

UNIVERSIDADE DE BRASÍLIA

Instituto de Psicologia

Programa de Pós-graduação em Psicologia Social, do Trabalho e das Organizações

Meta-avaliação das Auditorias de Natureza Operacional do Tribunal de  
Contas da União: Um estudo sobre auditorias de Programas Sociais

Helga Cristina Hedler

Brasília, DF

2007

Universidade de Brasília

Instituto de Psicologia - IP

Programa de Pós-graduação em Psicologia Social, do Trabalho e das Organizações

Meta-avaliação das Auditorias de Natureza Operacional do Tribunal de Contas da União: Um  
estudo sobre auditorias de Programas Sociais

Helga Cristina Hedler

Brasília, DF

2007

Universidade de Brasília

Instituto de Psicologia - IP

Programa de Pós-graduação em Psicologia Social, do Trabalho e das Organizações

Meta-avaliação das Auditorias de Natureza Operacional do Tribunal de Contas da União: Um estudo sobre auditorias de Programas Sociais

Helga Cristina Hedler

*Tese apresentada ao Programa de Pós-graduação em Psicologia Social, do Trabalho e das Organizações, como requisito parcial à obtenção do grau de Doutor em Psicologia Social, do Trabalho e das Organizações.*

Orientador: Prof. Dr. Cláudio Vaz Torres

Co-orientadora: Profa. Dra. Elaine Rabelo Neiva

Brasília, DF

Junho de 2007

Meta-avaliação das Auditorias de Natureza Operacional do Tribunal de Contas da União: Um  
estudo sobre auditorias de Programas Sociais

Tese defendida e aprovada pela banca examinadora constituída por:

---

Prof. Dr. Cláudio Vaz Torres (Presidente)

Programa de Pós-graduação em Psicologia Social, do Trabalho e das Organizações,  
Universidade de Brasília

---

Profa. Dr. Elaine Rabelo Neiva (co-orientadora)

Programa de Pós-graduação em Psicologia Social, do Trabalho e das Organizações  
Universidade de Brasília

---

Profa. Dr. Eda Castro Lucas de Souza (membro)

Programa Pós-graduação em Administração, Universidade de Brasília

---

Profa. Dr. Maria das Graças Torres da Paz (membro externo)

Universidade Católica de Brasília

---

Prof. Dr. Tomás de Aquino Guimarães (membro)

Programa de Pós-graduação em Administração, Universidade de Brasília

## Agradecimentos

A Deus por permitir que eu chegasse até o final deste trabalho com saúde mental e disposição de aplicar o conhecimento adquirido.

Ao prof. Dr. Cláudio V. Torres por me aceitar como doutoranda e por me ajudar a vencer este desafio mais difícil que o do mestrado e alcançar minha grande realização pessoal.

A profa. Dr. Elaine co-orientadora que soube conduzir meus trabalhos, desde a Banca de Qualificação até a análise dos resultados da tese, pelo seu tempo e disponibilidade essenciais para a conclusão deste trabalho.

A profa. Dr. Graça por participar como membro da Banca de Qualificação do Doutorado e da Banca Final do Doutorado, pelos seus sábios comentários e orientações em relação à avaliação de programas na organização pesquisada.

Ao prof. Dr. Tomás pela confiança em minha competência, por seu estímulo, apoio, participação como juiz das categorias da análise de conteúdo e também como membro da Banca de Doutorado.

A profa. Dr. Eda pela confiança no meu trabalho e participação na Banca de Doutorado.

A profa. Dr. Iône Vasques Menezes por participar como suplente.

Ao amigo-irmão, mestre e agora doutorando Onofre por me ajudar a passar pelos momentos difíceis da minha vida e importantes orientações sobre a análise de conteúdo e interpretação de textos teóricos sobre meta-avaliação.

A mestre Ariane pela preciosa amizade, conselhos sempre pertinentes e análise como juiz das categorias da análise de conteúdo.

A mestre Danúzia pela revisão da tese, correção e dicas importantes sobre forma e estruturação de cada parágrafo.

A amiga e servidora do Instituto de Psicologia Edna pelas orientações necessárias em relação aos caminhos burocráticos nesses anos de Curso de Pós-graduação.

Ao Instituto de Psicologia e professores que ajudaram na minha formação acadêmica.

Ao Tribunal de Contas da União e seus representantes na pessoa da Sra. Selma e Glória Bastos e suas equipes de trabalho.

Ao prof. Dr. Günther por participar como juiz das categorias do estudo-piloto e compartilhar seu conhecimento sobre metodologia científica.

Ao Maxwell pela alegria, carinho e descontração, compartilhados comigo nesses anos, ajudando-me a superar os momentos de estresse.

Aos meus familiares mesmo distantes, em especial Tia Victória, Ruth e Randolph pelo estímulo, carinho e apreço.

Aos meus pais Cristina e Ludwig e irmãos Rosemeri e Roberto, seus filhos Robert e Luciano, Roberta e Evelim pelo incentivo e confiança na minha capacidade de concluir esse trabalho.

A Edileuza e filhas lindas que são uma inspiração para mim para futuros planos.

Aos servidores do IP pela amizade e estímulo, em especial o Basílio, Joyce e Margareth.

Ao servidor do IP Luciano Roberto Sousa em memória como servidor dedicado e sorridente.

Aos Secretários da pós-graduação em Psicologia em especial o Fábio pela prontidão e ajuda com as documentações necessárias à qualificação e defesa do doutorado.

A Mestre Cilene Braga Lins amiga pessoal e incentivadora por me introduzir no mundo da Assistência Social e ampliar minha visão a respeito das políticas públicas de direito.

A Dr. Namara Ribeiro por sua disponibilidade e gentileza, pelas sugestões em relação à análise de conteúdo, síntese de categorias e comparação final das análises.

Aos colegas do MDS em especial a Sra.Margarida Munguba pela oportunidade de ampliar meus conhecimentos sobre o Programa de Erradicação do Trabalho Infantil – PETI.

A todos(as) os colegas do DPSE em especial a Adriana, Cristiane, Vera e Gustavo, pelos ensinamentos sobre política pública e sobre a operacionalização do PETI.

A todos que me ajudaram direta ou indiretamente e que não estejam nomeados nesta relação, o meu sincero agradecimento.

## Resumo

O objetivo geral desta pesquisa foi produzir uma meta-avaliação da auditoria de natureza operacional do Tribunal de Contas da União, de seus critérios e padrões, verificando-se a validade e confiabilidade dos procedimentos e resultados. Na revisão de literatura delimitaram-se esses conceitos, fazendo-se a diferenciação entre meta-avaliação e meta-análise, além disso, propôs-se um modelo de meta-avaliação. Uma meta-avaliação é uma avaliação de uma avaliação considerando-se aspectos como metodologia, seleção dos sujeitos, objetivos, critérios e análise dos resultados. Analisaram-se auditorias das áreas de atuação do governo: agricultura; assistência social; cidadania; comércio e serviços; educação energia; habitação; meio ambiente; saúde e trabalho. Como método utilizou-se a abordagem qualitativa através dos procedimentos: análise de conteúdo, síntese de categorias da análise de conteúdo; checagem de critérios do Joint Committee, agrupamento dedutivo de critérios Joint Committee; comparação dos resultados das análises e Síntese de Estudos Qualitativos. Os principais resultados demonstram que o processo de auditoria está associado a variáveis como contexto político, características dos programas sociais, enfoque, métodos e técnicas das auditorias. Por fim, foram descritas as dificuldades metodológicas para a realização das auditorias e feitas sugestões de melhorias ao modelo de auditoria do tribunal.

Palavra-chave: auditoria, avaliação de programas, programas sociais.

## Abstract

The general purpose was evaluate an audit based on its natural operation, criteria and patterns and validates procedures and results' validity as evaluation practice from Tribunal de Contas da União. These concepts had been presented and operacionalized by literature review, and it also showed differences between meta-evaluation and meta-analyses. Moreover, it was proposed a model of meta-evaluation, which is an evaluation of evaluations that concern methodology, subjects, objectives, and the results analyzed. An audit from government performance had been developed from areas as agriculture; social assistance; citizenship; commerce and services; education, energy; habitation; environment; health and work. As method it was used qualitative approach, its procedures were: analysis of content, synthesis of categories of the content analysis, verifying the criteria of the Joint Committee, deductive grouping of criteria Joint Committee; comparison of the results of the analyses and the Synthesis of Qualitative Studies. The main results reveal that the audit process is associated the multifaceted aspect as political context, dimensions of social programs, approach, methods and techniques from audit. Finally, limitations concerning audits methodologies and agenda for futures audit on Tribunal were described.

**Key-word:** auditing, program evaluation, social programs.

## Sumário

Capítulo 1: Contextualização da meta-avaliação .....	18
Capítulo 2: Revisão de Literatura.....	22
2.1 Contextualização Histórica da Avaliação de Programas.....	22
2.2 Avaliação de programas: conceitos.....	35
2.3 Avaliação de Programas: abordagens .....	44
2.4 O planejamento da avaliação.....	51
2.2.1 Avaliação de programas e auditorias: contrastes e interface.....	54
2.2.2 Meta-avaliação, avaliação da avaliação ou auditoria de avaliação .....	58
2.2.3 Meta-análise: diferenças e semelhanças em relação à meta-avaliação .....	68
2.5 Panorama da Avaliação de Programas no Brasil .....	72
2.6 Questões Metodológicas na Avaliação de Programas.....	75
2.6.1. Delineamentos de Pesquisa .....	83
2.6.2 Delineamentos não-experimentais ou pré-experimentais .....	83
2.6.3 Delineamentos experimentais .....	86
2.6.4 Delineamentos quasi-experimentais.....	87
Capítulo 3: Metodologia.....	93
Visão geral.....	93
3.1 Problema de pesquisa .....	93
3.2 Estudo-piloto .....	94
3.3 Delineamento da pesquisa.....	96
3.4 Universo e Amostra.....	96
3.4.1 Descrição da Amostra .....	97
3.5 Procedimento de coleta de dados .....	101

3.6 Procedimento de análise dos dados .....	102
Capítulo 4: Resultados e Análise .....	108
4.1 Descrição do processo de trabalho de ANOP no TCU .....	108
4.2 Descrição das Técnicas Utilizadas pela Organização .....	115
a) Indicadores de desempenho e mapas de produto .....	115
b) Mapa de processos .....	118
c) Análise RECI .....	119
d) Marco lógico .....	119
e) Análise de Stakeholder.....	120
f) Análise SWOT e Matriz de Verificação de Risco .....	121
g) Técnica de entrevista.....	122
4.3 Meta-avaliação de acordo com Normas do Joint Committee .....	125
1. Programa Novo Mundo Rural: Assentamento de Trabalhadores Rurais; Novo Mundo Rural: Consolidação de Assentamentos .....	128
2. Programa Combate ao Abuso e à Exploração Sexual de Crianças e Adolescentes - Sentinela.....	131
3. Programa Reinserção Social do Adolescente em Conflito com a Lei.....	136
4. Programa Desenvolvimento do Turismo no Nordeste – Prodetur/NE.....	142
5. Programa de Desenvolvimento Sustentável de Recursos Hídricos para o Semi-Árido Brasileiro – Pró-Água/Semi-Árido.....	148
6. Programa Energia as Pequenas Comunidades .....	152
7. Programa Morar Melhor .....	155
8. Programa Nacional do Livro Didático – PNLD.....	160
9. Monitoramento e Controle da Mortalidade Materna.....	164

10. Programa Novo Emprego e Seguro-Desemprego .....	167
4.4 Análise dos Conteúdos .....	171
1. Avaliação de Processo.....	172
2. Avaliação de Necessidades .....	173
3. Avaliação de Impacto/Resultado.....	175
4. Avaliação da Eficiência.....	177
5. Boas Práticas do Programa.....	178
6. Monitoramento do Programa .....	180
7. Política do Programa .....	182
8. Fontes de dados da auditoria .....	184
9. Respondente da auditoria .....	186
10. Planejamento da Auditoria .....	187
11. Critério de Seleção da Amostra da Auditoria.....	189
12. Técnicas de Coleta de Dados da auditoria .....	191
13. Procedimento de Coleta e Tabulação dos Dados .....	193
14. Análise dos Dados.....	194
15. Métodos da Auditoria.....	196
16. Dificuldades da Auditoria .....	198
17. Responsabilização .....	201
18. Comunicação Institucional.....	203
19. Objetivos dos programas auditados.....	205
4.5 Síntese de Resultados das Categorias da Análise de Conteúdo .....	208
4.6 Síntese de Estudos Qualitativos: Critérios do Joint Committee e Síntese de Categorias da Análise de Conteúdo.....	215

Capítulo 5. Discussão da meta-avaliação .....	228
5.1 Síntese de estudos qualitativos .....	228
5.1.1 Contexto político .....	229
5.1.2. Características dos programas sociais auditados.....	230
5.1.3 Enfoques das ANOPs .....	230
5.1.4 Métodos e técnicas .....	232
5.1.5 Dificuldades para a realização de auditorias .....	234
5.2 Respostas às perguntas de pesquisa .....	237
5.3 Limitações da presente meta-avaliação.....	243
5.4 Agenda de pesquisa.....	244
Apêndice.....	245
Referências bibliográficas .....	250

## Lista de Tabelas

Tabela 1. Resumo das normas para avaliação de programas .....	64
Tabela 2. Fontes de ameaça Interna e Externa para Delineamentos .....	82
Tabela 3. Critérios para avaliação das auditorias, baseado nas normas do <i>Joint Committee...</i> .....	127

## Lista de Quadros

Quadro 1. Eventos que influenciaram o desenvolvimento da avaliação de programas .....	34
Quadro 2. Comparação entre meta-análise e meta-avaliação .....	71
Quadro 3. Variáveis do estudo-piloto .....	<a href="#">94</a>
Quadro 4. Auditorias de Programas Meta-avaliados .....	<a href="#">98</a>
Quadro 5. Descrição dos programas auditados pelo TCU .....	<a href="#">99</a>
Quadro 6. Categoria Avaliação de Processo .....	<a href="#">172</a>
Quadro 7. Avaliação de Necessidades .....	<a href="#">174</a>
Quadro 8. Avaliação de Impacto/Resultados .....	<a href="#">175</a>
Quadro 9. Avaliação da eficiência .....	<a href="#">177</a>
Quadro 10. Boas Práticas do Programa.....	<a href="#">179</a>
Quadro 11. Monitoramento do Programa .....	<a href="#">181</a>
Quadro 12. Política do Programa .....	<a href="#">183</a>
Quadro 13. Fonte de dados de auditoria.....	<a href="#">185</a>
Quadro 14. Respondentes da auditoria.....	<a href="#">187</a>
Quadro 15. Planejamento da auditoria .....	<a href="#">188</a>
Quadro 16. Amostra da auditoria .....	<a href="#">190</a>
Quadro 17. Técnica de coleta de dados.....	<a href="#">191</a>
Quadro 18. Procedimento de coleta e tabulação de dados .....	<a href="#">193</a>
Quadro 19. Análise dos dados.....	<a href="#">194</a>
Quadro 20. Delineamento da Auditoria .....	<a href="#">197</a>
Quadro 21. Dificuldades da Auditoria .....	<a href="#">199</a>
Quadro 22. Responsabilização .....	<a href="#">202</a>
Quadro 23. Comunicação Institucional.....	<a href="#">204</a>
Quadro 24. Frequência da Classificação dos Objetivos dos Programas Avaliados.....	<a href="#">205</a>

Quadro 25. Resumo das Categorias da Análise de Conteúdo.....	<a href="#">207</a>
Quadro 26. Síntese de Categorias: Características dos Programas .....	<a href="#">209</a>
Quadro 27. Síntese de Categorias: Enfoque da Auditoria.....	<a href="#">210</a>
Quadro 28. Síntese de Categorias: Contexto Político .....	<a href="#">211</a>
Quadro 29: Síntese de Categorias: Métodos e Técnicas .....	<a href="#">212</a>
Quadro 30. Síntese de Categorias: Dificuldades para realização da auditoria.....	<a href="#">214</a>
Quadro 31. Classificação dos Objetivos dos Programas Auditados .....	<a href="#">245</a>

## Lista de Figuras

Figura 1. Contextualização da Avaliação de programas .....	40
Figura 2. Tipologia de auditoria no setor público .....	56
Figura 3. Modelo de meta-avaliação .....	61
Figura 4. Procedimento de análise de dados .....	102
Figura 5. Ciclo de trabalho da Auditoria de Natureza Operacional .....	113
Figura 6. Exemplo de Mapa de Processo de expedição de ofício de diligência .....	118
Figura 7. Sinopse das estratégias metodológicas utilizadas pelo TCU .....	124
Figura 8. Comparação de Estudos Qualitativos: conceitos .....	200
Figura 9. Síntese de Estudos Qualitativos .....	201
Figura 10. Variáveis associadas ao processo de realização de ANOP .....	237

## Capítulo 1: Contextualização da meta-avaliação

Esta tese meta-avalia a Auditoria de Natureza Operacional – ANOP do Tribunal de Contas da União – TCU. Para tal lançou mão de análises qualitativas para compreender a realização da ANOP a partir de uma amostra de dez auditorias de Programas Sociais de Governo. Antes de adentrar na explicação do processo partiu-se de algumas indagações para contextualizar o processo de ANOP e de meta-avaliação, comumente confundida com meta-análise.

Por que as avaliações precisam ser avaliadas? Há inúmeras respostas a essa pergunta, entre elas, o fato de que tanto o avaliador quanto o cliente, o avaliado ou outro interessado na avaliação, terão prejuízo financeiro, moral ou outro, se a avaliação for ineficiente. As meta-avaliações são importantes porque melhoram a qualidade das avaliações em termos metodológicos e práticos, essas melhorias foram buscadas no presente trabalho ao se avaliar os procedimentos de ANOP.

Assim, como ponto de partida se esclarece sobre as auditorias de programas utilizadas nesta pesquisa. Primeiro, tratam-se de programas sociais de governo. Tais programas incidem sobre diferentes problemas sociais. Por exemplo, o Programa de Enfrentamento ao Abuso e Exploração Sexual de Crianças e Adolescentes - Sentinela, oferece serviços especializados para atender crianças e adolescentes vítimas de abuso ou exploração sexual. Segundo, o governo brasileiro executa suas políticas por meio de programas e serviços em dez áreas de atuação diferentes, sendo analisado nesta tese o resultado de uma auditoria de programa de cada área de governo, compondo a meta-avaliação.

Tais programas, às vezes, perpassam governos e podem se transformar em serviços de ação continuada, isto é, tornam-se independentes da política de governo para a sua execução, transformando-se em direito assegurado pela Lei e Política de Estado. Dessa forma, passam a ser oferecidos como serviços regulares à população. Tais programas públicos recebem dotações orçamentárias, atingindo milhões de reais por ano e atendem a milhões de brasileiros. Por exemplo, o Programa Bolsa Família – PBF atende a 8.789.538 famílias e teve um orçamento anual de R\$543.779.956,00 em 2006 (Brasil, 2006).

Assim, inicia-se o terceiro aspecto deste trabalho, a auditoria de programas sociais é tratada pelo TCU como avaliação de programas sociais. No Brasil, como em outros países, a avaliação de programas sociais é responsabilidade do governo e também recai sob

instituições da sociedade civil que possuem interesse no funcionamento e acompanhamento dos programas, além de tribunais, auditorias, institutos e associações. De acordo com Stufflebeam e Shinkfield (1987) uma parte do trabalho avaliativo, que consiste na regulação e proteção dos interesses públicos é dirigida por organismos independentes como agências governamentais e por associações profissionais de avaliação.

A importância da avaliação de programas está inserida em diversos âmbitos, como o acadêmico e científico, além de forma mais ampla, está inserida nas políticas públicas. No caso acadêmico e científico, a avaliação de programas fomenta os estudos e as pesquisas desenvolvidos nas Universidades, desenvolvendo o conhecimento sobre avaliação de programas – AP.

A avaliação dos programas sociais é fundamental e necessária, pois seu uso possibilita aos gestores governamentais a prestação de contas e verificação da eficiência dos serviços ofertados, subsidiando a decisão sobre a continuidade dos programas. Tanto no caso acadêmico, como no caso das políticas públicas, a melhoria do programa é um objetivo importante a ser alcançado a partir do uso dos resultados das avaliações. Argumenta-se que, além da necessidade de desenvolver uma cultura de avaliação nas organizações, é primordial avaliar políticas e programas de governo. Há necessidade de transformar a avaliação em instrumento de conhecimento sobre o programa e sobre a realidade na qual intervêm, sobre os impactos diretos e indiretos dos programas e utilizá-la para auxiliar a tomada de decisão dos gestores, rompendo com a descontinuidade dos programas.

No Brasil, não existe ainda uma tradição em avaliação de programas como em países onde essa prática é mais antiga e coincide com o seu desenvolvimento sócio-cultural e histórico, como por exemplo, na Inglaterra, no Canadá e nos Estados Unidos. A década de 1980 no Brasil foi profícua para a luta dos movimentos sociais populares e urbanos, pois a AP passou a ser priorizada, divulgada e publicada, (Barreira, 2002).

A descontinuidade dos programas de governo é um fator que interfere na trajetória de desenvolvimento da avaliação. Assim, torna-se importante a responsabilização e a prestação de contas dos recursos utilizados em programas sociais, divulgando de forma transparente e massiva seus resultados, formas de acesso e impactos nas populações atendidas.

Um dos fatores que tem levado ao desenvolvimento da avaliação de programas são as demandas de responsabilização do governo e da Lei. No caso do Brasil, por exemplo, as pressões provêm dos Poderes Legislativo e Judiciário, dos Tribunais de Contas, pressões

intra-governamentais para maior transparência dos processos decisórios e também da mídia de comunicação (Silva, 2002a) e acrescenta-se, da sociedade civil.

A avaliação de programas sociais poderia ser incorporada ao cotidiano da administração pública, no ciclo da gestão dos programas. Entretanto, isso não tem ocorrido, devido à “complexidade inerente às metodologias de avaliação, em parte devido ao desinteresse dos formuladores de políticas e à desconfiança dos responsáveis pelos programas” (Cotta, 1998, p. 103). Concorde-se que independentemente da vontade individual dos formuladores de política e tomadores de decisão, do panorama de crise fiscal que enseja a utilização eficiente e eficaz dos recursos nas ações governamentais, os programas necessitam atuar sobre a realidade para a qual foram desenhados. Daí a necessidade de melhorias nos processos de avaliação de programas, mas principalmente programas sociais.

Os governos sempre tiveram a necessidade de avaliar suas atividades e tanto os governantes como os cidadãos comuns têm formado juízos de valor sobre o sucesso de programas específicos (Lerner, 1977, citado por Oskamp, 1984).

No intuito de contribuir para o crescimento dos estudos na área de Avaliação de Programas e melhoria das suas práticas, buscou-se neste trabalho investigar como ocorre o processo de ANOP, que possui característica de AP, no TCU.

Assim, o objeto de estudo desta tese de doutorado é formado pelas auditorias de natureza operacional - ANOP. Analisou-se sua metodologia, fundamentos e aplicação com base em 20 critérios do *Joint Committee*. Procedeu-se a meta-avaliação de dez ANOPs realizadas no período de 1999 a 2005, por meio da análise de conteúdo, síntese de categorias e no final integrou-se as análises com os critérios do *Joint Committee*, produzindo a Síntese de Estudos Qualitativos.

Utilizou-se como **método a meta-avaliação**. Assim, selecionou-se a auditoria do Programa de Erradicação do Trabalho Infantil – PETI, para testar a metodologia por meio de um estudo-piloto e através dos resultados direcionar as análises das demais dez avaliações que foram escolhidas intencionalmente para corresponder a uma avaliação em cada área de governo - agricultura, assistência social, cidadania, comércio e serviços, educação, energia, habitação, meio-ambiente, saúde e a área de trabalho.

Nesse sentido, **o objetivo geral** desta pesquisa é *meta-avaliar a auditoria de natureza operacional do TCU, seus critérios e padrões*, verificando-se a validade e confiabilidade dos procedimentos e resultados. Para consecução desse objetivo geral, será necessário alcançar os seguintes objetivos específicos:

- descrever os conceitos, metodologias e técnicas utilizadas pelo TCU para os seus trabalhos de ANOP;

- analisar a metodologia avaliativa do TCU e a extensão com que serve à análise dos diferentes programas nas áreas de governo;

- identificar os pontos fortes a serem ressaltados e os deficientes a serem melhorados no processo de ANOP mediante a checagem de critérios de qualidade de avaliações do *Joint Committee*;

- descrever as variáveis associadas a ANOP nos dez casos analisados mediante análise de conteúdo e síntese de categorias;

- caracterizar as práticas auditoria e avaliação de programas; e,

- propor contribuições da área de avaliação de programas, em termos do processo de ANOP realizado pelo TCU.

## **Capítulo 2: Revisão de Literatura**

Uma vez que o objetivo do presente trabalho é meta-avaliar a ANOP, seus critérios e padrões, a validade e confiabilidade dos procedimentos e resultados produzidos mediante checagem de critérios *do Joint Committee* e identificação de variáveis associadas ao processo por meio de análise de conteúdo e síntese de categorias, estruturou-se o presente capítulo em quatro tópicos. No primeiro apresenta-se a contextualização histórica da avaliação de programa, destacando-se os eventos que contribuíram para a sua formação. Na segunda parte, os conceitos e terminologias, envolvendo o planejamento da avaliação. A terceira parte fornece um panorama da Avaliação de Programas - AP no Brasil. E a quarta e última, questões metodológicas, envolvendo a subdivisão em causalidade em pesquisa, os delineamentos de pesquisa e a contribuição da psicologia na avaliação de programas sociais.

### ***2.1 Contextualização Histórica da Avaliação de Programas***

A AP data do século XVII, com a revolução científica iniciada na Europa, alicerçada nas teorias de Kepler, Bacon, Galileu e Descartes, sendo que os primeiros estudos avaliativos se espelhavam no paradigma da pesquisa experimental, influenciada pelas ciências naturais, de probabilidade, dedução e matemática.

Foi na área da saúde pública que os precursores da estatística Graunt em 1662 e Holley em 1963, realizaram seu trabalho de pesquisa avaliativa. Eles propuseram que os planos governamentais americanos, na área de serviços públicos, fossem baseados nos dados estatísticos coletados através de programas de avaliação. As primeiras tentativas de avaliação de serviços e programas no campo social, segundo Suchman, foram realizadas no setor da administração pública. [...] desenvolvendo-se paulatinamente até atingir sua maior evolução no século XX, foram realizadas sucessivas pesquisas avaliativas visando monitorar e melhorar a eficácia dos serviços oferecidos à população (Barreira, 2002, p. 18).

Para compreender os avanços teóricos e as práticas da AP, Stufflebeam e Shinkfield (1987) dividem a história da AP em cinco grandes períodos: o período pré-tyleriano, a época tyleriana, a época da inocência, a época do realismo e a época do profissionalismo. A referência a Ralph W. Tyler é devida a este ser considerado o *pai* da avaliação educacional.

O **período pré-tyleriano** foi o período até 1930. Neste período a AP existia, entretanto, não era uma prática claramente identificada. Por exemplo, a tarefa de avaliar indivíduos e programas, surgiu, no século 2000 aC (antes de Cristo), quando oficiais chineses realizaram algumas investigações sobre os serviços civis. No século V a.C., Sócrates e outros filósofos gregos utilizaram questionários avaliativos como parte de sua metodologia didática. No século XIX na Inglaterra, utilizaram-se Comissões Reais para avaliar os serviços públicos. Em 1845, nos Estados Unidos, Horace Mann, realizou uma avaliação baseada em testes de rendimento para verificar se as escolas de Boston produziam ensino satisfatório aos seus estudantes. Entre 1887 e 1898, Joseph Rice estudou os conhecimentos em ortografia de 33.000 estudantes de diversas escolas, e concluiu que o ensino da ortografia, que era um modismo na época, não produzia avanços substanciais na aprendizagem. Este estudo foi reconhecido como a primeira avaliação formal sobre um programa educativo realizado na América. Uma aproximação maior com o que se chama avaliação de programas surgiu no início deste século com o surgimento dos testes padronizados (Stufflebeam & Shinkfield, 1987).

A **época Tyleriana**, data do início dos anos 1930, quando Ralph Tyler cunhou o termo Avaliação Educacional, publicando uma ampla e inovadora visão do currículo e da educação. Durante um período de 15 anos, diversas instituições educacionais confirmaram tal método, que parecia ser superior aos demais. A principal característica do método era centrar-se nos objetivos claramente definidos, assim, a avaliação era definida como algo que determina se certos objetivos foram alcançados. Os currículos deveriam ser construídos determinando os rendimentos que se esperavam pelos estudantes, e os testes de rendimento determinariam o alcance dos rendimentos relacionados aos objetivos, pelos estudantes. No decorrer dos anos 30, com a Grande Depressão, Tyler dirigiu o estudo conhecido com “Estudo dos Oito Anos”, no qual se buscava avaliar a efetividade de currículos educacionais e estratégias didáticas utilizadas em 30 escolas americanas. Através desse estudo conseguiu difundir o seu método, a avaliação Tyleriana, que incluía comparações internas entre os resultados e os objetivos e não necessitava recorrer às comparações entre grupos de controle e grupos experimentais. Seu método se centrava em determinar o grau de êxito, opondo-se a métodos indiretos que determinam a qualidade do ensino, como o número de livros que há numa biblioteca, os materiais e participações da comunidade. As avaliações tylerianas não eram afetadas pela diferença entre as pontuações individuais dos alunos, já que geralmente cobriam uma gama mais ampla de variáveis

exógenas que as dos testes tipificados. O método de Tyler foi bem aceito pela comunidade americana e assim difundido durante os próximos 25 anos.

**A Época da Inocência** refere-se ao período compreendido entre os anos 1940 e 1950, quando a sociedade americana buscou esquecer a Guerra, superar a Grande Depressão, adquirir recursos e desfrutar da *boa vida*. Foi um período que apresentou um paradoxo de plenitude nas cidades e pobreza nas áreas rurais. Essa época foi marcada por profundos preconceitos e segregação raciais, embora parecesse despercebido pela população branca. Os recursos naturais eram desperdiçados; houve grande desenvolvimento da indústria e da capacidade militar, sem preocupação com os impactos ambientais que paralelamente produziam. Quanto à área da educação houve uma expansão das ofertas educacionais, da formação de professores e instituições educativas. Assim, a avaliação educacional expandiu-se, avaliadores escreviam sobre avaliação e faziam recompilação de dados, mas não conseguiam traduzir em seus esforços a melhoria do sistema educativo. Além disso, as avaliações educacionais dependiam dos orçamentos e interesses locais, pois as agências federais e estaduais estavam pouco envolvidas nesse processo. A ausência de financiamento externo para a avaliação levou à chegada do próximo período (Stufflebeam & Shinkfield, 1987).

**A Época do Realismo** refere-se ao período no início dos anos 1960 e foi marcada pelas avaliações em grande escala de projetos e currículos, financiada por órgãos governamentais. Assim, o progresso das metodologias avaliativas está relacionado com conceitos como utilidade e relevância. Nesse sentido, ocorreram profundas mudanças em decorrência dos interesses públicos e dos contribuintes que financiavam as avaliações, a avaliação transformou-se em uma indústria e em uma atividade profissional. Os métodos de avaliação utilizados nesse período eram: o método de Tyler, para definir os objetivos dos novos currículos e para avaliar o grau em que esses objetivos eram alcançados; os testes padronizados em nível nacional, para refletir os objetivos e conteúdos dos novos currículos e o método do critério profissional, usado para avaliar as propostas e verificar periodicamente os resultados.

No início dos anos 1960, alguns dos mais importantes avaliadores educacionais perceberam que seus trabalhos e suas conclusões não ajudavam aos elaboradores de currículos e não respondiam às perguntas sobre os programas que necessitavam comprovar o grau de sua efetividade. Stufflebeam e Shinkfield (1987) referem-se ao artigo de Cronbach que criticava fortemente a relevância e a utilidade das avaliações realizadas até então.

Nos primórdios da avaliação de programas pairava certa ingenuidade, apontada por Cook e Shadish (1986), na concepção dos programas e no uso das avaliações. Os avaliadores estavam desatentos às intransigências dos programas sociais; à inexatidão descritiva das teorias de tomada de decisão, e do limitado papel da avaliação como um dos muitos modelos formais de resolução de problemas sociais. Os cientistas sociais esperavam que os resultados das avaliações melhorassem o funcionamento dos programas. Nesse sentido, é preciso considerar que as avaliações não existem no vácuo, pois estão sujeitas principalmente à influência política dos tomadores de decisão e criadores das políticas e programas públicos. Por isso, muitas vezes os programas não refletem soluções para os programas sociais e sim a visão política adequada ao momento, em detrimento à realidade apresentada. A visão dos avaliadores muda quando verificam que os criadores de políticas e programas e funcionários do programa utilizam as suas metas pessoais para administrar os programas e não as melhores alternativas descobertas pelos avaliadores.

Assim, verifica-se que a avaliação de programas está estritamente relacionada a uma mudança histórica e ideológica que ocorreu com a transferência da responsabilidade ambiental, social e da qualidade de vida dos cidadãos comuns, da instância individual para a instância governamental (Rossi & Freeman, 1993).

Essa mudança histórica de transferência de responsabilidade pode ser percebida desde antes da **Primeira Guerra Mundial** (1914 – 1918), quando a responsabilidade pelos serviços prestados aos cidadãos era individual ou de organizações voluntárias, o que mudou durante os anos 1960 e 1970, quando então os governos sócio-liberais passaram a se engajar em programas objetivando a solução dos problemas sociais (Ala-Harja & Sigurdur, 2000). A avaliação de programas sociais ganhou maior visibilidade e importância nos Estados Unidos, embora tenha sido realizada também em outros países (Barreira, 2002).

Considerando-se tal raiz histórica, pode-se dizer que a Pesquisa Avaliativa teve um desenvolvimento relativamente recente, pois antes de 1930 havia poucos programas de saúde, educação, bem-estar e emprego, sob responsabilidade governamental e pública. E foi *na área de educação*, especificamente nos processos de alfabetização *na área do emprego e da saúde pública*, que a avaliação tornou-se um empreendimento sistemático de verificação o desempenho dos programas.

No início da Primeira Guerra Mundial, os programas para treinamentos de soldados desencadearam avaliações. Teve destaque o desenvolvimento de testes para avaliar perfil e para recrutar soldados pelo exército estado-unidense (Oskamp, 1984). No período da

Primeira Guerra Mundial, o acontecimento que se tornou o primórdio para o desenvolvimento da avaliação de programas foi a necessidade de suprir o exército americano com soldados para o *front* de guerra. De acordo com Rossi e Freeman (1993), na área da saúde pública, os esforços de avaliação se concentraram nas iniciativas de programas para reduzir a mortalidade e morbidade por doenças infecciosas. Nesse contexto, foram realizadas diversas avaliações menores, para verificar o êxito de tais campanhas ou propagandas dos programas de alfabetização e de saúde pública. Na época, essas campanhas objetivavam modificar os hábitos alimentares da população, devido à escassez de certos alimentos durante o período da guerra.

Segundo Rossi e Freeman (1993), durante a Segunda Grande Guerra (1938-1945), o exército americano e inglês possuía preocupação semelhante à da Primeira Guerra Mundial, de monitorar o moral dos soldados e também dos civis e avaliar as políticas de seus governos. Um pouco antes da Segunda Guerra Mundial, havia forte pressão para que os conceitos e técnicas do então chamado gerenciamento científico, que fora bem aplicado às indústrias e aos programas e atividades do governo, fossem também aplicados aos programas para o bem estar social. Assim de acordo com Oskamp (1984), a ciência de avaliação teve um grande desenvolvimento durante a Primeira e Segunda Guerra Mundial, produzindo conhecimentos para ajudar os governos a desenvolver e implementar sistemáticas de avaliação de seus programas públicos. Buscava-se aferir os processos e resultados de ações sociais para “estabelecer padrões de atendimento na oferta dos serviços sociais públicos” (Barreira, 2002, p. 21).

Ocorreu uma “explosão” da Pesquisa avaliativa no período subsequente a Segunda Grande Guerra, devido às necessidades de desenvolvimento de programas na área de habitação, educação, tecnologia, cultura, treinamento profissional e desenvolvimento de atividades preventivas de saúde. Aguilar e Ander-Egg (1995) destacam que na medida em que a metodologia avaliativa era aperfeiçoada, sua aplicação iria além dos programas econômicos, centrando-se principalmente na avaliação de programas educativos.

Vários estudos foram realizados nos Estados Unidos e na Europa, na Ásia, na América Latina e na África. Na Ásia, por exemplo, foi introduzida a avaliação nos programas de planejamento familiar, na América Latina os programas de nutrição e saúde e na África os programas de agricultura rural (Rossi & Freeman, 1993).

Em suma, as mudanças na estrutura política, econômica e social, ocasionadas pela Primeira e Segunda Grande Guerras, ocasionaram mudanças na prática de construção da avaliação de programas. Além do aumento dos investimentos em programas sociais,

creceu também a necessidade de avaliar seus resultados (Rossi & Freeman, 1993). Sendo que a avaliação quantitativa e qualitativa dos processos e resultados sociais refletia fortíssima influência das ciências econômicas, matemáticas e biológicas. De acordo com Barreira (2002, p. 22). [...] o rigor lógico e metodológico que formava os estudos das ciências exatas e biológicas era transferido para o campo da avaliação de políticas e programas sociais [...]", buscava-se determinar as leis e normas e variáveis explicativas das necessidades individuais e sociais a serem atendidas pelos serviços públicos.

Esse cenário começou a mudar em 1930, juntamente com as mudanças geopolíticas que modificaram os Estados Unidos e a Europa. Em 1930, as ofertas de serviços públicos, em forma de Programas Sociais, começaram a crescer paralelamente à Grande Depressão de 1930, pois havia uma preocupação dos governos em recuperar a economia, o emprego e a indústria prejudicados pela Guerra. De acordo com Rossi e Freeman (1993), mesmo considerando-se os passos iniciais da AP durante a 1ª Guerra, foi a partir dos anos 1930, que um número significativo de cientistas sociais adotou a aplicação rigorosa de métodos científicos da pesquisa social para avaliar as ações dos programas nas comunidades e avaliar a sua implantação, coincidindo com o refinamento dos seus próprios métodos. Um elemento importante nesse período foi o desenvolvimento das ciências e pesquisas sociais nas universidades, que continuam até hoje a influenciar esse campo de estudos.

Em 1932 foi lançada a Política do *New Deal* (Novo Acordo) do Presidente dos Estados Unidos, Franklin Roosevelt. Essa política englobava um conjunto de reformas econômicas com a finalidade de fazer com que o país superasse a crise desencadeada em 1929. Nessa política foram criados programas de combate ao desemprego e fortalecimento da indústria nacional, financiados com dinheiro público. Surgiu então a necessidade de verificar a efetividade e o uso do dinheiro público investido nesses programas. O objetivo era que os laboratórios criados pela política do *New Deal* permitissem o uso mais efetivo dos métodos experimentais nos projetos de pesquisa dos cientistas sociais, dessa forma, as pesquisas subsidiaram a fiscalização e o planejamento dos programas implementados pelas autoridades governamentais. Verificou-se também nesse período o desenvolvimento científico em termos de métodos de coleta e análise de dados por meio de técnicas estatísticas.

Outro elemento importante que levou ao desenvolvimento das políticas universalistas, asseguradoras dos direitos dos cidadãos foi o Welfare State (Estado de Bem Estar Social), que teve amplo desenvolvimento nos países desenvolvidos da Europa e nos Estados Unidos, após a Segunda Guerra Mundial. O *Welfare State* concebido como um

pacto social buscava proporcionar os serviços de segurança socioeconômica, promovendo os serviços sociais dentro das políticas que visavam atender à pobreza e assegurar os mínimos sociais. De acordo com Barreira (2002, p. 22), “... foi enfatizada a importância da execução de programas de ajuda internacional, ao chamado Terceiro Mundo, voltados para combater a pobreza pela via do planejamento familiar, saúde, nutrição e desenvolvimento de comunidades rurais”.

Outros fatos ocorreram na década de 1930, no campo científico, e configuram a trajetória histórica da AP. Por exemplo, o aumento dos estudos nas universidades, o que levou ao amplo desenvolvimento das ciências sociais, oferecendo suporte à pesquisa social. Nesse sentido, Shadish, Cook, e Levinton (1995) informam em seu trabalho que a avaliação de programas desenvolveu-se em três estágios: 1) ao longo da década de 1960, caracterizando-se pela aplicação rigorosa dos métodos científicos à avaliação de programas; 2) a partir da década de 1970, com o enfoque no uso dos resultados das avaliações pelos administradores e gestores públicos e 3) ainda na década de 1970, a partir da síntese dos trabalhos anteriores e da avaliação dos aspectos positivos e negativos dos métodos e técnicas de avaliação. Enfatizava-se a escolha de determinada prática a partir das circunstâncias e do contexto político que envolvesse a implantação de políticas.

Quanto aos trabalhos pioneiros e publicações de AP, pode-se citar na educação o trabalho de abordagem experimental de Donald Campbell (1969, citado por Cano, 2002), quando a avaliação de programas começou a se firmar como uma disciplina. Nesse campo, tiveram destaque os textos publicados por Campbell e Julian Stanley em 1963, *Desenhos experimentais e quasi-experimentais para a pesquisa sobre o ensino*, e outro artigo escrito juntamente com Thomas Cook em 1979, intitulado *Quasi-experimentação. Questões de desenho e análise para contextos naturais* (Cano, 2002).

Nos anos 1950, a avaliação de programas tornou-se amplamente utilizada abrangendo tópicos como programas de planejamento familiar, nutrição, saúde e programas de desenvolvimento da agricultura em diversos países. Nos Estados Unidos, a pesquisa avaliativa foi conduzida por meio de procedimentos e tratamentos psicoterápicos e médicos, programas de atendimento familiar, de prevenção à criminalidade e muitos outros programas locais e nacionais (Oskamp, 1984). No final dos anos 1950, a avaliação de programas em larga escala tornou-se comum. Os cientistas sociais estavam engajados em avaliações de programas de prevenção à delinquência, projetos de reabilitação de presos, tratamentos psicoterápicos e psicofarmacológicos, programas habitacionais e atividades da sociedade organizada. Esses estudos aconteceram nos Estados Unidos,

Europa e em outros países industrializados e também em países menos desenvolvidos (Rossi & Freeman, 1993).

Em síntese, no período de 1950, no campo das Ciências Sociais, a avaliação de programas surgiu como uma necessidade a partir da qual se desenvolveu metodologicamente, sendo esse conhecimento aplicado à avaliação da política social e pública da administração do bem estar. O foco das avaliações direcionava-se para programas para diminuir a mortalidade infantil, diminuir a gravidez entre adolescentes e programas para diminuir os acidentes de trânsito (Cano, 2002). De acordo com Barreira (2002), os anos 1950 a 1970 podem ser considerados como período de maior desenvolvimento ou período áureo do campo da pesquisa avaliativa de programas sociais.

Conforme observado, a partir dos exemplos fornecidos até o momento, a avaliação de programas é uma área de estudos ampla, constituindo-se em uma disciplina, dentro da qual convergem várias abordagens, o que pode ser verificado pela diversidade de projetos ou programas sociais avaliados dentro do âmbito da política dos governos.

