir da perspectiva do Marco Lógico, a avaliação que ocorre durante a execução do programa, é chamada de *evaluación concurrente* e tem uma função formativa, porque

> *"...su propósito es respaldar la mejora contínua de la ejecución del proyecto". [...] Adopta a forma de monitoreo continuo aunque en ciertos casos el banco, y algunas veces los prestatarios, tamben hacen evaluaciones formales operativas durante la ejecución del proyecto"* (BID, 1997, p.5).

Com esta afirmação, o BID (1997), estabelece a existência de duas formas de avaliar o processo de execução do projeto e que ele chama de ferramentas: o *seguimiento de ejecución y monitoreo* e a *evaluación operativa (evaluación intermedia ou evaluación de ejecución.*

Para o BID (1997, p.25), o *monitoramento* é o

> *"...procedimiento mediante el cual verificamos la eficiencia y eficacia de la ejecución de un proyecto mediante la identificación de sus logros y debilidades y en consecuencia, recomienda medidas correctivas para optimizar los resultados esperados del proyecto".*

Já a avaliação operativa é uma atividade levada a cabo quando existem indícios de que o projeto apresenta falhas na gestão e até mesmo na sua programação, podendo esta ser realizada a qualquer momento durante o seu processo de execução. Essa ferramenta de avaliação serve "*para analizar más a fondo los problemas de ejecución, o para ayudar a tomar decisiones sobre la programación del proyecto*" (BID, 1997, p. 33-34).

A partir dessas afirmações pode-se dizer que o seguimento de execução e o monitoramento são uma atividade rotineira e contínua de acompanhamento do processo de execução de um programa, enquanto a

avaliação operativa ocorre após ter sido detectado algum problema de gestão que compromete o desenvolvimento do programa.

Na perspectiva da Gestão Integral, o SIEMPRO (1999), afirma que a avaliação que ocorre durante a execução do programa é chamada de evaluación de la ejecución del programa. Portanto, esse tipo de avaliação deve auxiliar a tomada de decisão através da

> *"...indagación y valoración de como se está ejecutando el programa en todos sus elementos metodológicos: componentes, resultados, actividades, tareas, recursos, presupuesto, etcétera"* (SIEMPRO, 1999, p.60).

Como ferramentas para operacionalizar a avaliação da execução do programa, cita: o *monitoreo, evaluación diagnostica* e a *evaluación desde la perspectiva de los beneficiarios (evaluación participativa).* Segundo o SIEMPRO (1999, p.87), o monitoramento,

> *"...consiste en el análisis periódico de la ejecución del programa social con el objeto de establecer el grado en que las tareas y las actividades de cada componente, el cronograma de trabajo, las prestaciones y los resultados se cumplen de acuerdo con lo planificado. Gracias a este control, se puede detectar deficiencias e incongruencias, y corregir o replanificar oportunamente".*

Em relação às outras duas ferramentas que são utilizadas para realizar a avaliação da execução do programa, o SIEMPRO (1999, p. 94) afirma que a avaliação diagnóstica, *"debe realizar-se durante la ejecución del programa cuando se detectan problemas relacionados con la planificación, la ejecución y el logro de los resultados esperados".* E a avaliação a partir da perspectiva dos beneficiários, nada mais é do que a incorporação das impressões da população beneficiada no processo de *"análisis de las fortalezas y debilidades de los programas y de los impactos*

que generan las condiciones de vida de esas poblaciones" (SIEMPRO, 1999, p.97).

Após as definições dos tipos de avaliação que podem ser realizados durante a execução do programa, propostos pelo BID (1997) e SIEMPRO (1999), pode-se afirmar que esses organismos ora se assemelham ora se diferenciam.

Como exemplo dessa constatação pode–se apontar que, enquanto o BID (1997) afirma que existem duas ferramentas para realizar a avaliação que ocorre durante a execução do projeto — monitoramento e avaliação intermédia de execução — o SIEMPRO (1999) aponta para três ferramentas: monitoramento, avaliação diagnóstica e avaliação com a perspectiva dos beneficiários (avaliação participativa). Sendo que em relação à concepção do monitoramento, ambos se assemelham, pois afirmam que essa ferramenta tem como objetivo possibilitar correções no decorrer do desenvolvimento do projeto. Já em relação à segunda ferramenta, cada um utiliza um termo. No entanto, ao analisarmos as definições, percebe-se que elas são, na verdade, equivalentes.

Martinic (1997) estabelece os tipos de avaliação em dois momentos distintos. No primeiro, esse autor utiliza os critérios do momento que se avalia e o aspecto que é avaliado.

Desta forma, afirma que a avaliação ao ser realizada durante a execução do programa é chamada de *evaluación ex-post*. Em relação a essa avaliação, Martinic (1997, p. 20-21), afirma que:

> *"En los proyectos que están en marcha, las decisiones se relacionan con el cambio o no del proyecto en virtud de los resultados alcanzados. En esta fase se hace una evaluación de procesos que consiste en una evaluación periódica de la programación, administración, control del proyecto, con el objetivo de superar las dificultades y rescatar las potencialidades del mismo".*

Esta citação, aponta-nos para um tipo de avaliação que Martinic (1997), classifica a partir do aspecto que é avaliado dentro do processo de execução de um projeto, denominada de *evaluación de proceso ou intermedia.*

Em relação a essa avaliação, Martinic (1997, p.21) afirma:

> *"La evaluación de procesos estudia la relación existente entre las actividades realizadas y los productos o resultados alcanzados. Por lo general se concentra en la eficiencia en el uso de recursos e insumos para alcanzar los resultados más inmediatos del proyecto".*

Já no segundo momento, Martinic (1997) considerando a avaliação como um processo integrado ao desenvolvimento do projeto, relaciona as avaliações com o ciclo de vida do projeto (*diseño, ejecución y evaluación*).

Com base neste pressuposto, este autor, afirma que durante a fase execução de um projeto é possível realizar dois tipos de atividades avaliativas: *el seguimiento o monitoreo e la evaluación de proceso.*

O seguimento ou monitoramento é definido por Martinic (1997, p. 24), como uma

> *"...actividad de registro y verificación de lo que se realiza en el transcurso del proyecto con el fin de mantener la dirección del trabajo; realizar las correcciones o cambios que se consideren pertinentes para garantizar los resultados esperados."*

Segundo Martinic (1997, p. 119) a avaliação de processo ou intermédia[124], agora um pouco mais detalhada

> *"...estudia la relación existente entre las actividades y los productos o resultados alcanzados por las mismas actividades. El objetivo de análisis son los resultados de las actividades realizadas por el proyecto y que se expresan en cambios en la realidad material; en los conocimientos o información que se tiene sobre un tema o problema, en las relaciones sociales y sistemas de comunicación entre quienes por el programa y los actores o grupos que causan tales problemas".*

Dentro dessa perspectiva, pode-se apontar dois tipos de processos que são avaliados: o *proceso de mantenimiento o funcionamiento* e os *procesos interactivos*, a partir dois quais se transmitem novos conhecimentos, prática e habilidades.

Partindo dos conceitos propostas por Martinic (1997), pode-se inferir que o seguimento ou monitoramento diferencia-se da avaliação de processo no que diz respeito ao objetivo que cada uma delas possui. No entanto, ao compararmos as definições de avaliação de processo propostas por esse autor e Briones (1998), pode-se afirmar que ambas são equivalentes.

Já Faria (1998) ao distinguir os tipos de avaliação não utiliza critérios. Mesmo assim pode-se considerar que a autora faz um percurso mais simples do que Martinic (1997).

A partir das definições propostas por Faria (1998), pode-se inferir que os tipos apresentados pela autora ocorrem durante a execução de um programa e se relacionam com os aspectos/atividades avaliados.

---

[124] Termos também utilizados por Briones (1998).

Assim, as avaliações realizadas durante a execução do programa, apontadas por Faria (1998) são o acompanhamento e monitoramento (avaliação da eficiência[125]) e a avaliação formativa (avaliação de processo ou avaliação de eficácia[126]).

Para Faria (1998, p.44) acompanhamento e monitoramento (avaliação da eficiência) busca a adequação entre o plano do programa e sua execução. Essa atividade avaliativa tem como objetivo intervir no processo de execução corrigindo os rumos dos desvios detectados, através da modificação dos "*cenários jurídico-administrativo, financeiro, organizacional*'".

Já a avaliação de processo (avaliação formativa ou avaliação da eficácia), segundo Faria (1998, p.45), tem como objetivo:

> *"...acompanhar, observar e testar o desempenho do programa para aprimorá-lo. Este acompanhamento inclui o diagnóstico, das eventuais falhas dos instrumentos, procedimentos conteúdos e métodos, bem como adequação do público-alvo e do impacto do programa, aumentando sua adequação aos seus objetivos e metas".*

Ao comparar os tipos de avaliação propostos por Martinic (1997) e Faria (1998), mas sem considerar a maneira como foram estabelecidos, pode-se afirmar que ambos compartilham da mesma perspectiva, no que diz respeito às distinções entre seguimento e monitoramento e avaliação de processos.

Figueiredo & Figueiredo (1986), sem se preocuparem em estabelecer critérios para apontar os tipos de avaliação, afirmam que a

---

[125] Será abordada no Capítulo V.
[126] Este termo será abordado no Capítulo V.

avaliação que ocorre durante a execução (implementação) é chamada de avaliação de processo.

Esse tipo de avaliação tem como objetivo aferir a eficácia[127]. Sendo mais específica, Figueiredo & Figueiredo (1986, p. 112), afirmam que a avaliação de processo tem como objetivo *"acompanhar e aferir se os propósitos, estratégias e execução do programa estão sendo realizados segundo as definições previamente estabelecidas"*.

Os autores afirmam que dentro da avaliação de processos, é possível encontrar três linhas distintas: a avaliação de metas (o mais difundido), a avaliação de meios (com três critérios de eficácia assim estabelecidos: funcional, administrativa e contábil) e avaliação da relação custo-benefício e custo-eficiência[128].

Cohen & Franco (1999) diferentemente de Faria (1998) e Figueiredo & Figueiredo (1986) utilizam o critério do momento em que se avalia para estabelecer sua tipologia de avaliação.

Desta forma, afirmam que a avaliação que ocorre após a avaliação ex-ante e durante a execução (implementação) do projeto é denominada de avaliação *ex-post*. Mas como o termo ex-post também é utilizado para representar a avaliação que ocorre após o término do programa, Cohen & Franco (1999) utilizam o termo avaliação de processos ou concomitante como sinônimo dessa avaliação.

Esse tipo de avaliação, segundo Cohen & Franco (1999, p.109),

> *"Procura detectar as dificuldades que ocorrem durante na programação, administração e controle, etc., para serem*

---

[127] Este termo será detalhado no Capítulo V.
[128] Para maiores detalhes FIGUEIREDO & FIGUEIREDO (1986).

> *corrigidas oportunamente, diminuindo os custos derivados de sua ineficiência. Não é um balanço final e, sim uma avaliação periódica. Diferencia-se da retroalimentação que é uma atividade permenente de revisão, realizada por aqueles que estão implementando o projeto. Sua função central é medir a eficiência*[129] *de operação do projeto".*

Já Tenório (2000), ao estabelecer os tipos de avaliação não se preocupa em estabelecer critérios. No entanto, tendo em vista que o projeto é constituído de fases, o autor sinaliza para a necessidade de realizar uma avaliação que possibilite o conhecimento acerca do desenvolvimento do projeto. Para ele, esse tipo de avaliação é chamado de acompanhamento e deve ser uma atividade contínua até o término do programa.

Esse tipo de avaliação, segundo o autor, tem como objetivo "*apoiar a implementação do projeto identificando desvios entre a execução das ações e a programação estabelecida*" (TENÓRIO, 2000, p.48).

Assim, para Tenório (2000, p.47) o acompanhamento realizado durante a execução do projeto:

> *"Procura verificar as ações e medir o desempenho com a finalidade básica de auxiliar a tomada de decisão e corrigir eventuais desvios na programação, durante execução do projeto, assegurando que os objetivos estabelecidos sejam alcançados".*

Com o objetivo de proporcionar um melhor entendimento a respeito dos tipos de avaliação que podem ser realizados durante a execução do programa, foi elaborado o **Quadro 10** que resume e

---

[129] Este conceito será tratado no Capítulo V.

possibilita uma visualização sobre as semelhanças e diferenças entre os autores analisados.