Na continuidade do devir histórico da avaliação de programas, a maior parte dos acontecimentos a seguir apresentados e que marcaram o período de 1960 a 1990, foram descritos por Rossi e Freeman (1993), exceto quando se destacar outro autor.

Nesse período, de acordo com Ala-Harja e Sigurdur (2000, p. 14) “o processo de planejamento e programação foram introduzidos para melhorar a formulação de políticas”. Além disso, melhores condições fiscais e o conhecimento dos métodos das ciências sociais, incluindo a aplicação de *surveys* em grandes amostras e sofisticadas análises estatísticas, ampliaram ainda mais o seu uso em avaliação.

Durante os anos 1960, o número de livros e artigos sobre a prática da pesquisa avaliativa cresceu grandemente. Nos Estados Unidos, o motivo principal para esse interesse em avaliação durante o período foi a guerra do governo contra a pobreza (Rossi & Freeman, 1993). Em 1961, houve o empenho do presidente Lyndon Johnson e depois do presidente John F. Kennedy no desenvolvimento da Grande Sociedade Americana, oferecendo suporte social e financiamento de bilhões de dólares para as Reformas Sociais e programas comunitários na área da saúde, educação e assistência social. O objetivo dos Estados Unidos era reforçar o seu domínio, por meio da política social através de um forte marketing nacional e internacional (Barreira, 2002).

Até o final de 1967, a definição de Pesquisa Avaliativa de Schuman (citado por Rossi & Freeman, 1993) era uma definição útil e suficiente para o campo, pois envolvia a aplicação de técnicas de pesquisa social ao estudo em grande escala dos programas de

intervenção social. A mudança qualitativa que ocorreu em relação à definição foi originada pela lacuna que essa mudança trazia em relação à influência dos consumidores da pesquisa avaliativa - que eram os criadores das políticas, os planejadores dos programas e os gerentes dos programas. Nesse sentido, a análise de Cook e Shadish (1986) é pertinente quando afirmam que o que mudou ao final da década de 1960 foi a perda da inocência em relação à conceituação e realização da AP. Conforme Cook e Shadish (1986), os avaliadores emergiram desse período “mais sofisticados” sobre a complexidade de sua tarefa, sendo mais realistas sobre a política nas quais os programas sociais existem e sobre como usar as informações da ciência social na resolução dos programas.

Esses atores utilizavam os resultados das avaliações e suas percepções eram afetadas pelas organizações que encomendavam as avaliações. O foco das avaliações era a melhoria dos programas e os seus resultados eram empregados como mecanismos de *feedback* para os gestores dos programas (Rossi & Freeman, 1993; Ala-Harja & Sigurdur, 2000).

Nesse sentido, a crítica que Rossi e Freeman (1993) fazem à avaliação refere-se ao ímpeto da aplicação, pura e simples, de métodos, pois, segundo os autores, a avaliação envolve mais aspectos que apenas a sua aplicação, ou seja, sua concepção, planejamento, execução e monitoramento. Segundo esse pensamento a avaliação é concebida como parte do ciclo da política e atividade gerencial, e como tal, serve como subsídio à entrada (*inputs*) de informações dentro de um complexo mosaico. Dentro desse mosaico emergem decisões políticas e alocações de recursos para o planejamento, delineamento, implantação e continuidade dos programas, que por sua vez, propõem melhores condições de vida às pessoas. Assim, conclui-se nesse contexto, que a pesquisa avaliativa precisa ser concebida como parte integral do ciclo de gestão que envolve formular, implementar e controlar a política social e pública das administrações governamentais. Cook e Shadish (1986) já apontavam um cuidado a ser tomado em relação ao uso da avaliação, que ela não fosse apenas justificativa de decisões repentinas tomadas em curto-prazo, quando a decisão nem sequer estivesse relacionada aos seus resultados, ou seja, fazer um uso político da avaliação.

Conforme Faleiros (2000) no Brasil, dos anos 1930 aos anos 1960, a política trabalhista desenvolvida por Getúlio Vargas procurou controlar as greves e os movimentos operários e estabelecer um sistema de seguro social. Assim, foram criados o Ministério do Trabalho, em 1930; e o sistema de seguros foi sendo implantado através do Instituto de Previdência Social, para uma parcela da população (marítimos, estivadores, bancários e industriários). Sendo que os trabalhadores rurais passaram a ser atendidos pela

Previdência, somente a partir dos anos 1970. Em 1942, foi criada a Legião Brasileira de Assistência – LBA, para atender as famílias de pracinhas envolvidos na Guerra. Em 1946 começaram a ser implantados serviços sociais, envolvendo a maternidade e a infância; entretanto, predominada o assistencialismo, envolvendo obras de caridade, ações das primeiras-damas. Em 1960, houve a aprovação da Lei Orgânica da Previdência Social – LOPS, definindo a unificação de vários institutos, organizando cerca de 300 leis e decretos referente à previdência social.

Ao longo dos anos 1970 e 1980, o Estado Brasileiro tentou satisfazer as demandas da população desprotegida. Por exemplo, em 1966 foram unificados os institutos de previdência; criando-se assim o Instituto Nacional de Previdência Social – INPS. Em 1974, os idosos, com mais de 70 anos, foram beneficiados por lei que implementou a renda mensal vitalícia no valor de um salário mínimo, para os que tinham contribuído, pelo menos por um ano, com a Previdência Social. Também em 1974 foi criado o Ministério da Previdência e Assistência Social; com a incorporação da LBA, Fundação para o Bem-Estar do Menor - FUNABEM, Central de Medicamentos – CEME, Empresa de Processamento de Dados da Previdência Social - DATAPREV (Faleiros, 2000).

No Brasil, a problemática social tem sido prioridade de atuação do governo até a atualidade, como pode ser observado por meio da criação de programas sociais de grande escopo como o Programa Bolsa Família, já no governo Luis Inácio Lula da Silva.

No contexto mundial, no início dos anos 1970, uma grande variedade de livros surgiu, incluindo críticas no que diz respeito à qualidade metodológica das avaliações de vários estudos. Entre as diversas publicações da época, foi criado o jornal *Evaluation Review*, em 1976. Essa publicação é amplamente lida por pesquisadores avaliadores de diversas disciplinas até hoje. A proliferação de publicações e conferências, a formação de associações profissionais e seções especiais de estudos de avaliação nos encontros acadêmicos e nos grupos de consultorias testemunharam o rápido desenvolvimento do campo.

Em suma, pode-se afirmar que dois fatos contribuíram para o desenvolvimento do campo: a) a criação de uma sistemática de coleta de dados e o refinamento de procedimentos de análise de dados, tipo *survey*; e b) o desenvolvimento de computadores com capacidade para examinar uma grande quantidade de variáveis utilizando-se estatística multivariada.

Nos anos 1980 a avaliação de programas enfrentou problemas, no âmbito dos governos, que na época eram mais conservadores, de tentar lidar com a escassez de

recursos destinados a programas públicos. Essa escassez fiscal fez com que a avaliação passasse a ser vista como “uma ferramenta útil na reconsideração da justificativa de políticas e na racionalização da alocação de recursos orçamentários” (Ala-Harja & Sigurdur, 2000, p.14). Voltando ao contexto do Brasil e a política de bem estar social; com a Constituinte de 1988, conseguiu-se, “no plano social, o avanço dos direitos das mulheres, das crianças, dos índios e a inclusão do conceito de Seguridade Social que compreende direitos universais à saúde, direitos à previdência e à assistência social” (Faleiros, 2000, p. 49). Em 1985 o Instituto Nacional de Previdência Social – INAMPS, foi transferido para o Ministério da Saúde, configurando-se no Sistema Único de Saúde – SUS. Sendo que o papel do governo federal era a coordenação, e não execução das políticas sociais, que ficou a encargo dos municípios e estados. Em 1993 “foi aprovada a Lei Orgânica da Assistência Social - LOAS, que torna a Assistência Social um dever do Estado e um direito do cidadão” (Faleiros, 2000, p. 51).

Por fim, nos anos 1990, a crise fiscal e econômica que os países enfrentavam e que começou desde a década de 1980, com a crise fiscal e econômica, tornam necessário o estabelecimento de prioridades na alocação dos recursos para que os programas mais importantes sejam contemplados.

Em Cartagena em 1994 o *Office of Evaluation and Oversight* – OVE/BID - Banco Interamericano de Desenvolvimento, realizou um encontro entre Brasil, Argentina e Chile, priorizando a discussão e a construção de uma nova visão para o exercício da *função avaliação*, que entre o período de 1980 e 1990 era concebida como mecanismo de controle dos governos em relação à alocação e uso dos recursos em programas públicos. Scriven, (1991, citado em Barreira, 2002, p. 40-41) esclarece que função avaliativa ou medição avaliativa é uma transposição da matemática. Refere-se ao “processo de calcular a soma de vários valores de peso, como na avaliação de um polinômio. Essa é a chave do processo lógico que distingue avaliação de mensuração”. Assim, a avaliação torna-se “distinta das atividades de controle e acompanhamento da execução de projetos setoriais, enfocados isoladamente [...] passa a se constituir no objeto de acompanhamento dos resultados e dos processos de sua consecução [...] das estratégias de política pública” (Ala-Harja & Sigurdur, 2000, p. 7). Assim, tornou-se crescente a participação da sociedade, dos grupos políticos, das organizações e das associações, nas discussões sobre a destinação dos recursos públicos. Silva (2002a) refere que os estudos e pesquisas de avaliação para a gestão das políticas públicas, em diferentes países, vêm sendo desenvolvido principalmente a partir da década de 1990. Especificamente nos EUA, no final do governo Clinton, o

financiamento para os programas sociais aumentou, assim como as atividades de avaliação dos mesmos. Dentro do contexto que era de escassez fiscal, os programas passaram a receber financiamento.

Rossi e Freeman (1993) comentam que independentemente da postura política dos governos, seja liberal ou conservadora, a avaliação dos programas desencadeada por governos, dentro de suas políticas, precisa ser concebida como uma constante necessária e integradora do seu ciclo de gestão. E complementando-se ainda, realizada, planejada e com destinação de verbas previstas nos orçamentos dos programas para a realização de avaliações.

As evoluções e transformações tratadas até este ponto foram resumidas e descritas sinteticamente no quadro a seguir, que retratou os principais eventos históricos que marcaram o desenvolvimento da avaliação de programas no mundo. Tal esquematização teve como objetivo tornar claro os principais eventos que colaboraram para o desenvolvimento da AP, porém, não tem a pretensão de ser a melhor ou mais abrangente.

### **Quadro 1. Eventos que influenciaram o desenvolvimento da avaliação de programas**

Período	Eventos Históricos
Século XVII	Os primeiros estudos avaliativos se espelhavam no paradigma da pesquisa experimental, modelada nas ciências naturais, de probabilidade e dedução, na “matematização” do conhecimento (influência da teoria de Kepler, Bacon, Galileu e Descartes)
1ª Guerra Mundial 1914-1918	Desenvolvimento de testes para avaliar o perfil e recrutar soldados pelo exército americano. Treinamento de soldados. Na área da saúde pública tomaram-se iniciativas para reduzir a mortalidade e morbidade por doenças infecciosas.
	Em 1932, Política do <i>New Deal</i> (Rosevelt - EUA), desenvolvimento de programas de combate ao desemprego e fortalecimento da indústria.

Década de 1930	Cientistas sociais adotaram a aplicação rigorosa de métodos científicos da pesquisa social para avaliar as ações dos programas nas comunidades e avaliar a sua implantação.
	Desenvolvimento dos métodos estatísticos e de pesquisas aplicáveis ao estudo dos problemas sociais e das relações interpessoais. Aperfeiçoamento da sistematização da coleta de dados, sofisticação das técnicas de pesquisa, desenvolvimento de computadores capazes de processar grande quantidade de variáveis estatísticas.
2ª Guerra Mundial	Desenvolvimento do método científico para avaliar diversos tipos de treinamento e propagandas.
1939-1945	O aumento do investimento financeiro em programas de bem estar sociais gerou uma forte demanda em relação aos resultados desses programas.
Década de 1950	A ampliação da prática da avaliação de programas nos países industrializados e nos países em desenvolvimento.
	Surgimento do <i>Welfare State</i> na Europa e Estados Unidos. Execução de programas de ajuda internacional ao Terceiro Mundo, voltados para combater a pobreza pela via do planejamento familiar, saúde, nutrição e desenvolvimento de comunidades rurais.
	Nos Estados Unidos, desenvolvimento de procedimentos e tratamentos médicos, de atendimento familiar e prevenção à criminalidade.
Década de 1960	Lançamento, nos Estados Unidos, de um amplo conjunto de programas sociais, iniciados pelo presidente Lyndon Johnson.
	Criação de novas formas de ajuda aos países em desenvolvimento, mediante financiamento de projetos econômicos.
	O número de livros e artigos sobre a prática da Pesquisa avaliativa cresceu (ocorreu o <i>boom</i> da avaliação de programas).
Década de 1970	Uma grande variedade de livros surgiu, incluindo críticas em relação à qualidade metodológica das avaliações de vários estudos.
	Lançamento do jornal <i>Evaluation Review</i> em 1976, amplamente lido por pesquisadores avaliadores de diversas disciplinas.
Década de 1990	Crise fiscal e econômica que os países enfrentam torna necessário o estabelecimento de prioridades na alocação dos recursos para que os programas mais importantes sejam contemplados.
	Em Cartagena em 1994, o <i>Office of Evaluation and Oversight – OVE/BID</i> realizou um encontro entre diversos países, com prioridade à discussão e a construção de uma nova visão para o exercício da função avaliação.
	Crescente participação da sociedade, grupos políticos, organizações e associações, nas discussões sobre a destinação dos recursos públicos.

De acordo com as informações sintetizadas no Quadro 1 e detalhadas anteriormente, o que se verifica tanto na revisão dos eventos históricos, quanto de sua síntese, é que a evolução histórica se deu tanto em termos científicos, de pesquisas em avaliação, quanto em acontecimentos que marcaram a ampliação do seu uso. Entretanto, há que se considerarem ainda as necessidades existentes, refinamento, difusão e apropriação da avaliação nos diversos campos das políticas e programas de governo. Considera-se também que o uso da AP é relativamente recente no Brasil, apesar de já estar consolidado em países da Europa e no Canadá, Reino Unido e Inglaterra (Barreira, 2002).

Apesar desta contextualização histórica mostrar a evolução em termos de transformação e produção de conhecimento, de práticas, pesquisa e aplicação, não se poderia deixar de considerar alguns mitos associados à avaliação o que prejudica o seu uso e difusão (Ala-Harja & Sigurdur, 2000; Posavac & Carey, 2003; Rossi & Fremann, 1993).

Assim, por exemplo, alguns tomadores de decisão, responsáveis pela criação de políticas, consideram a avaliação de programas como um modismo sem aplicação prática e que produz burocracia, ou ainda uma atividade teórica que produz dados vagos e pouco aplicáveis na prática. Esse quadro de descrédito na avaliação existe na concepção de alguns gestores que receiam pela responsabilização que esse processo pode gerar o que lhes implica pensar o uso que se faz da avaliação.

Podem-se contestar esses mitos à medida que se avança no conhecimento sobre avaliação de programas, tanto por aqueles que produzem conhecimento nesta área, quanto por aqueles que fazem do seu uso uma ferramenta de trabalho. Ferramenta que apóia a tomada de decisão, a promoção, o conhecimento e a melhoria dos programas, no processo de gestão de organizações públicas ou privadas. Além da melhoria do processo de gestão organizacional, concorda-se com Ala-Harja e Sigurdur (2000, p. 6) que existem qualidades no processo de avaliação de programas e que tais qualidades “podem complementar e desenvolver esforços de mensuração, monitoramento e auditoria de desempenho”.

Para o esclarecimento dos termos, caracterização e diferenciação na prática em avaliação de programas buscaram-se ilustrar no próximo tópico os principais conceitos e terminologias associadas à avaliação de programas.

## ***2.2 Avaliação de programas: conceitos***

Neste tópico abordaram-se alguns conceitos centrais para a compreensão do que vem a ser avaliação de programas. Primeiro, abordou-se a diferença entre avaliar e medir; segundo, apresentou-se o conceito de programa, projeto e plano; terceiro, apresentou-se o conceito de avaliação de programas.

### *Avaliar e medir*

Segundo Cano (2002), o termo avaliação pode assumir diversas conceituações *latu senso*, por exemplo, as avaliações que geralmente se fazem no cotidiano, a respeito de coisas, pessoas ou situações, em tais avaliações são emitidos juízos de valor. Então, avaliar consiste em emitir julgamento de valor ou atribuir valor a algo.

Pode-se dizer que avaliar é uma forma de estimar, apreciar e calcular. É um processo no qual emitimos juízos de valor de acordo com determinados critérios, no qual nos baseamos para tomarmos decisões sobre algo. Essa definição genérica pode ser aplicada a diversas ponderações que se realiza no dia-a-dia e refere-se à avaliação no sentido informal. Já a avaliação formal, a avaliação sistemática, é aquela utilizada para avaliar

serviços ou atividades profissionais e utiliza os mesmos métodos e técnicas da pesquisa social (Aguilar & Ander-Egg, 1995).

Sobre a *diferença entre avaliar e medir*, os dois termos estão relacionados, entende-se que a noção de medição de programas está associada ao paradigma positivista, conforme verificado na quantificação dos resultados e metas atingidas. Está relacionada com a sua origem histórica associada às ciências exatas e a contabilidade, por isso alguns autores às vezes utilizam o termo função avaliação, que provêm da função matemática. Embora o processo avaliativo possa conter mensuração, crê-se que vai além dele. De acordo com Aguilar e Ander-Egg (1995, p. 19), “a medição é o ato ou processo de determinar a extensão e/ou quantificação de alguma coisa; a avaliação, ao contrário, faz referência ao ato ou processo de determinar o valor dessa coisa”, pois é possível medir sem valorar, e pode-se valorar algo sem medi-lo. Os autores destacam diferenças entre avaliação e alguns outros termos como: seguimento e controle.

Seguimento refere-se ao exame contínuo e periódico de uma unidade administrativa, com o objetivo de garantir o cumprimento do calendário de trabalho e a entrega dos insumos (bens, fundos, serviços, mão-de-obra, tecnologia, etc.) para alcançar os objetivos previstos no programa. O seguimento fornece informação que permite modificar e reorientar os aspectos operativos do programa e é parte da função gerencial que a ser realizada periodicamente. Aguilar e Ander-Egg (1995, p. 21) destacam que o seguimento pode contribuir na tarefa de avaliar “[...] mas pode haver avaliação sem ter havido seguimento e vice-versa: pode haver seguimento de um programa ou projeto sem que isto implique necessariamente em uma avaliação sistemática”.

Já a diferença entre avaliação e controle consiste em que o controle é uma verificação de resultados e a avaliação é uma ponderação ou julgamento desses resultados, “o controle é uma verificação de algo, e a avaliação é um juízo acerca desse algo” (Aguilar & Ander-Egg, 1995, p. 21).

Avaliação é uma estimativa das características e efeitos de uma política ou de um programa ou projeto, concluído ou em fase de execução, da sua concepção, implantação e resultados. Ela visa determinar a relevância e realização dos seus objetivos e sua eficiência, efetividade, impactos e sustentabilidade do ponto de vista do desenvolvimento. Uma avaliação supre de informações críveis e úteis e permite a absorção dos resultados pelo processo decisório e pelos beneficiários das políticas e programas (Silva, 2002b).

O enfoque da avaliação geralmente está na eficiência, efetividade; impacto e sustentabilidade do programa, visando por meio da avaliação, a sustentação do processo de

tomada de decisão. Estimar as características de um programa em qualquer uma de suas fases é uma das atribuições básicas da avaliação. Essa forma de conceber avaliação assemelha-se à de por Fernández-Ballesteros, Vedung e Seyfried, (1998) e a de Posavac e Carey (2003).

Entende-se que a avaliação de políticas e programas sociais envolve além da medição um processo amplo de julgamento de mérito, valor e validade (Barreira, 2002), idéia associada a diversos autores (Aguilar & Ander-Egg, 1995; Fernández-Ballesteros, Vedung & Seyfried, 1998; Posavac & Carey, 2003). Avaliar um programa implica uma operação mais ampla, que abrange diversos aspectos do programa, e medir, seria uma atividade mais específica e pontual que envolve a determinação de critérios específicos, sua quantificação, mensuração. Nos programas de escopo social, a avaliação também pode ser realizada por meio de números, mas também pelo estabelecimento de indicadores, no caso de caso da circunstância de objetivos mais complexos, encontrados comumente nos programas sociais.

#### *Programa, projeto e plano*

Inicialmente, considera-se o que vem a ser um *programa*. Entre os diversos tipos de programas existentes considerou-se um tipo de programa específico – o programa social – que pode ser compreendido como “uma intervenção sistemática planejada com o objetivo de atingir uma mudança na realidade social” (Cano, 2002, p.9).

Uma idéia adotada neste trabalho e que pode distinguir programa de outras atividades ou ações é a de que um programa surge para suprir uma necessidade detectada em um ambiente em determinada população (Posavac & Carey, 2003).

A existência de problemas e necessidades em uma população pode ser atendida por meio de programas e no caso do presente trabalho, no âmbito das políticas públicas. Considerando-se que os problemas sociais geralmente são tratados pelas políticas públicas de governo, justifica-se, o uso da literatura de avaliação conforme tratada nos programas com contexto de política pública. Assim, considerou-se a realidade social como fluída e mutável e que os programas são desenhados dentro desse contexto, para intervir sobre determinada situação na busca de solucionar, aprimorar ou prover soluções. Pode-se traçar uma comparação entre o conceito de programa enquanto alternativa ao problema da população e o conceito de organização do trabalho; ambos surgem para suprir o anseio de determinado sujeito ou agrupamento de pessoas. Os programas e as organizações se diferenciam no seu ciclo de vida, as organizações possuem aspiração de perdurar indefinidamente, e os programas nascem com escopo e tempo de existência determinado,

mesmo os de longa duração. Não se está afirmando que seja impossível transformar um programa em algo mais duradouro, como um serviço de prestação continuada, e sim, que a partir das avaliações que são realizadas sob os mesmos, existem prioritariamente três possibilidades: manutenção, modificação ou encerramento. Associada ao ciclo de vida dos programas tem-se a avaliação.

Cotta (1998, p.104) faz a distinção entre *projeto*, *programa* e *plano*. Para a autora, essas três modalidades de intervenção social diferem em escopo e duração. Sendo que o projeto seria “uma unidade mínima de destinação de recursos, e por meio de um conjunto integrado de atividades, pretende transformar uma parcela da realidade, suprimindo uma carência ou alterando uma situação problema”. Já um conjunto de projetos visando os mesmos objetivos formariam um programa. Por fim, o plano agregaria programas afins, estabelecendo diretrizes para a intervenção social.

A concepção de plano remete a uma abrangência maior quando se trata de intervenção social. Por exemplo, nas políticas públicas brasileiras os planos são elaborados para estabelecer as diretrizes de uma política, assim são criados pelos governos, os Planos Pluri-anuais. Tais planos têm um grande escopo, prevendo diretrizes, custos, orçamentos para áreas de atuação do governo, e podem ser desdobrados em programas de diversas áreas.

#### *Avaliação de Programas*

A profusão de conceitos e definições que ocorre em ciências humanas, também, é característica da área de avaliação de programas. Pode-se afirmar que não há consenso em relação a uma única definição nessa área. Argumenta-se que uma definição apenas também não seja o ideal, em uma área onde há confluência de conhecimentos de diversas disciplinas como economia, política, administração, estatística, sociologia, psicologia e outras. A confluência de conhecimentos e seus diferentes usos associados à avaliação, tornam indesejável e não aplicável na prática “o melhor método ou técnica” de avaliação a qualquer programa.

A avaliação de programas refere-se ao processo de determinar o mérito, custo ou valor de programas sociais, como também de programas de políticas públicas, projetos privados, regulamentos públicos e intervenções públicas ou privadas (Fernández-Ballesteros, Vedung e Seyfried, 1998). Sendo que o valor, mérito ou custo de um dado objeto possui conotações filosóficas e éticas e na literatura que trata de programas, depende de várias ideologias de avaliação. No entanto, há um consenso no campo de avaliação de programas: valores, custo ou mérito estão associados ao atendimento de metas e objetivos

pré-estabelecidos e resultados testados depois da implantação do programa (Fernández-Ballesteros, Vedung & Seyfried, 1998).

Uma das formas de determinação do valor dos programas é por meio do estabelecimento de critérios e definição de padrões. Nesse sentido, Posavac e Carey (2003) destacam que essa é uma tarefa árdua, de disputa e controvérsia entre as partes envolvidas na avaliação do programa. Os critérios são estabelecidos com base nos valores das pessoas, nos seus julgamentos e preferências.

Para ilustrar a diversidade de métodos, termos e conceitos apresentaram-se neste trabalho a terminologia utilizada dentro da área da Psicologia, também por ser este um trabalho da área de Psicologia Organizacional. Tem-se claro que a avaliação de programas se faz pela atuação conjunta de especialistas e que a responsabilidade é, em geral, distribuída entre muitas pessoas, com papéis e especialidades próprios (Posavac & Carey, 2003). Portanto, não se supõe ser essa uma área exclusiva da Psicologia.

Fernández-Ballesteros, Vedung e Seyfried (1998) compreendem a *avaliação de programas*, como uma disciplina metodológica, exigindo investigação sistemática, ferramentas, instrumentos e estratégias científicas. Em termos práticos, pode-se falar em avaliação como uma tarefa empreendida por pessoas ou por instituições em relação a programas específicos.

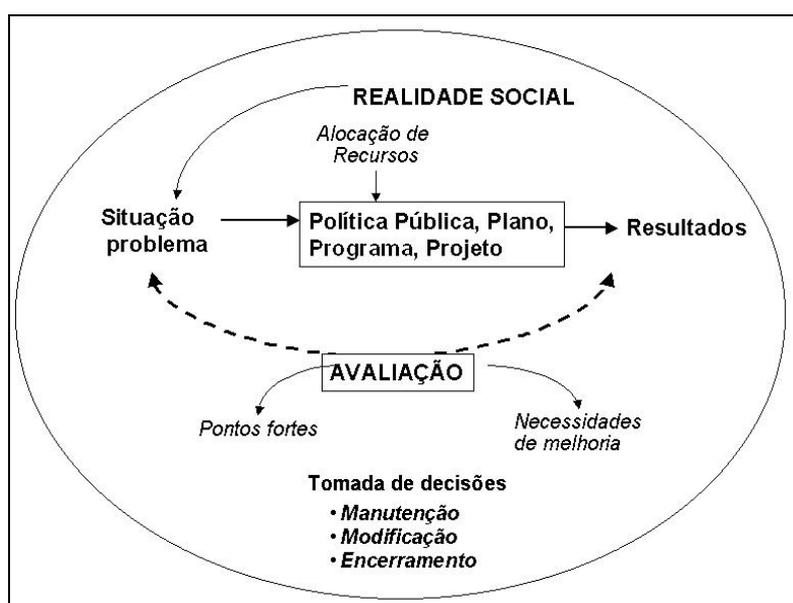
Considerou-se na presente pesquisa, que um programa pode fornecer serviços e produtos, que podem ser tangíveis ou intangíveis. Destaca-se que o foco desta pesquisa recaiu sobre programas sociais e públicos, que atendem a necessidades existentes na população. Assim, os conceitos a seguir, remetem à literatura de avaliação de programas existente na área da Psicologia, e enfatizam o entendimento da avaliação de programas enquanto avaliação de um serviço prestado à população.

Para proceder à avaliação são necessários determinados procedimentos. Tais procedimentos são necessários para determinar se um serviço é: a) necessário e utilizável; b) oferecido de maneira suficientemente intensiva para suprir às necessidades identificadas e não atendidas; c) oferecido tal como foi planejado; e d) efetivo, no sentido de atender as necessidades das pessoas dentro de um custo razoável e sem outros tipos de efeitos não planejados ou desejados. Entende-se que esses quatro pontos convergem para a avaliação que compara os objetivos do programa e os seus resultados.

Em uma visão sistêmica, inter-relacionada ao ambiente apresenta-se na Figura 1, a avaliação como um processo que interage com o ambiente no qual está ou será inserida estando sujeita às suas influências. De acordo com a Figura 1, a avaliação pode dar-se

antes, durante ou depois da instalação de um programa. Ocorre antes, quando do diagnóstico da Situação Problema; durante, para verificar se os passos necessários ao funcionamento do Programa/Política/Plano ou Projeto, estão sendo cumpridos e após, quando analisa os impactos do programa para balizar a tomada de decisão.

**Figura 1. Contextualização da Avaliação de programas**



Fonte: adaptado de Posavac e Carey (2003, p. 14).

Posavac e Carey (2003) e outros autores (Cotta, 1998; Cano, 2002) concebem a avaliação dentro de um ciclo de *feedback*, que atende às necessidades humanas, por meio dos programas. Destaca-se que a concepção de avaliação no contexto ambiental envolve e gera a necessidade da criação do programa, e posteriormente, a necessidade de avaliação do programa.

Avaliar determinando o mérito, custo e valor, também é uma tarefa necessária que se estende aos programas; políticas públicas; projetos privados; regulamentos públicos; intervenções públicas ou privadas. Acrescenta-se que a *avaliação de programas* – AP, tratada neste trabalho como *pesquisa avaliativa*, vai além desse conceito e envolve uma discussão do conceito de avaliação enquanto método, disciplina e estabelecimento de padrões científicos. De acordo com Barreira (2002), “[...] O desenvolvimento da pesquisa avaliativa traz no cerne não apenas a importância da avaliação como o modo de julgar processos ou ações, mas traz a apreensão de que a avaliação é produção de conhecimento”.

No caso de políticas públicas, que concretizam planos e metas por meio da ação dos programas, a avaliação é uma ferramenta, que propicia informações sobre os resultados alcançados pelos programas, isto é, “[...] um conjunto de atividades organizadas para realização dentro de cronograma e orçamento específicos do que se dispõe para a implantação de políticas, ou seja, para a criação de condições que permitam o alcance de metas políticas desejáveis” (Ala-Harja & Sigurdur 2000, p. 8). Os autores ressaltam ainda que a avaliação [...] “pode comportar atividades relativamente menos abrangentes ou programas de amplo escopo”. Sendo assim, há de acordo com Rossi e Freeman (1993) uma aceitação geral de que a pesquisa avaliativa deve se apropriar do método científico como um meio para tratar os problemas sociais.

Para Ala-Harja e Sigurdur (2000) o termo avaliação de programa compreende a avaliação dos resultados de um programa em relação aos objetivos propostos e abrange qualquer estágio do ciclo de vida de um programa, nesse caso incluindo avaliação *ex post*, avaliação intermediária e avaliação *ex ante*. Essa compreensão está de acordo com o entendimento que se pode obter do esquema da Figura 1, originado de Posavac e Carey (2003).

Pode-se complementar o conceito de avaliação de programas, com a noção de critério e de padrão. Posavac e Carey (2003) dedicam um capítulo em sua obra *Evaluation*, para esclarecer o estabelecimento de critérios e padrões na AP. Assim, avaliar envolveria a emissão de um julgamento, com base em valores, e também estabelecimento de critérios e padrões. Segundo Posavac e Carey, os critérios e padrões estabelecidos precisam ser claros e explícitos para que subsidiem avaliações úteis e precisam apresentar quatro características importantes: primeiro, representar os objetivos do programa; segundo, refletir os esforços institucionais que estejam de acordo com o propósito do programa, ou seja, são os critérios com influência institucional; terceiro, os critérios devem ser mensuráveis e confiáveis e quarto; critérios selecionados junto com os *stakeholders*. No que diz respeito ao terceiro ponto, supõe-se uma diferença entre a avaliação em si, como método de investigação e análise, e a avaliação de um programa.

Posavac e Carey (2003) ressaltam algumas tarefas fundamentais na avaliação de programas. Tarefas como: a) alocar recursos para identificar as necessidades não atendidas; b) verificar se o programa planejado realmente oferece os serviços, e se as suas atividades foram iniciadas; c) examinar os seus resultados, pois o programa conduziria a melhoria do nível de satisfação das necessidades identificadas; d) determinar quais os serviços que produzem os melhores resultados; e) selecionar os tipos de programas que oferecem os

serviços mais demandados; f) oferecer informações que sejam úteis para a melhoria da qualidade do programa e g) o caso de verificar a ocorrência dos efeitos não planejados refere-se à possibilidade de se resolver um problema criando-se outros efeitos não desejados.

Oskamp (1984) caracteriza a avaliação de programas como uma tentativa de avaliar a operação, o impacto e a efetividade de programas de organizações públicas ou privadas. Considera também que é uma disciplina, que desde 1970 possuía procedimentos e padrões amplamente conhecidos, conforme pode ser também verificado na Figura 1 que caracteriza a origem e o desenvolvimento da avaliação de programas.

Entre as muitas definições de avaliação de programa pertencentes a outras áreas do conhecimento, parece haver aspectos comuns. A definição de Ala-Harja e Sigurdur (2000) sintetiza alguns deles.

*A avaliação de programa pode ser definida como uma análise sistemática de aspectos importantes de um programa e seu valor, visando fornecer resultados confiáveis e utilizáveis. A avaliação de programa deve ser vista como mecanismo de melhoria no processo de tomada de decisão, a fim de garantir melhores informações aos governantes, sobre as quais eles possam fundamentar suas decisões e melhor prestar contas sobre as políticas e os programas públicos. Não se destina a resolver ou substituir juízos subjetivos envolvidos na tomada de decisão, mas permite um certo conhecimento dos resultados de um dado programa — informação que pode ser utilizada para melhorar a concepção de um programa e sua implantação (Ala-Harja & Sigurdur, 2000, p. 5).*

Observa-se na definição os aspectos citados referem-se à geração de dados e/ou informações, emissão de juízo de valor subjetivo a respeito de um programa ou política para subsidiar a tomada de decisão, permitindo o conhecimento sobre os resultados do programa (Posavac & Carey, 2003).

Ala-Harja e Sigurdur (2000) definem a avaliação de programa de forma abrangente: analisar organizações; funções; procedimentos; políticas e projetos. Compreendem a avaliação dos resultados de um programa em relação aos objetivos propostos, incluindo a avaliação *ex post*, *ex ante* e a avaliação intermediária. Conforme Ala-Harja e Sigurdur a avaliação contínua para o setor público foi eleita pela Comissão de Gestão Pública da *Organization for Economic Cooperation and Development* - OCDE como um indicador de bom funcionamento, daí a importância de avaliar utilizando informações obtidas, mesmo

diante da subjetividade que envolve o processo de emitir um juízo de valor, para garantir o êxito de um programa.

Aguilar e Ander-Egg (1995, p. 31-32), revisaram uma série de definições de avaliação de programas e também propuseram uma definição abrangente, compreendendo as definições de outros autores como: Stufflebeam e Shinkfield (1987); Fernández-Ballesteros, Vedung e Seyfried (1998); Cano, (2002); Posavac e Carey (2003); a saber,

*A avaliação é uma forma de pesquisa social aplicada, sistemática, planejada e dirigida; destinada a identificar, obter e proporcionar de maneira válida e confiável dados e informação suficiente e relevante 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 sirva 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 ao fracasso de seus resultados.*

De acordo com os autores a AP desenvolveu-se através da aplicação do método científico ao conhecimento da realidade e baseia-se nos passos e exigências do método científico. Ressalta-se ainda que para a coleta e sistematização dos dados para a AP é necessário que os procedimentos adotados cumpram os requisitos de confiabilidade e validade, para que seus resultados possam ser considerados válidos e utilizáveis (Aguilar & Ander-Egg, 1995). Por fim, a definição expressa o sentido de utilidade que a AP possuiria como uma prática conectada com a realidade e as necessidades de usuários, *stakeholders* e envolvidos com o programa, com finalidade última de melhorar a prestação do serviço.

Quanto à prestação de serviços, de acordo com Gray, Jenkins e Segsworth (1993, citado por Fernández-Ballesteros, Vedung & Seyfried, 1998), o controle dos gastos públicos e o gerenciamento de programas ou políticas de assistência, tem sido os dois principais enfoques para os quais se tem dirigido a avaliação de programas nas últimas três décadas. Portanto, haveria duas perspectivas na avaliação de programas; a primeira, visando à contribuição, planejamento e melhoria do programa e a segunda; a verificação de sua efetividade/impacto. Nesse sentido, além dos princípios de legalidade, regulamentação e gerenciamento financeiro das despesas públicas em auxílio de suas ações; as avaliações são instrumentos de controle das ações governamentais na esfera pública.

Conforme a OCDE, a avaliação de programas é uma atividade regulamentada por normas, servindo-se de alguns métodos usados pelos cientistas sociais em geral. Esses métodos são utilizados para coletar evidências confiáveis e válidas sobre os resultados das avaliações. Em suma, pode-se afirmar que a avaliação de programas compartilha o mesmo campo da metodologia da psicologia geral e da avaliação, enquanto disciplinas.

Nesse sentido, imprescindível não é distinguir as intervenções que sejam boas ou más, exitosas ou fracassadas e sim utilizá-las como um instrumento de apoio a um aprendizado contínuo, de busca de melhores decisões e amadurecimento da gestão dos programas sociais.

Rossi e Freeman (1993, p. 5) utilizam o termo pesquisa avaliativa de forma intercambiável com o termo avaliação e definem o conceito como “[...] Pesquisa avaliativa é uma aplicação sistemática dos procedimentos da pesquisa social para avaliar a conceituação, delineamento, implantação e utilidade de um programa de intervenção social”. Neste trabalho, adotou-se o termo pesquisa avaliativa de acordo com a definição desses autores, agregando à utilidade, o conhecimento do impacto dos resultados.

### ***2.3 Avaliação de Programas: abordagens***

Neste tópico explicitaram-se as abordagens em AP: Avaliação de Necessidades; Avaliação de Processos; Avaliação de Resultados e Avaliação da Eficiência. Tais modalidades de avaliação de programas são comumente empreendidas (Posavac & Carey, 2003).

De acordo com Posavac e Carey (2003), a Avaliação de Necessidades é realizada para identificar e mensurar o nível de necessidades existentes e não atendidas, sejam elas de uma organização ou de uma comunidade. Nesse sentido, acrescenta-se que é um tipo de avaliação *ex ante* e também uma avaliação de demanda. Uma definição de avaliação *ex ante* é dada pelo *Evaluation Research Society - ERS* (1998) Sociedade de Pesquisa Avaliativa. A análise de início-fim, pré-instalação, análise de viabilidade (*Front-end analysis, preinstallation, feasibility analysis*) – conhecida como análise de contexto. Inclui atividades de avaliação que antecedem a implantação de um programa. Possui a finalidade de confirmar, averiguar, ou emitir uma estimativa precisa da suficiência de concepção, viabilidade operacional, fontes de recurso financeiro e disponibilidade de outros tipos de apoio necessários como o organizacional. Os resultados deveriam fornecer uma direção útil

para refinar o planejamento do programa, determinando o nível apropriado de implantação, ou decidir sobre a instalação ou não do programa.

Há uma relação próxima entre a avaliação das necessidades e o planejamento da criação do programa. Neste caso, os avaliadores podem investigar diversos aspectos, por exemplo, o perfil socioeconômico da comunidade, o nível em que se situa o problema dentro da comunidade, além de tentar identificar agências governamentais e instituições que atendem à comunidade. Com o contato com a comunidade e seus líderes-membros, os avaliadores podem levantar quais aspectos do futuro programa poderá ser útil, e quais talvez sejam considerados inaceitáveis para a população a ser atendida. A avaliação de necessidade pode considerar aspectos desde concepção, até prováveis efeitos futuros do programa. Em suma, a avaliação de necessidades trata da avaliação e levantamento das demandas existentes, em determinada população, em relação a um dado problema existente.

A Avaliação de Processo fornece informações sobre a fase atual de funcionamento do programa, ou seja, de como o programa vem sendo implementado. Tal avaliação pode servir também como um processo de aprendizagem, no qual são observadas diversas questões, por exemplo: tipo da população atendida, nível de qualidade do funcionamento e pressupostos sob os quais foi planejado o atendimento das necessidades do público-alvo. Então pode-se afirmar que as informações obtidas geram aprendizagem sobre o programa configurando-se como uma avaliação formativa.

Uma Avaliação de Resultados assume diferentes níveis de complexidade. Os resultados podem se referir a serviços e produtos, entretanto neste estudo tratou-se da meta-avaliação de Programas que oferecem serviços públicos às comunidades. Para Posavac e Carey (2003) o nível mais elementar da avaliação de resultados refere-se a como o serviço é recebido. Nesse caso, o desafio é comparar as pessoas que receberam o serviço, com aquelas que não o receberam, verificando se realmente ocorreu melhoria na necessidade, a qual o programa visava atender. Embora os responsáveis pelos programas geralmente associem as melhorias observadas na população atendida, ao funcionamento do programas em si, tal melhoria pode ser atribuída a outros fatores, podendo ocorrer, inclusive, efeitos indesejados. Tais efeitos ou ameaças à validade interna e externa foram tratados neste trabalho no tópico de delineamentos de pesquisa.

A Avaliação da Eficiência trata de quando os avaliadores precisam demonstrar o quanto um programa ajudou aos participantes, quando defrontam-se com a questão dos

custos do programa. Assim, ponderam a questão da relação entre o custo e o benefício avaliando se resultados similares poderiam ser obtidos com menos recursos.

Os conceitos de eficiência, efetividade e eficácia, considerados em avaliação de programas, possuem diferentes significados. A **eficiência** de um programa estabelece a correlação entre os efeitos do programa e os esforços despendidos para obtê-lo. Envolve o montante de recursos, buscando medir a relação custo-benefício ou otimização *versus* desperdício na obtenção dos resultados.

A avaliação de **eficácia** não se restringe somente a verificar o alcance das metas propostas por um programa. Está associada ao alcance de objetivos gerais, específicos e metas do programa, estabelece em que medida os objetivos foram alcançados e as razões dos êxitos e dos fracassos, ou seja, relaciona as metas e objetivos propostos e as metas/objetivos alcançadas. A eficácia funcional, administrativa e contábil determina o grau de adequação dos meios aos objetivos do programa, envolve ainda a coerência entre os meios e os instrumentos utilizados.

Nesse sentido, podem-se caracterizar os tipos: a) avaliação de necessidades; e b) avaliação de processo como avaliações formativas e o tipo c) avaliação de resultados e d) avaliação da eficiência, como somativas.

Ala-Harja e Sigurdur (2000) consideram a avaliação como um mecanismo de *feedback*, entretanto, enfatizam a distinção entre avaliação e outros mecanismos de *feedback*. Compreendem mecanismos de *feedback* como: os estudos científicos; as auditorias tradicionais; o monitoramento; a mensuração de desempenho e a análise de políticas. Os autores afirmam que as avaliações podem ser classificadas em diversos tipos, e em diversas combinações, como é o caso da **avaliação somativa e formativa**. A terminologia – somativa e formativa – é utilizada por diversos autores, que enfatizam que geralmente há relação entre a abordagem e a técnica aplicada na avaliação. Pode-se afirmar que as diferenças entre avaliação somativa e avaliação formativa estão associadas ao processo ou ao propósito específico pelo qual a avaliação é empreendida, ou seja, a finalidade da avaliação.