Após as explanações acerca da tipologia de avaliação durante a execução do programa ou projeto, pode-se fazer algumas considerações.

A primeira é que mesmo os autores utilizando critérios diferentes para estabelecerem suas tipologias, eles referem-se à avaliação que ocorre durante a execução do programa.

A segunda é que, de forma geral, os autores aqui citados utilizam termos diferentes, mas se referem aos mesmos aspectos que podem ser avaliados durante a execução de um programa.

A terceira é que cada autor, devido a sua formação ou interesse desenvolve seus pressupostos em um determinado tipo de avaliação ou aspecto que pode ser avaliado, detalhando apenas aspectos teórico- metodológicos que considera mais importantes.

E, por fim, a quarta consideração é que a avaliação que ocorre durante a execução de um programa, independentemente do termo que seja utilizado ou do aspecto que seja objeto de avaliação, tem como objetivo principal a melhoria do funcionamento do programa.

Após apresentarmos a tipologia de avaliação, que se refere às avaliações que podem ser realizadas durante a execução do programa, a próxima seção deter-se-á sobre o **Grupo 3** — avaliações que ocorrem ao término do programa (ver **Quadro 8**).

Quadro 10

Nomenclatura utilizada para os tipos de avaliação do grupo 2 por autor

	Nomenclatura utilizada para as avaliações	
Autor	Critério A	Critério B
Alkin (1972)	N.C	avaliação da implementação; avaliação para melhoria do programa.
Holister (1972)	avaliação de processo	N.C
Worthen (1997)	avaliação formativa	N.C
Weiss (1998, 1999)	avaliação formativa	avaliação de processo
Viedma (1996)	avaliação de processo	N.C
Briones (1998)	evaluación intermedia evaluación formativa	evaluación de proceso; evaluación de objetivos internos
Ballesteros (1996)	evaluación formativa	evaluación de la implementación
Ander-Egg (1990)	evaluación parcial	evaluación de la ejecución; evaluación de la implemtación evaluación del perfil estructural y fimcional; evaluación de los procedimientos utilizados.
Lara (1996)	N.C	evaluación de proceso; evaluación de la implementación; monitoreo y seguimiento; evaluación de la cobertura; evaluación de la produtividade; evaluación del ezforzo.
Aguilar& Ander-Egg (1995)	avaliação de proceso; avaliação da gestão; avaliação concorrente; avaliação concomitante ou monitoramento.	avaliação da cobertura; avaliação da implementação; avaliação do ambiente organizacional; avaliação do rendimento pessoal.
BID (1997)	evaluación concurrente (monitoreo contínuo)	seguimiento de ejecución y monitoreo; evaluación operativa (intermedia o ejecución)
SIEMPRO (1999)	evaluación de la ejecución	evaluación diagnóstica; monitoreo; evaluación de la perspectiva de los beneficiarios (avaluación participativa)
Martinic (1997)	evaluación de proceso intermedia; avaliação *ex-post* seguimiento o monitoreo	evaluación de proceso; evaluación de la eficiencia
Faria (1998)	acompanhamento e monitoramento (avaliação da eficiência); avaliação formativa (processo ou eficácia)	avaliação de processo ou avaliação de eficácia
Figueiredo & Figueiredo (1986)	Avaliação de processo	avaliação de processo avaliação de metas; avaliação de meios avaliação da relação custo-benefício e custo resultado.
Cohen & Franco (1999)	avaliação *ex-post* (avaliação de processos)	N.C
Tenório (2000)	acompanhamento	N.C

4.3. Avaliação após o término do programa

A diversidade apontada no que diz respeito aos termos e conceituações acerca das avaliações que ocorrem no início do programa e avaliações que ocorrem durante sua execução pode também ser observada na tipologia das avaliações que ocorrem após o término do programa.

Dando início à análise, pode-se citar Alkin (1972). Este autor, em seu Modelo CEA[130] (Centro para o Estudo da Avaliação), da mesma forma que apontou para a necessidade de avaliar aspectos antes do início do programa e durante sua execução também sinaliza para a necessidade de avaliar aspectos/atividades após o seu término.

Sendo assim, Alkin (1972) partindo do pressuposto de que o programa possui áreas de necessidade de avaliação, afirma que, ao final de um programa, é necessário que se realize uma avaliação capaz de proporcionar informações para os *decision makers* e que possibilite a tomada de decisão sobre o programa como um todo. Essa avaliação é denominada de avaliação para certificação do programa[131].

Já Holister (1972) afirma que a avaliação que ocorre após a conclusão do programa é chamada de avaliação de resultados e também conhecida por avaliação de custo-benefício. Segundo esse autor, esse tipo de avaliação possibilita saber se o programa foi um fracasso ou um sucesso.

---

[130] Tradução do termo original em inglês, Model CSE (Center for the Study of Evaluation)
[131] Tradução do termo original em inglês.

Worthen (1997), assim como Alkin (1992) e Holister (1972), não utilizam critérios para estabelecer sua tipologia de avaliação, mas aponta para uma avaliação que ocorre após o término do programa. Este autor afirma que esse tipo de avaliação é denominado de avaliação somativa.

Worthen (1997), ao se referir à avaliação somativa, recorre aos pressupostos conceituais propostos por Scriven na década de 60. Essa avaliação tem como função (papel) principal auxiliar tomada de decisão no que se refere à continuação, finalização ou expansão do programa que está sendo avaliado.

Ao analisarmos os conceitos das avaliações que ocorrem ao término do programa proposto por Alkin (1972) Holister (1972) e Worthen (1997), mas sem considerar os termos utilizados por eles, pode-se inferir que em sentido estrito, o objetivo dos tipos de avaliação é emitir um juízo de valor acerca do programa e, assim, tomar decisões sobre sua continuidade ou o seu encerramento.

Weiss (1998, 1999) tampouco utiliza critérios ao especificar a sua tipologia, mas aponta para dois tipos de avaliação que ocorrem após o término do programa: avaliação somativa e avaliação de resultados.

Em relação ao primeiro tipo — avaliação somativa — a autora a relaciona com a função (papel) que a avaliação desempenha dentro do processo decisório. Segundo a autora, a avaliação somativa tem como principal função proporcionar informações para auxiliar na tomada de decisão acerca do término ou continuação de um programa. Neste sentido, a autora, assim como Worthen (1997), recorre aos pressupostos de Scriven.

A partir da definição do segundo tipo de avaliação — avaliação de resultados — pode-se inferir que essa avaliação se relaciona ao

momento em que se avalia e ao aspecto que é objeto dessa avaliação. A avaliação de resultados, segundo Weiss (1998), tem como foco de interesse as conseqüências do programa nos beneficiários. Ou seja, se os resultados esperados foram alcançados e se promoveram algum tipo de mudança nos beneficiários (em que proporção essas mudanças se relacionam com o programa).

Mais adiante, Weiss (1998) faz uma relação entre a avaliação somativa e avaliação de resultados. A autora afirma que as informações promovidas pela avaliação de resultados auxiliam a tomada de decisão acerca da continuidade ou não do programa avaliado.

Viedma (1996), diferentemente de Weiss (1998, 1999), faz sua tipologia utilizando como critério o momento em que ocorre a avaliação. Em relação à avaliação que ocorre depois, ou seja, ao término do programa, a autora afirma que, esse tipo de avaliação é chamado de *evaluación de resultados*, pois se preocupa com os efeitos e impactos do programa.

Sendo um pouco mais específico em relação à avaliação de resultados, Viedma (1996, p. 4-5), afirma que esse tipo de avaliação:

> *"...requiere normalmente la utilización de un marco experimental. Dicho marco, se aplica para estimar el impacto de un agente causal o tratamiento. En este sentido las unidades a las que se asigna un tratamiento tienen que ser definidas y observadas*[...] *una pregunta crucial para la evaluación de resultados es: Ha logrado el programa los efectos especificados en su diseño? Para responder esta pregunta hay que se dirigirse a los objetivos del proyecto, a los instrumentos de monitorización y medida".*

Briones (1998), diferentemente de Lara (1996),não se preocupa com critérios para estabelecer sua tipologia. Mas aponta para alguns tipos de avaliação que no nosso entendimento, ocorrem após o término do programa. A partir das definições, pode-se afirmar que para Briones (1998) a avaliação que ocorre ao término do programa é chamada de *evaluación terminal* o *evaluación ex-post*. Essa avaliação tem como *"finalidad establecer si se obtuvieron o no los resultados esperados y los factores que actuaran en una u otra dirección"* (BRIONES, 1998, p.23).

A citação acima remete-nos ao segundo tipo de avaliação apontada por esse autor, *a evaluación de resultados (logros)*[132].

> *"La evaluación de resultados se refiere: a) si el programa produjo o no cambios, es decir, si fue o no efectivo; b) al nivel o magnitud de los cambios producidos globalmente o en sub-grupos de la población; c) a la intensidad o porfundidad de los cambios; y d) a la caliadad de los mismos"*. (BRIONES, 1998, p.159)

Além dessas avaliações, Briones (1998) cita a avaliação somativa e afirma que esta não pode ser vista como sinônimo da avaliação terminal ou *ex-post*. Com esta afirmativa, Briones (1998) diverge de Worthen (1997) e Weiss (1998, 1999), uma vez que esses últimos autores, apontam a avaliação somativa como sendo um tipo de avaliação que ocorre ao final de programa, mesmo sabendo que a essa avaliação foi direcionada para a avaliação de programas curriculares.

---

[132] Para o autor, os resultados "son los cambios o modificaciones que produce en una población o en un cierto contexto"(BRIONES, 1998, p.159).

Ballesteros (1996) utiliza o critério do momento em que se avalia para estabelecer sua tipologia de avaliação. Para a autora, a avaliação que ocorre após o término do programa é denominada de avaliação somativa e pode também ser chamada como avaliação de resultados ou de impactos. Essa avaliação tem como objetivo avaliar os resultados e, desta forma, melhorar o programa.

Após ter definido esse tipo de avaliação, Ballesteros (1997), em um segundo momento, parte da lógica implícita do *Ciclo de Intervención Social proposto* por Palumbo (1987), e afirma que a avaliação somativa (resultados ou impactos) tem como objetivo analisar os efeitos ou resultados em função de sua eficácia, eficiência e efetividade[133].

Ao apontar apara a avaliação somativa, Ballesteros (1996) remete-nos para os pressupostos de Scriven, da mesma forma que Worthen (1997) e Weiss (1998, 1999). Em relação a essa avaliação como um tipo de avaliação que ocorre após o término do programa, pode-se afirmar que Ballesteros (1996) assemelha-se com Worthen (1997) e Weiss (1998, 1999) mas se diferencia de Briones (1998).

Já Ander–Egg (1990), da mesma forma que Ballesteros (1996), utiliza como critério para estabelecer sua tipologia de avaliação o momento em que se avalia o programa. Partindo desse pressuposto, Ander-Egg (1990) afirma que a avaliação que ocorre ao término do programa é chamada de *evaluación final*

---

[133] Por uma questão de organização, detalharemos o assunto no Capítulo V.

Segundo Ander–Egg (1990, p.37) *"esta modalidad de evaluación constituye la etapa final del proceso metodológico y está orientada a medir, analizar e estimar el desarrollo total de un proceso"*.

Como foi dito nas seções anteriores, segundo esse autor, os aspectos que são avaliados antes do início do programa ou durante a sua execução pertencem à área de coerência interna. Porém, uma outra distinção é feita quando se trata da avaliação ao término do programa. Neste sentido, o autor afirma que na avaliação final os aspectos que podem ser avaliados pertencem à área coerência externa.

Deste modo, para Ander- Egg (1990, p.46-47), dentro da área de coerência externa, pode-se realizar quatro subtipos de avaliação que foram assim definidos pelo autor:

- *evaluación de los resultados efectivos*, que busca saber se o programa respondeu ou não às necessidades e problemas reais;

- *evaluación de la coordinación externa del programa* que *"(...) tiene en cuente a integración del programa o proyecto a evaluar con unidades operativas externas a él y con los cuales necesariamente está vinculado";*

- *evaluación de los efectos indirectos que* mede os efeitos indiretos ao programa sejam eles negativos ou positivos.