É importante destacar que há um grande número de modelos e técnicas que podem ser utilizados pelos avaliadores e irão ser escolhidos de acordo com as especificidades do programa, os conhecimentos e as especialidades dos avaliadores.

De acordo Ala-Harja e Sigurdur (2000), a **avaliação somativa** é realizada quando o programa já foi implementado, ou seja, é uma intervenção em funcionamento, para atingir objetivos propostos. Por esse motivo ela também é chamada de *avaliação ex post*.

Avaliação *ex post* trata da avaliação de um programa implantado e em funcionamento, já apresentando condições de avaliabilidade.

As **Condições de avaliabilidade** (*Evaluability assessment*) de acordo com o *Evaluation Research Society* - ERS (1998) incluem as atividades empreendidas para avaliar se outros tipos de esforços de avaliação de programas (especialmente a avaliação de impacto) deveriam ser iniciados. O surgimento da avaliação da avaliabilidade ganhou destaque e legitimidade diante da crescente preocupação profissional com o custo de avaliação em relação aos seus benefícios, e também com a identificação das características gerais de programas (significância, abrangência, execução) que facilitam ou dificultam os esforços das avaliações formais. A avaliação da avaliabilidade pode abranger os questionamentos de viabilidade técnica (por exemplo, podem ser criados indicadores de desempenho válidos?), questões políticas (os diretores do programa entenderiam o tipo de informação produzida pela avaliação proposta?). Questões das agências de fomento (os interesses dos financiadores são apenas imediatos?). Há também as características do próprio programa (o programa foi instalado de fato?).

Ainda de acordo com o *Evaluation Research Society* – ERS (1998) avaliação somativa é conceituada como: **Avaliação somativa, de resultado, de efetividade** (*Impact evaluation*). Essa categoria de avaliação corresponde a uma das definições mais comuns de avaliação, e revela se o programa como um todo tem funcionado adequadamente. Os resultados da avaliação de impacto têm como objetivo, fornecer informação útil aos tomadores de decisão sobre a continuação de programa, expansão, ou reduções. O desafio para os avaliadores é estabelecer os indicadores de impacto apropriados, indicando os tipos de impacto que podem ser atribuídos ao programa e não a outras influências. Algum conhecimento ou estimativa das condições anteriores ao programa é necessário. As avaliações de impacto diferem no grau em que a busca por indicadores apropriados, ultrapassa os objetivos declarados ou expectativas dos formuladores, diretores, fundadores e outros patrocinadores do programa ou da avaliação. Porém, há consenso significativo que quanto mais independente o avaliador maior credibilidade os resultados da avaliação de impacto terão, desde que as expectativas das pessoas que o administram, controlam, ou influenciam encontrem-se refletidas na avaliação. Alcançar o equilíbrio entre os conflitos potenciais dos critérios constitui um desafio contínuo.

No que diz respeito a essa definição pode-se problematizar uma diferença entre resultado e impacto. O resultado de um programa pode envolver a geração de produtos e também o seu impacto. O produto é algo imediato, já o seu efeito em longo prazo pode ser

considerado impacto. Assim, nem todo resultado, do tipo produto, poderá eventualmente reverter-se em impacto de um programa.

A Avaliação somativa é conduzida geralmente para o estudo da eficácia ou um julgamento de valor de determinado programa, sendo utilizada para decidir sobre a alocação de recursos. Outra característica desse tipo de avaliação é que, tanto seus clientes, quanto as pessoas que a conduzem são externos à organização. São exemplos de clientes externos, políticos e cidadãos que recebem os serviços do programa.

A **avaliação formativa** é realizada quando da implantação de um programa e pode ser designada como um tipo de **avaliação intermediária**, como uma forma de obter conhecimento sobre o programa. O objetivo dessa avaliação é subsidiar o processo de gestão do programa, como uma forma de retroalimentação, em termos de sua implantação e desenvolvimento. Os avaliadores e os clientes do serviço são em geral internos, provavelmente os gerentes do programa. As questões geralmente abordadas referem-se às questões operacionais de monitoramento dos eventos e aspectos relacionados a impactos, entre outras (Ala-Harja & Sigurdur, 2000). Sua principal contribuição está na formulação do programa (Posavac & Carey, 2003).

De acordo com o *Evaluation Research Society – ERS* (1982, p. 10) avaliação formativa (*formative evaluation*) seria também denominada avaliação de processo, de um programa contínuo, para fazer modificações e melhorias. As atividades podem incluir análise das estratégias de gerenciamento e de interações entre as pessoas envolvidas no programa, avaliação dos recursos humanos, pesquisas de atitudes em relação ao programa e observação. Em alguns casos, a avaliação formativa significa pesquisa de campo em pequena escala, antes de sua implantação mais ampla. O avaliador formativo geralmente trabalha em conjunto com os formuladores e administradores do programa, e participa diretamente nas decisões para realizar as modificações necessárias.

Entre as modalidades de avaliação, há algumas que são confundidas freqüentemente com avaliação de programas. Em alguns casos, existem semelhanças entre alguns procedimentos científicos, como a Pesquisa Básica e a Pesquisa Avaliativa. Nesse caso é esclarecedora a distinção oferecida por Posavac e Carey (2003). Para esses autores há três atividades que são confundidas geralmente com a avaliação de programas; a saber: (1) pesquisa básica; (2) avaliação individual e (3) auditorias. Entende-se que a classificação dos autores foi elaborada considerando-se o objetivo e a finalidade da avaliação nos diferentes níveis.

No caso da **pesquisa básica**, o foco recai geralmente em questões de interesse teórico, sem necessariamente atentar para necessidades individuais ou organizacionais. Assim, os estudos em pesquisa básica são geralmente planejados para examinar questões teóricas sobre fenômenos específicos de uma área, buscando acumular conhecimentos que podem também levar futuramente a resultados “aplicados”. A discussão de pesquisa básica é comumente associada à discussão da pesquisa aplicada. No que diz respeito à pesquisa aplicada geralmente estão associadas implicações práticas imediatas, pois a pesquisa aplicada examina questões relativas a problemas práticos e suas possíveis soluções. De acordo com Patton (2001) a pesquisa básica contribui para o conhecimento fundamental e teórico e a pesquisa aplicada busca alcançar o que se refere às sociedades.

A avaliação de programas é uma área da pesquisa aplicada (Cozby, 2003). Na avaliação de programas, a geração de informação contribui para com a implantação de melhorias e exame dos resultados do programa pelo público, subsidiando a tomada de decisão sobre o programa.

A **avaliação individual** geralmente usa diferentes tipos de testes, como os de inteligência, atitude, personalidade e outros, com o objetivo de avaliar necessidades individuais de serviços ou medir qualificações para o trabalho ou fins de promoção. A diferença entre avaliação individual e avaliação de programas está no uso dos resultados: em avaliação de programas os resultados contribuem para uma comunidade atendida pelo próprio programa avaliado e não necessariamente para as pessoas individualmente.

Ainda de acordo com os mesmos autores, no caso das **auditorias**, estas avaliam aspectos legais e orçamentários ligados ao patrocínio de instituições governamentais, verificando se os programas estão transcorrendo de acordo com a legislação vigente. Na avaliação de programas as diferenças em termos de treinamento e de orientação entre avaliadores e auditores são estabelecidas em relação às variáveis examinadas, à ênfase em análise e aos tipos de recomendações oferecidas. Cook e Shadish (1986) trazem uma compreensão ampla do papel do avaliador e assinalam que as APs sociais produzem conhecimento que podem ser utilizados em curto prazo para tornar os programas mais rápidos na melhoria dos problemas sociais identificados. Para Cook e Shadish (1986), os avaliadores reconhecem que os programas sociais não respeitam os limites nos quais esses avaliadores foram formados ou treinados, pois os programas trazem lições práticas, que muitas vezes não são ensinadas na formação acadêmica dos avaliadores.

Por fim, merece destaque a questão do **monitoramento**. Posavac e Carey (2003) tratam do monitoramento e o concebem como um processo de *feedback* constante,

freqüente e necessário para manter a qualidade do programa. Destaca-se que pode haver feedback sem monitoramento. Entretanto, o feedback é o objetivo do monitoramento. A avaliação do tipo monitoramento, quando se trata de programas, projetos, sistemas e políticas públicas subsidiam a revisão dos procedimentos de tomada de decisão, gestão e aplicação de recursos em um governo, que estabelece programas e políticas mais eficientes e mais eficazes.

De acordo com o *Evaluation Research Society – ERS* (1998), **Monitoramento do programa** (*Program monitoring*) é a menos reconhecida, mas provavelmente a categoria de avaliação mais necessária, abandonando-se a noção de que o avaliador necessariamente entra, faz o trabalho, e depois vai embora. Inúmeros órgãos federais, estaduais e locais exigem que os programas sejam intensamente monitorados, mesmo os implementados há mais tempo. Esses programas podem ou não ter sido o assunto de análise de *ex-ante*, avaliação de processo, avaliação de impacto, e talvez até mesmo meta-avaliação. Os tipos de atividades envolvidas no monitoramento envolvem checagens periódicas de programas de políticas, em relação a como os serviços foram entregues aos clientes. Avaliações de monitoramento podem incluir propósito ou resultados, também achados dentro de outras categorias de avaliação. Podem também envolver reexame das necessidades para as quais o programa foi projetado, para verificar se ainda originalmente existem, ou podem sugerir modificação de sistema, atualizações e revitalização do programa.

Em síntese, quanto ao uso das avaliações de programas, pode-se considerar como referência o objetivo da avaliação desencadeada. Posavac e Carey (2003) asseguram que o objetivo geral das avaliações de programa tem sido contribuir para a garantia da qualidade dos serviços para as pessoas que dele necessitam. Dessa forma, pode-se investir em avaliações formativas, somativas ou de monitoramento. A avaliação formativa contribui para a formulação do programa; a somativa, para decidir se um programa será iniciado, continuado ou encerrado, e o monitoramento, fornece o *feedback* necessário para manter a qualidade do programa.

Em suma, com base nas definições anteriores (Rossi & Freeman 1993; Fernández-Ballesteros e cols.1998; Posavac & Carey, 2003), define-se avaliação de programas como **o processo de determinar o mérito, custo ou valor de programas sociais, políticas e regulamentos públicos, projetos e intervenções públicos ou privados, por meio de critérios e padrões e aplicação de procedimentos da pesquisa social as diferentes fases das intervenções sociais produzindo resultados sobre o programa e seus impactos.**

## 2.4 O planejamento da avaliação

Para planejar uma avaliação de programa, algumas considerações em relação à escolha da técnica, dos critérios e da abordagem utilizada. Há inúmeras técnicas na área de avaliação de programas.

De acordo com Krueger, Casey, Donner, Kirsch e Maack (2001) os analistas sociais defrontam-se com quatro desafios ao realizar a avaliação de programas. Esses desafios dizem respeito a: 1) levantar informações que possam fornecer entendimento sobre os grupos minoritários, a partir dos *stakeholders*, e identificar como o trabalho pode influenciar o contexto social; 2) verificar e identificar modelos, a partir dos dados, para entender a lógica subjacente que os resultados produzem, pois entendendo essa lógica o analista será capaz de aplicar os conjuntos de incentivos necessários para mudar os comportamentos. Isto no caso de programas que têm como foco a modificação do comportamento das pessoas como, programas ou campanhas publicitárias que visam induzir ao consumo de determinado produto; 3) traduzir para as agências de desenvolvimento os interesses expressos pelos atores, ou seja, os objetivos de desenvolvimento.

São consideradas três etapas para o planejamento. Primeiro, a **identificação** do programa, sua descrição e dos *stakeholders*. Segundo, a **caracterização das necessidades de informação**. Isto envolve questionar o que demanda a avaliação, que tipo de avaliação é apropriada, por que a avaliação é demandada e quais recursos estão disponíveis para dar suporte à avaliação. A terceira etapa é o **planejamento** em si. No planejamento é realizado o exame da literatura, é determinada a metodologia a ser empregada e é redigida e apresentada uma proposta escrita.

De acordo com Posavac e Carey (2003), pode-se planejar uma avaliação com foco em melhorias no programa. As melhorias devem ser efetuadas quando discrepâncias são observadas entre o que é observado e o que é planejado, projetado e a real necessidade do público-alvo. Os objetivos referem-se à necessidade da população-alvo; os serviços prestados precisam ser implementados pelo programa e, finalmente, os resultados projetados conduzir aos resultados obtidos.

Outra consideração importante diz respeito ao contexto do planejamento da avaliação. A demanda pela avaliação origina-se de diferentes entidades, por exemplo, uma iniciativa própria do pessoal da equipe do programa, ou ainda, uma demanda da

administração central ou agência de financiamento, ou a própria equipe interna de avaliação pode sugerir uma avaliação do programa.

A escolha de critérios para avaliação do programa, assim como a escolha da técnica é uma das tarefas mais difíceis do avaliador. Posavac e Carey (2003) argumentam que o estabelecimento de critérios é necessário para a avaliação, pois expressam a base valorativa sob a qual o programa será avaliado. Os critérios são elementos que geram controvérsia e disputa entre os atores envolvidos com o programa (*stakeholders*) e com a sua avaliação (avaliadores). Um fator a ser observado em sua escolha é que os critérios permitam sustentar avaliações úteis, ou seja, com resultados confiáveis. Dessa forma, refletem também os objetivos do programa e o seu sucesso.

Os critérios dependem, entre outras coisas, da etapa e do tipo do programa a ser avaliado. Posavac e Carey (2003) apresentam sete considerações gerais ou questionamentos a serem observados em relação ao planejamento da avaliação: 1) o programa ou o plano incorpora valores dos gestores; 2) incorpora as necessidades das pessoas atendidas; 3) o programa cumpriu o plano; 4) os resultados obtidos estão de acordo com as metas e especificam os níveis dos resultados esperados; 5) existe suporte de uma teoria para o programa; 6) o programa é aceito; e 7) os recursos estão sendo gastos apropriadamente pelo programa. O que os autores concluem em relação à escolha de critérios é que mesmo critérios “bem” definidos não garantem o sucesso da avaliação, porém, são imprescindíveis para se tentar realizar uma avaliação verdadeiramente útil.

Considera-se que o planejamento da avaliação é uma das etapas mais importantes e por isso, também, requer que se disponha de tempo suficiente e por isso pode parecer demorado. Tal tempo é necessário para esclarecer as dúvidas de quem contrata a avaliação e de quem a desenvolverá. Pois um planejamento em que se consideram esses critérios tem possibilidade de conduzir a um resultado que seja útil à organização.

Outro aspecto importante é a escolha da técnica de coleta de dados, envolvendo uma série de fatores, entre eles, o estágio do programa, o orçamento, o tempo disponível para a avaliação, a amplitude e a profundidade da análise. As técnicas podem ser qualitativas ou quantitativas, e dependendo da abordagem em que estão inseridas possuem vantagens e desvantagens. Essa visão é corroborada por Krueger, Casey, Donner, Kirsh e Maack (2001), que afirmam que quando a técnica é qualitativa tem como vantagem capturar informações de maneira detalhada, rica e contextual. Já a desvantagem seria a dificuldade em administrá-la de maneira consistente. No caso de técnicas quantitativas, podem ser

administradas de maneira clara e metodológica e serem, por outro lado, rígidas e de difícil controle.

O ponto de vista adotado por este trabalho, é que para justificar a escolha, por determinada técnica ou método, deve-se considerar prioritariamente fatores relativos ao programa, e não pautar a escolha somente no viés do pesquisador. Acredita-se que a adoção de apenas um dos métodos para avaliar é ineficiente e parcial.

De acordo com Cook e Shadish (1986) a avaliação com base nos objetivos do programa foi utilizada desde o início dos processos de avaliação, pois os objetivos do programa pareciam ser um bom critério para avaliá-lo. Mesmo com as críticas a respeito da fraqueza da forma com que algumas vezes são definidos os objetivos do programa, não se encontrou consenso na sugestão de outro método de construir critérios de avaliação. Não existem critérios de avaliação individuais.

Para suprir algumas deficiências no processo de avaliação com base nos objetivos do programa, Cook e Shadish (1986) sugerem que sejam agregados alguns passos como: 1) exame das reivindicações dos clientes, prestadores de serviço e gerentes do programa sobre o que o programa deverá alcançar e sobre como esses fatores conduzirão ao sucesso ou fracasso; 2) fatores que afetarão a tomada de decisão dos criadores de política ou gerentes no programa; 3) avaliar qualquer modelo de sistema que tenha sido desenvolvido para explicar como o programa produz *imputs* relacionados aos processos de resultados; 4) calcular o grau com que foram satisfeitas as necessidades materiais dos clientes; e 5) suprir as necessidades de informação com informações dos diferentes *stakeholders* ou tomadores de decisão da política.

Entretanto, nenhum desses critérios de avaliação merece ser julgado como melhor, pois não há razão para priorizar mais algum critério que outro, pois não existe método ou critério perfeito para qualquer avaliação. Embora os objetivos sejam rejeitados como único critério de mérito, nenhum outro critério ainda foi acordado a esse respeito.

### ***2.2.1 Avaliação de programas e auditorias: contrastes e interface***

Neste tópico abordaram-se as diferenças e semelhanças em relação a auditoria e a AP. Considera-se que auditorias são diferentes de avaliação em relação à sua origem histórica, método de trabalho, pressupostos teóricos e finalidades.

De forma semelhante à avaliação, auditoria é uma prática profissional desenvolvida por escritórios independentes, organizações governamentais, tribunais e organizações não governamentais. Conforme Chelimsky (1985), há uma tendência de aproximação entre as

práticas de auditoria e AP, por exemplo, o *General Accounting Office* – GAO, em Washington, realiza as duas atividades.

Para Woodside e Sakai (2001), há diferentes procedimentos de auditoria: auditoria de programa, operações de auditoria e gerenciamento de auditoria. Primeiro, a auditoria de programa enfoca como um conjunto de atividades alcança os objetivos. Pode-se realizar uma auditoria de programa ou combiná-la com outros procedimentos de auditoria. Uma operação de auditoria enfoca a eficiência e economia com que uma organização administra suas operações. No Havaí, o termo auditoria de gerenciamento é freqüentemente utilizado para referir à auditoria, que combina aspectos de avaliação de programa e operações de auditoria. Uma auditoria de gerenciamento analisa a efetividade do programa ou a eficiência da organização na implantação do programa, ou ambos.

De acordo com Woodside e Sakai (2001), os criadores das políticas precisam de evidências confiáveis sobre os programas, independente da terminologia adotada, se avaliação de programa ou auditoria de programa. Por isso, utilizam esses termos de forma intercambiável, pois a preocupação reside em obter respostas às suas perguntas urgentes sobre o desempenho do programas. Os autores também adotam o termo de forma intercambiável em seu artigo. Neste trabalho discorda-se da posição de Woodside e Sakai (2001) e adota-se a posição de Chelimsky (1985), que diferencia tais práticas. Para Chelimsky, auditoria e avaliação de programas possuem diferentes origens históricas; diferenças na análise dos objetivos do programa; enfatizam diferentes questões nas suas intervenções; e também possuem áreas de mútuo interesse e aspetos comuns.

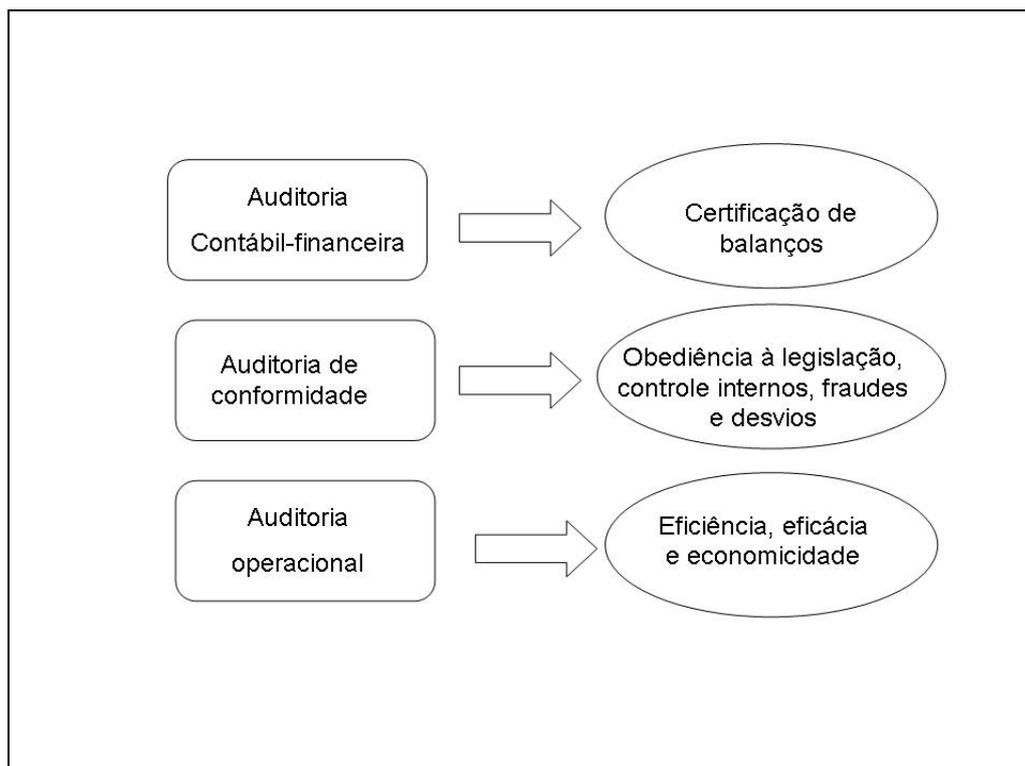
Conforme Chelimsky (1985), a auditoria se originou a partir do estabelecimento de procedimentos financeiros da área de contabilidade, e o primeiro registro de uma Associação de Contadores data de 1581, em Veneza. O início moderno da Contabilidade foi traçado na Inglaterra, sendo registrada a primeira Associação de Contadores em Edimburgo, em 1854. Nos Estados Unidos, o primeiro Curso de Graduação em Contabilidade foi iniciado em 1881 e o governo regulamentou a profissão em 1896. O desenvolvimento da contabilidade deu-se em face do desenvolvimento da legislação e da necessidade de conter os abusos das empresas na sonegação de impostos e falsificação de balanços financeiros. Historicamente a auditoria se desenvolveu como um procedimento de checagem, fortemente influenciado pelos procedimentos contábeis envolvendo o detalhamento e a verificação de informações. Atualmente, o escopo da auditoria ampliou-se. Além da análise financeira busca a compreensão de como os programas operam e atingem os seus objetivos. Assim, desenvolveu-se notavelmente durante os últimos 20 anos

o interesse no desempenho dos Programas Sociais, como resultado gradual de dois fluxos aparentemente discrepantes e independentes de investigação intelectual, a auditoria e a AP.

No início dos anos 1950, as auditorias buscavam racionalizar o gerenciamento e distribuição de recursos de Programas e Missões de Defesa do Governo, esforço que cresceu no Departamento de Defesa, Planejamento, Programação e Sistema de Orçamento (*Department of Defense's Planning, Programming, and Budgeting System – PPBS*). Entretanto, tal esforço era periférico e relacionado à verificação na aceção da contabilidade, cujo esforço principal era: 1) planejar o custo-efetividade do programa e então avaliar tal custo-efetividade; 2) checar se o custo-efetividade era resultante do planejamento e programação do processo. Apesar do enfoque restrito do modelo PPBS, também eram realizadas análises da área da Economia, como análise técnico-política, custo-benefício e custo-efetividade, na tentativa de compreensão das atividades dos programas. Entretanto, o enfoque principal das auditorias era planejar, de forma que as técnicas delineassem os efeitos prováveis e futuros dos programas, em vez de identificar os efeitos atuais da implantação dos programas ou políticas existentes (Chelimsky, 1985).

Destaca-se nesse sentido, a esclarecedora diferenciação realizada por Freitas (2005) em relação aos diferentes tipos de auditoria existentes no contexto público e que concerne diretamente à ANOP, objeto da meta-avaliação. De acordo com Freitas (2005) não há consenso quanto à definição de auditoria de natureza operacional, visto que a definição mais aceita no setor público, “associa auditoria operacional ao exame da economicidade, eficiência e eficácia de órgãos públicos ou programas governamentais” (Freitas, 2005, p. 1). O autor estabelece a seguinte diferenciação, auditoria de natureza operacional, auditoria contábil e financeira, auditoria de conformidade e auditoria operacional. 1) Auditoria contábil refere-se à certificação de balanços, 2) auditoria de conformidade refere à obediência à legislação, controle interno, fraudes e desvios e 3) auditoria operacional refere à eficácia, eficiência e economicidade. Essa tipologia foi ilustrada pelo autor da seguinte forma:

## **Figura 2. Tipologia de auditoria no setor público**



Fonte: Freitas (2005, p. 2)

Freitas (2005) descreve que historicamente a evolução da auditoria na área pública em diferentes países teve várias denominações, como auditoria de desempenho, auditoria pelo valor do dinheiro, auditoria administrativa, gerencial, operativa, auditoria de eficiência e efetividade e auditoria dos 3 Es (eficácia, eficiência e efetividade).

Conforme descrito no item 2.1 Contextualização Histórica da Avaliação de Programas, a AP teve forte desenvolvimento no início dos anos 1900 na área da pesquisa agrícola. Foram desenvolvidos delineamentos experimentais e técnicas de análise estatística para identificar quais práticas agrícolas produziriam os melhores rendimentos na colheita. Essas idéias foram desenvolvidas posteriormente por cientistas sociais, que contribuíram para esses avanços. Nos anos 1950 as avaliações eram executadas em grande escala, utilizando métodos da pesquisa social aplicada e softwares de análises estatísticas para a análise de dados (Rossi & Freeman, 1993). Entretanto, essa direção avaliativa foi dada somente aos programas sociais da área de educação, saúde pública, criminalidade e delinqüência e não aos Programas de Defesa.

De acordo com Chelimsky (1985), a auditoria e a AP analisam os Programas de forma diferente, entretanto, há sobreposições entre os dois processos. Chelimsky faz algumas comparações elucidativas sobre estes aspectos, resumidas a seguir:

Primeiro, em relação ao objetivo de auditorias e objetivos de avaliação de programas. Há algumas distinções entre auditoria e AP que surgem da análise de suas definições, objetivos e das perguntas elaboradas para o trabalho avaliativo.

- A auditoria é essencialmente normativa e analisa a discrepância entre um padrão e uma condição. Essa abordagem quando aplicada à análise de programas de Governo, resulta na verificação da atuação do programa em relação às normas existentes.
- A AP aplica os métodos sistemáticos da pesquisa para analisar o planejamento, implantação e efetividade do programa, fornecendo, sobretudo, informação empírica para a tomada de decisão dos gerentes dos programas na política de formulação; execução e responsabilização pelos programas.

Embora os objetivos da AP pareçam estar relacionados à checagem da correspondência entre “*o que é*” e “*o que deveria ser*”, a AP utiliza métodos de identificação empírica e mensuração dos resultados. Assim o enfoque recai sobre questões como: tamanho, frequência e direção, com achados empíricos e avaliação de impactos. Além disso, envolve a descrição e a caracterização sistemática da implantação do programa. Envolve também a estimativa de resultados que poderiam ser esperados na ausência do programa, bem como a comparação de mudanças observadas nos resultados, para estimar os efeitos atuais.

Em relação às questões investigadas por auditorias e APs:

- As APs investigam três tipos de questões: 1) questões descritivas; 2) perguntas normativas; e 3) estabelecimento de causa-e-efeito. APs buscam relacionar as atividades do programa aos efeitos encontrados, de forma que possam ser úteis ao programa.
- As auditorias usam questões normativas com o objetivo principal de responsabilização (*accountability*).

Ambos, AP e auditorias, analisam questões normativas. Questões descritivas são rotineiras na AP, mas são raras na auditoria. Porém, somente a AP busca questões para inferir relações de causa-e-efeito, pois os delineamentos utilizados pelos auditores são referenciados em critério e não permitem inferências causais dos resultados do programas.

A AP usa delineamentos diferentes para realizar inferências causais, tais como experimento, quasi-experimento e séries temporais interrompidas.

Embora a AP também analise questões normativas, adiciona questões descritivas e, especialmente questões de causa-e-efeito. Já o auditor assume que há um modo correto de

fazer as coisas conforme critérios estabelecidos, medindo e comparando as condições encontradas em relação aos critérios. O avaliador raramente utiliza critérios de referências, utiliza outros delineamentos e enfoca a relação entre as mudanças observadas e o programa, ao invés das mudanças que *deveriam* ser observadas. A avaliação é menos afetada que a auditoria, pela ausência de correspondência entre objetivos e resultados. Outra diferença é que a auditoria não produz estimativas do que poderia ter acontecido na ausência do programa que é avaliado, e a avaliação o faz.

Essas são as distinções fundamentais que influenciam os modos pelos quais as auditorias e avaliações são conceituadas, administradas, e avaliada sua qualidade.

O objetivo principal deste trabalho foi analisar o processo de realização da auditoria de natureza operacional desde sua concepção, métodos, técnicas, análises e recomendações oferecidas. Analisou-se o processo realizado pelo Tribunal de Contas da União, que tem como missão desenvolver avaliações e auditorias de programas públicos. Assim, foi empreendida a meta-avaliação detalhada no tópico a seguir.

### **2.2.2 Meta-avaliação, avaliação da avaliação ou auditoria de avaliação**

Das diversas terminologias usadas para expressar o conceito de meta-avaliação, Cook e Gruder (1978) afirmam que o termo avaliação da avaliação foi criado por Orata em 1940 e meta-avaliação, por Scriven em 1969.

De acordo com Patton (2001), uma meta-avaliação é uma avaliação de uma avaliação, ou seja, uma re-análise de um estudo avaliativo já concluído, considerando-se diversos aspectos do estudo anterior, seja metodologia, seleção dos sujeitos, objetivos, critérios adotados, resultados e análise.

A definição de Schwandt (1989) corrobora a definição de Patton (2001), uma vez que compreende meta-avaliação como um método para checar a qualidade de uma avaliação. Para tal, exige o exame do método da avaliação e dos seus procedimentos para chegar aos resultados e conclusões. Já para Woodside e Sakay (2001), a meta-avaliação inclui a avaliação da utilidade e validade de dois ou mais estudos que enfocam a mesma questão.

Considerando-se que a própria meta-avaliação pode ser objeto de outra meta-avaliação, há níveis de exigência para fazê-lo. Esses níveis de exigência referem-se aos procedimentos adotados e merecimento do avaliador. Conforme Schwandt (1989), a meta-avaliação exige um conjunto de procedimentos, padrões e critérios para julgamento da qualidade da avaliação.

A partir dos entendimentos desses autores (Patton, 2001; Schawandt, 1989; Woodside & Sakai, 2001) define-se meta-avaliação como **um método de pesquisa a partir do qual são re-analisadas uma ou mais etapas do(s) estudo(s) avaliativo(s) já concluído(s); compara-se a avaliação anterior com padrões de qualidade e validade aceitos na comunidade científica e ao final emite-se nova avaliação sobre o estudo avaliativo analisado.**

Na literatura específica sobre meta-avaliação e na busca de um modelo de como realizá-la, encontraram-se diferentes procedimentos de análise de dados para subsidiar seus resultados e conclusões. Por exemplo, Woodside e Sakai (2001) utilizaram como procedimento de análise na meta-avaliação, modelos teóricos específicos da área do programa. O estudo avaliativo meta-avaliado pelos autores foi um Programa de Marketing e Turismo Federal, assim utilizaram o modelo Kotler (citado em Woodside & Sakai, 2001) para checar os pressupostos do que seria esperado num programa de turismo. Para avaliar o planejamento do programa compararam o procedimento empregado na avaliação anterior com as premissas do que poderia ser planejado através de uma análise SWOT. Para avaliar a implantação do Programa, traçaram comparações do procedimento adotado na avaliação anterior, com os conceitos de Mintzberg (citado em Woodside & Sakai, 2001), sobre estratégia planejada e estratégia deliberada. Por fim, para avaliar os resultados anteriores da AP analisaram os indicadores de impacto do Programa de Marketing e Turismo Federal e, ao final do artigo, propuseram um modelo para futuras avaliações de programas dessa natureza.

Na literatura revisada (Cook & Gruder, 1978; Chelimsky, 1985; Stufflebeam, Shinkfield, 1987; Schawandt, 1989; Woodside & Sakai, 2001; Patton, 2001; Ashworth, Cebulla, Greenberg & Walker, 2004), encontraram-se diferentes procedimentos para a condução de meta-avaliações, assim como diferentes concepções acerca do processo. Nesse sentido, não se encontrou resposta direta à questão das etapas e técnicas necessárias à realização de um estudo meta-avaliativo.

Por exemplo, alguns autores confundem o método da meta-avaliação e o procedimento de análise meta-análise. No caso, Ashworth, Cebulla, Greenberg e Walker (2004) realizaram uma meta-avaliação utilizando o procedimento de meta-análise. Os autores defendem o ponto de vista que fundamentar a meta-avaliação no procedimento de meta-análise, favorece o poder explicativo de replicação, acumulação rigorosa de evidência, revisão e síntese. Além disso, criticam autores como Patton (2001), que caracterizam a meta-avaliação como um método qualitativo.

Em relação a críticas dos autores à abordagem qualitativa, acredita-se ser um ponto de vista já superado na literatura científica (Patton, 2001; Günther, 2006). O método de pesquisa deve ser escolhido considerando-se, entre outros fatores, as características do fenômeno abordado, e sempre que possível empregar ambas abordagens para produzir uma compreensão ampla e profunda dos dados e realidade em questão.

Entende-se que a adoção do procedimento de meta-análise para compor uma meta-avaliação só poderia ser considerada como um procedimento para realizar uma meta-avaliação se, ao final do mesmo, fosse esboçada uma nova avaliação sobre a AP analisada. Além disso, seria necessário agregar ao procedimento os critérios da comunidade científica de avaliação de programas. Devido a essas divergências conceituais e alguma confusão ainda existentes no campo de estudos, não se considera a adoção exclusiva desse procedimento de análise de dados como a melhor forma de realizar uma meta-avaliação e por isso ele não foi incluído no modelo que se apresenta a seguir. Conforme se argumentou, é possível o uso da meta-análise para proceder a uma meta-avaliação, desde que somado a outros procedimentos qualitativos validados pelas associações de avaliação de programa e produzindo-se ao final uma nova avaliação.

Apesar dessa controvérsia, o conjunto de pesquisas analisadas subsidiou a proposição de um modelo para caracterizar as etapas e procedimentos de uma meta-avaliação. Diante do exposto, apresentam-se seis proposições e a seguir o modelo:

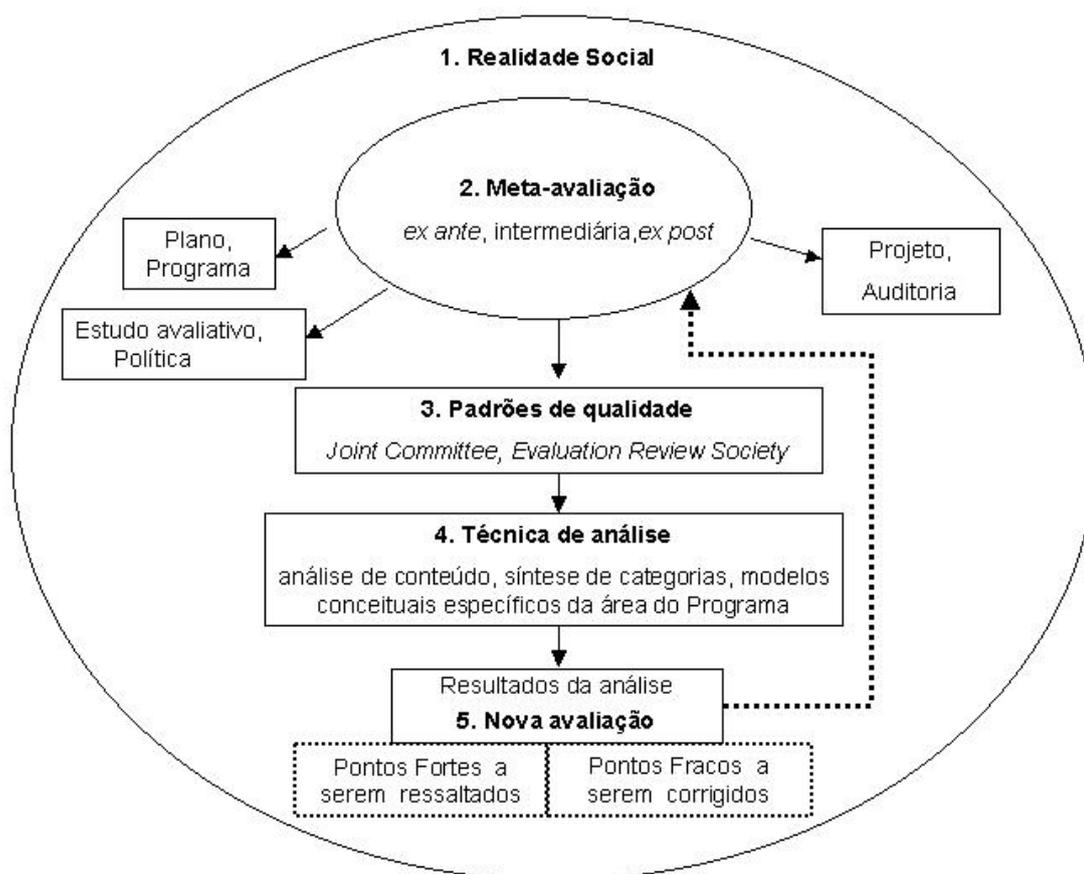
- 1) As meta-avaliações recaem sobre estudos avaliativos existentes dentro do contexto da realidade social. A meta-avaliação existe dentro desse contexto que pode influenciar a sua realização.
- 2) Estudos meta-avaliativos contemplam estudos anteriores, em quaisquer fases dos programas, sejam *ex-ante*, intermediários e *ex-post* e incidem sob qualquer estudo que tenha um desenho avaliativo (avaliação, política, plano, programa projeto, auditoria).
- 3) A meta-avaliação depende de um conjunto de critérios de qualidade que a tornem válida. Tais critérios de julgamento de valor são compartilhados pela comunidade internacional de avaliadores, pois são necessários para balizar a avaliação de uma outra avaliação. Algumas associações que fornecem esses critérios são o *Joint Committee* e a *Evaluation Research Society*.
- 4) A meta-avaliação pode comportar um conjunto de técnicas e procedimentos de análise para se chegar ao seu resultado final – a emissão de um novo juízo valorativo, uma nova avaliação. Essas técnicas podem ser análise de

conteúdo, síntese de categorias ou ainda modelos conceituais da área específica do programa.

- 5) Os resultados das análises de dados subsidiam ao final do processo o julgamento do estudo avaliativo, comparando-os com os critérios anteriormente estabelecidos e com as demais análises realizadas, permitindo a emissão de uma nova avaliação. A nova avaliação apresentará, em relação ao estudo avaliativo analisado, os pontos fortes a serem valorizados e os pontos fracos a serem corrigidos.

Considerando-se essas proposições, apresenta-se a seguir o modelo gráfico das mesmas.

**Figura 3. Modelo de meta-avaliação**



Para fins da presente meta-avaliação, considerou-se a auditoria de natureza operacional como um estudo avaliativo com natureza e objetivo próprio que a diferencia de uma AP, conforme explicado anteriormente, e suas características permitem, portanto, que seja objeto da meta-avaliação. No presente estudo, os padrões de qualidade e validade da meta-avaliação foram aqueles conhecidos e difundidos na comunidade científica de

avaliação, por associações idôneas como *Joint Committee* ou *Evaluation Research Society* - *ERS*, para atestar a qualidade global da avaliação e emitir uma nova avaliação em relação à anteriormente realizada.

De acordo com Schwandt (1989), o conceito de auditoria de avaliação ou meta-avaliação, modelado conceitualmente com uma auditoria de fiscalização, foi introduzido para estabelecer a validade das pesquisas naturalísticas (qualitativas) por Guba e Lincoln (citado em Schwandt, 1989).

A meta-avaliação pode ser formativa ou somativa. A meta-avaliação somativa é uma atividade retrospectiva desenvolvida por um agente independente e externo sobre o processo e produto da avaliação comparando-o a um conjunto de padrões para avaliação. Nesse sentido, no presente trabalho utilizou-se os padrões do *Joint Committee*. Realizar uma meta-avaliação envolve um posicionamento político por parte do avaliador, um comportamento extremamente prudente, em relação aos procedimentos de meta-avaliação. Os meta-avaliadores não analisam os dados coletados e sim as inferências que outros avaliadores fizeram acerca deles, e assim, posteriormente também emitem uma avaliação embasados em inferências pessoais sobre os resultados de outros.

A meta-avaliação pode possuir diversas formas e variar de críticas profissionais aos relatórios de avaliações, a procedimentos de re-análise de dados originais (às vezes com hipóteses diferentes), até a coleta de novas informações. No caso de programas que geram amplo interesse público, as meta-avaliações analisam os resultados de diferentes avaliações sobre os mesmos (inclusive avaliações de unidades ou componentes do programa) no sentido de verificar o impacto global. Esse tipo de avaliação num movimento relativamente novo em termos de estudos de avaliação chama a atenção para o fato de que os resultados da avaliação não estão necessariamente relacionados à falhas na avaliação.

Por fim, Schwandt (1989) aponta implicações políticas em dimensões comuns para o papel dos meta-avaliadores: o comportamento político do avaliador gera moeda política dos achados pelo uso do seu método. Outro aspecto mais profundo é que o ato de verificar a integridade ou qualidade de uma avaliação é um ato político, ou seja, é um modo de endossar uma política particular de significado com referência para uma comunidade particular de pesquisadores.

A meta-avaliação é estimulada por vários interesses como investigação acadêmica, exigências das agências que coordenam ou supervisionam o programa, sendo que não há necessidade do avaliado aceitar o resultado da avaliação original. Pode ser realizada também por interesse nos efeitos posteriores da avaliação de programa.

As associações internacionais oferecem uma padronização de procedimentos para a realização de APs; são conceitos e técnicas utilizadas pelos avaliadores, exemplo, o *Evaluation Research Society - ERS* (1998). Para o *Evaluation Research Society - ERS* (1998) a meta-avaliação é também conhecida como secundária, auditoria de avaliação e avaliação da avaliação (*Evaluation of evaluation*). Tais termos diferentes referem-se às atividades freqüentemente aplicadas à avaliação de impacto.

Conforme Worthen, Sanders, Fitzpatrick (2004), em termos históricos a meta-avaliação teve início em 1960 quando avaliadores como Scriven, Stake e Stufflebeam começaram a discutir os procedimentos e critérios formais de meta-avaliação. De acordo com Woodside e Sakay (2001), foi Scriven que criou o termo meta-avaliação.