- e por fim *a opinión de los beneficiários o destinatários que* tem como objetivo conhecer a opinião dos beneficiários sobre o programa .

Conforme foi apontado nas seções anteriores, o critério utilizado por Lara (1996) para realizar sua tipologia, foi considerado equivalente a critério aspectos/atividades que são avaliados.

Assim, o autor afirma que se podem avaliar os produtos de um programa. Essa avaliação é chamada de *evaluación del producto (evaluación de los logros o evaluación del impacto) a qual "se centra en los efectos generados por el programa en su proyección hacia exterior"* (LARA, 1996, p.115). Nesse tipo de avaliação se realiza um subtipo, a depender do aspecto ou atividade que será avaliado. Essas avaliações são assim denominadas: *evaluación de la eficacia, evaluación de la efectividad, evaluación de la eficiencia*[134], *evaluación del impacto e evaluación de la pertinencia.*

Em relação à avaliação de impacto, Lara (1996, p.117) afirma que esse tipo de avaliação de produto tem como foco de interesse os efeitos produzidos pelo programa e "*supone, pues, la formulación del juicio de valor global sobre la efectividad del programa*".

Já a avaliação da pertinência, segundo Lara (1996, p. 118):

> *"...se centra en la consideración de sí los objetivos del programa son o siguen siendo socialmente significativos y pertinentes, justifican los recursos empleados, teniendo en cuenta que podrían asignarse al cumplimiento de objetivos alternativos, y existe el adecuado consenso en torno a la utilidad del programa".*

Ao compararmos Lara (1996) e Ballesteros (1996), no que diz respeito aos aspectos/atividades que podem ser avaliados após o término do programa, pode-se afirmar que esses autores se assemelham.

---

[134] Os termos eficácia, efetividade e eficiência serão abordados no Capítulo V.

Enquanto Ballesteros (1996), Ander-Egg (1990) e Lara (1996) utilizam apenas um critério para estabelecer os tipos de avaliação que ocorrem após o término do programa, Aguilar & Ander-Egg (1995) utilizam dois critérios de avaliação — momento em que se avalia e aspecto/atividade que é avaliado — para estabelecer sua tipologia de avaliação.

Em relação ao primeiro critério, Aguilar & Ander-Egg (1995) afirmam que existem duas avaliações que são realizadas quando o programa chega ao fim: a avaliação do fim do projeto e a avaliação *ex-post*.

Segundo os autores, o primeiro tipo de avaliação — a avaliação do fim do programa - é realizado logo após a conclusão do projeto ou programa. E o segundo tipo de avaliação — a avaliação *ex-post*, também conhecida como avaliação de impacto ou avaliação pós-decisão —, "*é levada a cabo quando o programa ou projeto alcançou seu pleno desenvolvimento (meses ou até anos depois de finalizada a execução)*" (AGUILAR & ANDER-EGG, 1995, p.42).

Assim, os autores afirmam que essas avaliações são na verdade duas formas de avaliação de resultado "*que dão informação sobre a execução, funcionamento e resultados ou efeitos de um programa ou projeto*" (AGUILAR & ANDER-EGG, 1995, p.42).

No entanto, considerando o critério — aspecto/atividade que pode ser avaliado — Aguilar & Ander-Egg (1995, p. 55) afirmam que a avaliação realizada ao término do programa deve avaliar a eficácia (resultados) e eficiência[135] (rentabilidade econômica) de um programa.

---

[135] A ser detalhado no Capítulo V.

Apontando também para avaliações que ocorrem após o término do programa, pode-se encontrar os organismos supranacionais BID (1997) e SIEMPRO (1999).

Para o primeiro, na perspectiva do Marco Lógico, assim como se realizam avaliações antes do início ou durante a execução do projeto, também se faz uma avaliação após o término do projeto que é chamada de avaliação *ex-post* e tem uma função somativa. Essa avaliação "*reexamina la identificación y diseño original e indaga sobre su ejecución y desempeño de desarrollo*" (BID, 1997, p. 5).

No entanto, o BID (1997) diferencia dois tipos de avaliação ex-post: *a evaluación de terminación e a evaluación de pos-terminación o ex-post.*

Em relação a esses dois tipos de avaliação *ex-post*, o BID (1997, p. 37), afirma que ao término do projeto

> "*...se producen evaluaciones de terminación y de pos-terminación o ex-post. Tal continuidad en el ciclo de evaluación asume que la terminación de un proyecto no significa el final de las contribuciones del mismo al proceso de desarrollo*".

Assim, para o BID (1997), a *evaluación de terminación* é realizada imediatamente após a execução de um programa e esta é operacionalizada por três subtipos de avaliações assim chamadas: *informes de terminación de proyectos, análisis de ejecución de proyectos* e *evaluaciones ex-post del prestatario*[136].

---

[136] Para maiores detalhes ver BID (1997).

Já em relação à avaliação denominada de *evaluación de pos-terminación o ex-post*, o BID (1997) afirma que a mesma deve ocorrer preferencialmente após três anos do término do programa. Esta avaliação também é conhecida como *evaluación operativa y ex-post o evaluación de impacto.*

Para o BID (1999, p.45), esse tipo de avaliação em um projeto tem como objetivo "*entender sus efectos sobre el proceso de desarrollo y usar este conocimiento para confirmar sus éxitos, inclusos aprovechar la experiencia adquirida par aprender de proyectos cuyo impacto fue modesto o negativo*".

Ainda sobre a avaliação operativa e *ex-post* ou avaliação de impacto[137], o BID (1997) afirma que esse tipo de avaliação além de deter-se sobre os efeitos e impactos deve também considerar outros quatro "*conjuntos importantes de questiones generales: la eficiencia, la efetividad*[138] *la continuidad de la justificación*" (BID, 1997, p. 39).

Em relação ao último tipo apontado — *continuidad de la justificación del proyecto* —, o BID (1997, p.39) afirma que esta avaliação se "*refiere al grado hasta la cual las condiciones de arranque y el diseño del proyecto continúan siendo compatibles con las prioridades de desarrollo del país y sus políticas macro y de sector*".

Segundo o SIEMPRO (1999, p. 61) a avaliação que ocorre ao término do programa é chamada de *evaluación de la finalización del programa* e seu propósito é

---

[137] Segundo o BID (1997, p.47) "las evaluaciones de impacto u OERs *ex-post* examinan los efectos e impactos tecnológicos, institucionales, económicos, políticos, socioculturales y ambientales."

[138] Será detalhada no Capítulo V.

> *"...conocer la eficacia, la eficiencia y la efectividad de la inversión social y saber si el programa modifico las condiciones de vida de la población beneficiaria en relación con el problema que le dio origen".*

Para realizar essa avaliação são utilizadas duas ferramentas: *evaluación de resultados e evaluación de impacto.*

A primeira ferramenta, a avaliação de resultados, pode ser utilizada em um período durante a execução do programa (resultados em médio prazo) ou logo após a sua conclusão. Essa avaliação segundo o SIEMPRO (1999, p.101-102):

> *"Permite indagar y analizar el cumplimiento de las metas previstas para cada uno de los objetivos específicos del programa (metas de resultado) en términos de su efectividad, su eficácia y su eficiencia[139], para obtener conclusiones que sirvan tanto para retroalimentación de la gestión del programa como para la identificación, planificación y gestión de otros programas y políticas sociales".".*

Já a segunda ferramenta, a avaliação de impacto, é uma análise que *"permite conocer si el programa social ha generado cambios relevantes en la condiciones de vida de la población beneficiaria"* (SIEMPRO, 1999, p.106), e por isso deve ser executada só após um determinado tempo que o programa foi concluído.

Em relação às avaliações que ocorrem ao término do programa, proposto pelo BID (1997) e SIEMPRO (1999), *evaluación operativa y ex-post (evaluación de impacto)* e a *evaluación del impacto*

---

[139] Esses conceitos serão abordados no Capítulo V.

pode-se inferir que ambos compartilham de definições que podem ser consideradas equivalentes.

Já os tipos de avaliação propostos por Martinic (1997), como já foi dito nas seções anteriores deste capítulo, eles foram classificados de duas maneiras distintas: utilizando os critérios apontados no início deste capítulo e considerando a avaliação como parte integrante do ciclo de vida do projeto.

Desta forma, considerando o critério momento em que se avalia, ou seja, após o término do projeto, Martinic (1997) aponta para uma avaliação denominada de *evaluación ex-post*. Em relação a essa avaliação, o autor, afirma que:

> *"En los proyectos terminados, la evaluación ex-post tiene que ver con la decisión de continuar implementando el proyecto ya que hay más beneficios que costos o si se suspende la implementación por que ocurre la situación contraria"*. (MARTINIC, 1997, p. 21)

Já em relação ao critério aspecto/atividade a ser avaliado, Martinic (1997) aponta para a avaliação denominada de *evaluación del impacto*. Esse tipo de avaliação, segundo o autor, tem como objeto estudar as mudanças produzidas na situação ou problema que gerou tal projeto.

Nesse tipo de avaliação realiza-se a análise dos resultados com o objetivo de verificar se o projeto contribuiu para a solução do problema. Essa avaliação também pode ser chamada de avaliação somativa e "*suele hacerse una vez terminado el proyecto o algunos años después*". Portanto, segundo Martinic (1997, p.21), a avaliação de impacto[140]

---

[140] Este autor também aponta para uma relação direta entre a avaliação de impacto e os modelos que foram apresentados no Capítulo III.

> *"...estudia los cambios que produce el proyecto en la situación problema que aborda. Esta puede centrarse en el análisis del resultado intermedio o bien en los resultados finales o de impacto propiamente".*

Martinic (1997) em um segundo momento, desta vez considerando que o projeto tem um ciclo de vida com diversas fases, afirma que a avaliação é um processo integrado ao seu desenvolvimento.

Partindo desse pressuposto o autor, remete-nos para a avaliação de impacto, citada anteriormente, e detalha mais um pouco sobre alguns aspectos/atividades que podem ser avaliados após o término de um programa: a eficácia ou efetividade[141].

Em relação à avaliação de impacto e os aspectos/ atividades que são avaliados nesse tipo de avaliação, pode-se afirmar que o BID (1997), SIEMPRO (1999) e Martinic (1997) compartilham da mesma perspectiva, mesmo estabelecendo suas tipologias de maneira diferente.

Já Faria (1998) estabelece os tipos de avaliação sem mencionar que critérios foram utilizados. Mesmo assim, diante das definições apresentadas pela autora, é possível afirmar que, as avaliações apresentadas ocorrem ao término do programa.

Assim, esta autora aponta para uma avaliação que é realizada após o término do programa denominada de: avaliação *ex-post*, avaliação somativa ou avaliação de resultado.

---

[141] Esses aspectos serão abordados no Capítulo V.

Esse tipo de avaliação, segundo Faria (1998, p.45),

> *"Envolve estudos comparativos entre programas rivais, subsidia a decisão e avalia, principalmente a maior ou menor efetividade*[142] *de diferentes tratamentos oferecidos ao grupo –alvo. Seu objetivo primordial é permitir a escolha do melhor programa para alcançar as metas de médio e longo prazo".*

Figueiredo & Figueiredo (1986, p. 111) também apontam para uma avaliação que ocorre ao término de uma política ou programa, mas não se preocupam em estabelecer critérios. Segundo esses autores, a avaliação de impactos relaciona-se com os efeitos de uma política ou programa e tem como objetivo "*estabelecer uma relação de causa entre a política e as alterações nas condições sociais*".

Para os autores, como uma mesma política pode promover três tipos de impactos, cabe aferir a esta política ou programa, três tipos de efetividade[143]. Em relação a esses impactos, afirmam que:

> *"Uma mesma política pode causar impactos objetivos, gerando mudanças quantitativas nas condições da população-alvo, pode ainda gerar impactos subjetivos alterando o estado de espírito da população, e, finalmente, pode causar um impacto substantivo mudando qualitativamente as condições de vida da população"* (FIGUEIREDO & FIGUEIREDO, 1986, p. 116).

---

[142] Esse aspecto será abordado no Capítulo V.
[143] Serão abordados no Capítulo V.

Cohen & Franco (1999) utilizando o critério do momento em que se avalia, afirmam que a avaliação que ocorre após o projeto ter sido concluído é chamada de avaliação *ex-post* ou avaliação terminal. Nesse tipo de avaliação os objetos a serem avaliados são os impactos produzidos pelo projeto.

Para os autores, a avaliação de impacto tem como objetivo "*determinar em que medida o projeto alcança seus objetivos e quais são os seus efeitos secundários (previstos ou não previstos)*" (COHEN & FRANCO, 1999, p.109). Ou seja, o avaliador "olha para trás", observando se o projeto funcionou ou não descobrindo as causas.

Para Cohen & Franco (1999, p. 109) as informações obtidas através da avaliação de impacto de um projeto "*é utilizável para decidir sobre sua eventual continuação, para formular outros projetos futuros e, enfim, para tomar decisões sobre política*".

Ainda sobre a avaliação de impacto, os autores afirmam que as metodologias utilizadas para operacionalizar esse tipo de avaliação são: o método experimental, o quase experimental ou não–experimental [144].

Tenório (2000, p.63) sem se preocupar com critérios para estabelecer os tipos de avaliação, afirma, que após o término de um projeto, é necessário realizar uma avaliação com o objetivo de coletar e analisar informações para *"fazer um julgamento objetivo de um projeto concluído"* .

---

[144] para maiores detalhes ver COHEN & FRANCO (1999).

Segundo esse autor, essa avaliação é chamada de avaliação final ou avaliação de impacto e:

> *"...serve para conhecer como foi executado o projeto, dar valor aos resultados obtidos, tomar decisões sobre a continuidade ou interrupção das ações, eliminar ou diminuir os efeitos negativos, aumentar os efeitos positivos ou ainda, decidir sobre um novo projeto com base em informações seguras"* (TENÓRIO, 2000, p. 64).

Para Tenório (2000, p.73–75) existem várias formas de realizar uma avaliação final ou de impacto e que são assim definidas:

- avaliação do processo de gestão, que tem como objetivo verificar *"como foram tomadas as decisões de gestão";*

- avaliação de objetivos e metas tem como objetivo verificar se *"os objetivos e metas programados foram alcançados, com que velocidade e quais os custos para isso";*

- avaliação comparativa que tem como objetivo *"comparar a situação existente antes e depois da implementação do projeto";*

- avaliação *custo-benefício* tem como objetivo *"comparar os custos do projeto com os benefícios do projeto".*

Após as explanações acerca das avaliações que ocorrem após o término do programa podem ser feitas algumas considerações a respeito do tema.

A primeira é que também pode ser observado, na tipologia de avaliação que ocorre ao final do programa, a utilização de termos diferentes que, na verdade, representam a mesma coisa.

A segunda é que as avaliações que são realizadas ao final do programa têm como foco de interesse os resultados e efeitos do programa, e detêm-se sobre a eficácia, eficiência e efetividade do programa.

E, por fim, a terceira consideração é que as avaliações que ocorrem ao final do programa, também conhecida como *ex-post*, têm como objetivo obter informações acerca do sucesso ou fracasso do programa para que, desta forma, possam auxiliar na tomada de decisão no que diz respeito a continuar, cancelar ou reaplicar o programa.

O **Quadro 11** possibilitará uma melhor visualização dos tipos de avaliação apresentados nesta seção.

Após a apresentação dos tipos de avaliação, pode-se observar que os termos eficácia, efetividade e eficiência aparecem freqüentemente, quando são apresentados os conceitos e objetivos da avaliação. Estes termos se relacionam com os aspectos que podem ser avaliados ou são utilizados como critérios para avaliar uma determinada política, programa ou projeto.

Tendo em vista esta realidade, o próximo capítulo deter-se-á, nos critérios de avaliação: eficácia, efetividade e eficiência, apresentados pelos autores.

Quadro 11

Nomenclatura utilizada para os tipos de avaliação do grupo 3 por autor

	Nomenclatura utilizada para as avaliações	
Autor	Critério A	Critério B
Alkin (1972)	N.C	avaliação para certificação do programa
Holister (1972)	avaliação de resultados	avaliação de resultados (custo-benefício)
Worthen (1997)	avaliação somativa	N.C
Weiss (1998, 1999)	avaliação somativa avaliação de resultados	avaliação de resultados
Viedma (1996)	avaliação de resultados	N.C
Briones (1998)	evaluación terminal o expost evaluación *ex-post*	evaluación de resultados
Ballesteros (1996)	evaluación somativa	evaluación de resultados o impactos; evaluación de eficácia, eficiência e efetividade
Ander-Egg (1990)	evaluación final	evaluación de los resultados efectivos; evaluación de la coordinación; evaluación de los efectos indirectos; opinión de os beneficiários.
Lara (1996)	N.C	evaluación del producto (logros o impacto) evaluación de la eficacia, efectividad e eficiencia; evaluación del a pertinencia
Aguilar & Ander-Egg (1995)	avaliação de processo; avaliação da gestão; avaliação concorrente; avaliação concomitante ou monitoramento.	avaliação da cobertura; avaliação da implementação; avaliação do ambiente organizacional; avaliação do rendimento pessoal.
BID (1997)	evaluación ex-post (terminación o de pos-terminación); .	evaluación del impacto (operativa y *ex-post)* ; evaluación de los efectos y impactos; evaluación de la eficiencia e efectividad; continuidade de la justificación
SIEMPRO (1999)	evaluación de finalización	evaluación de impacto; evaluación de resultados
Martinic (1997)	evaluación ex-post	evaluación de impacto evaluación de los resultados; evaluación de la eficacia o efectividad
Faria (1998)	avaliação ex-post, avaliação somativa	avaliação da efetividade (medida de resultado)
Figueiredo & Figueiredo (1986)	N.C.	avaliação de impacto (efetividade objetiva, subjetiva substantiva)
Cohen & Franco (1999)	avaliação *ex-post* ou avaliação terminal	avaliação de impacto
Tenório (2000)	avaliação final ou avaliação de impacto	avaliação de processo de gestão; avaliação dos objetivos e metas; avaliação comparativa; avaliação custo-benefício.

CAPÍTULO V

Critérios de avaliação: eficácia, efetividade e eficiência

Após a análise da literatura acerca da avaliação de políticas, programas e projetos, pode-se perceber que os termos eficácia, efetividade e eficiência estão intimamente relacionados com o processo de avaliação. Esta relação pode ser vista no Capítulo II e no Capítulo IV, quando foram abordados respectivamente os conceitos e a tipologia de avaliação.

No Capítulo II, os termos eficácia, efetividade e eficiência estão presentes em muitos dos conceitos analisados e se relacionam com os objetivos e propósitos da avaliação. Como exemplo, pode-se citar alguns autores, que utilizam esses termos ao estabelecerem o objetivo da avaliação:

- *"Medir el grado de pertinencia, idoneidad, efectividad y eficiencia de un proyecto"*. (ANDER–EGG, 1990, p. 29);

- *"Estabelecer o grau de pertinência, idoneidade, efetividade/eficácia e eficiência e rendimento e um programa, projeto ou serviço"*. (AGUILAR & ANDER-EGG, 1996, p.62);

- *"...proporcionar información continua para tomar decisiones con respecto a la eficacia, la eficiencia y la calidad de los resultados e impactos que genera el programa"*. (SIEMPRO, 1999, p.56);

- *"Mejorar la efectividad cuantitativa del programa, es decir, el logro de sus objetivos;* [...] *aumentar la eficiencia interna del programa, vale decir, la relación entre insumos y resultados"* (BRIONES, 1996, p.17).

Segundo Figueiredo & Figueiredo (1996, p. 111), a relação entre os objetivos de avaliação e seus critérios (aferição de sucesso ou fracasso) pode ser descrito da seguinte forma:

> *"...a conexão lógica pode ser resumidamente delineada: a definição dos propósitos da avaliação da política ou programa condiciona os objetivos da avaliação; estes, por sua vez, condicionam o tipo de pesquisa empregado na avaliação; esta escolha, por sua vez, determina o*

critério de sucesso, que, por último, estabelece o Modelo analítico[145] *de aferição do sucesso".*

Portanto, os objetivos citados acima, relacionam-se com os modelos de avaliação apresentados no Capítulo III e com os tipos de avaliação abordados no Capítulo IV. Assim, de forma geral, os termos eficácia, efetividade e eficiência, em sua grande maioria, relacionam-se com a avaliação que ocorre ao término do programa e com os aspectos/atividades que são avaliados durante esta avaliação, podendo, no entanto, haver algumas exceções.

Como exemplo dessa relação, pode-se citar as avaliações que têm como objetivo avaliar a eficácia, efetividade e eficiência dos programas e projetos, tais como : *evaluación final* (ANDER-EGG, 1990); *evaluación de pós terminación o ex-post* (BID, 1997); avaliação *ex-post* (AGUILAR & ANDER-EGG, 1996); *evaluación de la finalización* (SIEMPRO, 1999); *evaluación del producto o impacto* (LARA, 1996); *evaluación ex-post* (MARTINIC, 1997), avaliação final ou de impacto (TENÓRIO, 2000); avaliação de processos e avaliação de impacto (FIGUEIREDO & FIGUEIREDO, 1986) e avaliação de processos e avaliação de resultados (FARIA, 1998).

Em relação aos critérios eficácia, efetividade e eficiência, Ballesteros (1996, p.68-69) afirma que estes termos são semanticamente e etimologicamente semelhantes, mais tecnicamente são diferentes.

> *"Así eficaz/eficacia procede de efficaxàcis significando aquello que es activo o poderoso para actuar. Por su parte, efecto / efectividad procede de effectus cuyo significado supone lo que sigue por virtud de una causa. Finalmente, eficiente / eficiencia procede del latín*

---

[145] O termo modelo analítico utilizado por FIGUEIREDO & FIGUEIREDO (1986) não se relaciona com os modelos analíticos, vistos no Capítulo III.

> *efficientia (ex-facère) significando aquello que tiene la virtud para lograr un efecto determinado" (...) con base en esa mínima diferenciación terminológica y más aun, "en la depuración técnica ocurrida a todo lo largo de la historia de la evaluación de programas, cada uno de estos vocablos se utiliza con una significación técnica relativamente precisa".*

No entanto, essa distinção parece não ser consensual, pois dentre os autores analisados, alguns utilizam eficácia e efetividade como sinônimos.

A partir das observações acerca do tema, é possível destacar os seguintes aspectos. O primeiro é que existem autores que fazem distinção entre eficácia e efetividade. O segundo é que existem autores que não fazem distinção e utilizam os termos eficácia e efetividade como sinônimos. O terceiro é que, independentemente dos autores fazerem ou não a distinção entre os termos eficácia e efetividade, eles conceituam o termo eficiência. O quarto aspecto é que nem todos os autores dos Capítulos III e IV conceituam os termos eficácia, efetividade e eficiência.

Considerando todos os aspectos acima mencionados, este capítulo tem como objetivo apresentar os diversos conceitos dos critérios de avaliação eficácia, efetividade e eficiência destacando as diferenças e ou semelhanças entre os autores.

Como foi dito anteriormente, alguns autores não fazem distinção entre eficácia e efetividade, portanto, estes termos serão o apresentados em uma única seção. Assim, este capítulo é composto de duas seções. A primeira deter-se-á sobre os aspectos relacionados à avaliação de eficácia e efetividade E a segunda seção apresentará os aspectos relacionados à avaliação da eficiência.

5.1. Eficácia e efetividade

Na literatura que trata da avaliação de política, programas e projetos sociais e educacionais, os termos eficácia e efetividade podem estar dentro de três perspectivas: a primeira quando os autores só se referem ao termo efetividade sem fazer menção ao termo eficácia. A segunda perspectiva é quando a eficácia e efetividade são utilizadas como sinônimos. Neste caso os autores ora utilizam o termo eficácia ora o termo efetividade. E a terceira perspectiva é quando os termos eficácia e efetividade são diferenciados e significam coisas diferentes.

Dentro da primeira perspectiva pode-se apontar Ander-Egg (1990). Este autor não faz distinção entre eficácia e efetividade de um programa e nem utiliza esses dois termos como sinônimos. O autor refere-se apenas ao termo efetividade.

Para Ander-Egg (1990, p. 30) a efetividade é a "*capacidad que tiene las actividades y tareas del proyecto para alcanzar los objetivos y metas que el mismo define en cantidad, calidad y tiempo*". Ou seja, um programa é considerado efetivo quando atinge os objetivos propostos.