Os avaliadores dessa época criaram listas para checagem do que poderia se constituir uma avaliação “boa” ou “ruim”. Tais listas de diretrizes para avaliação começaram a ser difundidas informalmente entre os avaliadores e também houve algumas publicações. Assim, foi lançado no final da década de 1970, um projeto que visava desenvolver um conjunto de diretrizes aplicadas às avaliações educacionais como consenso geral em termos da qualidade da avaliação. A formulação dessas diretrizes começou em 1975, coordenada por Stufflebeam, do *Western Michigan University's Evaluation Center*. “A orientação e a autorização foram dadas pelo *Joint Committee on Standards for Educational Evaluation*, conhecido desde então como *Joint Committee*” (Worthen, Sanders, Fitzpatrick, 2004, p. 595). Como resultado desse trabalho publicou-se o *Standards for evaluations of educational programs, projects, and materials*, em 1981.

As Diretrizes do *Joint Committee* abrangem 30 tópicos e apresentam definições, fundamentos lógicos, diretrizes, erros comuns, casos ilustrativos e descrição de práticas de avaliação, por meio de exemplos de casos. De acordo com Stufflebeam e Shinkfield (1987), as normas para a avaliação de programas são: utilidade, viabilidade, propriedade e precisão. Para conhecimento das normas e aplicação na análise dos dados desta tese, foram descritas a seguir, de acordo com o *Joint Committee on Standards for Educational Evaluation*, citado em Stufflebeam e Shinkfield (1987) o resumo das normas.

**Tabela 1. Resumo das normas para avaliação de programas**

<b>A</b>	<b><i>Normas de utilidade</i></b>
	As normas de utilidade foram concebidas para assegurar que a avaliação proporcione informações práticas para os usuários potenciais

<i>A1</i>	<i>Descrição dos potenciais usuários</i>
	As pessoas envolvidas ou afetadas pela avaliação devem ser identificadas para que suas necessidades sejam satisfeitas
<i>A2</i>	<i>Credibilidade do avaliador</i>
	As pessoas que realizam a avaliação devem ser fidedignas e competentes para realizar a avaliação e para que seus resultados tenham o máximo de credibilidade e aceitação
<i>A3</i>	<i>Alcance e seleção da informação</i>
	As informações coletadas devem possuir alcance suficiente, e ter sido selecionadas de forma que incluam questões sobre o objeto da avaliação e responda às necessidades e aos interesses dos clientes e outros envolvidos
<i>A4</i>	<i>Interpretação de valores</i>
	As perspectivas, procedimentos e bases lógicas utilizadas para interpretar os resultados, devem ser cuidadosamente descritas, para que as bases de juízo de valor sejam claras
<i>A5</i>	<i>Clareza do relatório</i>
	Os relatórios de avaliação devem descrever claramente o programa que está sendo avaliado, seu contexto, finalidades, procedimentos e resultados da avaliação, com a finalidade de que os interessados compreendam, o que se fez, como se fez, que tipo de informação se obteve e quais conclusões e recomendações foram obtidas
<i>A6</i>	<i>Disseminação do relatório</i>
	Os resultados da avaliação devem ser divulgados aos clientes e ao público interessado para que possam utilizar a tempo as suas conclusões
<i>A7</i>	<i>Oportunidade do relatório</i>
	A disseminação dos resultados deve ser oportuna, com a finalidade de que o público possa fazer o melhor uso possível da informação coletada
<i>A8</i>	<i>Impacto da avaliação</i>
	A avaliação deve ser planejada e realizada de forma que estimule o seu acompanhamento pelos interessados e com isto tenham mais probabilidade de usá-la
<b>B</b>	<b><i>Normas de viabilidade</i></b>
	As normas de viabilidade foram concebidas para assegurar que a avaliação seja realista, prudente, diplomática e moderada
<i>B1</i>	<i>Procedimentos práticos</i>
	Os procedimentos avaliativos devem ser práticos e interferirem o mínimo possível no programa durante a coleta de informações
<i>B2</i>	<i>Viabilidade política</i>
	A avaliação deve ser planejada e realizada conhecendo as diferentes posturas dos grupos de interesse, de forma a obter a sua colaboração fazendo com que todas as tentativas de restringir os trabalhos da avaliação, ou de refratar ou usar mal os resultados possam ser contornadas ou neutralizadas
<i>B3</i>	<i>Custo-efetividade</i>
	A avaliação deve ser eficiente e produzir informação de valor suficiente para justificar os recursos gastos
<b>C</b>	<b><i>Normas sobre ética</i></b>
	As normas sobre ética foram concebidas para assegurar que a avaliação seja realizada de forma legal e ética, com respeito ao bem-estar dos envolvidos na avaliação ou afetados pela avaliação
<i>C1</i>	<i>Acordo formal</i>
	As obrigações das partes formais da avaliação (o que deve ser feito, como, por que e quando) devem ser assumidas de comum acordo e descritas num documento escrito, com a finalidade de que as partes estejam obrigadas a cumprir todas as condições do compromisso ou renegociá-las formalmente
<i>C2</i>	<i>Conflito de interesses</i>
	O conflito de interesses, freqüentemente inevitável, deve ser enfrentado aberta e honestamente para não interferir no processo de avaliação e seus resultados
<i>C3</i>	<i>Informações completas e claras</i>
	As comunicações orais ou escritas devem ser francas, diretas e honestas na revelação das conclusões, incluindo as limitações da avaliação

(cont.) Tabela 1 – Resumo das normas para avaliação de programas

<i>C4</i>	<i>Direito do público a conhecer</i>
	As partes formais da avaliação devem respeitar e assegurar o direito do público, dentro dos limites

	dos princípios e estatutos, como por exemplo, a segurança do público
<i>C5</i>	<i>Direitos humanos</i>
	As avaliações devem ser planejadas e conduzidas de forma a respeitar e proteger o bem-estar dos envolvidos
<i>C6</i>	<i>Relações humanas</i>
	Os avaliadores devem respeitar a dignidade e a integridade humana no trato com as outras pessoas relacionadas à avaliação
<i>C7</i>	<i>Informação equilibrada</i>
	A avaliação deve ser completa e justa na apresentação das virtudes e defeitos do objeto da investigação (programa, projeto), para que as virtudes possam ser melhoradas e os problemas solucionados
<i>C8</i>	<i>Responsabilidade fiscal</i>
	A alocação e o gasto dos recursos devem refletir sólidos procedimentos contábeis, prudência, ética e responsabilidade para que as despesas sejam justificadas e apropriadas
<b>D</b>	<b><i>Normas de precisão</i></b>
	As normas de precisão foram concebidas para assegurar que a avaliação revele e transmita uma informação tecnicamente exata sobre as características do programa que está sendo avaliado, para poder determinar seu mérito e valor
<i>D1</i>	<i>Identificação do objeto</i>
	O objeto da avaliação (programa, projeto ou material) deve ser examinado em profundidade de forma que a sua estrutura possa ser claramente identificada
<i>D2</i>	<i>Análise de contexto</i>
	O contexto envolvido no programa, projeto ou material deve ser examinado de forma suficiente detalhada para que prováveis influências que houver sobre o programa sejam identificadas
<i>D3</i>	<i>Descrição das finalidades e procedimentos</i>
	As informações e finalidades dos procedimentos de avaliação devem ser concebidas e descritas com todos os detalhes, de forma que possam ser identificados e aferidos
<i>D4</i>	<i>Fontes de informação confiáveis</i>
	As fontes de informação devem ser descritas com todo detalhe, para que se possa avaliar se as informações são suficientes
<i>D5</i>	<i>Determinar a validade</i>
	Os instrumentos e procedimentos para a coleta de informações devem ser escolhidos ou criados, e utilizados de forma que assegurem uma interpretação final válida para seu uso
<i>D6</i>	<i>Determinar a fidedignidade</i>
	Os instrumentos e procedimentos de coleta de dados devem ser escolhidos ou criados, e implementados de tal forma que garantam informações suficientemente fidedignas para o uso que se pretende fazer delas
<i>D7</i>	<i>Controle sistemático dos dados</i>
	As informações coletadas, processadas e apresentadas numa avaliação devem ser sistematicamente revistas e todos os erros encontrados e corrigidos
<i>D8</i>	<i>Análise de informações quantitativas</i>
	A informação de uma avaliação deve ser adequada e sistematicamente analisada para que as interpretações sejam fidedignas e as perguntas sejam efetivamente respondidas
<i>D9</i>	<i>Análise da informação qualitativa</i>
	As informações qualitativas de uma avaliação devem ser analisadas de forma apropriada e sistemática para que as interpretações sejam fidedignas
<i>D10</i>	<i>Conclusões justificadas</i>
	As conclusões da avaliação devem ser explicitamente justificadas, para que os interessados possam ter acesso a elas

(cont.) Tabela 1 – Resumo das normas para avaliação de programas

<i>D11</i>	<i>Objetividade do relatório</i>
	Os procedimentos das avaliações devem proporcionar as garantias suficientes para proteger os

	resultados e os informes da avaliação contra distorções, através dos sentimentos e inclinações pessoais de quaisquer das partes formais possuam
<i>D12<sup>1</sup></i>	<i>Meta-avaliação</i>
	A própria avaliação deve ser avaliada de maneira formal e somativa de acordo com essas e outras diretrizes pertinentes para que a sua realização seja orientada e para que, depois do término, os interessados possam examinar de perto seus pontos fortes e fracos

Fonte: Stufflebeam e Shinkfield (1987, p. 27-31)

Conforme Stufflebeam e Shinkfield (1987) **normas de utilidade** são aquelas dirigidas a pessoas e grupos que estão relacionados com a tarefa de avaliar, ou seja, os responsáveis diretos pela avaliação. Assim as normas de utilidade devem ajudar a identificar o “bom” e o “mau” funcionamento do objeto avaliado, proporcionando informações claras e de forma oportuna, deve fornecer informações sobre as virtudes e defeitos e sobre a forma de melhorá-los.

As **normas de viabilidade** referem-se ao emprego de procedimentos avaliativos que possam ser utilizados sem gerar maiores problemas. Deve levar em consideração e aplicar os controles possíveis sobre as forças políticas que possam de alguma forma interferir na avaliação.

As normas sobre **ética na avaliação** dizem respeito a compromissos explícitos que assegurem a cooperação necessária, a proteção e os direitos dos participantes envolvidos na avaliação e a honestidade dos resultados.

Por fim as **normas de precisão** representam aquelas que descrevem o objeto avaliado com clareza na sua evolução e seu contexto, deve revelar as virtudes e os defeitos do planejamento da avaliação dos procedimentos e das conclusões.

Silva (2002b, p. 14) destaca que há necessidade de que os governos empreendam “rotinas de meta-avaliação sobre casos concretos de sistemas nacionais ou subnacionais de avaliação”.

Atendendo-se a essa recomendação, realizou-se neste trabalho a meta-avaliação dos procedimentos de avaliação das auditorias de natureza operacional utilizados pelo TCU em

relação a programas sociais. Adotaram-se os critérios do *Joint Committee* para realizar a meta-avaliação das ANOPs de dez programas de governo.

A meta-avaliação foi utilizada neste trabalho com o intuito de aperfeiçoar o modelo avaliativo existente no TCU.

De acordo com os Woodside e Sakay (2001) as meta-avaliações possuem três características principais: 1) são sínteses de achados e inferências da pesquisa avaliativa sobre o desempenho do programa ou ambos, relatam a efetividade do gerenciamento das metas alcançadas nos programas. As meta-avaliações aumentam o conhecimento e fornecem informações sobre dois extremos: programas bem administrados ou programas mal administrados; 2) informam sobre a validade e utilidade dos métodos de avaliação e oferecem direção sobre métodos úteis para avaliar e 3) fornecem forte evidência sobre o impacto do programa, subsidiando a tomada de decisão sobre o programa. Conseqüentemente, os resultados da meta-avaliação auxiliam e justificam o aumento da confiança dos interessados e gerentes dos programas nos resultados das avaliações.

Um exemplo de meta-avaliação aparece no artigo de Weber (2004), ao fazer considerações sobre a teoria do planejamento e monitoramento como subsídios para desenvolvimento de competências. Assim, as pressuposições básicas do planejamento e monitoramento servem de ponto de análise do conceito de avaliação, fornecendo um padrão para a meta-avaliação dos padrões de avaliação de desempenho. A avaliação tornou-se objeto de avaliação com a ajuda desses critérios. Pode-se deduzir a partir desse exemplo que a meta-avaliação seria parte de um diálogo aberto entre várias partes do processo de avaliação.

Outros exemplos de meta-avaliação podem ser encontrados na área da avaliação educacional, quando o foco é a própria avaliação, quer seja do aluno, professor ou avaliação enquanto prova aplicada aos alunos.

Assim, adotou-se o conceito de que o **objeto de avaliação** é a própria avaliação (**meta-avaliação**). Conforme Schawandt (1989), há diferentes padrões de conduta e qualidade para os auditores na área de auditoria governamental, financeira e de desempenho. Tais padrões, conhecidos como padrões gerais de auditoria (*Generally Accepted Auditing Standards* - GAAS) significam guiar a conduta do próprio auditor e fornecer diretrizes para a auditoria. Os GAAS são desenvolvidos e disseminados por

---

<sup>1</sup> Este item não consta em Stufflebeam e Shinkfield (1987) e foi obtido em Worthen, Sanders e Fitzpatrick (2004, p. 600) de acordo com os autores a meta-avaliação foi acrescentada às normas pelo *Joint Committee* em 1994.

associações de auditoria profissionais e renomadas como o Instituto Americano de Certificação de Auditoria Pública e Código de Ética.

Ainda conforme Schawandt (1989), na área de avaliação de programas governamentais, não há um único conjunto de padrões para a conduta do auditor como meta-avaliador.

Por fim, caracteriza-se a meta-avaliação como método de pesquisa. A seguir, foi feita a apresentação do procedimento de meta-análise e a sua contribuição para a AP.

### ***2.2.3 Meta-análise: diferenças e semelhanças em relação à meta-avaliação***

Neste tópico, descreveu-se a meta-análise, procedimento de análise de dados, que apresenta algumas semelhanças em relação ao método de pesquisa meta-avaliação e por isso confundido por alguns autores com meta-avaliação. A meta-análise pode ser empregada para diversos tipos de estudos, e torna-se útil à medida que em algumas áreas específicas de AP, a literatura científica está pulverizada em artigos, livros e não há sistematização, apenas produção de conhecimento dispersa em várias publicações. Com a ausência de sistematização desse conhecimento, torna-se necessário empreender novas pesquisas avaliativas acerca dos programas.

De acordo com Rossi e Freeman (1993), em áreas onde há grande quantidade de estudo sobre determinado objeto é possível realizar revisões quantitativas da literatura, tais estudos são conhecidos como meta-análise. Assim, considerando que em alguns casos estudos individualizados por si só não asseguram a validade e generalização dos resultados, o procedimento de meta-análise pode assegurá-los por meio da comparação dos diversos resultados, de acordo com o recorte escolhido.

Smith e Bond (1999, p. 15) apresentam uma definição clara e completa de meta-análise. De acordo com os autores,

*Meta-análise é uma técnica estatística muito usada por psicólogos na elaboração de resumos com conclusões gerais, acerca de diversos estudos que investigam questões semelhantes de pesquisas. Uma dificuldade para se chegar a resumos válidos é o uso de diferentes técnicas de medição e amostras de respondentes, que os pesquisadores geralmente usam, coletando seus dados sob circunstâncias diferentes. A meta-análise se apóia no pressuposto de que pode calcular o “tamanho do efeito” independente da natureza da medida real usada por determinado pesquisador. O tamanho do efeito de um estudo é calculado como a diferença entre os escores obtidos pelos sujeitos*

*experimentais e o grupo de controle, dividido pelo desvio padrão dos escores dos sujeitos do grupo de controle. O tamanho do efeito fornece a média através de diferentes estudos, determinando se ocorreu ou não o efeito experimental que está sendo investigado e se é consistente entre os achados. Se a amostra dos estudos for grande o suficiente, a influência da variação no delineamento experimental, localização geográfica, dados do estudo, sobre o tamanho do efeito, podem ser estimados [...]*<sup>2</sup>

Conforme se observa na definição a meta-análise utiliza procedimentos estatísticos para calcular e estimar o efeito de múltiplos estudos, sintetizando seus achados e agrupando os resultados de muitas avaliações de impacto. Embora no sentido exato meta-análise não seja um delineamento de pesquisa, é uma alternativa para projetar uma avaliação que pode ser útil em algumas situações. Por exemplo, quando se necessita de maior tempo para coletar dados originais. Os achados da meta-análise são particularmente úteis na fase de delineamento de um programa, pois resumem o conhecimento existente sobre programas semelhantes já implantados, fornecendo conhecimento e aprendizado para o novo programa.

Rossi e Freeman (op cit) sugerem a verificação de dez passos para realizar uma meta-análise;

1. Qual é o objetivo específico da revisão? Lembrar que as revisões com objetivos científicos enfatizam aspectos diferentes daqueles que as políticas públicas buscam.
2. Como os estudos foram selecionados? Verificar se a literatura em determinado campo foi representada corretamente na revisão.
3. Há viés na seleção das publicações?
4. Os tratamentos são semelhantes o suficiente para serem combinados? Não combinar achados que se referem às intervenções radicalmente diferentes para alcançar alguma conclusão global.
5. Os grupos de controle são semelhantes o bastante para serem combinados?
6. O que representa a distribuição dos resultados do estudo?
7. Os resultados estão relacionados ao delineamento da pesquisa?
8. Os resultados estão relacionados às características do programa, dos participantes e das situações?
9. A unidade de análise é semelhante entre os estudos?

---

<sup>2</sup> Tradução da autora.

## 10. Quais são as diretrizes para as pesquisas futuras?

Um exemplo de meta-análise fornecido por Rossi e Freeman (1993) é a do estudo pioneiro realizado por Smith, Glass e Miller, (1980, citado em Rossi & Freeman, 1993), que resumiram quantitativamente 375 resultados de estudos sobre o efeito da psicoterapia, e ao final do estudo, elaboraram conclusões acerca da eficiência dos tratamentos psicoterápicos. Entretanto, não confirmaram a superioridade de uma terapia em relação à outra. Nesse exemplo, embora fossem estudos diferentes, testando abordagens diferentes de psicologia, foi escolhida uma variável comum a todos os estudos - o tratamento psicológico. Foi Glass que denominou o método empregado nesse estudo de meta-análise. Assim, entende-se que a meta-análise mencionada como exemplo por Rossi e Freeman (1993), foi realizada através da análise de diversos estudos, de diferentes abordagens psicológicas, entretanto possuíam uma característica comum: o tratamento psicológico.

Hunter e Schmidt (1990) abordam em seu livro os métodos de meta-análise e as formas de corrigir erros nos estudos evitando o viés nos resultados de meta-análises. Para os autores o objetivo da meta-análise refere-se à descrição de distribuições de correlações reais entre uma variável independente e uma variável dependente. Assim, se todos os estudos tiverem sido conduzidos corretamente, a distribuição das correlações pode ser usada diretamente para estimar a distribuição da correlação real ou precisará passar por correções, quando possível.

Em outros estudos os autores referem-se à meta-análise como análise de análises, com uso de análises estatísticas de um grande conjunto de resultados de estudos individuais, com o objetivo de integrar os achados. Ou seja, meta-análise emprega um conjunto de técnicas quantitativas para unir resultados entre múltiplos estudos (Schmidt & Hunter, 1996; 1999).

Conforme Rossi e Freeman (1993), há diferença entre meta-análise e revisão de literatura. Essa diferença pode ser observada quando no primeiro caso se realiza um recorte, selecionando-se variáveis para observação e comparação, entre os diversos estudos, para se chegar a alguma conclusão. Já no segundo caso, dependendo do propósito para qual foi empreendida a revisão de literatura o objetivo é rever, como o próprio nome afirma, os estudos de uma área do conhecimento, sem necessariamente fazer uma comparação ou análise.

Empreender uma meta-análise é uma tarefa complexa, que exige conhecimento técnico. Nesse sentido, alguns cuidados precisam ser tomados como combinar os estudos

corretamente selecionando aqueles a serem excluídos da análise. Precisa-se evitar o engano ou viés na seleção dos estudos dentro de um conjunto de estudos a serem analisados.

Ou ainda, quando há poucos estudos relevantes e representativos; nessa circunstância, os procedimentos quantitativos de meta-análise não podem ser empregados. Portanto, a meta-análise exige o uso considerável de julgamento por parte do pesquisador, bem como de habilidades técnicas.

Por fim, a abordagem da meta-análise contribui significativamente para os delineamentos de programas que podem obter consideráveis ganhos, levando-se em consideração a pesquisa da ciência social existente e relatórios profissionais de programas estabelecidos. Para fornecer uma comparação clara entre meta-análise e meta-avaliação, apresentou-se o quadro a seguir, caracterizando o objeto de estudo, procedimento de aplicação, análise de dados e distinção entre os usos da meta-análise e da meta-avaliação.

## Quadro 2. Comparação entre meta-análise e meta-avaliação

Características	Meta-análise	Meta-avaliação
Objeto de estudo	Qualquer tipo de estudos	Avaliação (ões) concluída(s)
Fonte de dados	Secundária	Secundária
Procedimento de aplicação	Organizam-se diferentes estudos, seguindo um critério ou variável, utiliza-se recorte temporal ou temático	Seleção da(s) avaliação (ões) concluída(s) sobre o estudo avaliativo ou diferentes estudos avaliativos com a mesma temática
Análise dos dados	Quantitativa (estatística), a partir de achados semelhantes sintetizam-se constatações, calcula-se o tamanho do efeito entre os estudos	Qualitativa (análise de conteúdo, análise de critérios) emite-se nova avaliação, comparando-se os métodos e procedimentos do estudo anterior com critérios pré-estabelecidos, propõe-se melhorias ou novo modelo
Usos	Em geral acadêmico, mas pode subsidiar a prática profissional	Acadêmica ou profissional. Serve de referencial para programas na área analisada

Conforme destacado no quadro, há semelhanças entre meta-análise e meta-avaliação. A semelhança ocorre no uso de fontes de dados secundários e nas suas aplicações.

As especificidades ocorrem em função do objeto de estudo e o procedimento de análise dos dados. Na meta-análise, o objeto pode ser qualquer tipo de estudo. Na meta-avaliação o objeto de estudo é exclusivamente avaliações já concluídas.

Por fim, na meta-análise o procedimento de análise dos dados é quantitativo e estatístico, calcula-se o tamanho do efeito e se fornece a média através de diferentes estudos, determinando se ocorreu ou não o efeito experimental investigado, e a consistência entre os achados. Na meta-avaliação, utilizam-se procedimentos de análise qualitativos, como a análise de conteúdo ou a checagem de critérios de organizações internacionais de avaliação como o *Joint Committee* ou o *Evaluation Review Society*.

Por fim, destaca-se que o processo utilizado na presente pesquisa configura-se como meta-avaliação e para sua realização utilizaram-se os critérios *do Joint Committee* para checagem da qualidade, validade e fidedignidade das ANOPs e levantaram-se as variáveis presentes no processo de ANOP mediante a análise de conteúdo e síntese de estudos qualitativos.

### ***2.5 Panorama da Avaliação de Programas no Brasil***

A avaliação tem evoluído significativamente em vários países da América Latina e do Caribe. Isso pode ser percebido por quatro fatores: a) a criação de organizações de avaliação profissional, b) a intensificação da difusão das idéias e do uso da avaliação profissional em três setores chave, o estatal, privado e o filantrópico; c) o aumento do número de publicações associadas à avaliação e d) o estabelecimento de programas de capacitação sobre avaliação, em curto prazo, para graduados (Chianca & Yourker, 2004).

De acordo com Barreira (2002), os problemas no Brasil e na América Latina são essencialmente sociais, econômicos e políticos, principalmente de desigualdade de renda, pobreza e exclusão social. Nesse sentido o contexto social necessita de métodos próprios de interpretação e avaliação.

Segundo Silva (2002b), os governos da Argentina, Brasil e Chile dentro do arcabouço de sua reforma estrutural e construção de novas relações com países vizinhos têm dado prioridade ao fortalecimento da avaliação, incentivando a adoção da avaliação de programas como instrumento de gestão pública. Ou seja, iniciativas relacionadas a sistemas nacionais de avaliação da educação, inovação nos sistemas administrativos do Estado e programas sociais apoiados por organismos de cooperação internacional (Chianca & Yourker, 2004).

De acordo com Chianca e Yourker (2004), o Brasil, Colômbia e Peru, em virtude de seus contextos específicos, desenvolveram redes de avaliação, ou seja, organizações menos formais e flexíveis do que as associações e sociedades de avaliação. Assim, a primeira organização profissional na região da América Latina e Caribe foi a Associação Centro Americana de Avaliação – ACE, em 1989, com sede na Costa Rica, depois em 2002 foram criadas outras organizações profissionais de avaliação. Sendo que os avaliadores que trabalham nessas organizações são na sua maioria procedentes de programas sociais e educativos nos quais trabalham, financiados pelo governo ou por fundações. No Brasil, em São Paulo, foi criada, em setembro de 2003, a Rede de Seguimento, Avaliação e

Sistematização da América Latina e Caribe – RELAC. De acordo com esses autores, atualmente há um aumento das redes de avaliação, principalmente em países como Argentina, Bolívia, Chile, Equador, Honduras, Nicarágua, Uruguai e Venezuela.

Segundo Silva (2002b), a função avaliação no Brasil desenvolveu-se segundo diferentes necessidades de acompanhamento da parte das entidades responsáveis pelo planejamento e acompanhamento da avaliação e/ou financiamento dos projetos. A função avaliação está associada historicamente à influência da concepção de ciência positivista e restringe a avaliação à medição e quantificação dos resultados do programa. Por isso neste trabalho o termo utilizado foi simplesmente avaliação ou pesquisa avaliativa, pois possuem conceituação mais abrangente. De acordo com Chianca e Yourker (2004), são os organismos de cooperação internacional como Banco Mundial, Banco Interamericano de Desenvolvimento – BID; Organização Mundial de Saúde – OMS e a Agência dos Estados Unidos para o Desenvolvimento Internacional – USAID, que exigem que os programas que apóiam sejam avaliados por avaliadores internos e avaliadores externos.

Segundo Holanda (1999, citado por Silva, 2002b), as áreas nas quais se desenvolveram capacidades em avaliação no Brasil foram planos nacionais de desenvolvimento, planos e programas regionais, políticas e programas setoriais, projetos financiados por bancos de fomento e sistemas estaduais de acompanhamento e avaliação. Uma lacuna apontada pelo autor é que apesar da criação de sistemas específicos de avaliação em diferentes áreas/setores, como por exemplo políticas de educação, saúde e saneamento, há ausência de integração entre os indicadores setoriais das diferentes áreas de ação pública e políticas integradas, no sentido de serem absorvidas como instrumento de gestão e acrescentem-se também instrumentos de avaliação. Segundo Silva (2002a, p. 34), os programas avaliados englobam “propostas de âmbito nacional dedicadas ao exame de política econômica ou social como um todo [...] à avaliação no nível setorial (política educacional, por exemplo), no nível multisetorial, territorial e de projetos específicos”. Podem-se citar quatro iniciativas de avaliação na área de educação no Brasil, pelo Ministério da Educação, a saber: avaliação institucional envolvendo a avaliação das condições gerais de todas as instituições de educação superior do país; uma avaliação da educação de estudantes recém formados, por exemplo, o exame da Ordem dos Advogados do Brasil – OAB. No nível do ensino básico, há duas avaliações, um sistema nacional de avaliação da educação básica, da qualidade das escolas, do jardim de infância até o início do ensino fundamental.

Na revisão de Silva (2002a) sobre a avaliação nas administrações públicas de programas, da América do Sul (Brasil, Argentina e Chile), o autor empregou como critérios: validade, utilidade, eficiência e efetividade de programas e políticas, dos seus resultados, impactos e repercussões econômicas e sociais. Quanto a um momento específico da história brasileira, Barreira (2002, p. 14) afirma “no Brasil, a década de 80 foi extremamente rica no que se refere à emergência e luta de movimentos sociais populares urbanos. A avaliação de programas sociais nesta época foi priorizada e publicizada”.

No caso brasileiro, Silva (2002a) analisou programas em diferentes níveis e abrangência, como planos nacionais, planos e programas regionais, programas setoriais, projetos financiados por bancos de fomento e o planejamento estadual. Os programas analisados foram: o Programa de Ações Prioritárias do Governo Federal do Nordeste; a primeira etapa do Plano Plurianual – PPA do período de 1996/1999; o planejamento para o Programa Brasil em Ação.

Mesmo diante da abrangência estadual dos programas citados, verificaram-se, segundo os autores, falhas comuns. As falhas referem-se à dispersão e descontinuidade, com evidências e resultados pouco documentados e sistematizados. Essa situação foi associada a duas características do planejamento governamental do país que são ênfases no processo de formulação de planos e de elaboração de programas e projetos; e a negligência nas etapas de acompanhamento e avaliação dos processos, resultados e impacto. Essas falhas foram associadas a gerenciamento difuso; descontinuidade; e implantação à margem da estrutura burocrática formal.

De acordo com Cook e Shadish (1986), esse cenário também era vivenciado nos Estados Unidos, na década de 60. Conforme Cook e Shadish, devido à constituição política dos programas sociais, a sua implantação direcionada por administração central não assegurava que fossem implantados em todos os distritos, de acordo com as diretrizes da coordenação central do programa. Pois não raro, programas sociais são tidos como meio para atingir objetivos pessoais de quem os administra. Dessa forma possuem características marcadamente diferentes em cada local onde são implantados. Por conseguinte, essas falhas podem justificar, no caso brasileiro, a meta-avaliação, empreendida neste trabalho, com o objetivo de aumentar a eficiência e eficácia dos processos de avaliação empreendidos no período de 1999 a 2005, e que se encontram descritos detalhadamente no capítulo 3, sobre a metodologia da presente pesquisa.

Em suma, o plano político-organizacional interno brasileiro, no processo de planejamento plurianual, pretende oferecer uma base estratégica para intervenção, acompanhamento e revisão dos programas e projetos em execução. Esse processo resultou numa nova classificação das ações e intervenções estatais: *programa, projeto, atividade e operações especiais*. Acredita-se que os casos discutidos servem de base para pensar a validade e fidedignidade das avaliações empreendidas em qualquer nível por uma organização ou por um governo.

As situações-problema foram expostas no nível macro das avaliações dos programas no Brasil. Diante delas, o tópico seguinte aborda a questão dos problemas metodológicos, ao empreender ou delinear uma avaliação. Enfoca-se agora o nível micro em relação à avaliação de um programa.

### ***2.6 Questões Metodológicas na Avaliação de Programas***

Ao se abordar as questões metodológicas em relação à avaliação de programas, os avaliadores defrontam-se comumente com diversos problemas que podem afetar a validade e aplicação dos resultados das avaliações pelos tomadores de decisão.

O debate entre os paradigmas metodológicos em avaliação representa a escolha por regras teóricas e metodológicas, instrumentos a serem utilizados, problemas a serem investigados, padrões de pesquisa a serem adotados, representando um modo de explicar os fenômenos reais. De acordo com Barreira (2002), o debate metodológico em pesquisa e por extensão à pesquisa avaliativa centra-se no valor de três paradigmas: o paradigma positivista, o paradigma fenomenológico e o paradigma marxista.

O paradigma positivista colocou-se em oposição à especulação pura, exaltando os fatos passíveis de observação, utilizando métodos quantitativos e experimentais para testar generalizações hipotético-dedutivas. Estabeleceu distinção clara entre valor e fato, valor como expressão cultural; e fato, o dado empírico da ciência capaz de ser mensurado. O paradigma fenomenológico dá importância ao sujeito no processo de construção do conhecimento, assim,

*Volta-se para a investigação do mundo vivido e utiliza aproximações qualitativas e naturalistas para compreender a experiência humana de forma indutiva e holística [...] privilegiou o enfoque antropológico, na busca do significado dos fenômenos* (Barreira, 2002, p. 39).

O paradigma marxista, fundamentado no materialismo dialético e histórico, utiliza a visão crítica da história, “na atuação dinâmica, social, transformadora e emancipatória do homem” (Barreira, 2002, p. 39).

Concorda-se com Barreira (2002) que a partir da adoção de um dos diferentes paradigmas metodológicos, que influenciam a visão e a prática da pesquisa avaliativa, têm-se diferentes propostas de atuação no campo da avaliação. As especificidades das políticas e programas sociais que são permeados por valores, pautas culturais, demandas e pressões dos diferentes atores envolvidos; caracterizam a avaliação na área social. Tudo isso mostra, que os resultados e impactos de um programa ou política podem ser múltiplos, derivados de múltiplas causas e interpretados sob diferentes paradigmas.

Compartilha-se do ponto de vista de Barreira (2002, p. 14) de que “há [...] necessidade de diferentes metodologias qualitativas e quantitativas ou mesmo uma síntese de métodos avaliativos que sejam mais adequados ao programa a ser avaliado”. Entende-se que cada método possui diferentes vantagens e desvantagens que estão relacionadas à qualidade dos dados obtidos, forma de obtê-los e uso e análise dos dados. Acrescente-se às vantagens e desvantagens características comuns e específicas de cada método (Günther, 2006). Um pesquisador pode utilizar dados mistos, trabalhando com evidências qualitativas e quantitativas, pois não requer um método único para a coleta de dados do seu processo de avaliação. Isto pode ser ilustrado, em casos de dados obtidos por meio de um mesmo instrumento, como um questionário, que contém questões abertas e fechadas. Ressalta-se que o método qualitativo pode lançar mão de fontes quantitativas, uma vez que informações coletadas de ambas as fontes fornecem maior confiabilidade às análises e resultados encontrados. Há possibilidade de dados qualitativos emergirem de experimentos, caso os participantes sejam divididos aleatoriamente entre o grupo de tratamento e o grupo de controle; porém, os dados qualitativos são coletados geralmente a partir de investigações naturalísticas, ou seja, aquelas que ocorrem em ambiente natural. O reconhecimento de diferentes abordagens avaliativas é didático e sugere formas distintas de conduzir a avaliação de programas. Em suma, as abordagens avaliativas, com técnicas específicas ou não, são complementares e não excludentes.

Diversos autores (Campbell & Stanley, 1979; Posavac & Carey, 2003; Rossi & Freeman, 1993; Patton, 2001) referem-se à problemática metodológica que cerca a avaliação de programas como diversos tipos de *ameaças à validade* dos resultados ou hipóteses rivais produzidas pelas avaliações (tratamento). Entre as ameaças pode-se citar:

validade da conclusão estatística, validade interna, validade do construto e validade externa.

Alguns exemplos podem ser encontrados para ilustrar a dificuldade de se atribuir os resultados encontrados à intervenção ou programa. Por exemplo: o caso de uma pessoa que procura o atendimento psicológico durante uma crise em sua vida. Se alguns meses depois de iniciada a terapia tal pessoa se sentir melhor, tal melhora pode ser atribuída à terapia, ou à resolução natural da crise, ou ainda, à combinação de ambos. Outro exemplo: pesquisas educacionais em que o grupo é estudado apenas uma vez, depois de receber o tratamento (unidade de ensino).

De acordo com Campbell e Stanley (1979) a evidência científica exige pelo menos uma comparação. Para que a comparação possa ser válida, ambos os lados da comparação precisam ser tratados com o mesmo cuidado e precisão, caso contrário, praticamente não possuem valor científico.

Outro aspecto importante quando se discute os aspectos metodológicos da pesquisa avaliativa é a questão do conhecimento e o uso dos delineamentos de pesquisa existentes.

Assim, um cuidado que diversos autores (Popham, 1977, Campbell, 1979, Rossi & Freeman, 1993; Cano, 2002; Posavac & Carey, 2003) ressaltam em relação ao planejamento de uma avaliação é a necessidade de procurar, sempre que possível, utilizar uma **distribuição randômica e de grupo de controle e/ou comparação, como uma tentativa de evitar as ameaças à validade interna.**

Campbell e Stanley (1979) esclarecem que a *validade interna* refere-se às condições que podem afetar os resultados de um experimento. Segundo Campbell e Stanley (*op cit.*), essas condições são diversas e dizem respeito à *história, maturação, testagem, instrumentação, regressão estatística, viés na seleção dos sujeitos, mortalidade experimental, e diversos tipos de interação seleção maturação*. Esses fatores por si só podem afetar diretamente os escores dos grupos observados e podem ser confundidos com os resultados do tratamento. Entretanto esses problemas podem ser solucionados utilizando-se a lógica da estatística e da probabilidade, conforme explicado a seguir.

No caso de *história*, eventos que tenham ocorrido entre a primeira e segunda medição podem ter ocasionado à mudança observada e não necessariamente o tratamento. Assim, se o pré-teste ( $O_1$ ) e o pós-teste ( $O_2$ ) forem aplicados em dias diferentes, outras explicações ou hipóteses rivais podem ter causado a diferença observada. A história torna-se uma hipótese ou explicação rival de mudança quanto maior o tempo entre  $O_1$  e  $O_2$ . Para tentar controlar a história pode-se tentar a condição de isolamento experimental. Em

síntese, história pode ser um grupo de possíveis efeitos advindos de tempo decorrido entre uma medição e outra, embora esses efeitos possam ser também chamados de maturação.

**Maturação** refere-se aos processos internos biológicos ou psicológicos que afetam os respondentes, como envelhecimento, fome, cansaço, entre outros, independentes de eventos externos específicos. Por exemplo, entre  $O_1$  e  $O_2$  os estudantes podem ter ficado mais velhos, com fome ou cansados, etc., e a diferença obtida pode refletir esse processo mais que o tratamento. Assim, as mudanças observadas podem ser frutos de efeitos cumulativos de processos de aprendizagem, pressões ambientais diárias que operam mesmo sem a introdução do tratamento.

**Testagem** são mudanças relacionadas ao método de observação, ou seja, é o efeito do próprio pré-teste. Por exemplo, “em testes de desempenho e inteligência, estudantes submetidos ao teste pela segunda vez ou submetidos a uma forma paralela do teste, conseguem melhores resultados que os submetidos ao teste pela primeira vez” (Campbell & Stanley 1979, p. 17). Para tentar controlar o efeito de testagem é necessário atentar para a *reatividade* do teste. O efeito de reatividade pode ocorrer quando o processo de testagem é intrinsecamente um estímulo à mudança e não apenas um simples registro de comportamento. Por exemplo, em um experimento de controle de peso, a primeira pesagem pode ser um estímulo à redução de peso, mesmo sem se ter iniciado o tratamento.

**Instrumentação** refere-se a mudanças nos instrumentos em termos das medidas aplicadas ocorre geralmente no caso de observadores ou aplicadores do teste. Por exemplo, quando pais de estudantes são entrevistados, a familiaridade do entrevistador com o roteiro da entrevista e com alguns pais pode produzir modificações na forma de conduzir a entrevista ou registro da informação. Pode ocorrer a instrumentação quando há mudança nos observadores que realizaram a primeira e segunda observação.

**Regressão estatística** ocorre quando os grupos são selecionados com base em escores extremos. Por exemplo, alunos que obtiveram notas baixas em um teste  $O_1$  e depois de receber o tratamento, aulas de recuperação, receberam um novo teste ou o mesmo teste  $O_2$ . O grupo  $O_2$  provavelmente alcançará médias maiores que  $O_1$ , e esse resultado não poderá ser atribuído ao tratamento e sim à correlação imperfeita entre  $O_1$  e  $O_2$ . Para tentar controlar esse fator, utilizam-se análises de controle temporal invertidas e o exame direto de mudanças em variabilidade da população. A regressão estatística é um fenômeno que está presente em toda parte, não somente no pré-teste e pós-teste com o mesmo teste ou teste semelhante. Destaca-se que a regressão em direção à média ocorre devido a grupos selecionados com base em escores extremos. No caso de grupos selecionados por motivo

independente, caso obtenham uma média extrema, existe menor expectativa de que num segundo teste à média regrida, pois as fontes de variância aleatórias provavelmente já afetaram seus escores iniciais em ambas as direções.

O *viés na seleção dos sujeitos* pode ser evitado realizando-se a seleção de forma randômica para o grupo de tratamento e o grupo de controle. Assim, assegura-se equivalência dos dois grupos para todas as variáveis medidas. É importante que os grupos tenham um tamanho grande o suficiente (cerca de pelo menos 30) para satisfazer o critério dos grandes números, ou da distribuição normal dos resultados (Pasqualy, 1997).

*Mortalidade experimental* é a perda ou evasão de sujeitos nos grupos. É expressa entre a produção da diferença entre  $O_1$  e  $O_2$  devido à saída de sujeitos em quantidades diferentes em ambos os grupos. Assim, representa casos em que se obtêm dados parciais do grupo. Embora os grupos possam ter sido idênticos no início, podem diferir ao final, não devido ao tratamento, e sim à evasão de pessoas de um dos grupos. Pasqualy (1997) assinala que se a perda for de até 5% pode-se desconsiderá-la, entretanto se for superior a esse número então esse fator não foi controlado.

De acordo com Campbell e Stanley (1979) pode ocorrer ainda o viés na seleção dos sujeitos e diversos tipos de interação seleção-maturação.

Por fim, na linguagem de avaliação de programas, as ameaças à validade interna são mudanças reais, entretanto não relacionadas ao programa (Posavac & Carey, 2003). Daí se torna necessário considerar outras variáveis fora do programa como possíveis explicações para os resultados. Somente quando essas explicações não forem plausíveis, os resultados obtidos poderão ser atribuídos ao programa.

Quanto à *validade externa*, refere-se à interação do tratamento (X) com alguma outra variável. As ameaças à validade externa afetam as condições de aplicação e generalização dos resultados ou efeitos encontrados para populações, conjuntos de variáveis e outras medidas do tratamento, no caso programa ou serviço recebido (Campbell & Stanley, 1979). A validade externa diz respeito aos fatores que podem afetar a generalização dos resultados alcançados. Quanto à generalização, é importante destacar que em ciência refere-se a antecipações que se faz em situações com condições experimentais mais próximas possíveis à realidade comparada, deduzindo-se que quanto mais semelhantes essas condições, mais provavelmente sigam as mesmas leis.

Alguns exemplos oferecidos por Campbell e Stanley (*op cit*) de ameaças à validade externa são: *o efeito da interação entre a testagem e a variável experimental; interação*

*entre a seleção e a variável experimental; o efeito reativo das condições experimentais e a interferência de tratamentos múltiplos.*

O efeito da *interação entre a testagem e a variável experimental* é verificado quando o pré-teste sensibilizou os respondentes para o tratamento, aumentando o efeito do tratamento (X). Nesse caso, ocorre que o efeito observado de X foi potencializado e só ocorreria em tal intensidade devido ao pré-teste. Esse efeito não ocorre quando os testes são usuais na área estudada ou para a população estudada.

A *interação entre a seleção e a variável experimental* pode ocorrer quando há muita dificuldade em obter sujeitos para o experimento. Se após diversas tentativas de recrutamento de colaboradores para o experimento, obtêm-se os sujeitos, pode ser que os mesmos apresentem características que os diferencie dos demais que não quiseram participar do experimento. Tal característica pode interagir com o tratamento. Para contornar essa ameaça, é preciso na medida do possível, buscar a representatividade amostral, para então depois do cumprimento dessa condição, tentar generalizar os resultados.