Portanto, para Ander–Egg (1990, p.30) avaliar a efetividade de um projeto é medir *"el grado de éxito del proyecto en la consecución de los objetivos y metas propuestos"*.

O BID (1997) da mesma maneira que Ander- Egg (1990) não estabelece uma definição para a eficácia e refere-se apenas a efetividade de um projeto. Dentro da perspectiva de Análise do Marco Lógico, o BID (1997, p. iv) afirma que a efetividade é o "*grado en el cual un proyecto*

logró los resultados previstos y, por tanto, alcanzó su propósito[146] *y contribuyó a su fin*[147]".

Desse modo, a avaliação da efetividade segundo o BID (1997, p.41)

> *"Se ocupa del grado hasta el cual el proyecto produjo los componentes esperados, los componentes obtenidos conducen al logro del propósito del proyecto y el logro del propósito contribuyó al fin".*

No que diz respeito à avaliação de efetividade proposta por Ander-Egg (1990) e BID (1997), pode-se inferir que as definições possuem diferenças, se considerarmos o significado dos termos objetivos e resultados[148]. Isto significa dizer que como esses termos não são utilizados para significarem a mesma coisa, em cada um dos autores, portanto, a definição de efetividade, representa a avaliação de aspetos diferentes.

Briones (1998), diferentemente de Ander-Egg (1990) e BID (1997), utiliza os termos eficácia e efetividade como sinônimos e afirma que estes termos se relacionam com o alcance dos objetivos propostos pelo programa.

---

[146] Para o BID (1997, p.17) "el **propósito** es el impacto directo a ser logrado como resultado de la utilización de los componentes producidos por el proyecto. Es uma hipótesis sobre el impacto o beneficio que se desea lograr".
Segundo o BID (1997, p. iv) "los impactos e efectos se refieren a las consecuencias planeadas y/o no previstas de un proyecto; los efectos geralmente se relacionan com el propósito, mientras que los impactos se refieren a su fin".

[147] Para o BID (1997, p. 17) "el **fin** es una definición de como el proyecto o programa contribuirá a la solución del problema (o problemas) del setor".

[148] Para Ander-Egg (1990) "un programa puede estar alcanzando sus objetivos, hacer una buena utilización de sus recursos, tener una base administrativa adecuada, etc., sin enbargo, podría no responder a las necessidades y problemas reales existentes". Os resultados são exatamente a resposta a essa necessidade e a resolução do problema, portanto, para Ander-Egg (1990), resultados e objetivos se relacionam mais significam coisas distintas.

Sobre a avaliação da eficácia/efetividade, Briones (1998, p. 93) afirma:

> *"...el objetivo básico de todo programa de acción es producir alguna modificación o cambio en una cierta población, la tarea principal de la investigación evaluativa consiste, paralelamente, en establecer si en los hechos el programa, mediante la metodología empleada se alcanzan o no el o los objetivos propuestos. Se trata, como se advierte, de verificar la a avaliação da efetividade".*

Segundo Briones (1998, p.165) a avaliação da eficácia/efetividade *"es equivalente a establecer los logros*[149] *obtenidos"*. Considerando a definição do termo *logros*, estabelecida pelo autor, pode-se afirmar que a avaliação da eficácia/efetividade, se preocupa com as mudanças produzidas pelo projeto.(que ele chama de resultado).

Aguilar & Ander-Egg (1995), assim como Briones (1998) utiliza eficácia e efetividade como sinônimos. Para os autores, esses termos relacionam-se com o alcance dos resultados previstos, efeitos esperados dentro de um determinado tempo previsto.

Deste modo, avaliar a eficácia ou efetividade de um programa é avaliar o seu efeito ou resultado, levando em consideração o contexto no qual está inserido e a população beneficiária. Ou seja, se o programa responde ou não às necessidades que deram origem a ele.

Partindo desse pressuposto, para Aguilar & Ander-Egg (1995, p.114) a avaliação da eficácia ou efetividade, também chamada de avaliação de resultados ou impacto, consiste em

---

[149] "Los resultados de um programa son los câmbios o modificaciones que produce en una población o en un cierto contexto. Si tales modificaciones son las esperadas según

> *"...avaliar os resultados obtidos em termos de efeitos (sobre a população destinatária) e impacto (sobre o conjunto da população e o meio), isto é, de consecução dos fins propostos para resolver os problemas".*

Considerando, portanto, as definições de eficácia/efetividade propostas por Aguilar & Ander Egg (1995) e Briones (1998), pode-se afirmar que ambos se assemelham.

Martinic (1997), fazendo parte também da segunda perspectiva, não diferencia os termos eficácia e efetividade, utilizando-os como sinônimos. Este autor, no que diz respeito ao conceito de eficácia/ efetividade, posiciona-se da mesma maneira que Aguilar & Ander-Egg (1995) e Briones (1998) ao afirmar que a eficácia/efetividade são termos que se relacionam com os resultados ou produtos alcançados.

Segundo Martinic (1997, p.120) avaliar a eficácia ou efetividade de um programa "*es el análisis de los efectos que resultan de la utilización del conjunto de productos en los beneficiários directos y actores que inciden en la solución del problema*".

Já fazendo parte da terceira perspectiva — distinção entre eficácia e efetividade — e posicionando-se diferentemente dos autores anteriormente mencionados, pode-se encontrar Ballesteros (1996).

Para esta autora, a eficácia relaciona-se diretamente com a consecução dos objetivos inicialmente propostos no momento do planejamento e elaboração do programa ou projeto. Isto significa dizer

---

las definiciones que proporcionan los objetivos, las llamamos de **logros**". (BRIONES, 1998, p.159).

que, se o programa alcançou os objetivos propostos, este pode ser considerado eficaz.

Ao se referir à efetividade, Ballesteros (1996) recorre ao que Scriven estabeleceu como a efetividade de um programa, e afirma que um programa é considerado efetivo quando este produz um efeito positivo, independentemente do alcance dos objetivos propostos inicialmente.

Seguindo nesta mesma linha de distinção entre eficácia e efetividade, pode-se apontar Lara (1996). Para este autor, a eficácia é a capacidade que o programa tem de induzir o efeito que pretendia. Avaliar a eficácia é determinar em que medida os objetivos pretendidos foram alcançados.

Em relação à efetividade, Lara (1996, p. 116) afirma que este termo se relaciona com a capacidade de um programa produzir efeitos em uma determinada situação como conseqüência de sua implementação. Portanto, avaliar a efetividade de um programa é "*determinar em que medida produce efectos sobre la realidad social en que actua*".

Cohen & Franco (1999) também afirmam que a eficácia e efetividade se referem a coisas diferentes. Para eles a eficácia relaciona-se com o alcance dos objetivos e metas dentro de um determinado período de tempo, enquanto a efetividade é estabelecida a partir da relação entre resultados concretos (medida de impacto[150]) e objetivos atingidos.

---

[150] "O impacto é definido como resultado dos efeitos de um projeto (ONU, 1984) (...) É a conseqüência de um projeto" (COHEN & FRANCO, 1999, p. 92-93). O "Efeito é todo comportamento ou acontecimento que se pode razoavelmente dizer que sofreu

Assim, Cohen & Franco (1999, p.102) definem a eficácia como "*o grau em que se alcançam os objetivos e metas do projeto na população beneficiária, em um determinado período de tempo, independentemente dos custos implicados*".

Detalhando um pouco sobre a efetividade de um projeto ou programa, esses autores, utilizando a definição de López (1995), afirmam que a

> "*...efetividade é um termo que se usa freqüentemente para expressar o resultado concreto ou as ações conducentes a esse resultado concreto – dos fins, objetivos e metas desejadas*" (LOPEZ, 1995, apud. COHEN & FRANCO, 1999, p.107)".

No entanto, essa efetividade, pode ser observada sob duas dimensões: a medida de impacto (avaliação do impacto) e o grau do alcance dos objetivos.

Trazendo uma definição bem mais detalhada a respeito dos critérios de eficácia e efetividade, pode-se apontar Figueiredo & Figueiredo (1986). Esses autores afirmam que dentro do processo de execução podem ser avaliados aspectos diferentes, portanto pode-se aferir três tipos de critério de eficácia:

> "*O critério de sucesso usado é o da eficácia objetiva, isto é, se as metas atingidas são iguais, superiores ou inferiores às metas propostas.*" [...] *Quando a intenção da avaliação está na moralidade*[151] *executória, os*

---

influência se algum aspecto do programa ou projeto" (COHEN & FRANCO, 1999, p. 91).

[151] Segundo os autores, o motivo de avaliar uma política ou programa pode estar pautado em uma motivação moral ou instrumental. "A motivação pode estar na moralidade positiva, legal, notadamente no que se refere à probidade dos gestores e usuários na gestão do programa e na apropriação dos seus benefícios." Esses tipos de avaliações são chamados de auditorias contábeis (financeiras ou materiais) (...) A

> *critérios da eficácia administrativa e contábil são usados. [...] no caso da intenção instrumental*[152]*, o critério adequado é o da eficácia funcional, cujos Modelos analíticos são construídos para aferir se os meios e a metodologia de implantação do programa estão sendo empregados de acordo com as estratégias previamente definidas".* (FIGUEIREDO & FIGUEIREDO, 1986, p. 112).

Esses autores relacionam a efetividade com os impactos, proporcionados pela política ou projeto. Neste caso, como uma política pode promover vários tipos de impactos (objetivos, subjetivos e substantivos) podem ser utilizados vários critérios de efetividade. Sendo assim, para Figueiredo & Figueiredo (1986, p. 116),

> *"Aefetividade objetiva, é o critério de aferição da mudança entre o antes e o depois da execução do programa.[...] A efetividade subjetiva torna-se o critério de aferição de mudanças psicológicas, de mudanças nos sistemas de crenças e valores e, notadamente, tem função de aferir a percepção da população sobre a adequação dos resultados objetivos dos programas aos seus desejos e aspirações e demandas* [...] *A efetividade substantiva é, por vez, o critério de aferição de mudanças qualitativas nas condições sociais de vida da população alvo".*

Tenório (2000) também diferencia eficácia de efetividade, porém traz uma definição menos detalhada do que a de Figueiredo & Figueiredo (1986). Para este autor, a eficácia relaciona-se com o alcance dos resultados previstos. Avaliar a eficácia de um projeto "*é verificar se*

---

motivação pode estar na moralidade política ou social. Ou seja se os propósitos da política ou programas são condizentes com princípios de justiça política ou social minimamente aceitos". (FIGUEIREDO & FIGUEIREDO, 1986, p.110).

[152] "A motivação instrumental está intimamente ligada com a implementação da política ou programa. Neste caso, o objetivo básico da avaliação é a geração de

os resultados previstos foram alcançados, em que quantidade e qualidade" (TENÓRIO, 2000, p.76).

Em relação à efetividade, o autor afirma que este termo se relaciona com o impacto do projeto. Este autor, sem entrar em detalhes, afirma que avaliar a efetividade é verificar se *"o projeto respondeu de forma adequada às necessidades"*. (TENÓRIO, 2000, p. 76)

A partir da perspectiva de Gestão Integral de Programas, o SIEMPRO (1999), assim como Tenório (2000), Cohen & Franco (1999), Lara (1996) e Ballesteros (1996), aponta diferenças entre eficácia e efetividade.

Em relação à avaliação da eficácia, o SIEMPRO (1999, p.104) afirma que:

> *"El análisis de la eficácia considera la relación existente entre las metas de resultados efectivamente alcanzadas y el tiempo previsto, respecto de las metas previstas de resultados y el tiempo real".*

Para o SIEMPRO (1999, p.104), a efetividade relaciona-se com o alcance dos resultados independentemente do tempo previsto. Então, *"el análisis de la efectividad considera la relación existente entre las metas de resultado efectivamente alcanzadas y las metas de resultados planificadas"*.

A partir das definições propostas pelo SIEMPRO (1999), pode-se inferir que tanto a avaliação da eficácia quanto à avaliação de efetividade analisam a relação entre resultados alcançados e previstos. No entanto, para avaliar a eficácia, o fator tempo é inserido como um

---

informações para monitorar a execução do programa" (HARGROVE, 1975, apud

determinante para esta análise. Isto significa dizer que, a diferença entre eficácia e efetividade proposta pelo SIEMPRO (1999) é estabelecida a partir do fator tempo.

Diante da conceituação de eficácia e efetividade feita pelo SIEMPRO (1999), Tenório (2000), Cohen & Franco (1999), Lara (1996) e Ballesteros (1996), pode-se inferir que esses autores, ao apresentarem as diferenças entre ambas, o fazem sob perspectivas diferentes.

Neste sentido, pode-se afirmar que, enquanto o SIEMPRO (1999) estabelece como diferença entre a eficácia e efetividade o fator tempo, os demais autores acima citados a distinguem não apenas pelo fator tempo, mas também pelo objeto (aspecto) que é avaliado.

Faria (1998), assim como SIEMPRO (1999), Tenório (2000), Cohen & Franco (1999), Lara (1996) e Ballesteros (1996), também estabelece diferenças entre eficácia e efetividade.

No entanto, apresenta uma conceituação de eficácia que não se assemelha com os autores acima relacionados. Isso se deve ao fato de que, para a autora, a eficácia é uma dimensão do processo de execução do programa. Isto significa dizer que, para Faria (1998, p.46) a avaliação da eficácia é a "*medida do processo que responde a questões sobre os acertos do programa em sua fase de elaboração e implementação*" enquanto para os demais a eficácia relaciona-se com o término do programa e com o atingimento de objetivos e metas.

Porém, ao realizarmos uma análise mais detalhada, considerando que o alcance dos objetivos está intimamente relacionado ao processo de implementação de um programa e, que o desenvolvimento deficiente das atividades acarreta a não consecução

---

FIGUEIREDO & FIGUEIREDO, 1986, p.110)

desses objetivos, pode-se dizer que Tenório (2000), Cohen & Franco (1999), Lara (1996) e Ballesteros (1996), ao estabelecerem a eficácia de um programa, estão, indiretamente, também se preocupando com aspectos relacionados ao processo de execução (implementação) do programa. Assim, Faria (1998) e esses autores aproximam-se, no que diz respeito ao conceito de eficácia.

Ainda em relação ao conceito de eficácia proposto por Faria (1998) pode-se afirmar que este se aproxima de uma das dimensões da aferição de eficácia, proposta por Figueiredo & Figueiredo (1986).

Já em relação à efetividade Faria (1998, p.46) afirma que esta se relaciona com os resultados do programa, ou seja, com os alcance de metas. Sendo assim, para esta autora a avaliação da efetividade é a medida de resultado, que *"responde as questões relacionadas com o alcance de metas"*.

Já em relação ao conceito de efetividade, pode-se dizer que tanto Faria (1998) quanto Tenório (2000), Cohen & Franco (1999), Lara (1996) e Ballesteros (1996), compartilham da mesma definição.

Tendo em vista, que os autores fazem distinção entre os termos abordados nesta seção e a eficiência de um programa, na próxima seção serão abordados os conceitos de eficiência.

5.2. Eficiência

Em relação ao termo eficiência, parece que existe um consenso entre os autores analisados, no que diz respeito à sua distinção em relação aos termos eficácia e efetividade. Ou seja, todos os

autores analisados possuem uma definição do critério eficiência, diferenciando-a dos outros termos.

Neste sentido, pode-se apontar Ander-Egg (1990). Para este autor a eficiência é a relação entre os resultados e os recursos aplicados. Ao realizar a avaliação da eficiência do projeto determina-se "*el indice de productividad del proyecto*" e assim pode-se estabelecer se este projeto pode ser realizado de maneira mais econômica ou não.

Para Ander-Egg (1990, p. 30) a avaliação da eficiência de um projeto é:

> *"...la relación existente entre los bienes y servicios finales (resultados obtenidos) y los insumos requeridos para su realización (esfuerzos desplegados)". Con ello, este tipo de evaluación permite establecer si un programa se puede realizar de manera más económica y en menor tiempo".*

O BID (1997) assim como Ander-Egg (1990) relaciona o critério de eficiência aos custos e resultados do programa. Assim, a eficiência de um projeto refere-se

> *"...al grado hasta el cual se suministraron y administraron los insumos a través de acciones que se organizaron de la manera más apropriada y se hizo lo anterior al menor costo posible para generar los componentes esperados*" (BID, 1997, p.40).

Briones (1998) não faz uma conceituação explicita do termo eficiência. Este autor, no decorrer do seu trabalho, cita algumas vezes este termo no momento em que dá exemplos de avaliações. Mesmo assim, é possível afirmar que para Briones (1998) a eficiência de um

programa é estabelecida através da relação entre resultados obtidos e o custo investido.

Aguilar & Ander-Egg (1995), na mesma linha de BID (1997) e Ander-Egg (1990), afirmam que a eficiência de um programa relaciona-se com os seus custos e benefícios.

Para Aguilar & Ander-Egg (1995, p. 115) avaliar a eficiência é estabelecer "*o grau de rentabilidade econômica do programa ou serviço comparando benefícios e custos, resultados e custos, ou utilidade e custos*". Essa avaliação é realizada através da análise de custo–benefício, análise de custo-efetividade e análise de custo-utilidade[153].

Para Martinic (1997), a eficiência relaciona-se com os recursos e insumos para atingir as atividades e obter os resultados esperados. Neste sentido, o autor assemelha-se com Aguilar & Ander-Egg (1995), Briones (1998), BID (1997) e Ander-Egg (1990). Para este autor, um projeto eficiente é aquele que possui um gasto adequado e que não proporciona um déficit em sua operação.

Segundo Martinic (1997, p. 119) a avaliação da eficiência de um projeto:

> *"...el análisis de los costos o recursos realmente utilizados para generar una unidad de productos en relación con el costo programado para la misma unidad de producto en un tiempo determinado".*

Ainda sobre a eficiência Martinic (1997, p.119) afirma que:

> *"El proyecto es más eficiente si, minimiza el costo total prometido por una unidad de producto o, si el gasto está*

---

[153] Para maiores detalhes ver AGUILAR & ANDER-EGG (1995).

> *prefijado, busca optimizar la combinación de insumos para maximizar el producto*" (UNICEF, 1991, apud. COHEN,& FRANCO, et al, 1992)

Para Ballesteros (1996, p. 71), assim como para os autores anteriormente citados, a eficiência está relacionada com os recursos investidos no programa ou projeto e os resultados obtidos a partir desse investimento. Para esta autora, "*la actuación o programa más eficiente será aquél que consiga mejores resultados en comparación con los recursos inversos*".

Já segundo Lara (1996, p. 117), avaliar a eficiência de um programa "*é determinar si los custos justifican los resultados esperados*". Um programa pode ser considerado eficiente, quando existe um equilíbrio entre os recursos aplicados e os resultados alcançados.

Para realizar este tipo de avaliação, é necessário utilizar algumas técnicas como a análise custo-benefício, custo-eficácia ou custo-utilidade[154]. Estes dois últimos, segundo Lara (1996), são mais indicados para a avaliar a eficiência dos programas sociais.

Para Cohen & Franco (1999), a eficiência de um projeto relaciona-se com custos e seus produtos. Para estes autores, o conceito de eficiência é utilizado quando se realiza a análise financeira do projeto e sempre está associada à quantidade de recursos para de obter os produtos[155] programados.

---

[154] Para maiores detalhes, ver LARA (1996).
[155] "Produtos são os resultados concretos das atividades desenvolvidas a partir dos insumos disponíveis" (COHEN & FRANCO, 1999, p. 93).

Segundo Cohen & Franco (1999, p.104) a eficiência é

> *"... a relação existente entre os produtos e os custos dos insumos . portanto, o resultado será sempre o custo de uma unidade de produto final por um beneficiário em uma certa unidade de tempo".*

Figueiredo & Figueiredo (1986) também apontam para a definição do critério de eficiência. Para esses autores, existem dois tipos de eficiência que podem ser atribuídos a uma política ou programa de cunho social: a eficiência instrumental e a eficiência política.

Em relação ao primeiro tipo de eficiência, os autores afirmam que *"esta é definida pela relação estrita entre os custos econômicos e os benefícios que são, em geral, tangíveis e divisíveis"*. Já em relação à eficiência política, *"esta é definida pela relação entre os" custos, sociais ou políticos, e os benefícios derivados*" (FIGUEIREDO & FIGUEIREDO, 1986, p. 114).

Para Tenório (2000, p. 76) a eficiência é a definida como o "*cumprimento de normas e redução de custos*". Portanto, avaliar a eficiência de um projeto é averiguar a relação entre custo e resultado obtido. Para tanto, é necessário analisar o custo-resultado e o insumo/produto do projeto. Um projeto é considerado eficiente quando seus resultados são alcançados dentro do menor custo possível.

O SIEMPRO (1999), ao estabelecer a definição de eficiência, pode-se encontrar aspectos que se assemelham e se diferenciam dos autores anteriormente citados.

A semelhança é que tanto o SIEMPRO (1999) quanto os outros autores, ao definirem o que é eficiência, estabelecem a relação entre custos e resultados alcançados. A diferença, no entanto, é que o SIEMPRO (1999), além de apontar para essa relação, agrega o fator tempo como determinante dessa eficiência.

Assim, para o SIEMPRO (1999, p.104) a avaliação da eficiência "*mide los resultados alcanzados en relación com los custos y tiempos previstos.*"

Faria (1998, p.45) diferentemente dos outros autores, associa a avaliação da eficiência com a modificação dos "*cenários jurídico-administrativos, financeiro, organizacional*". Para a autora, ao realizar o monitoramento, que ocorre durante a execução do programa, está se avaliando a eficiência, que significa "*fazer certo as coisas*".

Após a apresentação dos critérios de avaliação eficácia, efetividade e eficiência, pode-se fazer algumas considerações a respeito do tema. A primeira é que o conceito estabelecido pelos autores que utilizam eficácia e efetividade como sinônimos é semelhante ao conceito de efetividade, proposto pelos autores que diferenciam os dois termos.

A segunda consideração é que todos os autores estabelecem diferenças entre o termo eficiência e os outros dois termos–eficácia e efetividade.

A terceira é que, em relação ao termo eficiência, parece existir um consenso entre os autores no que diz respeito à sua definição e aspectos que se relacionam com a essa avaliação.

A quarta e última consideração diz respeito à realização da avaliação de eficácia, efetividade e eficiência. Neste sentido, os conceitos propostos pelos autores, demostra-nos claramente que para se avaliar a eficácia, efetividade e eficiência de um projeto ou programa, é imprescindível que os objetivos, as metas, os resultados, os custos, etc., estejam explicitamente indicados. Caso contrário, corre-se o risco de realizar análises equivocadas e fornecer informações incompletas.

CONSIDERAÇÕES FINAIS

Antes de fazermos algumas considerações a respeito dos aspectos da avaliação que foram abordados, acreditamos ser necessário enfatizar mais uma vez que o presente trabalho não teve a intenção de discutir questões políticas e ideológicas que permeiam as práticas avaliativas no contexto brasileiro ou internacional.

Nosso objetivo foi o de sistematizar os diversos conceitos, modelos, tipos e critérios de avaliação, que foram construídos, (re)aproveitados, (re)elaborados e (re)utilizados nas mais diversas práticas avaliativas no Brasil e no exterior, demonstrando as semelhanças e ou diferenças entre os autores analisados.

Agora, para iniciarmos propriamente as considerações a respeito da avaliação de políticas, programas e projetos sociais e educacionais, acreditamos ser indispensável retomar os dois aspectos que foram observados na literatura e que deram início a este trabalho: a diversidade analítico-conceitual da metodologia de avaliação de políticas e a escassez de estudos que busquem a sua sistematização.

Em relação à diversidade analítico-conceitual da metodologia de avaliação, é necessário enfatizar que no Capítulo I, na intenção de construir o desenvolvimento da avaliação de políticas, programas e projetos, foram abordados alguns aspectos relacionados à

avaliação de aprendizagem (avaliação de ensino, avaliação de rendimento).

No entanto, é necessário destacar que este trabalho não se deteve sobre as metodologias desse tipo de avaliação, por considerar que elas não se relacionam diretamente com o tema objeto de estudo, mesmo tendo conhecimento de que se encontram propostas de avaliação de políticas, programas e projetos educacionais que, através exclusivamente do rendimento do aluno, buscam aferir seu valor, mérito, eficácia, etc. Este recurso, no entanto, não nos parece muito apropriado, devendo o mesmo ter seu uso limitado quando da avaliação de uma política ou programa e sempre em conjunto com outros insumos de análise.