O *efeito reativo das condições experimentais* ocorre geralmente quando o sujeito tem conhecimento de que está participando de um experimento e que seus comportamentos estão sendo observados. Quanto mais evidente para o sujeito a relação entre o tratamento experimental e o pós-teste, mais provável é o efeito reativo da condição experimental. De acordo com Pasqually (1997), o alto grau de controle experimental pode produzir nos sujeitos a sensação de estar sendo *cobaia*. Uma alternativa a essa ameaça é o experimento duplo cego, na qual nem experimentador, nem participantes conhecem as condições de grupo de controle, grupo de tratamento e tratamento sob os quais estão participando. Na área educacional, por exemplo, Campbell e Stanley (1979) sugerem que as pesquisas sejam sempre aplicadas pelos próprios professores da escola aos alunos, para tentar minimizar esse efeito.

A *interferência de tratamentos múltiplos* ocorre quando a mesma amostra estudada é alvo de dois ou mais tratamentos de forma simultânea. Por exemplo, na área de avaliação de programas quando uma mesma população é beneficiária de dois ou mais programas sociais: o Programa Bolsa Família – PBF e o Programa de Erradicação do Trabalho Infantil – PETI (Brasil, 2005a), como ocorre em diversos municípios brasileiros. Quando são mensurados os “efeitos” do Programa Bolsa Família em relação à redução da desigualdade social, um de seus objetivos; os resultados encontrados podem estar associados, ora à transferência de renda do PBF, ora as atividades de inclusão produtiva e

geração de renda desenvolvidas com as famílias do PETI (Brasil, 2001c; Braga Lins, 2005).

Por fim, resguardadas as questões éticas e morais de cada ameaça à validade externa no delineamento e avaliação de um programa, o tratamento fornecido aos colaboradores participantes de experimentos os apropria de conhecimentos sobre o programa ou pesquisa. Tal apropriação leva a um fortalecimento de sua auto-estima, de sua capacidade de decisão e de transformação da vivência de seus problemas em busca de soluções (Barreira, 2002). Defende-se neste trabalho a posição de que a divulgação, dos resultados de uma pesquisa, seja avaliativa ou não, sua publicação ou publicização são a contrapartida que o cientista ou pesquisador pode e deve oferecer às pessoas que participam das pesquisas.

Uma forma importante de sintetizar as ameaças à validade interna e externa pode ser obtida em Campbell e Stanley (1979). Utilizamos a tabela dos autores para auxiliar na análise dos delineamentos que foram apresentados no próximo tópico. Ilustra-se com o sinal - (menos) aqueles delineamentos em que a presença da ameaça à validade está presente e com o sinal + (mais) aqueles delineamentos nos qual a ameaça pode ser controlada. Por fim o sinal de ? (interrogação) é empregado quanto o fator não é relevante. A notação “R” refere-se a randômico e (X) quando a aplicação não é relevante.

Dessa forma, apresentou-se a seguir a síntese das ameaças à validade interna e externa na Tabela 2 e em seguida, no próximo tópico, foram comentados os delineamentos citados na Tabela, no próximo tópico.

**Tabela 2. Fontes de ameaça Interna e Externa para Delineamentos**

	Ameaças a Validade											
	Interna							Externa				
	História	Maturação	Testagem	Instrumentação	Regressão	Seleção	Mortalidade	Interação de seleção e maturação, etc.	Interação de teste e X	Interação de seleção e X	Condições reais	Interferência de múltiplos X
<b>Delineamentos Pré-experimentais</b>												
1. Grupo único com pós-teste e sem grupo de controle	-	-										
		X	O									
2. Grupo único com pré-teste e pós-teste	-	-	-	-	?	+	+	-	-	-	?	
		O	X	O								
3. Grupo de comparação estático com pós-teste	+	?	+	+	+	-	-	-				
		X	O	O								
<b>Delineamentos experimentais</b>												
4. Pré-teste e pós-teste a grupos experimentais e de controle aleatórios	+	+	+	+	+	+	+	+	-	?	?	
		R	O	X	O							
		R	O		O							
5. Delineamento de quatro grupos de Salomon	+	+	+	+	+	+	+	+	+	?	?	
		R	O	X	O							
		R	O		O							
		R		X	O							
		R			O							
6. Delineamento com grupo de controle e pós-teste	+	+	+	+	+	+	+	+	+	?	?	
		R		X	O							
					O							
<b>Delineamentos Quasi-experimentais</b>												
7. Delineamento de série temporal	-	+	+	?	+	+	+	+	-	?	?	
	O <sub>1</sub>	O <sub>2</sub>	O <sub>3</sub>	O <sub>4</sub>	X	O <sub>5</sub>	O <sub>6</sub>	O <sub>7</sub>	O <sub>8</sub>			
8. Delineamento de amostras temporais equivalentes	+	+	+	+	+	+	+	+	-	?	-	-
		X <sub>1</sub> O	X <sub>0</sub> O	X <sub>1</sub> O	X <sub>0</sub> O <sub>0</sub>							
9. Delineamento com grupo de controle não equivalente	+	+	+	+	?	+	+	-	-	?	?	
		O		X	O							
		O			O							
10. Delineamento de pré e pós-teste de amostras distintas	-	-	+	?	+	+	-	-	+	+	+	
		R	O	(X)								
		R		X	O							

Fonte: adaptado de Campbell e Stanley (1979, p.14, 70)

No contexto de validade interna e externa e outras questões que afetam diretamente as pesquisas, existem alguns tipos de delineamento que buscam controlar e/ou minimizar esses problemas. Na seção seguinte, para esclarecer tais possibilidades, foram descritos os delineamentos de pesquisa científicos mais utilizados.

### **2.6.1. Delineamentos de Pesquisa**

Os delineamentos são classificados como não experimentais ou pré-experimentais, experimentais e quasi-experimentais (Campbell & Stanley, 1979). O delineamento pré-experimental é também denominado de não-experimental.

Em termos da notação, cada grupo experimental ou de controle é representado em linhas diferentes. Cada observação é registrada com a letra “O”. Havendo várias medições inclui-se o número e a letra indicando a ordem. O tratamento experimental é indicado com a letra “X” na linha correspondente. E se os grupos tiverem sido formados por designação aleatória, cada um estará precedido da letra R (*random*).

Foi detalhado com maior ênfase os delineamentos pré-experimentais e quasi-experimentais, uma vez que delineamentos experimentais dificilmente são aplicados a AP.

### **2.6.2 Delineamentos não-experimentais ou pré-experimentais**

Delineamentos pré-experimentais ou não-experimentais são mais flexíveis e, portanto, mais fácil de adaptá-los à realidade pesquisada (Cano, 2002). Outra razão forte para adotá-lo, seria o caso, no qual se torna antiética a escolha de pessoas para o grupo experimental e de controle do programa ou, por razões de ordem prática, na alocação randômica das pessoas (Posavac & Carey, 2003).

Apesar do uso comum desse tipo de delineamento em pesquisas educacionais e outras, considera-se a sua restrição em termos de cientificidade. Esse delineamento é o mais comumente utilizado para analisar o desempenho ou performance dos participantes após o fim de um programa. Por fim, é o delineamento mais empregado comumente para medir o impacto do programa, em termos dos resultados alcançados (Posavac & Carey, 2003).

Os delineamentos pré-experimentais podem ser: a) grupo único com pós-teste e sem controle; b) grupo único com pré e pós-teste e c) grupo de comparação estático, com pós-teste.

#### *a) Grupo único com pós-teste e sem grupo de controle*

No caso do grupo único, o que se busca observar é se os participantes concluíram o programa atingindo os objetivos do programa. Entretanto, esse delineamento não permite identificar se os participantes mudaram durante o programa ou se ocorreram outras situações que impediram a validade interna e a possibilidade dessa inferência causal.

Ilustração: X O

Conforme se observa na ilustração, não existe grupo de comparação e/ou outra referência com a qual o pós-teste possa ser comparado. Conforme Campbell e Stanley (1979), nesse delineamento não há praticamente controle e a inexistência de grupo de comparação impede qualquer inferência causal, pois a evidência científica exige pelo menos uma comparação. Para que tal comparação possa ser útil “ambos os lados dela precisam ser tratados com a mesma precisão” (p. 13). Entretanto, alguns estudos de caso seguem esse tipo de delineamento, embora não recomendável. De acordo com Posavac e Carey (2003), por exemplo, geralmente professores, médicos e administradores julgam o sucesso de um programa com uma avaliação simples, porque necessitam saber quanto os participantes foram afetados pelo serviço ou programa recebido. Outro exemplo, verificar se os membros de um grupo conseguiram um emprego, após três meses de treinamento de preparação para o mercado de trabalho. Assim, o primeiro passo para empreender tal avaliação é identificar qual o nível de apreensão dos participantes que finalizaram o programa em relação aos objetivos implícitos e explícitos do programa. Lembra-se que essa abordagem não assegura que os participantes tenham mudado durante o programa, pois conforme a Tabela 2, esse delineamento está sujeito a problemas de história, maturação, seleção, mortalidade e interação da seleção em relação ao tratamento (X) (Campbell & Stanley, 1979).

*b) Grupo único com pré-teste e pós-teste*

Neste caso, é possível comparar as duas observações e tentar controlar problemas de seleção ou mortalidade amostral no pós-teste. Porém, explicações alternativas como história, maturação, testagem, instrumentação e interação da seleção e maturação podem ocorrer (Campbell & Stanley, 1979). Em relação à história, outros eventos podem ter ocorrido entre a  $O_1$  e a  $O_2$ , além do X; nesse caso, quanto maior for o intervalo de tempo entre as observações, ou até mesmo a aplicação em dias diferentes para  $O_1$  e  $O_2$ , maior a possibilidade de uma explicação alternativa.

Ilustração:  $O_1$  X  $O_2$

Um exemplo desse tipo de avaliação é quando os *stakeholders* desejam saber se os participantes melhoraram enquanto foram atendidos por um programa. Ilustrando, se estudantes lêem melhor após um ano de instruções de leitura, ou se o nível de colesterol dos pacientes diminuiu depois de determinado tratamento. Destaca-se que dessa forma não se pode assegurar totalmente que o tratamento foi o responsável pela melhoria do participante, mesmo quando se comparam os resultados de antes e depois de um

tratamento, pois pode haver várias explicações alternativas possíveis para as mudanças observadas. Uma avaliação precisa ser delineada de acordo com a complexidade do programa. E as ameaças mencionadas referem-se à validade dos resultados e são: a história; maturação; testagem; instrumentação; e, interação da seleção e maturação, conforme comentado.

Embora as avaliações de programa empreendidas por governos não possuam a mesma finalidade que a pesquisa básica e nem o rigor científico de pesquisas acadêmicas, é importante nesses casos o exame de variáveis-chaves depois que os programas estão em funcionamento, principalmente no caso de grandes programas, esse é um modo pelo qual se pode avaliá-los (Posavac & Carey, 2003). Sugere-se no caso de grupos simples, o uso de delineamentos descritivos, nos quais se consideram questões como: a) os participantes possuem conhecimento dos critérios/objetivos da avaliação do programa, que assim pode influenciar o comportamento dos participantes. Entretanto se o programa é desconhecido por 95% da população-alvo e ineficiente entre os participantes, não é necessário conduzir uma avaliação sofisticada; b) os participantes realmente obtiveram algum aperfeiçoamento ou melhoria a partir da participação no programa, ou seja, nem sempre uma diferença estatística nos resultados de pré-teste e pós-teste fornece evidência de que o programa foi o responsável pela melhoria, pois a diferença estatística pode estar relacionada a erro na amostragem e não a causalidade; e, c) se a melhoria foi o suficiente, em alguns casos quando há melhoria estatística significativa entre o pré-teste e pós-teste, surge a questão, se a melhoria foi suficientemente necessária àquela para a vida cotidiana dos participantes.

Entre as diversas formas de verificar se ocorreram melhorias, Posavac e Carey (2003) sugerem verificar se: 1) os resultados foram significativos em si mesmos; e 2) o resultado que servem como um *proxy* para uma variável que seja difícil ou dispendiosa financeiramente para medir. O *proxy* oferece uma aproximação do que se pretende avaliar. Pode-se usá-lo no caso em que a avaliação não consegue medir o objetivo real do programa, devido à sua complexidade.

#### *c) Grupo de comparação estático, com pós-teste*

Neste caso, conta-se com um grupo de controle, mas a distribuição das unidades pelos grupos não foi feita de modo aleatório. Assim, podem-se controlar alguns fatores que ameaçam a validade interna (história, testagem, instrumentação, regressão em direção à média). Nesse delineamento persistem problemas como viés na seleção; mortalidade; interação da seleção e maturação e interação da seleção e o tratamento (Campbell &

Stanley, 1979). Além disso, a ausência de uma amostragem aleatória não permite uma inferência causal forte (Cano, 2002).

Ilustração:  $\frac{X \quad O_1}{O_2}$

Alguns exemplos desse delineamento podem ser dados: comparação de sistemas escolares que exigem professores portadores de determinada especialidade (X) *versus* sistemas que não possuem essa exigência. Comparação de pessoas que foram submetidas a determinado treinamento na empresa com as que não foram submetidas ao treinamento. Nesses exemplos, não há como verificar se os grupos eram equivalentes, conforme indicado pela linha pontilhada que separa as duas observações. Caso ocorra diferença entre as duas observações, esta pode ter sido ocasionada pela forma com que os grupos foram selecionados; sendo assim, ainda existe a possibilidade de que os grupos manifestassem diferença, mesmo não tendo recebido o tratamento (X). Esse delineamento também está sujeito à mortalidade experimental conforme Campbell e Stanley (1979).

Conforme explanação dos delineamentos pré-experimentais, ao utilizá-los na avaliação de programas, mister se faz, estar atento à incidência de deficiências ou ameaças desses delineamentos e aos cuidados para utilizá-los. A seguir a descrição dos delineamentos experimentais.

### 2.6.3 *Delineamentos experimentais*

Os delineamentos experimentais são: i) delineamento com pré e pós-teste e com grupo de controle aleatório; ii) delineamentos de quatro grupos de Solomon (1949, citado por Campbell & Stanley, 1979) e, iii) delineamento com pós-teste e grupo de controle.

Todos os três delineamentos permitem controlar as ameaças à validade interna: história, maturação, testagem, instrumentação, regressão em direção à média, seleção, maturação, interação da seleção e maturação. Em relação às ameaças à validade externa no Delineamento de quatro grupos de Salomon e o Delineamento com grupo de controle e pós-teste pode-se controlar a interação da testagem e o tratamento. Para maior aprofundamento dos mesmos sugere-se a leitura de literatura específica.

### 2.6.4 *Delineamentos quasi-experimentais*

Os delineamentos quasi-experimentais em geral descumprem alguma condição básica para serem considerados experimentais, em geral essas condições dizem respeito ao controle sobre a situação experimental, ou à distribuição randômica dos indivíduos nos

grupos, gerando grupos não equivalentes. Grupos equivalentes são aqueles nos quais ocorre a aleatorização das unidades (Popham, 1977; Cano, 2002). Os delineamentos quasi-experimentais são utilizados em contextos sociais em que o pesquisador pode manipular e controlar algumas variáveis, não todas, por exemplo, pode ter controle sobre a coleta de dados. Entretanto, falta o pleno controle da situação experimental. A ausência do controle experimental alerta o pesquisador para as variáveis específicas que o seu estudo deixa de controlar. Enfim, aconselham-se o uso desses delineamentos, quando não for possível a utilização melhores delineamentos, como os experimentais. Pasqually (1997) classifica os delineamentos quasi-experimentais como pesquisa *ex post facto*, ou seja, a pesquisa começa depois que a(s) variável(veis) independente(s) já ocorreu ou já ocorreram. É a pesquisa onde as variáveis independentes não são manipuladas. “Quanto maior o número de variáveis independentes nestes delineamentos que não forem manipuladas, tanto mais *ex post facto* se tornam estes delineamentos quasi-experimentais” (p. 58).

De acordo com Pasqually (1997), há muitos tipos de delineamentos quasi-experimentais, apresentou-se neste trabalho, quatro delineamentos quasi-experimentais, citados na literatura clássica por Campbell e Stanley (1979): 1) delineamento de séries temporais; 2) delineamento de amostras temporais equivalentes; 3) delineamento com grupo de controle não equivalente e 4) delineamento com pré-teste e pós-teste de amostras distintas.

#### 1) *Delineamento de séries temporais*

Esse delineamento é também chamado de séries temporais interrompidas, pois uma série histórica é afetada ou interrompida por uma intervenção, “cujos resultados são indicados por uma descontinuidade nas medidas registradas na série temporal” (Campbell & Stanley, 1979, p. 67). Consiste em proceder a medições repetidas no pré-teste e no pós-teste, em diferentes momentos. Assim as medições anteriores do próprio grupo funcionam como uma espécie de grupo de controle, e as medições posteriores, medem a evolução do seu impacto. Esse delineamento favorece a compreensão, se ocorre uma tendência histórica de aumento ou melhoria nas observações iniciais ou variações sazonais que poderiam ser confundidas com o efeito de X. Já as observações do pós-teste contribuem para verificar até que ponto o efeito da intervenção é transitório ou permanente, caso exista.

Ilustração:

O<sub>1</sub> O<sub>2</sub> O<sub>3</sub> O<sub>4</sub> X O<sub>5</sub> O<sub>6</sub> O<sub>7</sub> O<sub>8</sub>

O delineamento de séries temporais consegue controlar todas as ameaças à validade interna, exceto história e a ameaça à validade externa, interação da testagem e X. Para tentar equacionar esses problemas são possíveis outros tipos de combinações a esse delineamento simples de séries temporais. Na tentativa de superar o problema de história e ausência do grupo de controle, pode-se compará-lo com outras séries históricas que não receberam o tratamento. Para outros problemas como mortalidade amostral, pode-se calcular a média para cada momento, todos os períodos, considerando-se todas as pessoas que saíram dos grupos.

Em relação à instrumentação evita-se modificar o instrumento durante as medições. Um exemplo de Campbell e Stanley (1979) em organizações: ao empossar um novo diretor pode-se registrar um aumento de infrações disciplinares, devido à mudança no instrumento, mais que no próprio índice de infrações. Assim, por exemplo, o delineamento de séries temporais pode ser empregado para medir os efeitos de uma mudança política administrativa em uma organização, entretanto evita-se mudar ou substituir instrumentos de pesquisa ao mesmo tempo em que se muda uma política.

## 2) *Delineamento de amostras temporais equivalentes*

Emprega duas amostras equivalentes de ocasiões, em uma delas a variável experimental está presente, sem que esteja na outra. Ou seja, introduz-se de maneira repetida a variável experimental, alternando-se variável experimental, variável não experimental. As medidas em geral são emparelhadas, com as apresentações da variável experimental concomitante. Nos primeiros usos desse delineamento a variável experimental  $X_1$  era comparada com outra variável experimental  $X_2$ , ao invés de com o não tratamento  $X_0$ .

Nesse delineamento podem-se controlar todas as ameaças à validade interna, entretanto ele está sujeito à interação de testagem e X, condições reativas e interferência de múltiplos tratamentos. A interferência de X múltiplo é uma ameaça à validade externa e significa que “múltiplos níveis de X são apresentados ao mesmo conjunto de pessoas” (Campbell e Stanley, 1979, p. 77). Nesse caso, é possível a generalização de  $X_1$  somente se ocorrer sua apresentação em repetidas vezes e com um espaço de tempo entre uma apresentação e outra; além disso, precisa ser realizada em seguida ou intercalada a cada apresentação  $X_1$ , a apresentação de X ou  $X_0$ .

Ilustração

$X_1O \quad X_0O \quad X_1O \quad X_0O_0$

Por fim, considera-se a exigência de aleatoriedade de ocasiões, evitando-se a periodicidade para a aplicação de  $X_1$  ou  $X_0$ . Ressalta-se que esse delineamento é útil na área da educação. Por exemplo, quando um professor deseja testar dois métodos de ensino, assim se introduz esses dois processos ao longo de uma unidade de ensino, montando um experimento. Esse delineamento também pode ser aplicado a um único indivíduo, por exemplo na pesquisa fisiológica, onde um estímulo é aplicado repetidas vezes a um animal, evitando-se, entretanto, a periodicidade na estimulação, atendendo, assim, a exigência de aleatoriedade do delineamento.

### 3) *Delineamento com grupo de controle não equivalente*

Nesse delineamento, a linha tracejada indica a não equivalência entre os grupos, ou seja, são grupos não aleatórios ou não-equivalentes. Podem ser, por exemplo, classes escolares tão semelhantes quanto a situação permitir, entretanto por não serem equivalentes não se pode dispensar o pré-teste. O pré-teste e o pós-teste ocorrem simultaneamente. Isso ocorre quando a intervenção precisa ser aplicada a grupos pré-existentes que não se pode separar, por exemplo, empresas, escolas, beneficiários de programas sociais.

No delineamento com grupo de controle não-equivalente pode-se controlar história, maturação, testagem, instrumentação, seleção e mortalidade, entretanto permanece a dúvida sobre o controle de regressão em direção à média. As ameaças à validade externa em relação a esse delineamento referem-se à interação da testagem e o tratamento.

Ilustração	$O_1$	$X$	$O_2$
	.....		
	$O_3$		$O_4$

Quando o número de agrupamentos é suficientemente grande pode-se proceder à seleção aleatória para um grupo de controle e um grupo experimental. O agrupamento passa a ser a unidade de análise, e poderá garantir a equivalência probabilística dos grupos. Entre as tentativas de controlar as ameaças a validade do experimento, Campbell & Stanley (1979) desaconselham a distribuição das unidades por agrupamentos inteiros (turmas, escolas), pois poderia produzir nos participantes a sensação de que são *cobaias*. Das diversas possibilidades de tentar atenuar as ameaças a validade interna, considera-se a possibilidade de formar um grupo de controle o mais semelhante possível, nas dimensões relevantes, com o experimental.

### 4) *Delineamento de pré-teste e pós-teste com amostras distintas*

É empregado em situações em que nenhum dos grupos pode ficar sem avaliação, por exemplo, no caso de programas sociais, por determinação da Lei ou por decisão política. Nesse caso, não se pode utilizar o grupo de controle clássico, mas se podem fazer algumas

medições antes da intervenção e depois da intervenção. Então depois se pode comparar o desempenho dos grupos, que tiveram as observações, no caso, não-simultânea. Esse tipo de delineamento está sujeito a problemas de validade interna, como história, maturação, mortalidade e interação da seleção e maturação. Podem ser controlados os fatores de validade externa, interação da testagem e X e interação da seleção e X.

Ilustração	R	O <sub>1</sub>	(X)	
	R		X	O <sub>2</sub>

Caso o pesquisador possa controlar o momento da medição e da aplicação do programa, adiando a intervenção em algumas unidades, transformando-a em grupo de controle, tem-se um delineamento que permite compatibilizar o rigor metodológico com o objetivo político e social (Cano, 2002).

Ilustração	R	X	O <sub>1</sub>
	R		O <sub>2</sub> X

Existem outros delineamentos quasi-experimentais apresentados por Campbell e Stanley (1979). Nesse sentido, recomenda-se a leitura desses autores, pois a apresentação realizada não se constitui em um estudo exaustivo. Os delineamentos apresentados neste trabalho possuem o objetivo de ilustrar os possíveis delineamentos em pesquisa, suas forças e fraquezas, assim como suas possibilidades de aplicação em avaliação de programas, mencionando-se alguns exemplos em alguns casos.

Encerrada a descrição dos delineamentos de pesquisa clássicos, são apresentadas no próximo tópico as contribuições da Psicologia em relação à avaliação de programas, devido a seu amplo estudo e prática por psicólogos sociais e organizacionais envolvidos com a avaliação de programas.

Por fim, além dos aspectos metodológicos é importante destacar a contribuição da psicologia para a avaliação de programas. Nesse sentido, segundo Fernández-Ballesteros, Vedung e Seyfried (1998), a contribuição não se refere somente a questões técnicas e metodológicas, e tem sido definida como um processo para a determinação do mérito, custo ou valor de programas sociais. Esse processo exige uma atividade epistemológica e de avaliação envolve processos cognitivos de julgamento, resolução de problemas e tomada de decisão. A psicologia cognitiva e social, bem como a avaliação psicológica é pertinente por meio de diferentes paradigmas, como os passos e condições exigidos em um processo normativo/avaliativo. A psicologia contribui para a avaliação de programas

testando os processos e etapas de forma mais satisfatória, estabelecendo etapas, para que sejam atingidos os propósitos da avaliação.

Ainda segundo Fernández-Ballesteros e cols. (1998), tal contribuição refere-se a um objeto específico: *programas sociais*. Contudo esse termo é abrangente e pode incluir outros aspectos como: políticas públicas, projetos, intervenções públicas ou privadas, entre outros. Esse conceito geralmente envolve diferentes níveis de complexidade.

A psicologia está envolvida em todas as fases da avaliação de programas (anterior, intermediária e posterior), não somente em programas com metas psicológicas, mas também em outros programas próximos da psicologia, como programas de saúde e de educação; e até mesmo daqueles programas que parecem distantes da psicologia; como programa de emprego, proteção ambiental, despoluição da água, entre outros, nos quais o fenômeno psicológico ou comportamental é a variável interveniente ou causal.

De acordo com Fernández-Ballesteros e cols. (1998), a atividade profissional de avaliação de programa é regulamentada por normas (Comitê de Normas e Padrões para Avaliação, 1994), assegurando a qualidade da avaliação que fornece princípios relacionados à utilidade, viabilidade, compostura do avaliador e exatidão da avaliação. A importância da metodologia na avaliação de programa pode ser checada por essas normas; ressaltando-se que metade delas refere-se à metodologia, por exemplo: descrever objetivos e procedimentos, informação válida, informação confiável, informação sistematizada, análise quantitativa da informação, análise qualitativa, conclusões justificadas, relatório imparcial, entre outros. Esses padrões são exigências comuns de toda pesquisa social científica, inclusive da psicologia, ciência política e economia. Tal ocorre porque a avaliação é administrada por meio de alguns métodos usados pelos cientistas sociais em geral, incluindo entre outros, economistas, sociólogos, educadores, psicólogos, pesquisadores de serviços clínicos de saúde, entre outros, com o objetivo de colecionar evidências confiáveis e válidas.

Por fim, contextualizando a tradição da avaliação de programas a partir da Europa, que possui uma longa tradição de gerenciamento financeiro e de responsabilização no setor público e privado, foi recentemente que a avaliação de programas penetrou no setor público. Há duas razões para tal penetração: os esforços realizados pelos cientistas sociais em suas pesquisas e esforços dos escritórios governamentais, na Europa, em nível nacional e a institucionalização da avaliação de programas na Comissão Européia. Não há dados publicados que permitam examinar a contribuição de diferentes disciplinas sobre avaliação de programas, embora o perfil dos participantes fundadores da Conferência da Sociedade

de Avaliação Européia (EES), seja oriundo de universidades (23%), institutos de auditoria públicos e privados (21%), governo central (20%), escritórios de auditoria nacionais (15%), consultorias e departamentos de contabilidade (12%) e organizações públicas nacionais e supranacionais (6%). Esses dados mostram que o envolvimento da psicologia e dos psicólogos ainda é pequeno (Fernández-Ballesteros & cols., 1998).

Finalizando a Revisão de Literatura sobre AP, espera-se ter contribuído conceitualmente para o conhecimento da avaliação de programas conceitualmente, tanto em termos históricos, quanto do seu delineamento de pesquisa. Após esta etapa, procurou-se alcançar os objetivos deste trabalho por meio da meta-avaliação de auditorias de programas concluídas. Retomando os objetivos descritos inicialmente neste trabalho, esperou-se atingir como *objetivo geral* a identificação de critérios e padrões utilizados nas avaliações do Tribunal de Contas da União e a partir dos mesmos propor contribuições ao modelo de avaliação, conforme descrito no capítulo de Metodologia.

## Capítulo 3: Metodologia

### *Visão geral*

Neste capítulo foram abordados os seguintes tópicos: 1) apresentação do problema de pesquisa; 2) estudo-piloto; 3) delineamento da pesquisa; 4) universo e amostra; 5) procedimento de coleta de dados e 6) procedimento de análise de dados. Foram utilizados diversos procedimentos de análise de dados como: descrição e análise da organização e de suas técnicas; meta-avaliação de acordo com os critérios do *Joint Committee*; análise de conteúdo; síntese de categorias (Ribeiro, 2004) e Síntese de estudos qualitativos.

### *3.1 Problema de pesquisa*

O objeto de estudo desta tese é a auditorias de natureza operacional, seus procedimentos e resultados. Assim, compreender os resultados do processo de trabalho das ANOPs levou à compreensão do que a organização trata como avaliação de programas, visto que concebe a ANOP como uma modalidade de AP. As ANOPs utilizam um conjunto de técnicas de análise, mesclando métodos e técnicas de pesquisa para a coleta e análise dos dados das auditorias. As técnicas utilizadas pela organização são: indicadores de desempenho e mapa de produtos; mapa de processos; análise RECI, marco lógico; análise *stakeholder*; análise SWOT; matriz de verificação de risco e técnica de entrevista.

A natureza exploratória da pesquisa levantou os seguintes problemas:

- A **metodologia avaliativa** do Tribunal de Contas da União, ou seja, a ANOP, é efetiva quanto à validade e confiabilidade de seus achados em relação aos programas sociais auditados?
- A **finalidade** da avaliação configura-se como avaliação ou auditoria?
- **Há regularidade** na forma de conduzir as avaliações nas diferentes áreas, com um processo avaliativo consistente com a **abordagem avaliativa** da organização?

As respostas a essas questões poderiam ajudar a organização a atingir a sua missão de “assegurar a efetiva e regular gestão dos recursos públicos em benefício da sociedade” (Brasil, 2003, p. 54), por meio de resultados válidos, confiáveis e que subsidiem a tomada de decisão dos gestores dos programas sobre as correções de curso necessárias aos Programas.

### 3.2 Estudo-piloto

#### Objetivo

Foi realizado um estudo-piloto que teve como objetivo planejar e validar a metodologia a ser utilizada na pesquisa. Buscou-se por meio do estudo-piloto a verificação da suficiência das fontes de pesquisa e de fonte de dados secundários, necessários para a compreensão do processo de trabalho e da metodologia avaliativa do Tribunal de Contas da União. Nesse sentido, foi verificada a suficiência das fontes.

#### Auditoria Meta-avaliada

Assim, realizou-se a meta-avaliação Programa de Erradicação do Trabalho Infantil – PETI (Brasil, 2002c), da área de governo Assistência Social. O programa foi selecionado por acessibilidade e de forma intencional. Entretanto, a metodologia empregada para o estudo-piloto não se mostrou adequada para explorar todas as informações do processo de ANOP. Assim, o estudo-piloto não foi relatado na íntegra nesta tese. A seguir apresenta-se a descrição dos seus principais achados e a metodologia empregada.

#### Metodologia do Estudo-piloto

Utilizou-se uma análise qualitativa dedutiva. Buscou-se a verificação da existência e desempenho de variáveis-chaves, previamente definidas, foram elas: 1) abordagem avaliativa; 2) metodologia de avaliação; 3) objetivo do programa considerado na avaliação e 4) usos da avaliação. Essas variáveis estão sintetizadas no quadro a seguir.

#### Quadro 3. Variáveis do estudo-piloto

Variáveis		Definição
1. Abordagem Avaliativa		Tipo de abordagem utilizada na avaliação. Pode ser: quantitativa/experimental; qualitativa ou naturalista e abordagem pluralista.
2. Metodologia de avaliação	Método	Estudo de caso, <i>survey</i> , delineamento experimental, quasi-experimental e não experimental.
	Critério de seleção dos sujeitos	Critério intencional e seleção aleatória.
	Natureza dos dados	Quantitativo, Qualitativo ou a combinação de ambos.
	Instrumento de coleta de dados	Entrevistas (estruturadas, semi-estruturadas); questionários; fontes de dados secundários (índices, documentos internos da organização), observações <i>in loco</i> e testes.
	Tratamento dos dados	Análises estatísticas (descritiva, indicadores, entre outros) Análises qualitativas (análise de conteúdo, de discurso, categorização, entre outros).
Resultados		Natureza dos resultados: psicológica, econômica, socioeconômica, responsabilização e modificação do processo de trabalho.

(cont.) Quadro 3. Variáveis do estudo-piloto

Variáveis		Definição
3. Objetivos	Implantação	Descrevem as atividades necessárias para que o programa esteja instalado e funcionando, para posteriormente atingir os resultados.
	Intermediários	Descrevem conseqüências da implantação do programa.
	Finais	Descrevem os resultados finais que se espera alcançar com o programa. Refletem a modificação da linha de base, antes da instalação do programa, ao momento da avaliação.
4. Usos da avaliação		Desenvolvimento de um racional para a ação. Legitimidade e responsabilidade social da ação organizacional. Estudo de processos organizacionais. Estudo de modelos de avaliação.

#### *Principais resultados da meta-avaliação do PETI*

De acordo com a meta-avaliação do PETI, verificou-se que a ANOP possui características mais próximas à auditoria de programas, com forte enfoque sobre a responsabilização dos gestores e responsáveis nas recomendações finais. Os resultados da avaliação não aprofundaram as análises realizadas e não esclareceram aspectos importantes, como o impacto no resultado último do programa, ou seja, a erradicação do trabalho infantil.

#### *Sugestões em relação aos resultados encontrados na meta-avaliação da auditoria do PETI*

- 1) Sugere-se que para uso interno da organização os pontos identificados sejam revistos para que haja maior efetividade no processo de avaliação dos programas quanto às recomendações oferecidas às organizações avaliadas.
- 2) Outro aspecto a destacar é a necessidade de clareza metodológica dos termos empregados para a divulgação do trabalho de auditoria realizada. Equívocos na concepção de “estudo de caso” precisam ser evitados. No referido caso foi utilizada pesquisa por amostragem, configurando-se em um estudo do tipo *survey*. A denominação de entrevista na auditoria do PETI (Brasil, 2001c), na referência aos instrumentos utilizados para a coleta de dados, na verdade caracterizou-se como questionário enviado aos respondentes pelo correio.
- 3) Sugere-se que os dados coletados devido ao seu alcance e alta taxa de retorno fossem convertidos em indicadores de desempenho e apresentados à organização avaliada, sugerindo-se formas de construí-los.

- 4) Em relação aos *objetivos da auditoria* sugere-se que sejam trabalhados em função do estágio de evolução do programa, para que não haja ênfase em objetivos intermediários somente, não chegando a uma avaliação global dos efeitos finais do programa, a saber, a erradicação do trabalho infantil (Brasil, 2002d; Brasil 2002e; Brasil, 2004b).
- 5) Em relação aos usos da avaliação, considerando-se que os objetivos de uma avaliação pública e de uma avaliação acadêmica são distintos, as afirmações feitas em relação aos resultados encontrados precisam ser cuidadosas, pois no caso do delineamento não experimental utilizado a validade externa é baixa ou inexistente. Isso significa que a generalização dos resultados para toda a população atendida pelo programa foi comprometida. Pode ter ocorrido interferência de tratamentos múltiplos ou interação, visto que a população atendida era alvo de outros programas governamentais. Para que essa limitação seja minimizada, em auditorias futuras, devem ser realizadas medições antes da implantação do programa, e quando não for possível, deve ser usado pelo menos um grupo de controle e os participantes selecionados de forma aleatória. Enfim, deve-se tentar controlar as ameaças à validade interna e externa das avaliações empreendidas.

### ***3.3 Delineamento da pesquisa***

O delineamento do presente trabalho configurou-se como uma pesquisa de natureza exploratória, pois investigou e analisou a metodologia avaliativa utilizada pela organização e o tipo de suporte oferecido pelas ANOPs aos programas auditados nas diferentes áreas de governo.

**O delineamento desta pesquisa na área de avaliação de programas** foi a **meta-avaliação**. Esse delineamento implica que os resultados da auditoria, objeto da meta-avaliação, não estão necessariamente relacionados à falhas, e sim à busca de melhoria e compreensão do processo e da prática de auditoria. Sendo que a meta-avaliação foi estimulada pelo interesse de investigação acadêmico e prático na área de AP, com a proposição da melhoria de um processo de trabalho existente.

### **3.4 Universo e Amostra**

Estão disponíveis para consulta na página eletrônica do TCU, entre os anos de 1999 e 2005 (posição em janeiro de 2006), 48 avaliações. Tais avaliações perpassam dez áreas de governo: agricultura, assistência social, cidadania, comércio e serviços, educação, energia, habitação, meio ambiente, saúde e trabalho. Cada área possui, por ano, quantidades diferentes de avaliações de programas. As avaliações escolhidas abrangem todas as áreas para que fosse possível assegurar melhor entendimento da auditoria e seus procedimentos.

A base empírica da meta-avaliação da pesquisa final foram **10 auditorias concluídas pela organização, sendo uma de cada área de governo, ou seja, 10 auditorias de governo**. A amostra é estatística e intencional, o total de avaliações foi escolhido de forma a assegurar a representatividade de todas as áreas de governo. Tal escolha se justifica pelo cuidado em não retirar uma amostra aleatória que poderia causar viés, como por exemplo, avaliações somente de uma ou duas áreas, caracterizando de forma parcial os procedimentos avaliativos utilizados pela organização nas dez áreas. Portanto, a amostra é estratificada e representativa do universo estudado.

Assim, utilizou-se esse critério para obter uma visão abrangente e completa de como as estratégias metodológicas são utilizadas nas diferentes áreas de governo. Para a meta-avaliação das dez ANOPs realizadas após a qualificação do projeto de doutorado, escolheu-se como marco zero, o ano de 1999, a partir do qual, havia dados disponíveis para a consecução da pesquisa.

#### **3.4.1 Descrição da Amostra**

As dez ANOPs que formaram a pesquisa final, escolhidas intencionalmente dentro de cada área apresentam, a homogeneidade necessária para a formação do *corpus*, apontada por Bardin (1977) para que ao final seja possível traçar comparações entre os resultados individuais e globais. A amostra da meta-avaliação está caracterizada no Quadro 4, no qual se apresenta o nome da ANOP, a área a qual pertence o programa, o ano de realização da ANOP e o total de páginas dos documentos analisados.

#### Quadro 4. Auditorias de Programas Meta-avaliados

ANOP	Área	Período	Total de páginas
1 - Programa Novo Mundo Rural - Consolidação de Assentamentos	Agricultura	2004	Planejamento = 75 Relatório = 65
2 - Programa Combate ao Abuso e à Exploração Sexual de Crianças e Adolescentes	Assistência Social	02 a 13/08/2004	Planejamento = 61 Relatório = 98
3 - Programa de Reinserção Social do Adolescente em Conflito com a Lei	Cidadania	03/10/2003 07/11/2003	Planejamento = 87 Relatório = 104
4 - Programa Desenvolvimento do Turismo no Nordeste - Prodetur/NE	Comércio e Serviços	2003	Planejamento = 60 Relatório = 96
5 - Programa Nacional do Livro Didático - PNLD	Educação	2000	Planejamento = 13 Relatório = 11
6 - Energia das Pequenas Comunidades	Energia	2004	Planejamento = 35 Relatório = 60
7 - Programa Morar Melhor	Habilitação	2004	Planejamento = 77 Relatório = 138
8 - Programa de Desenvolvimento Sustentável de Recursos Hídricos para o Semi-Árido Brasileiro	Meio Ambiente	2001	Planejamento = 31 Relatório = 113
9 - Monitoramento e Prevenção da Mortalidade Materna	Saúde	2001	Planejamento = 14 Relatório = 16
10 - Programa Novo Emprego e Seguro-Desemprego	Trabalho	2002	Planejamento = 39 Relatório = 13
Dez Auditorias de Programas	10 áreas	2000-2004	1.206 páginas

Características que descrevem o escopo das auditorias da organização foram sintetizadas no Quadro 5 são elas: critério de seleção do programa, público-alvo do programa, equipe de auditoria, período abrangido pela auditoria e período de realização do planejamento e execução da auditoria.

A organização segue os princípios do Termo de Cooperação com o Reino Unido para a realização das ANOPs, e visa contribuir para a melhoria de desempenho de instituições governamentais brasileiras, assim, como a melhor utilização dos recursos públicos.

As auditorias da organização fundamentam seus procedimentos nas premissas do Projeto CERDS - Projeto de Aperfeiçoamento do Controle Externo com Foco na Redução da Desigualdade Social (Brasil, 1999a; Brasil, 1999b; Brasil, 2000d; Brasil, 2000e; Brasil, 2000f; Brasil, 2001c).

De acordo com as premissas do Termo de Cooperação com o Reino Unido e do Projeto CERDS, a organização seleciona os programas a serem auditados e os submete ao estudo de viabilidade (Bastos, 2004). A seleção envolve a importância da auditoria para a imagem da organização e temas que chamam a atenção da mídia, da sociedade civil organizada e do Congresso Nacional. Além disso, considera-se a possibilidade do impacto favorável da auditoria na melhoria do desempenho do programa (Brasil, 1992).

Conforme sumarizado no Quadro 5, nos programas auditados o critério de seleção do TCU em três auditorias foi relevância, risco e materialidade; e integrante do Projeto

CERDS característica de outras três auditorias. Ações do programa previstas na Lei de Diretrizes Orçamentárias – LDO foi critério de escolha de três auditorias de programas. Ações do programa contempladas no Plano Pluri-anual foi critério de escolha de programas para as auditorias.

Quanto ao público-alvo dos programas auditados foi característica comum, por serem programas sociais, os segmentos desfavorecidos da sociedade, incluindo crianças e adolescentes (dois programas); adolescentes (um programa). Uma auditoria que se diferenciou das demais foi a do Prodetur/NE, que beneficia o setor turístico, segmentos A e B da população, mas indiretamente gera empregos na região para a classe C.

Já os períodos de funcionamento do Programa são extensos, variando de três a nove anos. Apesar da limitação de tempo, citada nos relatórios, para a realização do planejamento, execução da auditoria e confecção do relatório final, verificou-se na amostra meta-avaliada que o período variou no total, entre planejamento e execução da auditoria, de no mínimo 9 e no máximo 39 dias úteis. Verificou-se que a amplitude e abrangência das etapas do processo das ANOPs exigem maiores períodos de tempo para sua realização, e de acordo com citações dos relatórios de ANOP – maior quantidade de auditores para a realização do trabalho de auditoria.