Como pôde também ser observado neste trabalho, a diversidade analítico-conceitual apontada parece ter origem na forma como foi construído o corpo de conhecimento da avaliação, resultando em uma pluralidade de conceitos, modelos e tipos de avaliação de políticas públicas.

Já em relação aos conceitos apresentados no Capítulo II, pode-se afirmar que, independentemente de sua origem, todos os autores consideram a avaliação como um instrumento capaz de proporcionar informações para facilitar a tomada de decisão, seja ela durante a elaboração, execução ou após o término de um programa ou projeto.

Esse objetivo ou propósito da avaliação, remete-nos ao uso das informações da avaliação. Essa questão, mesmo não tendo sido desenvolvida explicitamente no decorrer do trabalho, parece ser um dos problemas enfrentados pelos avaliadores, pois freqüentemente são encontrados questionamentos a respeito da utilização das informações,

com a ressalva de que estas pouco servem para melhorar o que está sendo avaliado. Desta forma, perde-se um dos objetivos principais da avaliação que é o de fornecer informações para uma retroalimentação.

Após a sistematização acerca dos modelos de avaliação no Capítulo III, é possível apontar algumas correntes que esses modelos representam. A primeira pode ser denominada de tradicional e reflete a preocupação de administradores. Nessa corrente, a avaliação preocupa-se com as metas e objetivos programados *versus* metas e objetivos alcançados, instrumentos previstos *versus* instrumentos utilizados, etc.,

A segunda corrente pode ser chamada de econômica e utiliza técnicas de custo-benefício, rentabilidade-custo, etc., respondendo aos interesses, em geral de economistas que buscam "a melhor opção e o menor custo". Esse tipo de avaliação é difícil de ser realizada na área social e educacional, devido às especificidades dos programas e projetos. Por exemplo, os objetivos dos projetos educacionais são, muitas vezes, difíceis de serem quantificados. Então, como avaliar custo–benefício?

E a terceira, considerada como "alternativa", sinaliza para uma avaliação que enfatiza a participação dos atores sociais envolvidos no processo de formulação e implementação e dos beneficiários do programa ou projeto a ser avaliado. Essa corrente destaca ainda que, independentemente da compatibilização entre objetivos e metas ou projeto elaborado dentro dos critérios estabelecidos, etc., a ação dos atores é fundamental para o desempenho do programa ou projeto em foco, pois eles influenciam o processo de implementação e os seus resultados.

Ainda, em relação à diversidade analítico-conceitual, no Capitulo IV, pode-se observar os diversos tipos de avaliação, com suas nomenclaturas e aspectos a serem avaliados. A esse respeito, é possível apontar para a importância do processo de formulação de um programa ou projeto, inclusive para o estabelecimento de seus objetivos, metas e resultados. Essa inicial preocupação parece ser o ponto de partida para que o processo avaliativo seja minimamente possível. Além disso, aponta–se para a necessidade de estabelecer claramente o que se pretende ao se realizar uma avaliação.

No Capítulo V, também se relacionando a diversidade analítico-conceitual, foram apresentados os conceitos de eficácia, efetividade e eficiência. Neste sentido, consideramos que esses critérios ao serem utilizados não devem ficar restritos em sua dimensão instrumental. Esses critérios devem ser construídos coletivamente, devendo ser introduzido como critério o nível de satisfação da população beneficiária.

Em relação, à escassez de estudos no Brasil que buscassem sistematizar alguns aspectos acerca da avaliação de políticas, programa e projetos sociais e educacionais, tendo em vista a diversidade observada, é raro encontrar estudos e sistematizações que tragam á tona a discussão dos diversos aspectos analíticos, conceituais, metodológicos da avaliação de políticas, programas e projetos. No entanto, a literatura revela um número bastante significativo de estudos que avaliam e analisam as ações governamentais.

Tendo em vista que a literatura indica a avaliação como sendo um instrumento cada vez mais incorporado à agenda governamental dos países em desenvolvimento, é necessário, portanto, que se discuta e se demonstre a diversidade dos pressupostos teórico-metodológicos da avaliação.

Além disso, essa literatura aponta para a avaliação na perspectiva de ser um instrumento 'democrático de controle das ações governamentais' que proporciona informações para planejamento e gestão; de feedback, para possíveis intervenções e correção das ações e atividades desempenhadas; que possibilitam o controle dos recursos financeiros, cada vez mais escassos.

Adiciona-se a isso fato de que a literatura analisada neste trabalho pressupõe um modelo de *policy making* racional e um processo de implementação linear, em que são desencadeados passos seqüenciais e um perfeito entrosamento entre os órgãos e atores envolvidos com a política, programa, projeto social e educacional.

Surge, então, questionamentos a respeito do real objetivo da avaliação no contexto brasileiro, tais como: será que as avaliações são utilizadas para melhorar o que está sendo avaliado? Ou são apenas mecanismos burocráticos com a intenção de saber e controlar a aplicação dos recursos? Será que as informações das avaliações são discutidas e integralizadas com a intenção de proporcionar uma melhoria no processo de gestão ou são mecanismos de controle? Será que a avaliação realmente serve como fonte de retroalimentação? As políticas estão sendo avaliadas considerando a forma como são elaboradas? Será que as avaliações abordam os aspectos que consideramos relevantes para o entendimento destas, tais como:

indicação do problema a solucionar a através de uma política (programa, projeto), formulação da agenda, escolha do programas, etc.? Será que as avaliações consideram o jogo de interesses entre os *policy makers* e a população beneficiária.

Esses questionamentos indicam para a necessidade de um estudo ainda mais aprofundado acerca da avaliação de políticas públicas, a fim de proporcionar insumos analítico-conceituais, que ampliem a visão acerca do processo avaliativo no contexto brasileiro.

Por fim, mesmo diante de tantas perguntas, acreditamos que o presente trabalho cumpre o seu papel ao sistematizar alguns dos diversos estudos a cerca da avaliação, mas não esgota as múltiplas dimensões inerentes a este tema, oferecendo no entanto, insumo metodológico para iniciantes em avaliação de políticas, programas e projetos sociais e educacionais.

REFERÊNCIAS BIBLIOGRÁFICAS

ANDER-EGG, E. *Evaluación de programas* de trabajo social. 2ª ed. Buenos Aires, Argentina, 1990.

AFONSO, A. J. *Avaliação educacional: regulação ou emancipação.* São Paulo, Cotez, 2000.

AGUILAR, M e ANDER-EGG, E. *Avaliação de serviços e programas sociais.* 2ª ed. Trad. Jaime A. C. e Lucia M E. Orth. Petrópolis, RJ, Vozes, 1995.

ALKIN, M. C. Evaluation theory development. In: WEISS, C. *Evaluating action programs: readings social action and education,* Boston, USA, Allyn and Bacon, 1972.

ARRETCHE, M.T. Tendências no estudo sobre avaliação. In: RICO, E. M. (org.), *Avaliação de políticas Sociais: uma questão em debate.* São Paulo, SP, Cortez, Instituto de Estudos especiais, 1998.

BALLESTEROS, R. F. Cuestones conceptuales básicas en evaluación de programas. In: BALLESTEROS, R. F. et.al. *Evaluación de programas: una guía práctica en ámbitos sociales, educativos y de salud.* Madrid, Spain, Ed. Rocio Fernández-Ballesteros, 1996.

_____. El ciclo de intervención social y evaluación. . In: BALLESTEROS, R. F. et.al. *Evaluación de programas: una guía práctica en ámbitos sociales, educativos y de salud.* Madrid, Spain, Ed. Rocio Fernández-Ballesteros, 1996.

BID (Banco Interamericano de Desenvolvimento). *Evaluación: una herramienta de gestión para mejorar el desempeño de los proyectos, 1997.*

BRIONES, G. *Evaluación de programas sociales.* México, Trillas, 1998.

BROWN, D. R. Evaluación de políticas e programas en la región del caribe de habla inglesa: problemas conceptuáis y prácticos. In: *Serie Gestión pública,* Santiago, Chile, Publicación de las naciones unidas, enero, 2001.

CANESQUI, A.M. Avaliação de políticas e programas sociais: conceito e tipos de pesquisas. In: *Cadernos de Serviço Social.* Ano IX, n 17, PUC – Campinas, Capinas, 2000.

CARVALHO, M. do C. B. de. Avaliação participativa uma escolha metodológica. In: RICO, Elizabeth Melo (org.), *Avaliação de políticas Sociais: uma questão em debate.* São Paulo, SP, Cortez, Instituto de Estudos especiais, 1998.

COHEN, D. K. Politics and research: evaluation of social action programs in education. In: WEISS, C. *Evaluating action programs: readings social action and education.* Boston, USA, Allyn and Bacon Inc, 1972.

COHEN, E. & FRANCO, R. *Avaliação de projetos sociais.* 3ª ed. Petrópolis, Rio de Janeiro, RJ, Vozes, 1999.

FARIA, R. M. Avaliação de programas sociais: evoluções e tendências. In: RICO, Elizabeth Melo (org.), *Avaliação de políticas Sociais: uma questão em debate.* São Paulo, SP, Cortez, Instituto de Estudos Especiais, 1998.

FERNANDÉZ, G. D. *Evaluación y educación: modelos e propuestas.* Buenos Aires, Argentina, FUNDEC, 2000.

FIGUEIREDO & FIGUEIREDO. Avaliação política e avaliação de políticas: um quadro de referência teórica. *Análise & Conjuntura.,* Belo Horizonte, Fundação João Pinheiro, v.1, n.3, set/dez 1986.

FRANCO, M. L. P. B. Pressupostos epistemológicos da avaliação educacional. *Cadernos de Pesquisa,* n.74, São n.68, São Paulo, Fundação Carlos Chagas, agosto/ 1990.

HOLISTER, R. e CAIN, G. G. The methodology or evaluating social action program. In: ROSSI, P. e WILLIAMS, W. *Evaluating social program: theory, practice, and politics,* New York, USA, SEMINAR PRESS, 1972.

HOUSTON, T.R. The behavioral sciences impact-effectiveness model. In: ROSSI, P.; WILLIAMS, W. Evaluating soail program: theory, practice, and politics. New York, USA, 1972.

LARA, A. J. Recogida y análisis de datos. In: VIEDMA, C. A. e LARA, A. J. et. al. *Evaluación de programas sociales.* Madrid, España, Diaz de Santos, 1996.

LOBO, T. Avaliação de processo e impactos em programas sociais: algumas questões para reflexão. In: RICO, Elizabeth Melo (org.), *Avaliação de políticas Sociais: uma questão em debate.* São Paulo, SP, Cortez, Instituto de Estudos especiais, 1998.

MARTINIC, S. *Diseño y evaluación de proyectos sociales: herramienta para el aprendizaje.* Mexico, COMEXANI-CEJUV, 1997.

MATUS, C. *Política planejamento e governo.* Brasília, IPEA, 1996.

MÉNDEZ, J. M. *Evaluar para conocer, examinar para excluir.* Madrid, Sapin, Morata, 2001.

PALUMBO, D. Politics and evaluation. In: PALUMBO, D. The politics program evaluation, Newbury Park, Sage Pub, 1987.

PÉREZ, S. Investigación y evaluación de los sistemas educativos. In: *Cuadernos de la OEI / Educación comparada.* Madrid, España, 1998.

PESTANA, M. I. G. de Sá. Avaliação educacional – o sistema nacional de avaliação da educação básica. In: RICO, Elizabeth Melo (org.), *Avaliação de políticas Sociais: uma questão em debate.* São Paulo, SP, Cortez, Instituto de Estudos especiais, 1998.

RICO, E. M. (org.). *Avaliação de políticas Sociais: uma questão em debate.* São Paulo, SP, Cortez, Instituto de Estudos especiais, 1998.

RISTOFF, D. I. Avaliação institucional: construindo relatórios.In: SOBRINHO, José Dias, RISTOFF, Dilvo (org). *Universidade desconstruída: avaliação Institucional e resistência,* Florianópolis, Insular, 2000.

SANTACANA, M. F., BENITO, J. Evaluación de programas en educación. In: BALLESTEROS, R. F. et.al. *Evaluación de programas: uma guia práctica en âmbitos sociales, educativos y de salud.* Ed. Rocio Fernández-Ballesteros, Madrid, Spain, 1996.

SAUL, A. M. Avaliação participante — uma abordagem crítico-transformadora. In: RICO, E. M. (org.), *Avaliação de políticas Sociais: uma questão em debate.* São Paulo, SP, Cortez, Instituto de Estudos especiais, 1998.

_____. *Avaliação emancipatória: desafio à teoria e a pratica de avaliação e reformulação do currículo.* 5ª ed. São Paulo, SP, Cortez, 2000.

SIEMPRO. *Gestión Integrada de programas Sociales orientada a resultados.* Buenos Aires, 1999.

SOBRINHO, J. D. Campos e caminhos da avaliação: a avaliação da educação superior no Brasil. In: FREITAS, L. C. de (org). *Avaliação construindo o campo e a crítica.* Florianópolis, Insular, 2002.

SCRIVEN, M. The methodology of evaluation. In: WEISS, C. *Evaluating action programs: readings social action and education.* Boston, USA, Allyn and Bacon, 1972.

STAKE, R. E. The countenance of educational evaluation. In: *Evaluating action programs: readings social action and education.* Allyn and Bacon, Boston, USA, 1972.

SUCHMAN, E. A. Evaluating educational programs. In: CARO, F. G. *Readings in evaluation research.* 2° ed. New York, USA, Russell Sage Foundation, 1977.

SIEMPRO. *Manual metodológico para la planificación y evaluación de programas sociales: Gestión integral de programas sociales orientada a resultados*, 1999.

SULBRANDT, J. A avaliação dos programas sociais: uma perspectiva crítica dos modelos usuais. In: KLIKSBERG, B. (org.) *Pobreza uma questão inadiável.* Brasília, ENAP, 1994.

TENÓRIO, F. G. (coord.). *Avaliação de projetos comunitários: uma abordagem prática.* 2ª ed, São Paulo, SP, Loyola, 2000.

TRIPODI, T. et.al. *Avaliação de programas sociais: diretrizes para administradores nos campos da educação, saúde e assistência social.* 2ª ed., Trad. Tainá da C. Souza. Rio de Janeiro, RJ, Francisco Alves,1984.

VIANNA, H. M. A prática da avaliação educacional.: algumas colocações metodológicas. *Caderno de Pesquisa*, n.68, Fundação Carlos Chagas, São Paulo, maio/1989.

_____. Avaliação educacional nos cadernos de pesquisa. *Caderno de Pesquisa*, São Paulo, n.80, p. 100-105, fev/1992.

_____. *Avaliação educacional e o avaliador.* São Paulo, IBRASA, 2000.

VIEDMA, C. A. e LARA, A. J. et.al. *Evaluación de programas sociales.* Madrid, España, Diaz de Santos, 1996.

VIEDMA, C. A. Introducción a la evaluación. In: VIEDMA, C. A. e LARA, A. J. et.al. *Evaluación de programas sociales.* Madrid, España, Diaz de Santos, 1996.

WEISS, C. *Evaluating: methods for studying programs and policies.* 2ª ed. New Jersey, USA, Pretice Hall, 1998.

_____. *Investigación evaluativa: métodos para determinar la eficiencia de los programas de acción.* 2a. ed. México, Trillas, 1999.

WIESNER, E. Función de evaluación de planes, programas, estrategias y proyectos. In: *Serie Gestión pública,* Santiago, Chile, Publicación de las naciones unidas, maio/ 2000.

WORTHEN, B. R.; SANDERS, J. R.; FITZPATRICK, J. L. *Program evaluation: alternative approaches a practical guidelines.* 2a ed., New York, USA, LONGMAN, 1997.

BIBLIOGRAFIA CONSULTADA

ABRANCHES, S. H. et.al.. *Política social e combate à pobreza.* 3ª ed, Rio de Janeiro, RJ,Jorge Zahar Editor, 1994.

ANDRÉ, M. E. D. A. A avaliação da escola e a avaliação na escola. *Cadernos de Pesquisa*, n.80, São Paulo, Fundação Carlos Chagas, fev/1992.

AZEVEDO, J. M. Lins de, *A educação como Política Pública*, Campinas, SP, Autores Associados,1997. Coleção polêmicas do nosso tempo, v.56.

BARRETO, E. S.de Sá. Políticas de educação: atuais marcos de análise. *Caderno de Pesquisa*, n.90, São Paulo, Fundação Carlos Chagas, ago/1994.

BELLONI, I. A função da avaliação institucional. In: In: *Avaliação-Revista da Rede Nacional avaliação Institucional da Educação Superior*, v.3, n.4, Capinas, São Paulo, dez/1998.

BELLONI, I. et.al. *Metodologia de avaliação em políticas públicas: uma experiência em educação profissional*, São Paulo, Cortez, 2000. Questões da nossa época, v.75.

BIANCHETTI, R. G. *Modelo Neoliberal e Políticas educacionais.* 2ª ed. São Paulo, São Paulo, Cortez, 1999. Questões de nossa época, v.56.

BONANINO. A. C. de. Avaliação e política Educacional: o processo de institucionalização do SAEB. *Caderno de Pesquisa*, n.108, São Paulo, Fundação Carlos Chagas, nov/1999.

CANDAU, Vera. Avaliação no Brasil: uma revisão bibliográfica. In: *Caderno de Pesquisa*, n. 95, São Paulo, Fundação Carlos Chagas , nov/1995.

CASTRO, M.H. G. Interesses, organizações e políticas sociais. In: *BIB*, Rio de janeiro, n.31, 1991.

DEMO, P. *Mitologias da avaliação: de como ignorar, em vez de enfrentar problema.* Campinas, São Paulo, Autores Associados, 1999. Polêmicas de nosso tempo, v.68.

DRAIBE, S. M. O. Welfere State no Brasil: características e perspectivas, In: *Cadernos de Pesquisa do NEPP*, n. 8, Campinas, UNICAMP/NEPP, 1998.

FAGNANI, E. Avaliação do ponto de vista do gasto e financiamento das políticas públicas. In: RICO, E. M. (org.) *Avaliação de políticas Sociais: uma questão em debate.* São Paulo, Cortez, Instituto de Estudos especiais, 1998.

FREITAG, B. *Política educacional e industria cultural*, São Paulo, Cortez, Autores Associados, 1989.

FREITAS, L. C. de (org). *Avaliação construindo o campo e a crítica.* Florianópolis, Insular, 2002.

GEENBERG. B. G. Evaluation of social programs. In*: Readings in evaluation research.* 2 ed., New York, USA, Russell Sage Foundation, 1977.

GERSTON, L. N. Evaluation: assessments and directions. IN: GERSTON, L. N. *Public Policy Making.* New York, USA, Sharpe, 1997.

HADJI, C. *A avaliação, regras do jogo: das intenções aos instrumentos.* Trad. Júlia L.Ferreira e José M. Claudio. Porto, Portugal, Porto Editora, 1994.

HORN, R. et.al. Sistemas de evaluación educacional en América Latina.: reseña temática y experiencias recientes.. *Boletín 27- Proyecto Principal de Educación en América Latina y el Caribe*, Santiago, Chile, abril, 1992.

IPEA. *Planejamento, e políticas públicas.* n.23, Instituto de Pesquisa Econômica Aplicada, Brasília, 2001.

KING, D. O estado e as estruturas sociais de bem-estar. In*: Novos Estudos CEBRAP,* 22, São Paulo, 1988.

LEVIN, H. M. Cost-effectiveness analysis in evaluation research. In: GUTTENTAG, M e STRUENING, E. *Hanbook of evaluation research.* V.2, California, USA, Sage publications, 1975.

MANN, J. The outcome of evaluative research. In: WEISS, C. *Evaluating action programs: readings social action and education.* Boston, USA, Allyn and Bacon Inc, 1972.

MEDINA, A. Modelos e lentes: uma discussão sobre a análise da implementação de políticas públicas. In: *Análise e Conjuntura* 2, Fundação João Pinheiro.

MELO, A. A. S. de. *Educação e hegemonia no Brasil de hoje.* Maceio, EDUFAL, 1998.

MENDIOCA, G. E. *La planificación de las políticas sociales.* Buenos Aires, Argentina, Espacio Editorial, 1997.

NEPP. Brasil: *Sistema de proteção social no Brasil e suas transformações recentes,* UNICAMP/NEPP, Campinas, 1992.

PALFREY, C. Policy *Evaluation in the public sector: approaches and methods.* Aldershot, Avebury, 1992.

PEREZ, J. R. R. Avaliação do processo de implementação: algumas questões metodológicas. In: RICO, E. Melo (org.), *Avaliação de políticas Sociais: uma questão em debate.* São Paulo, Cortez, Instituto de Estudos especiais, 1998.

_____. *Avaliação e desafios da educação básica.* Campinas, São Paulo, Editora da Unicamp e Annablume Editora, 2000.

PHILP, J. La evaluación de proyectos educativos en América Latina. *Boletín 13 - Proyecto Principal de Educación en América Latina y el Caribe.* Santiago, Chile, UNESCO, agosto, 1987.

PIÑON, F. J. La educación como herramienta Estratégica. *Boletín 45 - Proyecto Principal de Educación en América Latina y el Caribe,* Santiago, Chile, UNESCO, abril, 1998.

RIOS, T. A. Avaliar: ver mais claro para caminhar mais longe. In: RICO, E. M. (org.) *Avaliação de políticas Sociais: uma questão em debate.* São Paulo, Cortez, Instituto de Estudos especiais, 1998.

VIANA, A.L. Abordagens metodológicas em políticas públicas. In: *Revista de Administração Pública.* v.30, São Paulo, Fundação Carlos Chagas, , n.2, 1996.

VIEIRA, L. S, Neoliberalismo, privatização e educação no Brasil. In: Oliveira, R. (org). *Política educacional: impasses e alternativas. Cortez, SP, 1995.*

SANTOS, W. G. *Cidadania e Justiça: a política social na ordem brasileira,* Rio de Janeiro, Ed. Campus, 1979.

SCHIELBEIN, E. et.al. El costo-efectividad de las políticas de educación primaria en América Latina: estudio basado en la opinión de expertos. *Boletín 49, Proyecto Principal de Educación en América Latina y el Caribe*, Santiago, Chile, UNESCO, abril, 1999.

SOARES, L. T. *Os custos sociais do ajuste neoliberal na América Latina.* São Paulo, Cortez, 2000. Coleção Questões da Nossa Época, v.78.

SOBRINHO, J. D. Avaliação Institucional da Educação superior: fontes externas e fontes internas. In: *Avaliação- Revista da Rede Nacional Avaliação Institucional da Educação Superior*, v.3, n.4, Capinas, São Paulo, dez/1998.

_____.Concepções de Universidade e de avaliação institucional. In: *Avaliação - Revista da Rede Nacional Avaliação Institucional da Educação Superior*, v.4, n.3 (12), Capinas, São Paulo, set/1999.

_____. Exames gerais, provão e avaliação educativa. In: *Avaliação- Revista da Rede Nacional Avaliação Institucional da Educação Superior*, v.4, n.3 (13), Capinas, São Paulo, set/1999.

SOUZA, S. M. Z. Avaliação de aprendizagem nas pesquisas no Brasil de 1930 a 1980. In: *Cadernos de Pesquisa*, n.94, São Paulo, Fundação Carlos Chagas, 1995.

_____. Avaliação do rendimento escolar como instrumento de gestão educacional. In: OLIVEIRA, D. A. (org.) *Gestão democrática da educação: desafios contemporâneos.* Petrópolis, RJ; Vozes, 1998.

TELES, V. S. Sociedade Civil e a construção de espaços públicos. In: Dagnino, E. (org) *Anos 90- Política e Sociedade no Brasil.* São Paulo, Ed. Brasiliense, 1994.

WEISS, Carol H. Evaluating Educational and social Action Programs: a "Treeful of Owls". In: WEISS, C. *Evaluating action programs: readings social action and education.* Boston, United States of America, Allyn and Bacon Inc, 1972.

_____. The politicization of evaluation research. In: WEISS, C. *Evaluating action programs: readings social action and education.* Boston, United States of America, Allyn and Bacon Inc, 1972.

_____. Utilization of evaluation: toward comparative study. In: WEISS, C. *Evaluating action programs: readings social action and education.* Boston, United States of America, Allyn and Bacon Inc, 1972.