#### **Quadro 5. Descrição dos programas auditados pelo TCU**

<i><b>Programas auditados</b></i>	<i><b>Características do Programa e da auditoria</b></i>				
	Critério de seleção do programa	Público-alvo do programa	Equipe de auditoria	Período abrangido pela auditoria	Período de planejamento e execução da auditoria
Programa Morar Melhor	Relevância, risco, materialidade e impacto na redução da pobreza PPA 2000/2003 Parte do Projeto CERDS	Famílias de baixa renda, com até 3 salários mínimos submetidas a condições precárias de moradia	4 pessoas	2000 a 2003	04/ago a 05/set/03 (24 dias úteis)
Programa de Desenvolvimento de Turismo do Nordeste Prodetur/NE	Relevância, risco e materialidade. Prioritário para LDO 2003.	Fortalecimento da capacidade municipal para a gestão do turismo. Planejamento Estratégico, treinamento e infra-estrutura para o crescimento turístico. Promoção e investimento do setor privado	7 pessoas (4 subequipes)	1994-2003	10/08/03 a 29/08/03 (15 dias úteis)

(cont.) Quadro 5. Descrição dos programas auditados pelo TCU

<b>Programas auditados</b>	<b>Características do Programa e da auditoria</b>				
	Critério de seleção do programa	Público-alvo do programa	Equipe de auditoria	Período abrangido pela auditoria	Período de planejamento e execução da auditoria
Programa Reinscrição Social do Adolescente em Conflito Com a Lei	Parte do projeto CERDS	Adolescentes em conflito com a lei	5 pessoas	jan/2000 a julho/2003	04/08 a 05/09/03 (24 dias úteis)
Programa Novo Emprego e Seguro-Desemprego	Orientação da LDO para apreciação de contas do governo da república referentes ao exercício financeiro de 2000	Desempregado dispensado sem justa causa; que recebeu salários no período de seis meses anteriores à data da dispensa; empregado de pessoa jurídica ou física equiparada á jurídica durante pelo menos seis meses nos últimos 36 meses que antecedem a data da dispensa.	5 pessoas	2000 a 2001	23/10 a 05/12/2000 (5 dias úteis para planejamento, 29 para trabalho de campo, 5 dias para confecção do relatório)
Programa Pró-Água Semi-Árido	Projeto de Cooperação Técnica Reino Unido	Populações rurais e das cidades de pequeno e médio porte da região semi-árida do NE e dos estados do MA e MG. População com renda <i>per capita</i> anual de aproximadamente US\$230,00	7 pessoas	Não consta nos relatórios	04/09 – 13/10/00 (27 dias úteis)
Programa Combate ao Abuso e a Exploração Sexual de Crianças e Adolescentes	relevância, risco e materialidade	Crianças e adolescentes vítimas de violência, abuso e exploração sexual, bem como suas famílias	5 pessoas	2001 a 2003	02/08 a 13/08/04 (9 dias úteis)
Programa Energia das Pequenas Comunidades	Previsto no PPA 2000-2003	Estimado em 20 milhões de brasileiros sem acesso à energia elétrica distribuídos em mais de 100 mil comunidades, além de 58 mil escolas e 3 mil comunidades indígenas.	18 pessoas	1996 a 2002	13/09/0205/08 a 13/09/2002 (28 dias úteis)
Programa Novo Mundo Rural	Relevância, risco e materialidade. Parte do projeto CERDS	Família de trabalhadores assentados	5 pessoas	2000-2003	Planejamento: 04/08-05/09 (33 dias) Execução: 29/09 a 10/10 (12 dias) + 12 dias elaboração do relatório

(cont.) Quadro 5. Descrição dos programas auditados pelo TCU

<i>Programas auditados</i>	<i>Características do Programa e da auditoria</i>				
	Critério de seleção do programa	Público-alvo do programa	Equipe de auditoria	Período abrangido pela auditoria	Período de planejamento e execução da auditoria
Programa Nacional do Livro Didático – PNLD	Projeto de Cooperação Técnica Reino Unido. Relevância social. Volume de recursos Mensuração da eficiência, eficácia e economicidade	Todos os alunos matriculados em escolas públicas do ensino fundamental de 1ª a 8ª séries	4 pessoas	1997-1999 incluindo planejamento do programa até 2002	Não consta nos relatórios
Programa Monitoramento e Controle da Mortalidade Materna	Projeto de Cooperação Técnica Reino Unido. Relevância social - instauração da Comissão Parlamentar de Inquérito CPI da mortalidade materna em 27/04/2000	Gestantes, mães, nutrizes	5 pessoas	Não consta nos relatórios	03-18/08/2000 (11 dias úteis planejamento) 04/09-13/10/2000 (28 dias úteis período de execução)

### *3.5 Procedimento de coleta de dados*

A presente meta-avaliação empreendida caracterizou-se pelo exame do método da ANOP e dos seus procedimentos para chegar aos resultados e conclusões. A re-análise considerou aspectos como metodologia, seleção dos sujeitos, objetivos, critérios adotados, resultados e recomendações. Foram analisadas as etapas de planejamento, execução e resultados. A auditoria do Programa de Erradicação do Trabalho Infantil foi meta-avaliada no estudo-piloto e dez auditorias de programas foram meta-avaliadas no trabalho final. Utilizou-se documentos de circulação restrita, como relatório final e relatório de planejamento da auditoria. Os relatórios finais foram obtidos após solicitação formal escrita à organização.

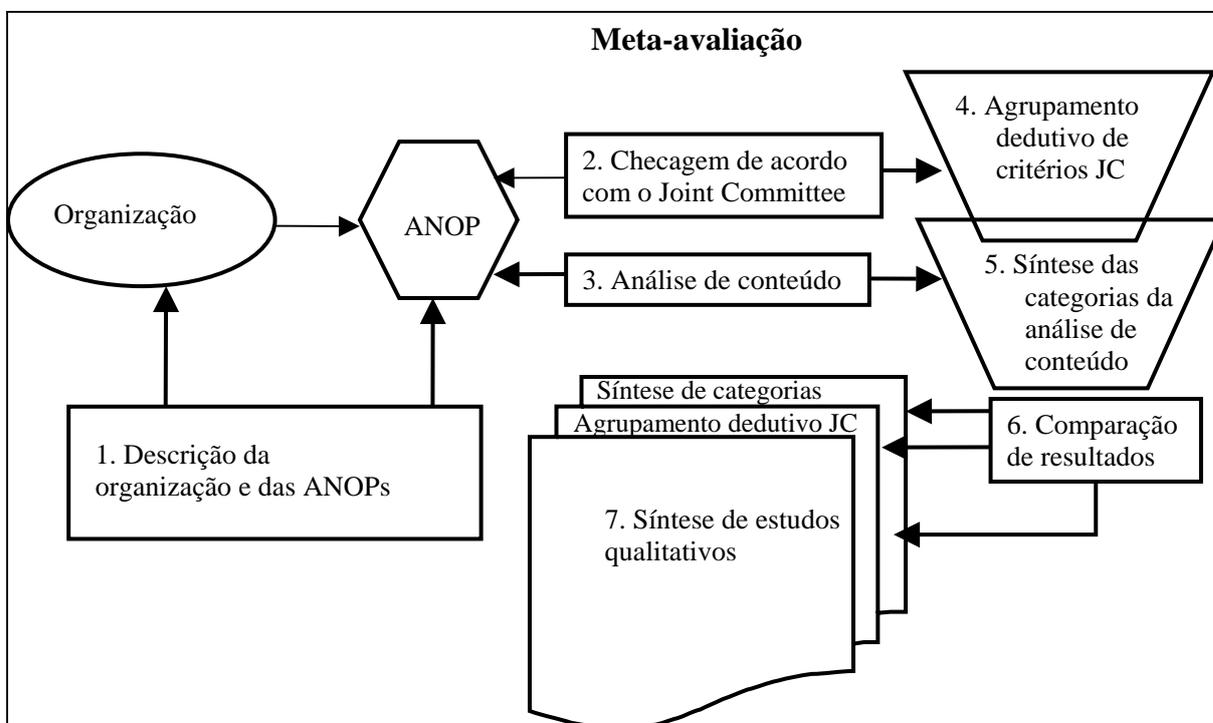
### 3.6 Procedimento de análise dos dados

Foram realizados cinco procedimentos para a análise de dados da meta-avaliação: 1) descrição e análise da organização e das técnicas utilizadas; 2) checagem de critérios de acordo do *Joint Committee*; 3) análise de conteúdo; 4) síntese de categorias; 5) agrupamento dedutivo dos critérios do *Joint Committee* e 6) síntese dos estudos qualitativos.

Esclarece-se que a **análise por juízes** foi realizada somente para a análise de conteúdo. Após a criação de 19 categorias, de forma indutiva, pela autora, estas foram submetidas a três juízes para análise semântica. A síntese de categorias da análise de conteúdo contou com a participação de um juiz para analisar e opinar sobre os agrupamentos das categorias.

A seguir é apresentada a ilustração das etapas dos procedimentos de análise que compuseram meta-avaliação.

**Figura 4. Procedimento de análise de dados**



1. Descrição da organização e das ANOPs

A organização e as técnicas de auditoria foram descritas com base na leitura e resumo de documentos internos da organização. Os principais documentos consultados foram: relatório de planejamento das auditorias e relatório final.

Tais documentos possuem tópicos semelhantes e tópicos específicos. Os relatórios de planejamento estruturam-se da seguinte forma: *descrição do programa; questões de auditoria e planejamento da auditoria*. A matriz de planejamento das auditorias dos programas possui oito colunas de informações: *informações requeridas; fontes de informação; estratégia metodológica; coleta de dados e análise de dados; limitações e o que a análise vai permitir dizer*.

Os relatórios finais estruturam-se da seguinte forma: *descrição do programa, perguntas de auditoria, descrição dos resultados, recomendação e acórdãos*. As perguntas de auditoria são norteadores do trabalho de planejamento da auditoria e são criadas com base nas informações levantadas na fase de planejamento da auditoria. O acórdão é a decisão final da organização, na qual o Ministro Relator da Matéria determina quais recomendações deverão ser cumpridas pelo programa auditado.

## 2. Checagem de critérios de acordo com *Joint Committee*

Para proceder à meta-avaliação construiu-se um quadro com base nos critérios a ser contemplados numa avaliação conforme orientação do *Joint Committee*.

Os critérios do *Joint Committee* foram abordados em Stufflebeam e Shinkfield (1987), e estão detalhados no tópico 2.2.1 *Modalidades de Avaliação*. Não se utilizaram todos os 30 critérios do *Joint Committee*, pois há questões não passíveis de resposta somente com base na leitura de fonte de dados secundários (relatório de planejamento de auditoria e o relatório final).

Escolheram-se critérios referentes à metodologia da pesquisa avaliativa, para a qual não seria necessário consultar *experts* em temas de domínios específicos dos programas. Entretanto, buscaram-se na literatura artigos que tratassem de avaliação de programas semelhantes aos meta-avaliados. Em alguns casos, encontrou-se o tópico na literatura e foram feitas comparações dos achados.

Cabe destacar que algumas questões específicas dos critérios do *Joint Committee* como as questões sobre Normas sobre Ética, no que se refere à assinatura de um documento escrito, contendo as obrigações formais das partes da avaliação, não foi incluída; por ser a auditoria um tipo de avaliação externa independente e autônoma. Entretanto, o estabelecimento deste acordo entre as partes – organização e programa

poderia ser um elemento a ser inserido na fase de planejamento da auditoria. Tal comprometimento inicial poderia ser um estímulo para a adoção das recomendações, assim como um fator facilitador da criação de grupos de contato de auditoria, recomendação recorrente em todas as auditorias analisadas.

Para diferenciar a forma do uso dos critérios do *Joint Committee*, indo além da indicação do uso ou não do critério (marcação de sim e não), conforme foi encontrado no exemplo de Worthen, Blaine, Sanders, Fitzpatrick e Jody (2004), adotou-se uma escala com quatro pontuações. Na escala “zero” significa não se aplica; “não” significa que o critério não foi utilizado; “parcialmente” significa que foi aplicado, mas não completamente ou não atendeu completamente aquele critério, e “sim”, significa que foi adequadamente utilizado.

Quanto à análise da Questão 2, sobre o treinamento adequado da equipe para a realização da auditoria, considerou-se as pontuações recebidas nas demais 19 questões, foi analisado que tratam do uso adequado e pertinência de procedimentos metodológicos, tipos de análises, resultados, conclusões e recomendações produzidas. Consideraram-se equipes com treinamento adequado, aquelas que produziram pontuações positivas (sim = 3), em pelo menos 70% das questões (13 questões); com treinamento parcial pontuações entre 50 e 69 %, (entre 9 e 12 questões) e sem treinamento pontuações maiores ou iguais a 49% dos itens (menos de 9 questões).

Na análise da Questão 3, considerou-se que houve participação quando consultores foram contratados ou consultados, como no caso de consulta a especialistas. Quando a participação foi somente em uma fase (planejamento ou execução) a participação foi parcial e “não houve participação” quando não houve contratação ou consulta a especialistas.

Na análise da questão 18 as medidas para garantir o mínimo de erros referem-se a aplicação de teste-piloto, adoção de grupo de controle, coleta antes e depois, escolha aleatória da amostra, ou outro procedimento que sirva de controle as ameaças à validade interna e externa à pesquisa avaliativa.

As demais questões foram respondidas de acordo com a escala de 1 a 3, onde: 0 = não se aplica, 1 = não; 2 = parcialmente e 3 = sim.

### *3. Análise de conteúdo*

Para analisar os documentos, utilizou-se a técnica de análise de conteúdo, por meio da análise temática, conforme proposta metodológica de Bardin (1977). A análise de

conteúdo é uma técnica de análise de textos que visa obter por procedimentos sistemáticos e objetivos, temas recorrentes agrupados para compor uma categoria empiricamente definida, que permita a interpretação de conhecimentos relativos ao objeto de pesquisa.

De acordo com Bardin (1977), o processo de *análise temática* é uma das técnicas da análise de conteúdo. A análise temática foi empreendida através dos seguintes passos:

- a) Leitura flutuante das partes comuns dos relatórios. Por exemplo, leitura da apresentação ou da descrição do programa de cada relatório analisado.
- b) Estabelecimento de regras de contagem de recorrências das palavras ou frases fundamentadas no tema.
- c) Análise temática para identificar os núcleos de sentido das frases do texto e cuja presença ou frequência de aparição pode possuir significação para as perguntas de pesquisa.

Assim, utilizou-se a análise categorial temática, ou seja, o desmembramento do texto dos relatórios em unidades e posterior organização de categorias com as respectivas frequências de citação dos temas.

Os critérios para a criação das categorias e dos temas foram: exaustividade, representatividade, homogeneidade, pertinência e exclusividade (Bardin, 1977). As categorias e temas foram submetidos à análise de *três juízes* para a análise semântica.

A regra da exaustividade refere-se à leitura de todo o material coletado, sem exceção de nenhuma parte do *corpus*. Assim, foram lidos na íntegra os relatórios de planejamento e relatórios finais. A representatividade foi assegurada, escolhendo-se uma auditoria de cada área de governo. A homogeneidade foi obtida através de uma característica comum dos programas auditados, trata-se de programas sociais. A pertinência foi assegurada obtendo-se dois documentos de cada auditoria de programa, o seu planejamento e o relatório de sua execução (relatório final). A exclusividade foi obtida por meio de temas que são os únicos pertinentes a cada categoria. Quanto à frequência dos temas, consideraram-se como temas as recorrências iguais ou maiores que três. O conjunto de informações analisadas deu origem às categorias criadas de forma indutiva.

O procedimento de tabulação dos dados envolveu digitação de grande parte dos materiais (textos). A leitura flutuante originou a percepção de temas que se sobressaíam dos significados das frases (Bardin, 1977). As frases foram classificadas por meio do editor de texto, após a verificação da similaridade de significados dos temas foram criadas categorias. Assim, os temas deram origem a categorias, que foram definidas a partir do significado temático dos textos.

As categorias foram criadas de forma indutiva, exceto a última categoria (categoria 19. Objetivos do Programa), que foi criada de forma dedutiva. Tal diferenciação justifica-se em função do pressuposto segundo o qual classificar e entender os objetivos do programa é tarefa indispensável ao planejamento da avaliação do programa ou da auditoria de um programa social. A classificação dos objetivos adotada foi a de Posavac e Carey (2003). Pois, a partir desta, é possível verificar o estágio de implantação do programa e verificar as condições de avaliabilidade do Programa.

#### *4. Agrupamento dedutivo de critérios Joint Committee*

Com o objetivo de demonstrar a existência de áreas de interseção entre os achados das diferentes análises qualitativas; comparou-se, através dos aspectos metodológicos o conjunto de 20 questões do *checklist* com as 19 questões que formaram os cinco agrupamentos da Síntese de Categorias. Os 20 critérios do *Joint Committee* verificados nas avaliações foram agrupados em cinco tópicos semelhantes ao agrupamento da Síntese de Categorias. Assim, as questões foram agrupadas conforme demonstrado na Figura 8. Comparação de estudos Qualitativos: conceitos.

#### *5. Síntese de categorias da análise de conteúdo*

A síntese de categorias foi elaborada conforme o procedimento adotado por Ribeiro (2004). O procedimento consiste em reagrupar as categorias em eixos temáticos que reflitam a Teoria e o Problema de Pesquisa abordados e emergentes a partir da análise de conteúdo.

Na presente pesquisa resultaram 19 categorias, de acordo com o procedimento clássico de Bardim (1977), que foram reagrupadas em cinco Categorias Sínteses. A justificativa para o uso desse procedimento foi a ampliação da compreensão dos resultados, comparação entre as categorias e organização facilitadora para a apreensão das informações geradas. Assim, as categorias sínteses facilitaram a apresentação e condensação de resultados, fornecendo uma visão ampla do significado dos dados no contexto da meta-avaliação.

#### *6. Comparação de resultados*

Procurou-se estabelecer uma metodologia de análise comparativa. Primeiro, foram organizados dedutivamente os 20 critérios do *Joint Committee* para compará-los com a síntese das categorias. Foram organizados sob os mesmos títulos da Síntese de Categorias,

para assinalar a correspondência entre o que foi avaliado enquanto critério *Joint Committee* e o que foi encontrado na análise de conteúdo.

Para facilitar a comparação o conjunto de pontos dos 20 critérios do *Joint Committee* (0 = não se aplica, 1 = não; 2 = parcialmente e 3 = sim) foram transformados nas seguintes categorias de avaliação:

EXCELENTE  $\geq 90 \leq 100$  de pontos atribuídos como SIM (3), ou seja, entre 9 e 10 pontos.

BOM  $\geq 70\% \leq 80\%$  de pontos atribuídos como SIM (3), ou seja, entre 7 e 8 pontos.

PARCIAL  $\geq 50\% \leq 60\%$  de pontos atribuídos como SIM (3), ou seja, entre 5 e 6 pontos.

INSUFICIENTE  $\leq 50\%$  de pontos atribuídos como SIM (3), ou seja, até 4 pontos.

Esse resultado (excelente, bom, parcial, insuficiente) foi interpretado individualmente dentro de cada agrupamento, com referência as frequências da síntese de categorias.

### *7. Síntese de Estudos Qualitativos*

Na síntese dos estudos qualitativos, foram resumidos e cruzados os resultados dos agrupamentos das questões do *Joint Committee* com os resultados da Síntese de Categorias, considerando-se, as pontuações das auditorias no *checklist* do *Joint Committee*, com a síntese da análise de conteúdo e suas frequências de recorrência.

Assim, os resultados foram discutidos em relação aos achados específicos de cada um, e também em relação às semelhanças e diferenças entre os agrupamentos dos mesmos.

A interpretação dos resultados foi realizada mediante a congruência das explicações dos resultados das duas análises mediante suposições básicas sobre avaliação de programas e auditorias, demonstradas na revisão de literatura. A síntese interpretativa final corresponde às conclusões da presente meta-avaliação.

## Capítulo 4: Resultados e Análise

Neste capítulo foram apresentados os resultados e as análises seguindo-se os procedimentos: 1) descrição da organização e das técnicas; 2) checagem de acordo com os critérios do *Joint Committee*; 3) análise de conteúdo; 4) síntese de categorias da análise de conteúdo; 5) agrupamento dedutivo dos critérios do *Joint Committee*; 6) comparação dos achados e 7) síntese de estudos qualitativos.

### 4.1 Descrição do processo de trabalho de ANOP no TCU

Em relação à descrição da organização, priorizou-se o processo de trabalho de *auditoria de natureza operacional* foco da meta-avaliação.

Freitas (2005) explica em seu trabalho a introdução e o desenvolvimento da ANOP na organização, por meio da teoria institucional. Nesse sentido, o autor argumenta que a introdução da auditoria operacional no TCU seria um fenômeno isomórfico, aquele onde ocorre a homogeneização das práticas, processos e estruturas através de um processo de legitimização. Conforme Freitas (2005, p. 1),

*A auditoria operacional constitui uma atividade de auditoria que difere do conceito tradicional relacionado à auditoria contábil ou financeira. Enquanto a auditoria contábil objetiva emitir um parecer sobre a adequação das demonstrações contábeis e financeiras de organizações, a auditoria operacional está direcionada para a avaliação de desempenho das operações, sistemas de informação e organização, métodos de administração e políticas administrativas.*

Conforme Mendes et al. (1999) a organização possui uma função social associada ao controle e zelo da “coisa pública”, seus aspectos patrimoniais e econômicos sobre a administração pública.

A organização utiliza a auditoria de natureza operacional e a avaliação de programa de forma complementar e refere-se à avaliação de programa como uma modalidade da auditoria de natureza operacional. Figuram nos relatórios e sumários executivos os termos avaliação e auditoria de forma intercambiável, referindo-se ora a uma ou a outra prática.

Para o TCU, as modalidades de avaliação são: a auditoria de desempenho operacional e a avaliação de programa (Brasil, 2000c). Assim, a definição de auditoria de natureza operacional e avaliação de programa seriam “[...] uma modalidade de controle que consiste

---

<sup>3</sup> Foi omitido o nome da organização.

na avaliação sistemática dos programas, projetos, atividades e sistemas governamentais. Abrange duas modalidades, a auditoria de desempenho operacional e a avaliação de programa” (Brasil, 2000c, p. 15). A auditoria de desempenho operacional objetiva examinar a ação do governo quanto aos aspectos da economicidade, eficiência e eficácia, buscando o grau de cumprimento das metas, pela comparação das metas previstas com metas realizadas, enquanto a avaliação de programas centra-se na efetividade.

Para a *International Organization of Supreme Audit Institutions* – INTOSAI a avaliação de programas distingue-se da auditoria de desempenho operacional porque é mais ambiciosa em relação à análise dos objetivos e dos resultados da intervenção governamental. A avaliação de programas utiliza estratégias metodológicas raramente utilizadas em auditoria de desempenho operacional, tais como *survey* e delineamentos quasi-experimentais (Brasil, 2000c).

Conforme Brasil (2000c, p. 109), a efetividade refere-se à “relação entre os produtos (bens e serviços) gerados por uma atividade e os custos dos insumos empregados em um determinado período de tempo. O resultado expressa o custo de uma unidade de produto final em um dado período de tempo. Analogamente ao conceito de eficácia, pode-se estabelecer o grau de eficiência relativa de um programa, projeto ou atividade, comparando-o com as metas e custos programados ou com dados de outras iniciativas que perseguem os mesmos objetivos. Nesse caso, somente serão comparáveis iniciativas que, além dos objetivos, possuam as mesmas características de programação”. Entretanto, em Brasil (2003a, p. 38), lê-se “Os objetivos da avaliação são verificar os padrões de economicidade, eficiência e eficácia empregadas no desempenho dos órgãos e entidades jurisdicionais e nos sistemas, ações e programas do governo federal [...]”. Ou seja, utilizou-se a definição de finalidade da auditoria de desempenho organizacional para a avaliação de programa.

Os objetivos dessa organização dizem respeito às auditorias, por meio das quais pretende “contribuir para a solução dos problemas sociais identificados” (Brasil, 2003a, p. 54), sendo que sua missão é “assegurar a efetiva e regular gestão dos recursos públicos, em benefício da sociedade” (Brasil, 2003a, p. 54). A avaliação de programas é considerada pelos governos estado-unidense, canadense, governo do Reino Unido e australiano, como uma modalidade de controle (Brasil, 2003a). Nesse sentido, apesar de o Tribunal de Contas da União caracterizar sua atuação como auditoria e avaliação de programas, no entanto, verificou-se que as suas atividades se caracterizam principalmente como auditoria, visto que estão ligadas à própria natureza da organização na modalidade de controle e

fiscalização. De acordo com Freitas (2005) que investigou a origem e desenvolvimento da ANOP no TCU, o tribunal realiza as ANOPs como auditoria de desempenho. Porém, encontraram-se nos documentos analisados, definições que sustentam que perpassa na ANOP a concepção de auditoria de conformidade, paralelamente ao de auditoria de desempenho. Auditoria de conformidade seria aquela relacionada à regularidade, legalidade, cumprimento da legislação, controles internos e responsabilização do gestor do programa.

As funções do Tribunal de Contas da União podem ser compreendidas como funções: fiscalizadora, consultiva, informativa, judicante, sancionadora, corretiva, normativa e de auditoria, podendo também assumir caráter pedagógico. Para realizar a sua missão, o Tribunal de Contas da União dispõe de uma Secretaria Técnico-administrativa, à qual estão subordinadas unidades técnico-executivas sediadas em Brasília e nos demais Estados da Federação (Brasil, 2003). A avaliação de programa não se restringe à eficiência operacional, conforme destacado anteriormente, “[...] o programa busca apurar em que medida as ações implementadas lograram produzir os efeitos pretendidos pela administração” (Brasil 2000c, p.17). Entende-se que a definição, de acordo com a INTOSAI, enfatiza a diferença entre AP e auditoria, entretanto, a organização considera a avaliação de programa como uma modalidade de auditoria de natureza operacional, afirmando que AP é parte da auditoria. Esse uso intercambiável é ambíguo, pois há diferença entre as duas abordagens e, conforme enfatizado, a auditoria é uma modalidade de controle que atende ao objetivo principal da organização: “(...) julgar as contas dos administradores e demais responsáveis por dinheiros, bens e valores públicos federais, bem como as contas de qualquer pessoa que der causa a perda, extravio ou outra irregularidade de que resulte prejuízo ao erário” (Brasil, 2003a, p. 9).

No trabalho sobre diagnóstico da cultura da organização, realizado por Mendes et. al. (1999, p. 14), a partir da Teoria do Poder nas Organizações, que possui a premissa “que o comportamento organizacional é um jogo de poder no qual, vários jogadores e influenciadores procuram controlar as ações organizacionais”, verificou-se que a organização apresenta uma configuração predominantemente de Autocracia, porém, com desenvolvimento da configuração de Arena Política. A Autocracia mostra que há centralização de poder nos altos chefes da organização e a Arena Política “forças de grupos ou mudanças no ambiente externo que pressionam a organização e estimulam a instabilidade” (p.15), essas mudanças podem estar associadas a mudanças no

funcionamento organizacional, no caso presente também no processo de realização de auditoria.

Conforme Freitas (2005), há na organização um conflito entre a adoção e a prática da auditoria de conformidade *versus* auditoria de natureza operacional. Esse conflito se dá desde a realização da primeira ANOP na Empresa Brasileira de Pesquisa e Agropecuária - EMBRAPA, no início dos anos 1990, quando o relatório de auditoria foi bem aceito na empresa, “constituindo-se fonte de legitimidade pragmática em relação à organização auditada e, ao mesmo tempo, objeto de crítica interna [na própria organização] quanto à atuação [...]” (Freitas, 2005, p. 51). Assim, iniciou-se a polêmica interna em relação ao uso da ANOP. Deve-se destacar que, segundo o autor, a auditoria operacional, com os postulados de economicidade, eficiência e eficácia foi institucionalizada pela Constituição Federal em 1988, auferindo competência legal à organização para desenvolver essa atividade. Inicialmente, as auditorias operacionais sofriam rejeição pelo tamanho de seus relatórios extensos, aos poucos mudando para a avaliação de programas, com outras metodologias e relatórios menores.

Historicamente, o processo de trabalho que envolve a auditoria de natureza operacional tem ampliado o seu foco, abrangendo aspectos freqüentes em AP. Entretanto, essa ampliação de foco é fonte de conflito internamente para a organização, conforme também verificado no trabalho de Mendes et al. (1999).

De acordo com Lima (2004), atualmente as auditorias e avaliações realizadas pela organização inserem-se no contexto de performance *accountability*, ou seja, prestação de contas por desempenho e responsabilização por desempenho. Segundo o autor, a contribuição da organização será na responsabilização pública, atingindo os seguintes objetivos: a) garantir para o Poder Legislativo e para a sociedade; b) que as contas sejam fidedignas, c) as receitas e as despesas estejam de acordo com o determinado pela Lei; e d) que os recursos públicos estejam sendo usados com eficácia, eficiência e efetividade.

A organização possui um termo de cooperação com o Reino Unido, no qual utiliza a metodologia para a realização das auditorias – tal projeto chama-se projeto de Aperfeiçoamento do Controle Externo em Foco na Redução da Desigualdade Social – CERDS.

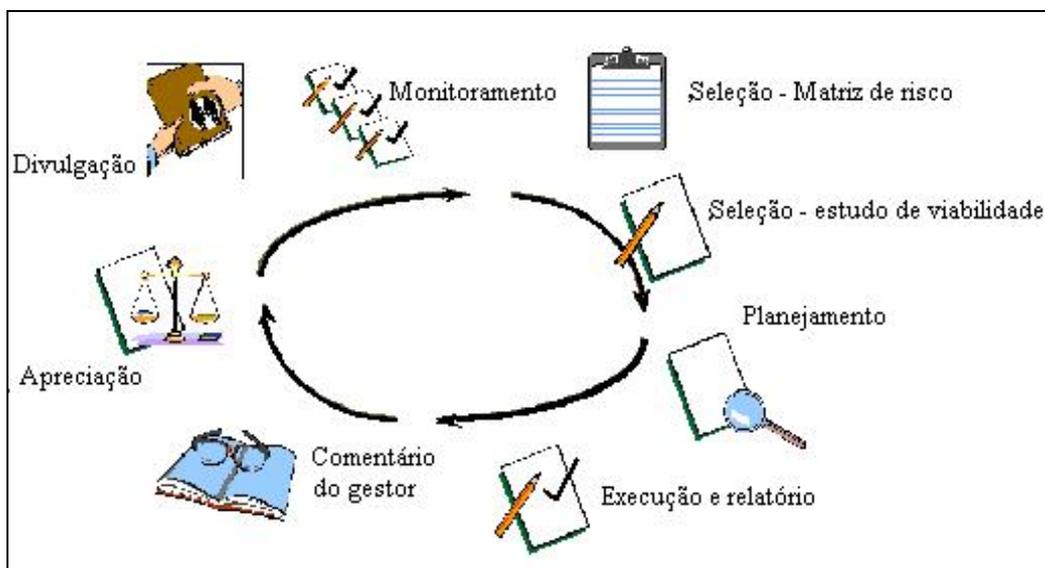
Nesse contexto, parece haver um conflito auditoria *versus* avaliação de programas, que a organização busca por meio de sua prática de auditoria. Tal mudança pode estar ocorrendo por diversos fatores. Uma hipótese seria a mudança no ambiente, no caso a sociedade, que historicamente busca se apropriar do seu estado de direitos. Mendes et al.

(1999) afirmam que há uma dicotomia entre o tradicional e o moderno, associados ao funcionamento da organização, o conflito entre fiscalizar e julgar. Conforme os autores “Está claramente posto o conflito entre duas concepções de controle externo: um modelo considerado moderno, o fiscalizador, e um outro identificado com o tradicionalismo, com conservadorismo, que é o modelo corretivo, julgador” (p.17).

Mendes et al. (1999) verificaram que há uma demanda de mudança cultural no TCU, um potencial para mudanças associados aos resultados da percepção dos valores reais e desejados para a organização. Essas percepções referem-se às prioridades axiológicas, onde se destacam na organização os pólos de Hierarquia e Conservação. Isso significa que a vida organizacional é percebida como orientada pela presença de níveis hierárquicos, pela manutenção de tradições, normas e exercício de autoridade entre os indivíduos. Porém, os indivíduos não estão necessariamente insatisfeitos com essa realidade, e sim esperam aperfeiçoar o que já existe.

A auditoria de natureza operacional obedece às seguintes etapas: 1) planejamento da avaliação; 2) coleta de dados; 3) confecção de relatório com os resultados para o relator da matéria (circulação interna); 4) relatório divulgado externamente; 5) implantação das recomendações formuladas e monitoramento das ações implementadas. Caracteriza-se o processo de trabalho do setor como uma atividade cíclica que envolve as etapas: a) seleção; e ainda dentro da seleção o estudo de viabilidade; b) planejamento da avaliação; c) execução e relatório; d) comentário do gestor; e) apreciação, f) divulgação; e g) monitoramento (Bastos, 2004). A seguir apresenta-se a representação gráfica do processo de ANOP.

**Figura 5. Ciclo de trabalho da Auditoria de Natureza Operacional**



Fonte: Bastos (2004, p. 5)

*a) Seleção por meio da matriz de risco*

Nessa fase, é realizado o exame das prioridades do governo por meio da matriz de verificação de risco para a seleção das auditorias em programas de governo e depois é realizado o estudo de viabilidade da auditoria do programa. Para a escolha das áreas nas quais se localizam os programas a serem objeto de auditoria, consideram-se os critérios de relevância, risco e materialidade. A **relevância** trata da importância relativa das ações em estudo. O **risco** refere-se à possibilidade da ocorrência de eventos inesperados e a **materialidade**, ao valor ou volume de recursos envolvidos.

*b) seleção - estudo de viabilidade*

O estudo de viabilidade envolve, de acordo com Bastos (2004, p. 3), “a manifestação valorativa, substantiva e convincente sobre a conveniência e a oportunidade de se realizar avaliação de programa ou auditoria de natureza operacional”. Envolve a relevância da auditoria para a imagem da organização e temas que chamam a atenção da mídia, da sociedade civil organizada e do Congresso Nacional. Por fim, o estudo envolve o impacto potencial da auditoria na melhoria do desempenho do programa e a observância do artigo 238, inciso III do Regimento Interno da Instituição, que reza em relação aos instrumentos de auditoria e levantamentos “avaliar a viabilidade da realização de fiscalizações” (Brasil, 1992, p. 60).

*c) Planejamento da avaliação*

No planejamento da auditoria, ocorre a definição do problema e das questões a serem auditadas, a identificação das fontes de informação, a definição de estratégia metodológica para a obtenção e análise de dados e a elaboração do projeto de auditoria com base na matriz de planejamento. Envolve as etapas de: planejamento das atividades, análise preliminar dos objetos da auditoria, coleta de dados sobre o desempenho, especificações dos critérios de auditoria, elaboração de instrumentos de coleta de dados, um teste-piloto desses instrumentos, a validação da matriz de planejamento e, por fim, a elaboração do relatório de planejamento.

As técnicas utilizadas pelo Tribunal de Contas da União são indicadas na fase do planejamento da auditoria. Utilizam-se as seguintes técnicas: mapa de processo, mapa de produto, entrevistas, análise RECI (Responsável, Executor, Consultado, Informado), análise SWOT (*Strengths* – Forças, *Weaknesses* – Fraquezas, *Opportunities* – Oportunidades, *Threats* – Ameaças), Matriz de Verificação de Risco e Análise de *Stakeholder* (atores envolvidos). Tais técnicas foram descritas no item 4.2 Descrição das técnicas utilizadas pela Organização.

*d) Execução e relatório*

A etapa de execução envolve: a coleta de dados, a análise dos dados coletados e a elaboração do relatório preliminar. Assim, após a execução do relatório tem-se: *i) comentários do gestor à primeira versão do relatório e, ii) elaboração do relatório final.*

*e) Apreciação pelo Tribunal de Contas da União.*

Após a aprovação pelo Tribunal de Contas da União é realizada a divulgação da auditoria, por meio de um *briefing* (notícia breve) pela mídia. São também lançadas publicações escritas denominadas “Série Sumário Executivo”, criada em 2002, dirigida a parlamentares, tribunais de contas estaduais e municipais, bibliotecas, instituições de ensino e demais atores interessados nas ANOPs.

Além das auditorias, é realizado também o seu monitoramento. O monitoramento realizado pela unidade técnica competente refere-se ao acompanhamento e cobrança de providências adotadas pelo órgão/programa auditado e implantação das recomendações do Tribunal de Contas da União (Bastos, 2004). O monitoramento visa aumentar a

probabilidade da adoção das recomendações acompanharem as providências adotadas, garantir o alcance dos efeitos desejados e verificar a adequação das recomendações em face das condições vigentes. A seguir, para a melhor compreensão do processo de trabalho do Tribunal de Contas da União, foram descritas as técnicas utilizadas na organização.

#### ***4.2 Descrição das Técnicas Utilizadas pela Organização***

As técnicas descritas a seguir são utilizadas no processo de ANOP do Tribunal de Contas da União. Cada auditoria de programa pode utilizar uma ou mais técnicas, formando combinações de técnicas, de acordo com a natureza do programa e fontes de dados disponíveis para a auditoria. Ressaltaram-se as diferenças na concepção da técnica de indicadores de desempenho e dos conceitos correlatos em pesquisa avaliativa para demonstrar que essa técnica associa-se à concepção de auditoria de programas e não de avaliação de programas como utilizado em pesquisa avaliativa.

*As estratégias metodológicas utilizadas* de acordo com Bastos (2004) são: estudo de caso, pesquisa, delineamento experimental, delineamento quasi-experimental e delineamento não experimental.

Utilizam-se dez técnicas, a saber: a) Indicadores de desempenho e Mapas de produtos; b) Mapa de processos; c) Análise RECI (responsável, executor, consultado, informado); d) Marco lógico; e) Análise de *Stakeholders*; f) Análise SWOT e Matriz de verificação de risco; e g) Técnica de entrevista.

##### ***a) Indicadores de desempenho e mapas de produto***

A construção e utilização dos ***indicadores de desempenho*** estão relacionadas ao conceito de gerenciamento voltado para resultados (*results oriented management – ROM*). O indicador de desempenho é compreendido como um instrumento de mensuração quantitativa ou qualitativa de aspectos do desempenho do programa. Brasil (2000a, p. 9) aponta que “Um indicador de desempenho é um número, percentagem ou razão que mede um aspecto do desempenho, com o objetivo de comparar esta medida com metas pré-estabelecidas”. Observa-se que embora haja duas acepções e possibilidades de mensuração, qualitativa e quantitativa, a organização a define de acordo com a concepção quantitativa.

Quanto à diferença entre *medidor de desempenho e indicador de desempenho*, no primeiro caso, a mensuração pode ser realizada de forma direta e no segundo caso, torna-se

necessário utilizar outras medidas ou índices para formar o indicador, que não pode ser mensurado diretamente. Os indicadores indiretos são considerados na avaliação de programas como indicadores *proxis*, ou seja, refere-se a um conjunto de outras medidas que podem mensurar um resultado. Os indicadores possuem uma natureza comparativa, a partir da qual se pretende comparar os resultados com as metas ou objetivos pretendidos e geralmente são formados a partir das variáveis de custo, tempo, quantidade e qualidade.

A seguir, apresenta-se a descrição das características necessárias para a construção de indicadores e comparação com critérios de pesquisa científica, quando pertinente. As qualidades desejáveis alcançadas na construção de indicadores seguem os critérios de representatividade, homogeneidade, praticidade, validade, independência, confiabilidade, seletividade, simplicidade, cobertura, economicidade, acessibilidade e estabilidade.

A **representatividade** (expressão dos produtos essenciais de uma atividade ou função, medindo-se aquilo que é produzido: produtos intermediários e finais e seu impacto ou *outcomes*). Entende-se esse critério em pesquisa avaliativa como característica representativa do fenômeno a ser avaliado, não recaindo necessariamente sobre os resultados do programa, dada a possibilidade de avaliação nas diversas fases do programa.

Para se obter a **Homogeneidade** na elaboração do indicador “deve-se considerar apenas variáveis homogêneas. Por exemplo: estabelecer o custo médio por auditoria, pois cada tipo tem uma composição de custo diversa” (Brasil, 2000a, p.12). Variáveis homogêneas em pesquisa avaliativa estão ligadas a características de uma amostra que seja semelhante o suficiente para possibilitar a comparação do comportamento das variáveis entre si. Portanto, essas variáveis assumem em AP uma diferente concepção daquela adotada pela organização no quesito homogeneidade para a construção de indicadores.

**Praticidade** refere-se à garantia de que o indicador funciona na prática e permite a tomada de decisão gerencial.

A **Validade** “deve refletir o fenômeno a ser monitorado” Brasil (2000a, p. 13). Em pesquisa avaliativa, refere-se à validade interna ou externa. Validade interna refere-se às mudanças reais, entretanto não relacionadas ao programa. Já a validade externa representa as condições de generalização que os resultados ou efeitos encontrados possam ser aplicados (populações no caso programa ou serviço recebido).

**Independência** “deve medir os resultados atribuíveis às ações que se quer monitorar evitando resultados que possa ser atribuídos a fatores externos” Brasil (2000a, p. 13). A independência em pesquisa avaliativa relaciona-se ao conceito de validade interna.

**Confiabilidade:** “a fonte de dados deve ser confiável da forma que diferentes avaliadores possam chegar aos mesmos resultados”; seletividade; “estabelecer um número equilibrado de indicadores, referentes a aspectos essenciais do que se quer mensurar”; simplicidade; “deve ser de fácil compreensão e não envolver dificuldades de cálculo ou de uso”; cobertura; “devem representar a amplitude e a diversidade de característica do fenômeno monitorado”; economicidade; “as informações necessárias ao cálculo do indicador devem ser coletadas e atualizadas a um custo razoável”; acessibilidade; “facilidade de acesso às informações primárias e de registro e manutenção do cálculo dos indicadores”, Brasil (2000a, p. 13 e 14).

**Estabilidade** “estabilidade conceitual dos componentes e do próprio indicador é condição necessária para avaliar o seu desempenho ao longo do tempo” Brasil (2000a, p. 14). Em pesquisa avaliativa, entende-se a estabilidade conceitual como validade do construto.

Para compreender a técnica de *mapas de produtos*, deve-se considerar inicialmente a definição de produto. De acordo com Brasil (2000a, p. 17), “[...]. Produto para a organização refere-se a um bem, um serviço ou uma condição resultante de atividade(s) de uma organização ou de um indivíduo”. São exemplos de produtos “benefícios de um programa, estradas conservadas, correspondência datilografada, pacientes atendidos. Implementar um programa; conservar estradas; datilografar a correspondência e atender pacientes [...]” (p. 17). Um mapa de produtos possui três componentes principais: insumos, produtos intermediários e produtos finais. São exemplos de insumos em relação a um Ministério, órgão ou um programa do governo, as políticas, o orçamento, as instalações, equipamentos, pessoal, programas de treinamento, entre outros. Esses insumos dão origem a produtos. Já um subproduto ou produto intermediário “seriam obtidos a partir dos insumos já existentes e que, juntos, darão origem ao produto final”, Brasil (2000a, p. 18).

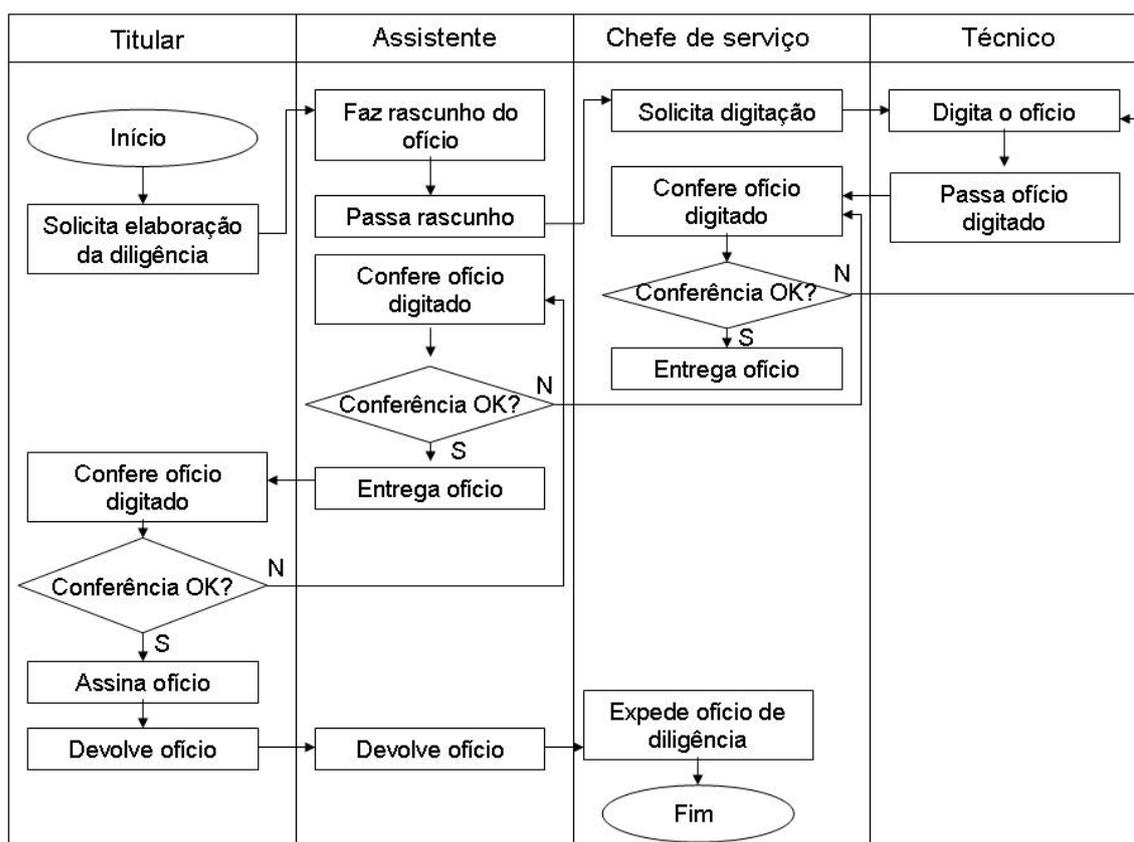
O mapa de produto pode ser construído com auxílio do pessoal da instituição, pois conhecem bem a organização. Em suma, a finalidade do mapa de produtos é conhecer os objetivos do órgão ou entidade auditada, mostrar as relações de dependência entre os produtos, identificar as áreas ou setores responsáveis pelos produtos críticos. Por fim, a

relação entre essas duas técnicas é que a partir do mapa de produto os indicadores de desempenho podem ser construídos.

### *b) Mapa de processos*

A técnica de mapa de processos concebe processo como “a seqüência de passos necessários à consecução de uma atividade”, ou ainda “a maneira pela qual se realiza uma operação” Brasil (2000b, p. 7). Ela utiliza a representação gráfica (diagrama de mapa de processo) para ilustrar como ocorrem os passos de uma tarefa, partindo do início ao seu final. Considera as atividades, os agentes envolvidos, os prazos e o fluxo de documentos. Na Figura 6, as setas indicam a seqüência e direção das operações, de acordo com o exemplo:

**Figura 6. Exemplo de Mapa de Processo de expedição de ofício de diligência**



Fonte: Brasil (2000b, p. 10)

A principal característica de um mapa de processo é a descrição da tarefa passo a passo. Por ser um método visual interativo, favorece o trabalho em grupo e a participação da organização auditada. As principais diferenças entre mapa de processo e mapa de

produto são as seguintes. O mapa de processo “representa graficamente uma série de atividades interligadas que dão origem a determinados produtos”, enquanto o mapa de produto “representa os produtos gerados por aquelas atividades” (Brasil, 2000b, p. 9). O mapa de processo é utilizado quando “se deseja obter melhorias de desempenho da organização enfocando processo de trabalho” (Brasil, 2000b p. 15).

### *c) Análise RECI*

As iniciais RECI correspondem, em inglês, à sigla RACI, proveniente das iniciais das seguintes palavras: *responsible, accountable, consulted e informed* (Brasil, 2001a). Assim, a letra **R** refere-se a quem é **responsável** (responsabilidade final, poder de veto). A letra **E** a quem **executa**, ou seja, o encarregado de desenvolver a atividade ou implementá-la. A letra **C** a quem é **consultado**, ou seja, a pessoa que é consultada antes ou depois da decisão ou ação ser implementada. E por fim a letra **I** refere-se a quem é **informado**, aquele a ser informado depois que a decisão ou ação for implementada. Em suma, tem-se **Responsável, Executor, Consultado e Informado**. Por meio dessa técnica é possível identificar os responsáveis pelas atividades desenvolvidas em termos de sua execução e suas inter-relações com todas as outras pessoas responsáveis para realizar as tarefas, sejam internas ou externas às organizações. A análise RECI é construída a partir de *workshops* com a participação das pessoas envolvidas. Algumas das vantagens dessa técnica são: clara definição de responsabilidades, redução da necessidade de repetição de tarefas, eliminação de superposição e de trabalho redundante e tentativa de reduzir os níveis desnecessários das estruturas organizacionais. Por fim, a técnica analisa os níveis de impacto político e psicológico das mudanças organizacionais.

### *d) Marco lógico*

Essa técnica refere-se a um modelo analítico que orienta **a formulação, a execução, o acompanhamento e a avaliação** de programas ou de projetos governamentais. De acordo com Brasil (2001b, p. 7), “foi criado pelo Banco Interamericano de Desenvolvimento - BID é um instrumento de planejamento obrigatoriamente adotado por todas as organizações públicas que postulam financiamento junto àquele Banco”. Esse modelo oferece uma análise completa do programa ou projeto, porém, devido à sua complexidade, exige maior disponibilidade de tempo e um sistema de informações gerenciais para o seu uso. Por meio do modelo, é possível responder questões acerca dos objetivos do programa, sobre a clientela beneficiada, a causalidade dos níveis hierárquicos

e o impacto do programa. Para responder a essas questões, torna-se necessário trabalhar com as atividades (tarefas executadas para a geração do bem ou serviço), produtos (bens ou serviços ofertados), objetivos (resultados diretos associados à geração de bens e serviços) e finalidades (benefícios almejados). O marco lógico ordena os processos organizacionais utilizando indicadores. Esses indicadores precisam ser objetivos e mensuráveis, expressos em termos quantitativos ou qualitativos para cada nível de análise, além de indicar fatores externos que possam afetar o objeto da auditoria, avaliando riscos e propondo ações alternativas. Conforme Brasil (2001b, p. 9), “Sua estrutura corresponde a uma matriz de quatro por quatro, onde as linhas apresentam informações acerca dos objetivos, indicadores, meios de verificação e pressupostos em quatro momentos diferentes da vida de um projeto ou programa governamental”.

#### *e) Análise de Stakeholder*

A análise de *stakeholder* pode ser aplicada no planejamento ou execução das auditorias. Quando na fase de planejamento, utiliza-se como complemento a técnica SWOT. “A análise *stakeholder* consiste na identificação dos principais atores envolvidos, dos seus interesses e do modo como esses interesses irão afetar os riscos e a viabilidade de programas ou projetos” (Brasil, 2002a, p. 8). Tal identificação por meio de entrevista ou consulta, se faz criando uma tabela para aplicar e levantar as seguintes informações: 1) identificação nominal do *stakeholder*, suas necessidades e direitos; 2) interesse do *stakeholder* no programa e seu poder de influência, estimando o seu grau de interesse, que pode variar de altamente favorável a altamente desfavorável; 3) o impacto no programa ou projeto, caso o interesse do *stakeholder* não seja atendido; e 4) o papel do *stakeholder* no programa. Os *stakeholders* são definidos como “pessoas, grupos ou instituições com interesse em algum programa ou projeto e incluem tanto aqueles envolvidos quanto os excluídos do processo de tomada de decisão” (Brasil, 2002, p. 8).

Os *stakeholders* podem ser primários e secundários. Os primários são aqueles mais afetados, tanto positivamente (os beneficiários) como negativamente (aqueles re-allocados involuntariamente); e os secundários são os intermediários envolvidos no processo de prestação de serviço. Pode-se ainda coletar as informações enviando o questionário pelo correio, através da participação em *workshops* e reuniões, ou participação dos membros em grupos focais. A técnica oferece ainda listas de verificação para identificação dos *stakeholders*; identificação dos mais importantes para o sucesso do projeto; identificação dos seus interesses e explora os possíveis impactos no programa ou projeto.

Essa técnica pode ser utilizada como subsídio para a construção e desenvolvimento da matriz do marco lógico.

#### *f) Análise SWOT e Matriz de Verificação de Risco*

A análise SWOT e a Matriz de Verificação de Risco são utilizadas para a verificação do ambiente interno e externo e também na fase de planejamento da auditoria. O modelo SWOT foi desenvolvido pela Escola do Design e “consiste em avaliação de pontos fortes (*strenghts*) e dos pontos fracos (*weaknesses*) da organização, com relação às oportunidades (*opportunities*) e as ameaças (*threats*) em relação ao ambiente” (Mintzberg, Ahlstrand & Lampel, 2002). A técnica SWOT é uma ferramenta utilizada no planejamento estratégico e no diagnóstico organizacional. No caso da avaliação de programa ela é aplicada pelo Tribunal de Contas da União na fase de planejamento da avaliação. De acordo com Brasil (2002, p. 9), “[...] a técnica permite à equipe identificar as ações a serem implementadas para corrigir problemas detectados a partir da análise das forças e fraquezas internas do objeto de auditoria e das oportunidades e ameaças do ambiente externo, buscando um novo equilíbrio entre essas variáveis”. De acordo com Mintzberg e cols., (2002) as forças e fraquezas são medidas em relação ao ambiente interno e as oportunidades e ameaças em relação ao ambiente externo. Um exemplo de variáveis ambientais a serem cheçadas poderiam ser as mudanças na sociedade, governamentais, econômicas, na competição, nos fornecedores, no mercado. O exemplo poderia ser de checagem de pontos fortes e pontos fracos seria: nas áreas de marketing, pesquisa e desenvolvimento, sistema de informações gerenciais, time gerencial, operações, finanças e recursos humanos.

De acordo com Brasil (2002b), a *matriz de verificação de risco* tem como objetivo levantar os riscos associados ao desempenho do programa. O exame é realizado para identificar a capacidade de gestão do risco pela administração do programa. A análise é realizada de acordo com os seguintes passos:

a) clareza dos objetivos do programa (definição da missão, metas, objetivos, se os objetivos estão claros para a organização, **definidos na legislação** e se são divulgados internamente). Em relação a esse aspecto, observa-se que desde o planejamento da avaliação, enfatiza-se o cumprimento da regulação ou lei, aspecto característico da ANOP. A matriz de verificação de risco é utilizada na fase de planejamento da avaliação; b) identificação do risco (se os departamentos identificam os riscos pelos quais são responsáveis e como podem causar impacto na organização); c) verificação do risco (envolve a análise e a avaliação do risco identificado no período pelos quais os riscos serão

gerenciados); d) tratamento de risco (controle pelas áreas responsáveis pelo risco no programa e priorização e identificação dos geradores da situação de risco); e) monitoramento e revisão (atividade contínua que permite a identificação de novos riscos e as oportunidades decorrentes).

***g) Técnica de entrevista***

São três os tipos de entrevistas utilizadas nas ANOPs: entrevista de apresentação; entrevista de coleta de dados (estruturada ou não estruturada) e entrevista de encerramento (Brasil, 1998).

A entrevista de apresentação ocorre com a apresentação da equipe de auditoria aos responsáveis pelo “objeto auditado, apresentando os objetivos do trabalho e seu planejamento global, bem como os procedimentos técnicos e administrativos que serão adotados pela equipe no cumprimento das suas atribuições” (Brasil, 1998, p. 9). A entrevista de coleta de dados é aquela na qual a equipe de auditoria busca informações específicas sobre o programa auditado junto ao entrevistado. A entrevista de encerramento ocorre na finalização do processo de auditoria e quando possível são apresentados alguns resultados preliminares dos trabalhos que constarão do relatório de auditoria, assim, como para obter críticas e sugestões do auditado (Brasil, 1998).

De acordo com Brasil (1998), as entrevistas de coleta de dados estruturada e não estruturada diferenciam-se em relação à análise das informações coletadas. A entrevista não estruturada contempla questões abertas e questões fechadas. Nessas entrevistas, utiliza-se a técnica do funil, que é a avaliação das respostas mediante sua interpretação e devolução ao entrevistado no momento da entrevista, quando então o entrevistado confirma, ou não, o resumo de suas respostas pela equipe que realizou a entrevista. Em relação à forma de análise dos dados coletados na entrevista não estruturada, utilizam-se algumas orientações, como “a regra de 75/25, a qual estipula que não mais do que 25% do tempo da entrevista será utilizado na confirmação das declarações do entrevistado, de tal forma que o tempo restante deve ser reservado à escuta e anotações”, Brasil (1998, p. 19). Na “técnica do funil”, são seguidos cinco passos, a saber:

- a) o entrevistado responde a uma questão aberta;*
- b) um membro da equipe de auditoria interpreta a resposta para o entrevistado;*
- c) o entrevistado esclarece e complementa as informações;*

*d) um outro membro da equipe de auditoria faz um resumo da declaração do entrevistado;*

*e) o entrevistado confirma, ou não, retornando à etapa “c”, caso não haja confirmação.*

A entrevista de coleta de dados estruturada caracteriza-se pelo desenvolvimento sequencial das questões, não havendo possibilidade de modificação na seqüência de perguntas e nem a inclusão de novas questões. As perguntas são do tipo fechado, com escalas de respostas, seja de múltipla escolha, simples, dicotômicas, preenchimento de lacunas, possibilitando ao final uma análise estatística das respostas.

Aqui se ressalva o seguinte, o que diferencia um tipo de entrevista do outro não é a forma de análise dos dados, e sim a forma de coleta dos mesmos. Esse tipo de entrevista assemelha-se mais a um formulário, uma vez que o principal diferencial da entrevista estruturada é sua seqüência de perguntas previamente estabelecida e seguida rigorosamente em sua aplicação.

As entrevistas podem ser aplicadas por meio de contato direto, contato telefônico e pelo correio. Devido às vantagens e desvantagens de cada tipo de entrevista, para a sua escolha, considere-se aquela que melhor se adaptar ao fenômeno considerando as variáveis: tempo disponível; custo e fidedignidade para a avaliação.

Para finalizar a descrição e comentários em relação às técnicas utilizadas pelo Tribunal de Contas da União, as informações selecionadas dos documentos referenciados neste tópico tiveram por objetivo familiarizar o leitor com as técnicas utilizadas. Para aprofundamento e detalhamento dessas técnicas, sugere-se a leitura dos documentos citados. A seguir, apresentou-se na Figura 5 o resumo das estratégias metodológicas utilizadas pela organização.

**Figura 7. Sinopse das estratégias metodológicas utilizadas pelo TCU**

ESTRATÉGIA METODOLÓGICA	TIPOLOGIA	MÉTODO DE SELEÇÃO DA AMOSTRA	MÉTODO DE COLETA DE DADOS
Estudo de caso	- ilustrativo - exploratório - caso crítico - implantação do programa - efeitos do programa - cumulativo	- conveniência - finalidades - casos extremos - melhores casos - piores casos - agrupamento - casos representativos - casos típicos - interesse especial - probabilístico	- entrevista - questionário enviado pelo correio - dados secundários - observação direta
Pesquisa	- corte temporal - painel	- probabilístico	- questionário enviado pelo correio - entrevista estruturada
Delineamento Experimental	- com pré e pós-testes - com pós-teste	- probabilístico ou não	- teste - questionário enviado pelo correio
Delineamentos Quase-experimentais	- grupos não-equivalentes com pré e pós-teste - séries temporais com grupo de controle - séries temporais com grupos de controle	- probabilístico ou não.	- teste - questionário enviado pelo correio - dados secundários.
Delineamentos não-experimentais	- antes e depois - somente depois com grupos de comparação	- probabilístico ou não	- teste - questionário enviado pelo correio - dados secundários.

Fonte: adaptado de Brasil (2000c, p. 77).

Pode-se analisar e resumir o seguinte em relação às técnicas e modelos descritos: Indicadores de desempenho, mapas de produtos, marco lógico e matriz de verificação de risco são técnicas claramente associadas a processo de auditoria e não de avaliações de programas, no sentido utilizado dentro da área da pesquisa avaliativa. Transparece nas técnicas e modelos a concepção positivista na avaliação dos programas, o que liga o conceito de avaliação à idéia de controle e quantificação de resultados.

O *indicador de desempenho* (conceito de gerenciamento voltado para resultados) utiliza a definição remetendo à quantificação do resultado do programa e utiliza Indicador de desempenho quando não é possível a medição direta. Os Mapas de Produtos têm como objetivo conhecer os objetivos do programa e os responsáveis pela execução. O mapa de processos, compreendido como a seqüência de passos necessários à consecução de uma atividade, consiste na construção de um diagrama para entendimento dos fluxos das tarefas relacionados aos produtos. E a análise RECI possibilita identificar os responsáveis pelas

atividades desenvolvidas, sua execução e inter-relações com todas as outras pessoas responsáveis pelas tarefas, sejam internas ou externas às organizações. Tais itens apresentados são técnicas e são utilizadas no planejamento da avaliação. Já o Marco Lógico é um modelo criado pelo BID, por meio do qual é possível responder questões sobre os objetivos do programa, clientela beneficiada, causalidade dos níveis hierárquicos e impactos do programa. Por meio de sua estrutura apresentam-se informações dos objetivos, indicadores, meios de verificação e pressupostos em quatro momentos diferentes da vida de um projeto ou programa governamental. É um modelo completo de avaliação e não uma técnica. A análise *stakeholder* consiste na identificação dos principais atores envolvidos, dos seus interesses e do modo que afetarão os riscos e a viabilidade de programas ou projetos, pode ser utilizada como subsídio para a construção e desenvolvimento da matriz do marco lógico; configura-se como uma técnica. O modelo SWOT foi desenvolvido pela Escola do Design e consiste em avaliação de pontos fortes (*strenghts*) e dos pontos fracos (*weaknesses*) da organização, com relação às oportunidades (*opportunities*) e as ameaças (*threats*) em relação ao ambiente. O modelo SWOT é uma ferramenta utilizada no planejamento estratégico e no diagnóstico organizacional; no caso da avaliação de programa é aplicada pelo Tribunal de Contas da União na fase de planejamento da avaliação. A análise SWOT é um modelo e não uma técnica.

Organização auditada é a forma com que a organização que recebe a ANOP é denominada, entende-se que o termo *auditada* estaria relacionado à auditoria e não à avaliação. Nos textos que descrevem as técnicas utilizadas nas avaliações, o termo que surge é auditoria e não avaliação. Os termos são utilizados de forma intercambiável, como se possuíssem o mesmo significado.

Após a descrição e comentários das técnicas e modelos usados nas ANOPs, descrevem-se os resultados da meta-avaliação das dez auditorias, conforme critérios do *Joint Committee*.

#### ***4.3 Meta-avaliação de acordo com Normas do Joint Committee***

Elaborou-se a Tabela 3. Critérios de avaliação para avaliação das auditorias, a partir das normas do *Joint Committee*. Assim, as auditorias foram analisadas conforme critérios escolhidos e receberam uma pontuação em cada item, conforme Tabela 3. Após a Tabela, as pontuações recebidas pela auditoria, no item, foram justificadas de acordo com o contexto da auditoria do programa.

Na primeira coluna, os critérios foram listados conforme codificação do *Joint Committee* (ver 2.2.2 Meta-avaliação). Na segunda coluna, os itens do *checklist* foram transformados em perguntas para a meta-avaliação das ANOPs. Na terceira coluna, A1 até a A10 representam as Auditorias de Programas. Cada auditoria de programa recebeu pontuações baseadas em uma escala de quatro pontos; onde zero significa não se aplica; “não” significa que o critério não foi utilizado; “parcialmente” significa que foi aplicado, mas não completamente ou não atendeu completamente àquele critério, e “sim”, significa que foi adequadamente utilizado.

Na quarta coluna, as questões foram respondidas de acordo com a escala, onde: 0 = não se aplica, 1 = não; 2 = parcialmente e 3 = sim.

Após o quadro, são comentadas as pontuações dos critérios para cada auditoria de programa.

Ao final da apresentação da meta-avaliação das dez auditorias de natureza operacional, apresentou-se o resumo dos achados da meta-avaliação.

No final do quadro, destacou-se o total de pontos e o percentual dos mesmos em relação às pontuações totais de todas as auditorias (200 pontos) e as pontuações máximas (60 pontos) por Programa.

Os resultados numéricos foram analisados por Programa, na análise do *Joint Committee* e por agrupamentos na Síntese de Categorias.

No quadro segue-se a seguinte codificação para os nomes dos programas auditados:

A1 Programa Novo Mundo Rural – Consolidação de Assentamentos;

A2 Programa Combate ao Abuso e a Exploração Sexual de Crianças e Adolescentes – Sentinela;

A3 Programa Reinserção Social do Adolescente em Conflito com a Lei;

A4 Programa Desenvolvimento do Turismo do Nordeste – Prodetur/NE;

A5 Desenvolvimento Sustentável de Recursos Hídricos para o Semi-Árido Brasileiro Pró-Água/Semi-Árido;

A6 Programa Energia das Pequenas Comunidade;

A7 Programa Morar Melhor;

A8 Programa Nacional do Livro Didático – PNLD;

A9 Monitoramento e Controle da Mortalidade Materna;

A10 Programa Novo Emprego e Seguro Desemprego.

**Tabela 3. Critérios para avaliação das auditorias, baseado nas normas do *Joint Committee***

Critérios <i>Joint Committee</i>		Questões adaptadas à meta-avaliação das ANOPs										0	1	2	3
		A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	N/A	não	Parc	Sim
<i>A Normas de utilidade</i>															
<i>A1 Descrição dos potenciais usuários</i>	1. Os públicos do programa e participantes da auditoria foram identificados?	3	3	3	3	3	3	3	3	2	3	0	0	1	9
<i>A2 Credibilidade do avaliador</i>	2. A equipe de auditoria possui treinamento adequado para realizar a auditoria?	2	1	3	1	3	2	3	3	3	2	0	2	3	5
<i>A3 Alcance e seleção da informação</i>	3. Consultores externos, de áreas específicas, participaram da auditoria ?	3	1	3	3	3	3	2	3	3	1	0	2	1	7
	4. As informações coletadas possuem alcance suficiente para entender o funcionamento do programa?	3	2	3	3	3	2	3	2	2	2	0	0	5	5
	5. As informações coletadas foram suficientes para refletir os objetivos da auditoria?	3	2	3	2	2	2	3	2	2	2	0	0	7	3
<i>A4 Interpretação de valores</i>	6. As técnicas de análise de dados foram explicitadas?	2	3	3	3	3	2	3	2	2	2	0	0	5	5
<i>A5 Clareza dos Relatórios</i>	7. O relatório descreve claramente o programa?	3	3	3	3	3	2	3	1	1	3	0	2	1	7
	8. O relatório descreve claramente o contexto do programa?	3	3	3	2	3	1	3	1	3	3	0	2	1	7
	9. O relatório descreve claramente os objetivos do programa?	3	3	3	3	3	2	3	2	1	3	0	1	2	7
	10. O relatório descreve claramente os procedimentos metodológicos da auditoria?	2	2	2	2	2	2	2	2	2	2	0	0	10	0
	11. O relatório descreve claramente os resultados da auditoria?	3	2	2	3	3	3	3	3	2	2	0	0	4	6
	12. O relatório descreve claramente as conclusões da auditoria?	3	2	3	2	3	3	3	3	3	3	0	0	2	8
<i>B Normas de viabilidade</i>	13. O relatório justifica claramente as recomendações da auditoria?	3	3	3	3	3	3	3	3	3	3	0	0	0	10
<i>B2 Viabilidade política</i>	14. A auditoria considerou as posturas dos diferentes grupos de interesse no programa?	3	1	3	1	2	1	2	1	2	2	0	4	4	2
<i>B3 Custo-efetividade</i>	15. Os recursos da auditoria (tempo, dinheiro e funcionários) foram adequados para a realização das atividades previstas?	3	3	3	3	2	2	3	3	3	1	0	1	2	7
<i>D Normas de precisão</i>	16. Os procedimentos de coleta de informações foram descritos com clareza?	3	3	3	3	3	1	3	2	1	1	0	3	1	6
<i>D5 Determinação da Validade</i>	17. Os instrumentos para a coleta de informações eram válidos?	3	1	3	2	3	0	3	0	0	0	4	1	1	4
<i>D8 Análise de informações quantitativas</i>	18. Foram tomadas medidas para garantir o mínimo de erros na coleta dos dados?	2	1	1	2	1	1	3	1	2	2	0	5	4	1
<i>D9 Análise da informação qualitativa</i>	19. A informação quantitativa foi adequadamente analisada?	2	3	2	2	3	3	3	3	2	3	0	0	4	6
	20. A informação qualitativa foi adequadamente analisada?	2	1	3	2	2	1	2	1	1	2	0	4	5	1
	Total de pontos (A1 até A10 no máximo 60 pontos por auditoria)	54	43	55	48	53	39	56	41	40	42	4	27	63	106
	A1 até A10 percentual arredondado	90	72	92	80	88	65	93	68	67	70	2	14	32	53

### ***1. Programa Novo Mundo Rural: Assentamento de Trabalhadores Rurais; Novo Mundo Rural: Consolidação de Assentamentos***

Na auditoria do Programa Novo Mundo Assentamento de Trabalhadores Rurais; Novo Mundo Rural: Consolidação de Assentamentos, foram identificados os públicos da avaliação (Questão 1), a saber: “gerentes, técnicos, associações, beneficiários, responsáveis pela prestação de assistência técnica e elaboração do PDA e membros de sindicatos rurais de trabalhadores” (Brasil, 2003h, p. 25).

A equipe possuía treinamento parcial para realizar a auditoria (Questão 2). Assim, utilizaram a colaboração de técnicos de instituições de notório saber sobre a questão como o INCRA, ARCO e EMATER (Questão 3). Entretanto, tais especialistas e técnicos não fizeram parte do grupo de auditoria, embora indiretamente tenham colaborado como consultores e especialistas.

Verificou-se que as informações coletadas foram suficientes para entender o funcionamento do programa (Questão 4). Foi apresentada uma visão geral, com informações de objetivos, legislação específica e contexto da reforma agrária. Entretanto, o Programa não foi avaliado como um todo. A auditoria se deteve em duas ações: Assentamentos de Trabalhadores Rurais e Consolidação de Assentamentos.

Por meio dos resultados encontrados, avalia-se que as informações levantadas foram suficientes para atender aos objetivos da auditoria (Questão 5).

Na questão 6, as técnicas de análise de dados foram explicitadas de forma parcial: “análise qualitativa (tipologia de situações, transcrição, narração, descrição) análise estatística” (Brasil, 2003h, p. 35). Conforme se nota, há inconsistência nessa descrição, as técnicas não estão claras, por exemplo, transcrição é considerada tabulação de dados e não análise. Nos documentos, não há citações literais que se assemelhem a narração ou descrição.

Verificou-se que os relatórios descrevem claramente o programa (Questão 7), apresentando seu contexto (Questão 8) em relação à reforma agrária no Brasil no período de 1995 e 1998. Incluem também a descrição das atribuições do Ministério do Desenvolvimento Agrário – DA; o processo de reestruturação do INCRA, órgão com atuação na área da reforma agrária; e a descrição da atuação do PRONAF, que presta assistência técnica e capacitação aos assentados.

As finalidades do programa (Questão 9) foram claramente descritas. Os objetivos do Programa Novo Mundo Rural – Assentamento de Trabalhadores Rurais foram:

*Concessão de crédito para a aquisição de imóveis rurais – Banco da Terra;*  
*Concessão de Crédito-Instalação às Famílias assentadas;*  
*Sistema de Informações de Projetos de Reforma Agrária – SIPRA;*  
*Discriminação de áreas Devolutas e Terras Irregularmente Ocupadas;*  
*Assistência jurídica e social às famílias acampadas;*  
*Titulação de terras;*  
*Identificação de imóveis passíveis de desapropriação e aquisição;*  
*Perícia judicial em ações de desapropriação de imóveis;*  
*Topografia em áreas de assentamento rural;*  
*Plano de Desenvolvimento do Assentamento Rural; e,*  
*Obtenção de terras.*

Do programa Novo Mundo Rural – Consolidação de Assentamentos:

*Concessão de crédito para implantação de infra-estrutura básica – Banco da Terra;*  
*Implantação e investimentos de infra-estrutura básica em assentamentos rurais – combate à pobreza rural;*  
*Acompanhamento da instalação de projetos de assentamentos rurais;*  
*Assistência técnica e capacitação de assentados (Brasil, 2003h, p. 10).*

Quanto à Questão 10, os procedimentos metodológicos da auditoria foram descritos parcialmente, uma vez que o delineamento indicado foi “estudo de caso e a pesquisa” (p. 25). Entretanto, verificou-se a realização da coleta de dados nos estados de São Paulo, Paraná, Mato Grosso, Ceará, Maranhão e Pará, num total de 7 assentamentos (Brasil, 2003h). Não há esclarecimento sobre o termo “pesquisa”, conforme relatório de planejamento e relatório final trata-se de aplicação de entrevistas ao público-alvo e envolvidos com o programa.

Verificou-se descrição clara dos resultados, conclusões e recomendações no relatório (Questões 11, 12 e 13 respectivamente). Quanto aos resultados (Questão 11), foram descritos metas quantitativas para os assentamentos rurais entre os anos de 1995 a 2003 e os percentuais de área de assentamento por Superintendência Regional do INCRA. Em relação aos Estados Brasileiros, apresentaram-se Indicadores de Desempenho e resultados da

execução orçamentária do programa. Foram listadas as responsabilidades dos agentes envolvidos nos programas nos níveis federal e estadual, a responsabilidade no processo de tomada de decisão, os processos organizacionais e os sistemas de controle. Também foram apresentadas as determinações, conclusões (Questão 12) e proposta de encaminhamento (Questão 13) em relação a todos os itens referenciados nos objetivos e perguntas de auditoria.

A auditoria considerou e avaliou as posturas dos diferentes grupos de interesse (Questão 14), o que pode ser verificado por meio da análise *Stakeholder*, relacionando atores, seus interesses no programa, bem como o impacto, se não atendido o interesse e a prioridade desses atores para a auditoria. Atores: INCRA, assentados, organização de assentados, movimentos sociais, ONGs de desenvolvimento rural sustentável, IBAMA, políticos locais (Prefeitos e Vereadores). Em relação à políticos locais obtiveram prioridade 7 (prioridades de 1 a 7) e não foram ouvidos os políticos locais no levantamento de informações da auditoria, embora tenham sido elaboradas conclusões da auditoria que evoquem os diferentes interesses políticos na reforma agrária, questão de assentamentos rurais e sua influência política nos mesmos.

Os recursos da avaliação (tempo, dinheiro e funcionários), Questão 15, foram considerados adequados, uma vez que não houve citação de restrições nos documentos consultados. Verificaram-se, entretanto, limitações para a realização da auditoria como: “baixa confiabilidade dos dados disponíveis (em razão da não alimentação dos sistemas pelas unidades executoras). Há também a dificuldade de acesso às associações, o que poderá dificultar a pesquisa” (Brasil, 2003h, p. 26).

Em relação à Questão 16, nesta e nas demais auditorias, os instrumentos estão anexos na maioria dos relatórios de planejamento. Entretanto, não há referência acerca do processo de construção ou validade dos instrumentos.

Entende-se que os instrumentos de coleta de dados, entrevistas-estruturadas e roteiro de entrevista para o grupo focal foram adequados (Questão 17), embora não tenha sido testada sua validade. Apesar de ter ocorrido estudo-piloto para teste do instrumental, não foi relatada modificação nos mesmos. Conforme anexos do relatório de planejamento, os questionários foram formados por questões abertas dirigidas a: beneficiários da reforma agrária e responsáveis pela elaboração do PDA. Apresentou-se: roteiro de grupo focal com técnicos do INCRA; entrevista com técnicos do INCRA responsáveis pelo acompanhamento; entrevista

com técnicos responsáveis pela assistência técnica e entrevista com as associações de assentados.

Adotaram-se medidas de forma parcial para garantir o mínimo de erros na coleta de dados (Questão 18), o estudo-piloto pode ser considerado uma estratégia para testar instrumentos e procedimentos, embora não tenha sido esclarecido o tipo de adequação dos instrumentos. Conforme citado “os levantamentos realizados contribuíram para o aperfeiçoamento dos instrumentos de levantamento de dados [...]” (Brasil, 2003h, p. 28). Também não foi realizada comparação de variáveis relativas aos assentados, trata-se de um estudo com delineamento não-experimental, por sua vez, sujeito as questões de validade interna e externa. O delineamento de séries temporais poderia fornecer exemplificação da evolução dos assentamentos.

Uma fonte de erro pode ter afetado os resultados da auditoria do Programa Novo Mundo Rural, no caso uma variável interveniente, o Projeto Lumiar, que começou a ser implantado em 1997 nos assentamentos rurais, embora o projeto tenha sido suspenso em 2000 (Brasil, 2003 h).

Quanto à análise da informação quantitativa (Questão 19), a descrição precisa de dados não foi adotada nos relatórios. Isso pode ser verificado na citação de frases vagas como: “Tanto os gerentes do Programa quanto os agentes envolvidos na prestação de serviços aos assentados *opinaram* no sentido de que a capacitação dos beneficiários é de tal importância [...]” (Brasil, 2003h, p. 20). Essa descrição de resultado parcial não indica percentuais em relação às demais alternativas da mesma questão.

Por fim, a análise da informação qualitativa foi realizada parcialmente (Questão 20), e relatou-se o uso da entrevista, mas os resultados não foram referenciados diretamente nos relatórios nem na fase de planejamento, quando também ocorre coleta de dados, nem na fase de execução. Não houve exemplificações dos resultados das entrevistas nos documentos consultados.

## ***2. Programa Combate ao Abuso e à Exploração Sexual de Crianças e Adolescentes - Sentinela***

Em relação ao Programa Sentinela, os públicos da auditoria foram descritos (Questão 1) a saber: gestor e técnicos do Centro de Referência (unidade de atendimento); Coordenação

Estadual do Programa, Secretaria Municipal de Assistência e mães de crianças e adolescentes atendidos pelo programa.

A equipe de auditoria não possuía treinamento adequado para realizar a auditoria (Questão 2). Não foram utilizados serviços de consultores da área abrangência temática do Programa (Questão 3). Verificou-se não compreensão dos pressupostos teóricos que embasam o programa, uma vez que foram traçadas recomendações não aplicáveis. Por exemplo, em Brasil (2004d, p. 59) foi sugerida adoção dos seguintes indicadores de desempenho do programa: “Proporção de encaminhamento para o PETI; proporção de famílias de baixa renda atendidas pelo programa”. É polêmica entre os estudiosos do assunto e criadores da política voltada para a criança e o adolescente, a concepção prostituição infantil/adolescente como trabalho infantil. O PETI, em 2006, passou a ampliar as situações de trabalho infantil visadas pelo programa, pautando-se na Classificação Brasileira de Ocupações – CBO. Tal classificação contempla somente atividades legais; as ilícitas, como trabalho infantil doméstico, tráfico de seres humanos e prostituição infantil não compõem a referida classificação. Há entidades como Organização Internacional do Trabalho – OIT e Fórum Nacional de Erradicação do Trabalho Infantil – FNPETI que consideram a prostituição infantil como trabalho infantil. Entretanto, até o presente momento, tal discussão e conceituação de atendimento não se encontram contempladas nos referidos Programas.

Verificou-se que as informações coletadas possuíram alcance parcial para entender o funcionamento do Sentinela (Questão 4). Conforme comentado, não houve compreensão suficiente da política e contexto do Programa. O Programa possui oito ações, porém, só foi auditada a ação Serviço de Proteção Socioassistencial às Crianças e aos Adolescentes Vítimas de Violência, Abuso e Exploração Sexual. A justificativa para tal foi “em razão da elevada materialidade, que corresponde à cerca de 75% dos recursos do programa” (Brasil, 2004d, p. 17). Sugere-se que em auditorias futuras se considerem as demais ações do programa para melhor compreensão de seu funcionamento e resultados.

Quanto às informações coletadas, elas refletiram em parte os objetivos da avaliação (Questão 5). A auditoria procurou responder às seguintes questões:

- a) *Em que medida o processo de atendimento é suficiente para assegurar a proteção das vítimas e o acesso às políticas públicas pertinentes?*
- b) *Em que medida os serviços do Sentinela asseguram o atendimento, encaminhamento e acompanhamento dos beneficiários?*

c) *Em que medida as ações contribuem para a melhoria da qualidade de vida dos beneficiários, fortalecendo sua auto-estima e convivência familiar?* Brasil (2004h, p. 3).

Os itens (a) e (b) foram respondidos satisfatoriamente, entretanto, não houve como avaliar a auto-estima e a convivência familiar, uma vez que somente mães foram entrevistadas, não foram ouvidos os pais e as crianças/adolescentes. Tampouco foram utilizados instrumentos válidos para mensurar auto-estima, além disso, não houve possibilidade de comparar o resultado com antes do acesso e após o acesso ao programa.

As técnicas de coleta de dados foram explicitadas (Questão 6), constou de entrevista estruturada para o gestor do Centro de Referência; para os técnicos; coordenação estadual e Secretaria Municipal de Assistência Social. Por fim, uma entrevista estruturada com treze questões abertas para as mães dos atendidos no programa. Na entrevista com as mães, foram abordadas dez questões sobre caracterização do atendimento e três questões sobre o convívio familiar, não houve na entrevista questões sobre auto-estima.

Os relatórios descrevem adequadamente o Programa Sentinela (Questão 7), com a preocupação em descrever a legislação específica do Programa.

O contexto do Programa foi descrito adequadamente (Questão 8). A partir da visão geral do programa, foi traçado um histórico e caracterização da ação auditada; seus objetivos, monitoramento e legislação. Alguns aspectos em relação ao contexto da exploração sexual são citados, por exemplo:

*A dominação de gênero e de raça se evidencia pelo fato de que a grande maioria das vítimas de exploração sexual é do sexo feminino, geralmente são mulheres, negras e mulatas (CECRIA, 2000).*

*A violência sexual contra crianças e adolescentes é um problema mundial. Por ser ilegal, clandestina e em grande parte doméstica, é uma questão ainda pouco visível e difícil de ser quantificada. O mais freqüente tipo de violência a que estão sujeitas crianças e adolescentes é aquele denominado estrutural, em função da precária situação socioeconômica das famílias das quais grande parte das crianças se originam (Brasil, 2004h, p. 10).*

Além desta contextualização, foram levantadas informações sobre o tema, apresentando-se um mapa de modalidades de violência sexual por região do país.

Quanto aos objetivos do Sentinela, foram descritos adequadamente (Questão 9) a saber: “prevenir e combater a violência, o abuso e a exploração sexual comercial de crianças e adolescentes [...]” (Brasil, 2004h, p. 9), sendo que o programa atua em dois eixos: prevenção e atendimento.

Em relação aos procedimentos metodológicos adotados (Questão 10), houve descrição parcial nos relatórios. O estudo-piloto foi adequadamente descrito, os instrumentos constam nos anexos do relatório de planejamento e os critérios de seleção da amostra foram descritos (distância geográfica da capital e quantidade de atendimentos). Porém, não foram descritos claramente os procedimentos de análise de dados.

Os resultados da auditoria e suas conclusões (Questões 11 e 12) foram descritos parcialmente. Foram analisadas e apresentadas as características orçamentárias e financeiras, indicando que o programa teve uma execução média de 90% no período de 2002 a 2003. Os resultados sobre o sistema de controle demonstraram que o programa não possuía recursos financeiros e humanos suficientes para o acompanhamento das ações. Foi indicada também limitada abrangência do programa em relação ao problema, pois atendia no período avaliado a 331 municípios, dos 5.561 existentes no país. A auditoria indicou falta de critérios equitativos para a definição do tamanho e número das equipes de atendimento do Programa nas unidades. Outro resultado apresentado foi a dificuldade de articulação nas esferas federal, estadual e municipal do governo, em relação às demais políticas sociais. Quanto ao atendimento, encaminhamento e acompanhamento destacaram-se:

*[...] insuficiência de estrutura e instrumentos para a realização da busca ativa;*  
*Alta rotatividade de pessoal;*  
*Capacitação pontual, não sistemática e genérica;*  
*Falta de comunicação entre as equipes do Programa;*  
*Ausência de advogados nas equipes de atendimento;*  
*Execução de atividades de prevenção sem a devida contrapartida (da comunidade e órgãos envolvidos na rede de serviços);*  
*Falta de uniformidade dos conceitos do Programa;*  
*Não encaminhamento das informações dos municípios aos estados;*  
*Especialização do atendimento (predominando o atendimento psicoterapêutico, voltado aos casos de abuso sexual), embora esteja previsto um atendimento mais*

*amplo, com encaminhamento, por exemplo, a serviços de saúde mental. Tal deficiência caracteriza as limitações ao encaminhamento à rede de atendimento; Insuficiência de infra-estrutura e material para a realização do atendimento; Funcionamento diurno com recursos para diuturno, Brasil (2004d, p. 39-48).*

Observe-se nessa descrição o destaque do aspecto contábil, o uso incorreto do recurso financeiro no Programa. Por fim, quanto à efetividade do Sentinela foram apontados três aspectos: aumento do número de denúncias; fortalecimento social e emocional dos beneficiários e outros achados. Esse último aspecto “outros achados” é o que mais carece de sustentação empírica e conceitual (Questão 12), o que transparece na frase “apresentam, por meio de expressões e indicadores diversos, o impacto positivo que as ações do Programa Sentinela tem apresentado” (Brasil 2004d, p. 55). Conforme comentado anteriormente, os instrumentos usados para verificar a relação Programa e auto-estima e Programa e convivência familiar não foram válidos (Questão 18), pois a estratégia de coleta e análise de dados para responder a essas questões foi inadequada.

Quanto às justificativas das recomendações (Questão 13), foram realizadas com base nos resultados da auditoria, mesmo que em alguns casos estejam sob bases conceituais não apropriadas ao programa, conforme comentado nas questões 2 e 3.

Não foram consideradas as posturas dos diferentes grupos de interesse (Questão 14). Os respondentes da auditoria resumiram-se a público-alvo do programa e gestores e técnicos do programa nos níveis federal, estadual e municipal.

Os recursos tempo, dinheiro e funcionários foram adequados (Questão 15), uma vez que não houve comentários a esse respeito no relatório de planejamento. Entretanto, conforme questão 3, o fato da auditoria não ter consultado consultores externos *experts* na questão da exploração e abuso sexual, pode ter levado ao entendimento equivocado de algumas questões.

Os procedimentos de coleta de informações foram descritos adequadamente (Questão 16), com os instrumentos anexados ao final do relatório de auditoria.

Os instrumentos não foram válidos e re-adequados após o teste-piloto, não se utilizaram questões específicas para mensurar auto-estima dos participantes do programa (Questão 17) Além de não ter sido utilizado grupo de controle, não foram tomadas medidas para assegurar o mínimo de erros na coleta de dados (Questão 18).

A informação quantitativa foi adequadamente analisada (Questão 19), a saber: análise orçamentária financeira, mapa de tipos de caso de exploração sexual no Brasil, por região e análises estatísticas descritivas dos atendimentos, além da proposição de indicadores de desempenho.

Considerando-se a quantidade de informações qualitativas coletadas, não se esclareceu o tipo de procedimento usado para analisá-la. Foi citado o *benchmarking*, que é uma análise comparativa, mas não houve nos resultados referências diretas com transcrições de pontos importantes das falas, ou comparações para embasar os resultados apresentados (Questão 20). Portanto a quantidade não foi adequadamente analisada.

### ***3. Programa Reinserção Social do Adolescente em Conflito com a Lei***

De acordo com os documentos consultados, os públicos da auditoria foram identificados (Questão 1) através da análise *Stakeholder*. Um aspecto positivo da auditoria deste programa foi a abrangência dos atores participantes da auditoria: gestores federais; estaduais; integrantes do poder judiciário; membros do Ministério Público; Defensores Públicos; Conselho Estadual; Diretores de Entidades de Atendimento; técnicos do Programa; adolescentes atendidos e egressos.

A equipe da auditoria possuía treinamento adequado para a realização do trabalho (Questão 2). Utilizaram-se os serviços de consultores externos através de parcerias (Questão 3). Assim, profissionais de áreas específicas foram consultados tanto na fase de planejamento quanto na execução.

As informações coletadas foram suficientes para entender o funcionamento do Programa (Questão 4). A auditoria apresentou o marco legal do Programa e atribuições da instância federal, estadual e municipal. O Programa era formado por 12 ações. Entretanto, a auditoria centrou seus esforços em seis delas, apresentando a justificativa de relevância, risco e materialidade das ações; as demais não foram analisadas devido à baixa implantação. O objeto da auditoria foi:

- a) *Implantação de serviços de atendimento a adolescentes com medidas socioeducativas não privativas de liberdade;*
- b) *Capacitação de recursos dos Sistemas de Segurança, Justiça e Atendimentos ao Adolescente em conflito com a lei;*

- c) *Implantação de serviços de acompanhamento do sistema socioeducativo;*
- d) *Implantação de serviços sociopsicopedagógicos destinados aos adolescentes em conflito com a lei e sua família;*
- e) *Atendimento socioeducativo ao adolescente em conflito com a lei* (Brasil, 2003d, p. 17).

Apesar da baixa taxa de retorno de questionários enviados para egressos e gestores estaduais, as informações coletadas forneceram um panorama do Programa e foram suficientes para atingir os objetivos da auditoria (Questão 5). Para alcançar os objetivos foram elaboradas as seguintes questões de auditoria:

- 1) *Em que medida normas, diretrizes e ações governamentais constituem uma política nacional integrada de atendimento ao adolescente em conflito com a lei?*
- 2) *Os recursos materiais, orçamentários, financeiros e humanos e de informações da SEDH, dos estados e das entidades de atendimento e a qualidade da formação profissional dos responsáveis pela implantação do Programa são suficientes e compatíveis com os objetivos traçados?*
- 3) *Em que medida as ações do Programa são consistentes com o Estatuto da Criança e do Adolescente, no que se refere à execução das medidas socioeducativas?*
- 4) *Qual o grau de implantação das atividades de apoio e acompanhamento dos egressos?* (Brasil, 2003d, p. 18).

Sugere-se para fins de clareza dos relatórios e compreensão dos resultados que se pretende obter, tanto para os auditores quanto para os demais interessados nos resultados do programa, que na elaboração das perguntas da auditoria se expresse uma pergunta de cada vez. Isso pode ser bem exemplificado nos itens 1, 2, e 4 da citação; já não foi o caso do item 3, que expressa dois objetivos. Essa sugestão, não possui somente um caráter didático ou de forma. Espera-se que ao expressar uma questão dessa maneira, se esteja direcionando a investigação para aspectos específicos do problema, expressados sob enfoques diferentes em cada questão.

As técnicas de análise de dados foram explicitadas adequadamente (Questão 6): análise qualitativa dos dados; tratamento estatístico; análise de conteúdo; *Benchmarking* das estruturas organizacionais dos Estados. Além dessas, verificou-se a análise orçamentária e

financeira do Programa. Um aspecto positivo a ser destacado foi que a auditoria buscou informação de todos os Estados.

O relatório descreveu claramente o Programa (Questão 7), explicitou seus objetivos, ações que o compõem, leis que o sustentam, responsáveis e suas atribuições.

O relatório descreveu adequadamente o contexto do Programa (Questão 8), citando opiniões de estudiosos do assunto e autoridades interessadas e mobilizadas pela questão do adolescente em conflito com a lei.

Quanto aos objetivos do Programa (Questão 9), foram descritos claramente. De forma geral, o programa possuía o seguinte objetivo: “articular e estimular os esforços do sistema socioeducativo instituído pelo Estatuto da Criança e do Adolescente – ECA, de forma a possibilitar a inclusão dos adolescentes em conflito com a lei no meio social” (Brasil, 2003d, p. 13)

Os procedimentos metodológicos foram descritos parcialmente (Questão 10). O delineamento da auditoria não foi explicitado, embora tenha sido realizado um estudo do tipo não experimental. Foi utilizado estudo-piloto, “estudos de casos” em seis Estados, aplicação de questionários e entrevistas e realizados grupos focais. Em relação aos “estudos de caso”, verificou-se tratar de análise comparativa, a análise *Benchmarking*. A citação de análise qualitativa não esclarece exatamente qual o procedimento adotado.

O relatório descreveu a maioria dos resultados obtidos; entretanto, em alguns casos de forma vaga e parcial (Questão 11). Por exemplo, foi apresentado um gráfico de distribuição de recursos de convênios por estado; porém, não houve explicações sobre o mesmo no relatório citou-se apenas que as informações foram repassadas pelos Estados. Algumas frases são usadas para reforçar os resultados, mas não esclarecem a informação. Por exemplo, ao afirmar que a articulação entre a União, Estados e Municípios é incipiente:

*Tais diagnósticos foram auferidos por meio de entrevistas, questionários e visitas em que foram evidenciadas as seguintes situações [...] (Brasil, 2003d, p. 30).*

Em relação à integração com outras áreas governamentais:

*Tal conclusão baseia-se nos estudos de caso realizados nos estados selecionados, nas informações registradas nas entrevistas com os operadores [...] e nos questionários aplicados aos gestores e a conselheiros de direitos da criança e do adolescente (Brasil, 2003d, p. 31).*

*No olhar dos juízes, promotores e defensores estaduais entrevistados, as ações oriundas do governo federal ainda não apresentam maior efetividade [...] (Brasil, 2003d, p. 31).*

Expressando a opinião dos técnicos:

*Foi bastante citado nos grupos focais que os adolescentes em conflito com a lei enfrentam discriminação na escola, por parte dos diretores, professores e famílias dos alunos (Brasil, 2003d, p. 40).*

Quanto aos recursos materiais, humanos, financeiros e orçamentários:

*Com base nos instrumentos de coleta de dados, constatou-se: baixa informatização das unidades de atendimento e dos conselhos estaduais; falta de recursos para o transporte dos adolescentes [...] (Brasil, 2003d, p. 41).*

Conforme exemplos os dados, não foram apresentados de forma que se entenda a opinião de todos os participantes. As informações parecem vagas devido ao tipo de descrição utilizada; embora talvez os dados não o fossem. O cuidado na apresentação dos dados precisa ser tomado para evitar o viés. Também em função da ausência de informações das gerências não foi possível descrever a quantidade de atendimentos, uma lacuna importante que prejudicou a compreensão do alcance do Programa.

Os resultados apresentados retratam sérias dificuldades do Programa: desarticulação da política de atendimento nos três níveis de governo; deficiência na estrutura física de atendimento e na capacitação de quem oferece as medidas socioeducativas; ênfase em medidas privativas de liberdade ao invés de medidas socioeducativas em meio aberto; fraco atendimento e acompanhamento em relação às famílias e dos egressos. Tais resultados se assemelham aos problemas encontrados em outras partes do mundo, em programas que visam, por exemplo, a redução da criminalidade. Na revisão de programas sobre prevenção de crimes, Dershen (1990), encontrou pontos importantes que corroboram os resultados da auditoria do Reinscrção Social. Por exemplo, importante a participação das pessoas afetadas pelo problema nos programas. Tal participação cria uma sensação de envolvimento, continuidade e maior probabilidade de êxito das ações. Assim, a participação e o enfoque nas famílias dos jovens atendidos tornam-se necessárias. Outro aspecto é que tais programas necessitam de gerenciamento profissional das ações, pois a desarticulação encontrada nos níveis gerenciais

não favorece a efetividade do programa. E por fim, outro aspecto reforçado por Dershen (1990) foi a necessidade de considerar as necessidades específicas da população, baseando o atendimento nas necessidades locais da população, ou seja, nas características demográficas e das infrações cometidas.

É utópico considerar que um Programa Federal poderia “dar conta da situação” das infrações praticadas por jovens, pois há muitos fatores estruturais, culturais e econômicos associados (Dershen, 1990). Entende-se que a política na qual o programa se insere necessita de diretrizes gerais. Entretanto, considerar as especificidades parece ser um diferencial de sucesso, conforme resultados encontrados em Programas que visam a segurança da população (Dershen, 1990).

As conclusões foram apresentadas de forma clara e sintética nos relatórios em item específico (Questão12). Foram expostos argumentos ressaltando-se os benefícios da implantação das recomendações geradas através dos resultados (Questão 13).

A auditoria considerou a opinião dos diferentes grupos de interesse no programa (Questão14).

Uma vez que não foram apresentados comentários no relatório e os objetivos da auditoria foram atingidos, avalia-se que os recursos para a auditoria foram adequados (tempo, dinheiro e funcionários) (Questão 15).

Os procedimentos de coleta de informação foram descritos com clareza (Questão 16) e apresentados nos anexo do relatório de planejamento. Utilizaram-se questionários e roteiro de entrevista estruturada, além de fonte de dados secundários a partir de documentos do Programa.

Os instrumentos de coleta de dados foram válidos (Questão 17). A auditoria tomou o cuidado de testá-los e modificá-los, após aplicação no estudo-piloto no Distrito Federal.

Infelizmente, não foram tomadas medidas para evitar o mínimo de erros na coleta dos dados (Questão 18), mesmo com a preocupação da auditoria em mapear a opinião de todos os envolvidos ocorreu mortalidade instrumental no caso das respostas dos gestores estaduais, com retorno de apenas 27 questionários dos egressos. Como não houve dados sobre o total de jovens atendidos no Programa, foi impossível estimar se 341 adolescentes (total de questionários enviados) seria uma amostra representativa. Além do mais, o delineamento adotado não subsidia quaisquer tipos de comparações e generalizações em relação aos resultados obtidos pelo Programa.

Em relação à Questão 19, a informação quantitativa foi analisada parcialmente, pois apresentou algumas deficiências na apresentação dos resultados no relatório. Por exemplo, em relação às atribuições dos Conselhos Estaduais e Municipais dos Direitos das Crianças e dos Adolescentes: *“Efetuada a consolidação dos dados registrados nos questionários (grifo nosso) endereçados aos conselhos estaduais [...] foram evidenciadas as seguintes situações”* (Brasil, 2003d, p. 35) e quanto à atuação das organizações não-governamentais – ONG: *“Nas respostas aos questionários, os gestores estaduais são quase unânimes em afirmar que a participação de ONGs tem contribuído para a melhoria da implantação das medidas socioeducativas”* (Brasil, 2003d, p. 33).

Deve-se ressaltar que apesar dessas falhas de comunicação dos resultados, a maioria dos resultados foi apresentada adequadamente, com explicações precisas a respeito das estatísticas utilizadas, como foi o caso dos resultados em relação à atuação policial na opinião dos adolescentes. Uma sugestão em relação à coleta de dados com os adolescentes e egressos, é que se explorem as necessidades apontadas por eles para o desenvolvimento das medidas socioeducativas, uma vez que eles são os sujeitos que vivenciam diretamente o problema da reintegração à sociedade.

Por fim, a análise da informação qualitativa (Questão 20) foi adequada. Dentre os relatórios analisados, este foi o que mais apresentou resultados de análise qualitativa das entrevistas expressados no texto por meio de citações de falas dos participantes. Por exemplo,

*Os conselheiros estaduais responderam [...] “Não há sintonia entre os poderes federais e há confusão na atribuição de cada poder. Há falta de divulgação e visibilidade das ações governamentais...” e “O foco da discussão estadual sobre o adolescente em conflito com a lei requer a construção de uma política estadual e nacional que incorpore ações intersecretarias [...]”.*

*Segundo um dos operadores de direito entrevistados, os cursos são uma “verdadeira enganação”. Sobre a qualidade dos cursos [...] “a nossa secretaria oferece uns cursos que na verdade não atende o interesse deles, e é de péssima qualidade, de péssima qualidade”* (Brasil, 2003d, p. 41).

Por fim, uma colaboração de Kanter (2001) em seu artigo sobre Programas de extensão de atividades após a escola regular. Apesar do enfoque do artigo não ser em adolescentes em conflito com a lei e sim em adolescentes em geral, algumas comparações são oportunas e possíveis. O autor apresenta dados estatísticos indicando que entre 15h e 18h ocorrem os

maiores problemas entre os adolescentes (entre 16 e 17 anos), sejam crimes, envolvimento com drogas ou prática de atividades sexuais. Kanter (2001) argumentou que o oferecimento de atividades culturais, esportivas e de desenvolvimento das potencialidades é um aspecto importante em defesa dos programas pós-escola, uma vez que pais (homens e mulheres) cada vez mais ingressam no mercado de trabalho e possuem menos tempo para acompanhar, compartilhar e educar seus filhos.

O autor apresentou também a opinião dos jovens em relação às atividades oferecidas por esses programas. De acordo com sua pesquisa, os jovens apreciam as atividades oferecidas, pois são mais interessantes que as oferecidas no horário regular das aulas; são mais divertidas (lúdicas). No entanto, acham necessário o apoio pedagógico oferecido para o “*dever de casa*”.

#### ***4. Programa Desenvolvimento do Turismo no Nordeste – Prodetur/NE***

O público da avaliação foi claramente identificado (Questão 1) conforme matriz de planejamento e relatório final. Verificou-se a consulta e utilização de fonte de informação como relatórios de instituições federais; por exemplo, o Ministério do Trabalho e Emprego – MTE, Relatórios internos do Prodetur/NE; Associação Brasileira da Indústria de Hotelaria – ABIH; Pesquisa Nacional por Amostra de Domicílio – PNAD; estudos desenvolvidos pelo Banco Nacional de Desenvolvimento Econômico e Social - BNDES; Banco Interamericano de Desenvolvimento - BID, Instituto de Pesquisa Econômica Aplicada - IPEA; Universidade Federal do Ceará – UFCE; Planos de Desenvolvimento Integrado de Turismo Sustentável da 1ª e 2ª fase do Prodetur/NE. O público consultado foram os gestores das Unidades Executoras Estaduais e os Conselheiros dos Pólos de Turismo por meio de questionários. Nesse sentido, Woodside e Sakai (2001) destacam a necessidade de programas de *marketing* e turismo federais coletarem dados diretamente com consumidores, clientes, competidores, fornecedores, empresas de marketing e públicos do turismo. Sugerem também o mapeamento dos demais *Stakeholders* envolvidos (mercados, indústria, clientes e competidores), o que pode permanecer como sugestão para futuras auditorias da organização nessa área.

Considerando-se todos os resultados da ANOP e pontuações nos demais critérios do *Joint Committee*, a equipe de auditoria não possuía treinamento adequado suficiente para realizar a auditoria (Questão 2).

Na fase de planejamento da auditoria houve participação de consultores externos (Questão 3), foram consultados técnicos e dirigentes do Ministério do Turismo, IPEA, Centro de Excelência em Turismo da Universidade de Brasília (CET/UnB), Secretaria Federal de Controle Interno (SFCI), Secretaria de Assuntos Internacionais (SEAIN), Ministério do Planejamento, Orçamento e Gestão (MPOG) e Banco do Brasil (BB) na fase de planejamento da auditoria e um especialista na área de ocupação produtiva e renda (Brasil, 2003b).

Quanto à Questão 4, as informações coletadas possuíam alcance suficiente para entender o funcionamento do programa. Porém, as informações coletadas foram parcialmente suficientes para refletir os objetivos da avaliação (Questão 5). De acordo com Brasil (2003b, p. 14), a auditoria teve como objetivos investigar as seguintes questões:

- a) *Os investimentos do Prodetur/NE I contribuíram para a melhoria dos indicadores socioeconômicos da região?*
- b) *O Prodetur/NE II contempla ações com vistas a mitigar os impactos ambientais observados no Prodetur I e evitar a ocorrência de novos?*
- c) *A estrutura organizacional, nos três níveis de governo, e a atuação dos conselhos de turismo são adequadas para a implantação do Programa?*
- d) *Os custos financeiros a que estão sujeitos os submutuários são adequados?*

O Prodetur/NE (1ª fase) investiu 5% do valor de suas despesas em estudos (Brasil, 2003b; porém, não produziu estudos sobre impactos ambientais do programa e indicadores. Conforme informações da auditoria, “Os trabalhos de auditoria sofreram restrições impostas pela inexistência de indicadores ambientais de desempenho e inexistência, ou baixa confiabilidade das informações ambientais dos PDITS, Planos de Desenvolvimento Integrado de Turismo Sustentável (Brasil, 2003b, p. 16). Mesmo utilizando o delineamento de séries temporais para o estudo do Prodetur/NE, a auditoria esteve sujeita a ausência de dados para responder de forma válida as questões (a) e (b) das perguntas de auditoria. Na tentativa de resposta ao impacto socioeconômico, buscou-se informação de fontes secundárias, dados intersetoriais, “uma vez que o turismo possui elevada relação com outros setores da economia” (Brasil, 2004a, p. 28), e da utilização da matriz de insumo-produto, com dados sobre emprego e renda. “Esse modelo fundamenta-se no equilíbrio entre oferta e demanda agregadas e leva em conta a importância das atividades econômicas dentro de um espaço econômico definido (país, região ou estado)”. A auditoria utilizou também um trabalho

acadêmico patrocinado pela Embratur, dados da Relação Anual de Informações Sociais – RAIS e Pesquisa Nacional por Amostra de Domicílio – PNAD, a partir dos quais fez algumas projeções que se mostraram favoráveis à elevação dos níveis de emprego e renda no período do Programa. Entretanto, conforme reconhecido pela própria auditoria “qualquer transposição dos seus resultados para o Nordeste seria temerária” (Brasil, 2003b, p. 30). Nesse sentido, é interessante a observação de Woodside e Sakai (2001), de que raramente as auditorias são planejadas e implementadas de forma que não confirmem as crenças positivas existentes sobre o programa.

As técnicas de análise de dados foram explicitadas claramente (Questão 6) e anexadas no relatório de planejamento.

O programa foi descrito claramente (Questão 7), a equipe de auditoria o descreveu por meio de normas, documentos internos e legislação. A descrição foi detalhada o suficiente para facilitar a compreensão do programa.

O relatório descreveu parcialmente o contexto do programa (Questão 8). Uma vez que o contexto do atual programa de turismo não foi totalmente investigado, embora, tenha sido citado em Brasil (2004a, p. 27) que:

*[...] os agentes econômicos (hotéis, restaurantes, bares, agências de turismo, empresas de transporte rodoviário e aéreo, empreendimentos recreativos) passam a perceber que a infra-estrutura posta à disposição do turista melhorou e assim sentem-se estimulados a investir na região, ampliando a oferta turística, tanto de bens e serviços específicos quanto dos relacionados.*

Verificou-se que não foram coletadas informações diretamente junto a esses agentes econômicos. De acordo com Woodside e Sakai (2001), a auditoria pode incluir métodos objetivos e subjetivos para identificar diferentes visões da realidade e do contexto. As estratégias avançadas de *marketing* turístico incluem uso e manutenção de banco de dados com endereços, nomes, telefones dos clientes, sistemas *on line* com informações para realizar os cruzamentos das preferências dos clientes com ofertas turísticas. Os autores destacam que os programas federais de marketing e turismo raramente utilizam firmemente essas estratégias.

O relatório descreveu claramente os objetivos do programa (Questão 9), que foram apresentados tanto no relatório de planejamento quanto no relatório de auditoria. Os objetivos do Prodetur/NE 1ª e 2ª fase foram:

*Prodetur/NE I*

*Reforçar a capacidade da Região Nordeste em manter e expandir sua crescente indústria turística, contribuindo para o desenvolvimento socioeconômico regional. O Programa pretende prover de infra-estrutura básica e serviços públicos a áreas atualmente de expansão turística e onde a capacidade do Estado não acompanha a demanda por tais serviços; os investimentos beneficiarão principalmente a população de baixa renda das áreas selecionadas. Com a melhoria das condições das áreas turísticas, busca-se também atrair atividades turísticas privadas adicionais e de melhor padrão, gerando assim oportunidades de emprego e aumento dos níveis de renda e das receitas públicas (Brasil, 2003b, p. 2).*

*Prodetur/NE II*

*A sustentabilidade dos investimentos realizados na primeira fase do programa, de forma a prioriza ações complementares aos projetos já financiados, segundo três componentes fundamentais: fortalecimento da capacidade municipal para a gestão do turismo; planejamento estratégico, treinamento e infra-estrutura para o crescimento turístico; e promoção de investimento do setor privado (Brasil, 2003b, p. 2).*

O relatório descreveu parcialmente os procedimentos metodológicos da auditoria (Questão 10); uma vez que foi citado em Brasil (2003b, p. 17):

*[...] os dados coletados foram analisados por meio dos seguintes métodos: análise de séries temporal, análise qualitativa de dados, análise de conteúdo das entrevistas e grupos focais, Benchmarking, análise comparativa, e tabulação de questionários e entrevistas estruturadas (Brasil, 2003b, p. 17).*

Nesse sentido, são necessárias três observações. Primeiro, na revisão de literatura de Woodside e Sakai (2001), não foi encontrada aplicação nos delineamentos das auditorias, de grupos de controle e grupos de pessoas expostas ao programa, para comparar o grupo de tratamento ao grupo de controle. Confirmou-se, dessa forma, que tais pesquisas avaliativas carecem de evidências válidas sobre os programas. Não se tomou nenhum cuidado para evitar erros. Segundo, a referência à análise de série temporal, foi realizada para demonstrar o crescimento médio da população ocupada no turismo no Nordeste e no Brasil não no Nordeste, no período de 1992 a 2001, com dados do IBGE/IPEA, ou seja, fonte de dados

secundários e sem cruzamento com dados produzidos pelo próprio Programa. E, terceiro, não há referências diretas com citações, aos resultados das entrevistas, por meio da análise de conteúdo.

O relatório descreveu claramente os resultados da auditoria (Questão 11). Pode-se verificar tais resultados através da apresentação de textos, análises estatísticas descritivas (percentuais, totais) dos indicadores de desempenho e gráfico de barras da série temporal de crescimento médio da população ocupada no turismo. Embora não seja uma pergunta direta da auditoria, foram analisados dados em relação à questão do gênero dos trabalhadores na área do turismo, com base nos dados da PNAD, embora “Assim, como nas análises anteriores, não se pode atribuir esse comportamento ao Prodetur” (Brasil, 2003b, p. 35).

O relatório descreveu parcialmente as conclusões da auditoria (Questão 12). Nas conclusões houve ênfase na afirmação que os impactos socioeconômicos não poderiam ser atribuídos ao Prodetur/NE I, o que foi explicitado pela ausência de dados que deveriam ser produzidos pelo programa no período.

Entretanto, há um aspecto que não foi esclarecido: ao contrário do que foi afirmado na segunda pergunta de auditoria no relatório de planejamento: “O Prodetur/NE II contempla ações com vistas a mitigar os impactos ambientais observados no Prodetur I e evitar a ocorrência de novo?” (Brasil, 2003b, p. 14). Tem-se como resposta no relatório final: “Constata-se que o Prodetur/NE I não provocou impactos ambientais significativos” (Brasil, 2004a, p. 39) e mais adiante no mesmo relatório, nas conclusões se lê,

*Verifica-se, ainda, que o passivo ambiental contemporâneo ao programa não pode ser atribuído exclusivamente ao Prodetur/NE I. Nesse sentido, os PDITS dos pólos de turismo e a opinião dos conselheiros convergem no sentido de que o passivo ambiental relevante da região foi provocado pela dinâmica socioeconômica local* (Brasil, 2004a, p. 61).

Não pareceu claro se o programa, na primeira fase, produziu impacto ambiental ou não. A conclusão se tornou dúvida, à medida que foi formada a partir da opinião dos conselheiros, conforme demonstrado em Brasil (2004a, p. 40), no gráfico de “Modalidades de passivo ambiental gerados nas áreas de investimento do Prodetur/NE I”, com base nos questionários respondidos pelos Conselheiros de turismo e não com base em um estudo de impactos ambientais. Apesar dessa inconsistência, as demais descrições de conclusões foram claras e cuidadosas. Por exemplo, no aspecto polêmico da inexistência da avaliação de impacto

ambiental por parte do Ministério do Turismo e que impossibilitou a verificação desse aspecto na auditoria, lê-se:

*Constata-se, entretanto, que, apesar da exigência contratual, o Ministério do Turismo não demonstra disposição para realizar avaliação ambiental estratégica para a região, em razão da complexidade, do elevado custo financeiro, da grande dimensão da região, da ausência de consultoria especializada no assunto e do longo tempo demandado para a realização* (Brasil, 2004a, p. 61).

Quanto à justificativa das recomendações a serem adotadas pelo programa (Questão 13), verificou-se que foram apresentadas de forma clara no relatório final.

A auditoria identificou as posturas dos diferentes grupos de interesse no programa (Questão 14), mas não considerou como participantes da pesquisa, conforme comentado na questão 1.

Os recursos da avaliação (tempo, dinheiro e funcionários) foram adequados para a realização das atividades previstas (Questão 15), pois não foram descritas limitações pela equipe de auditora nesse sentido. No relatório de planejamento, encontram-se descritos o plano de visitas em sete Estados, custo do planejamento da auditoria, execução e relatório. Ressaltou-se ainda que “[...] o custo da auditoria (R\$271 mil) representa 0,03% (três centésimos por cento) do valor total do Programa (3 milhões)”, (Brasil, 2003b, p. 18).

Os procedimentos de coleta de informação foram descritos com clareza (Questão 16). Os instrumentos (questionários enviados para unidades executoras estaduais UEEs, Conselheiros de Pólos de Turismo e roteiro de entrevista com UEEs) encontravam-se nos apêndices do relatório de planejamento.

Os instrumentos e procedimentos para a coleta de informações foram parcialmente válidos (Questão 17), já que foram elaborados com base em conhecimentos sobre o programa e auxílio de técnicos do Ministério do Turismo. Entretanto, não há indicação de uso de instrumentos similares utilizados na área, como o instrumento de Porter, sugerido por Woodside e Sakai (2001), ou outro, mesmo considerando que a auditoria consultou literatura acadêmica sobre o assunto.

Adotaram-se parcialmente medidas para assegurar o mínimo de erros na coleta de dados (Questão 18), uma vez que, dos 411 questionários enviados, retornaram 218 e não foi relatada a existência de dados nulos ou extremos na amostra, ou descrição da análise estatística realizada que permita conhecer a margem de erro da análise.

Quanto às Questões 19 e 20 sobre análise quantitativa e qualitativa, foi realizada parcialmente; principalmente em relação à análise de conteúdo não transparece claramente nos textos do relatório final. Por fim, os benefícios do Prodetur/NE I foram demonstrados por meio de fonte de dados secundários, com apresentação de gráficos de barras, gráfico de pizza e tabela.

### ***5. Programa de Desenvolvimento Sustentável de Recursos Hídricos para o Semi-Árido Brasileiro – Pró-Água/Semi-Árido***

O público-alvo do programa foi identificado, assim como os participantes da auditoria (Questão 1). Beneficiários do programa: populações carentes de comunidades rurais isoladas ou de cidades de pequeno e médio porte (20.000 a 50.000 habitantes) do semi-árido brasileiro (Alagoas, Bahia, Ceará, Minas Gerais, Paraíba, Pernambuco, Piauí, Rio Grande do Norte, Sergipe), incluindo o Maranhão.

Participantes da auditoria: gestores federais do Programa (MMA e MI); Unidade de Gerenciamento do Programa - Gestão (UGP-G); associação de usuários; comitês de bacias; Unidades Estaduais de Gestão do Programa (UEGPs); técnicos do Banco Mundial, representantes da Unidade de Monitoramento e Avaliação (UMA), consultores da KPMG e especialistas.

A equipe de auditoria possuía treinamento adequado para realizar a auditoria (Questão 2) e utilizaram-se os serviços de consultores externos na fase de planejamento (Questão 3). Os consultores eram do Banco Mundial, Curso de Ecologia da UnB e da Unesco (na fase de planejamento). Na fase de execução foi contratado um especialista para subsidiar a avaliação de aspectos ligados à equidade (Brasil, 2000d).

As informações foram suficientes para entender o funcionamento do Programa (Questão 4) e deve-se ressaltar, descritas com profundidade e detalhamento. Entretanto, as informações foram parcialmente suficientes para refletir os objetivos da auditoria (Questão 5). A pergunta da auditoria foi “o gerenciamento do Programa tem contribuído para o atingimento dos objetivos de fortalecimento institucional e de redução da pobreza?” (Brasil, 2000d, p. 9). De acordo com Brasil (2000e; 2000d) essa questão foi investigada através de:

*1.O desempenho da Unidade de Gerenciamento do Programa (UGP-G) e da Secretaria de Recursos Hídricos (SRH/MMA);*

*2.O desempenho das Unidades Estaduais de Gerenciamento do Programa (UEGP) e dos órgãos estaduais de administração de recursos hídricos (SE);*

*3.O aproveitamento das experiências bem sucedidas de participação de usuários na gestão da água (Brasil, 2000e, p. 10).*

As questões citadas abordam indiretamente o problema da redução da pobreza. Portanto, houve dificuldade em respondê-la, e o delineamento utilizado não possibilita essa resposta. Segundo, para a auditoria a diminuição da pobreza estaria associada ao abastecimento de água para consumo humano e conseqüente melhoria da qualidade de vida da população local; e a participação comunitária na gestão de recursos hídricos (Brasil, 2000e), assim como obras hidráulicas que apresentem retorno socioeconômico (Brasil, 2000d).

A auditoria ressaltou que o financiamento de projetos e a sustentabilidade dos sistemas de abastecimento estariam relacionados à redução da pobreza. Entretanto, aponta ainda na fase preliminar da auditoria que “o componente de financiamento de obras do Programa ainda não produziu resultados que permitam avaliar seus impactos, pois há apenas um projeto em um município em pleno funcionamento” (Brasil, 2000e, p. 9). A auditoria enfatizou que responderia à questão por meio da coerência e da lógica entre os objetivos de fortalecimento institucional e redução da pobreza e as ações implementadas em nível federal e estadual. Portanto, a forma buscada para responder a questão foi o enfoque na equidade, investigando-se a participação dos mais carentes na gestão de recursos hídricos.

As técnicas de análise de dados explicitadas (Questão 6) foram: mapeamento de produtos e processos, análise RECI, *Benchmarking*, análise *Stakeholder*, análise estatística e análise de conteúdo. A técnica de análise de conteúdo foi mencionada no relatório de planejamento, entretanto não consta no relatório final. Além disso, foram utilizadas entrevistas produzindo material qualitativo para análise na fase de execução da auditoria. Consta ainda o uso de “uma variação do Modelo de Condições Ambientais” (Brasil, 2000e, p. 8).

No relatório houve descrição longa e detalhada de cada etapa da auditoria. O relatório descreveu claramente o Programa (Questão 7), seus componentes (Gestão de Recursos Hídricos; Obras Prioritárias; Estudos e Projetos; Gestão de Recursos Hídricos na Bacia do Rio São Francisco e Componente de Gerenciamento, Monitoria e Avaliação); além dos seus objetivos, estrutura, estratégias e metas e aparato legal.

Entretanto, cabe uma sugestão quanto à forma do mesmo. No relatório final foi apresentada a descrição do programa, os resultados e depois a metodologia. Entende-se que a apresentação da Descrição do Programa, Metodologia e, por fim, Resultados e Recomendações, entende-se que devam ser mais claros. O formato sugerido é o adotado nas demais auditorias meta-avaliadas.

Dentre os componentes, os recursos do Pró-Água/Semi-Árido estão distribuídos em três programas: a) Pró-Água/Infra-estrutura; b) Pró-Água/Gestão e, c) Nossos Rios – São Francisco.

O relatório descreveu claramente o contexto do Pró-Água/Semi-Árido (Questão 8). Apresentou a visão geral da Política de Gestão de Recursos Hídricos no Brasil e as Leis que a fomentam. Nesse sentido, destaca-se que:

*Ao Ministério do Meio Ambiente – MMA, no papel de gestor da Política Nacional de Recursos Hídricos, cabe essencialmente a ação de implantação dessa Política, consubstanciada na Lei nº 9.433,97 (Lei das Águas), que estabelece os seguintes princípios:*

- *A água é um bem de domínio público;*
- *A água é um recurso natural limitado, dotado de valor econômico;*
- *Em situação de escassez, o uso prioritário dos recursos hídricos é o consumo humano e a dessedentação de animais;*
- *A gestão de recursos hídricos deve sempre proporcionar o uso múltiplo das águas;*
- *A bacia hidrográfica é a unidade territorial para implantação da Política Nacional de Recursos Hídricos e atuação do Sistema Nacional de Gerenciamento de Recursos Hídricos;*
- *A gestão dos recursos hídricos deve ser descentralizada e contar com a participação do Poder Público, dos usuários e das comunidades (Brasil, 2000e, p. 4).*

Os objetivos do Pró-Água/Semi-Árido foram descritos claramente (Questão 9), a saber:

- a) Promover o uso racional e sustentável e o gerenciamento participativo dos recursos hídricos; e,*
- b) Prover acesso confiável e sustentável a água para uso doméstico, municipal e outros (Brasil, 2000e, p.6).*

O relatório descreveu parcialmente os procedimentos metodológicos da auditoria (Questão 10). Essa auditoria, em comparação com as demais teve melhor detalhamento de seu planejamento e resultados; talvez isso se deva ao fato de configurar-se como uma auditoria-piloto do Projeto Organização - Reino Unido. Entretanto, não foi claro o tipo de delineamento adotado. Conforme análise, parece tratar de avaliação de impacto com medição somente após o programa e sem grupo de comparação.

O relatório descreveu claramente os resultados da auditoria (Questão 11) que foram agrupados em tópicos:

- Desempenho da Secretaria de Recursos Hídricos SRH/MMA e da UGPG, em que o Programa teve desempenho satisfatório no quesito de internalização dos objetivos e metas; entretanto apresentou sete aspectos que geraram recomendações para melhoria do seu funcionamento.

- Desempenho das Secretarias e Unidades Estaduais, em que é elogiada a Política de Recursos Hídricos e a gestão no nível federal, entretanto ressalta a necessidade de estabelecer melhoria no nível estadual e oferece recomendações nas demais 14 subquestões analisadas pela auditoria.

- Nas experiências de Gestão Participativa, foram destacadas as experiências positivas e recomendada sua disseminação e melhoria em alguns aspectos.

As conclusões da auditoria (Questão 12) são justificadas com base nos resultados e foram apresentadas claramente subsidiando as recomendações.

O relatório justificou de forma evidente suas recomendações (Questão 13), uma vez que para cada conclusão foram apresentadas recomendações, justificadas pelos resultados encontrados.

A auditoria considerou parcialmente a postura dos diferentes grupos de interesse (Questão 14), considerou, sobretudo a opinião de gestores envolvidos no programa, comitês de bacias e associação de usuários, além da consulta de especialistas na área. Entretanto, outras entidades poderiam ter sido consultadas, como a Associação Brasileira de Águas Subterrâneas – ABAS, ou outros grupos da Sociedade Civil impactados pelo Programa, sobretudo em Salvador e Fortaleza, que foram as melhores experiências do Programa.

Os recursos da avaliação (dinheiro e funcionários) foram suficientes. Entretanto, registrou-se como limitação (Questão 15) “a limitação de tempo para a realização dos trabalhos, que é uma das características das auditorias-piloto do Projeto Organização/Reino

Unido, não será possível uma avaliação completa do programa” (Brasil, 2000e, p. 11 e 12). Assim, considerou-se no geral que os recursos foram parcialmente suficientes.

Os procedimentos de coleta de informações foram descritos com clareza (Questão 16), e encontram-se em anexo no relatório de planejamento. Presume-se que os instrumentos foram válidos (Questão 17), considerando-se que houve ajuda de especialistas para sua construção na fase de planejamento. Entretanto, sugere-se que sejam descritos no relatório, quando houver modificações nos instrumentais, como no presente caso. No relatório de planejamento constou “Com base nas respostas, serão reavaliadas as minutas de questionários para os comitês e associações [...]” (Brasil, 2000e, p. 8). Houve indicação de modificação, porém, não foram explicitadas no relatório.

Não foram tomadas medidas para assegurar o mínimo de erros na coleta de dados (Questão 18), não foram descritas quaisquer medidas no relatório de planejamento ou final.

A informação quantitativa foi analisada de forma adequada (Questão 19), foram apresentadas estatísticas descritivas das respostas dos questionários aplicados aos usuários da água, comitês de bacias hidrográficas e entidades equivalentes. As estatísticas foram apresentadas por meio de gráficos de pizza e barra, com explicações no texto. Entretanto, possuem limitações para generalização dos resultados, conforme descrito em Brasil (2000d, p. 25) “Não se trata de amostra estatística, ou seja, os resultados que serão analisados a seguir não podem ser considerados como representativos da opinião dos responsáveis pelas entidades e usuários da água”.

Por fim, as informações qualitativas foram analisadas parcialmente (Questão 20). Foi citada análise de conteúdo, mas não há clareza do uso desse método de análise nos documentos consultados. Constam apenas três citações literais de respostas transcritas de questionários, que expressam opiniões sobre a eficácia das atividades de treinamento realizadas, porém, as demais partes da análise de conteúdo inexistem. Ressaltou-se, entretanto, que foram adequadamente apresentadas as informações da análise *Benchmarking* e das experiências exitosas do Programa.

## ***6. Programa Energia as Pequenas Comunidades***

Os públicos do programa e da auditoria foram identificados (Questão 1). Não obstante, os beneficiários diretos não foram ouvidos, foi citado que “não será possível fazer pesquisa

postal junto aos usuários em face da indisponibilidade de banco de dados com seus nomes e endereços e da dificuldade do próprio correio de chegar aos locais mais isolados” (Brasil, 2003e, p. 29). Foram ouvidas ONGs, Escolas Técnicas, Centros de Pesquisa, especialistas e Universidades.

A equipe de auditoria possuía treinamento adequado parcial para a realização dos trabalhos (Questão 2).

Foram contratados consultores para tarefas específicas da auditoria (Questão 3), como profissional em estatística para cálculo da amostra e consultora para auxílio na matriz de planejamento e criação de roteiro para o grupo focal.

As informações coletadas foram parcialmente suficientes para entender o funcionamento do programa (Questão 4). Devido à mortalidade instrumental dos questionários enviados às Prefeituras, não se obteve um panorama do programa nas comunidades pesquisadas. Além disso, havia ausência de informações confiáveis sobre os sistemas instalados, conforme relatou a equipe de auditoria.

As informações coletadas responderam parcialmente ao problema de auditoria (Questão 5), uma vez que avaliar o desempenho do Programa Energia das Pequenas Comunidades em relação a efetividades, equidade e sustentabilidade – que foi o objetivo da auditoria, sem ter uma amostra generalizável e dados suficientes do Programa, compromete a qualidade dos resultados obtidos.

As técnicas de análise de dados foram parcialmente explicitadas (Questão 6). Na fase de planejamento, foi utilizada a análise SWOT, *Benchmarking*, análise *Stakeholder*, matriz de risco, análise de dados secundários. Entretanto, não foram explicitadas as análises dos dados quantitativos do questionário, dados qualitativos das entrevistas aplicados na fase de execução da auditoria e das informações obtidas nos grupos focais. Os instrumentos de coleta de dados não se encontravam em anexo no relatório de planejamento.

O relatório descreveu parcialmente o Programa (Questão 7). Em comparação com as demais auditorias analisadas a descrição é breve, até mesmo na questão do embasamento legal do Programa foi sucinta. Seu público-alvo abrange cerca de 20 milhões de brasileiros sem acesso a energia elétrica, distribuídos em mais de 100 mil comunidades, 58 mil escolas e 3 mil comunidades indígenas. O programa Energia das Pequenas Comunidades é um dos 17 programas que compõem o Projeto Alvorada. Foram apresentadas brevemente as quatro fases

do programa (Fases I e II no período 1996 à 1997; Fase III em 1998; Fase IV entre 1999 e 2000).

O relatório descreveu de forma resumida e pouco clara o contexto do programa (Questão 8) uma vez que não foi feita abordagem mais ampla sobre a política que cerca a temática, destacando-se apenas que o Programa Energia de Pequenas Comunidades é parte do Programa Avança Brasil e Plano-Plurianual 2000-2003 (Brasil, 2003c).

Os objetivos do programa foram apresentados de forma parcial (Questão 9). Seu objetivo principal é permitir o acesso a energia elétrica às pequenas comunidades isoladas, não atendidas pela rede convencional de distribuição de energia elétrica, para a execução dos programas sociais nas comunidades remotas (Brasil, 2003c).

Os procedimentos metodológicos foram parcialmente descritos e não suficientes para representar a situação do programa, visto que foram verificadas inconsistências nas descrições de dados e resultados (Questão 10). Cita-se o uso de “estudo de caso” em 71 comunidades, distribuídas por 36 municípios. Essa informação é pertinente a amostra e não ao delineamento do tipo estudo de caso. O critério de escolha dos municípios foi intencional: estados que dispunham de Boas Práticas; localização geográfica; existência de sistema em funcionamento. O levantamento de dados se restringiu as organizações, especialistas e gestores.

O relatório descreveu claramente os resultados da auditoria (Questão 11). Os resultados indicaram a existência de falhas graves no funcionamento do programa, como ausência de controle patrimonial e inadequado uso dos recursos; inexistência de cadastro confiável de dados dos sistemas energéticos instalados; baixa integração com outros programas públicos; atendimento em comunidades onde há rede convencional de energia elétrica; inexistência de levantamento das necessidades das comunidades a serem beneficiadas e opção preferencial da tecnologia fotovoltaica. Apesar dos problemas encontrados foi recomendada a adoção de boas práticas identificadas em algumas localidades.

As conclusões foram descritas adequadamente (Questão 12) a partir dos resultados foram tecidas as conclusões e recomendações (Questão 13) pertinentes a cada questão de auditoria.

A auditoria não considerou as posturas dos diferentes grupos de interesse no programa (Questão 14). Conforme explicitado na Questão 1, os grupos de interesse foram identificados, porém, nem todos foram ouvidos, principalmente no que diz respeito ao beneficiário direto do Programa.

O desenvolvimento do trabalho de auditoria enfrentou problemas de tempo e recursos (Questão 15), conforme se verificou “A decisão de restringir o estudo às comunidades selecionadas segundo sua acessibilidade foi norteadada pela escassez de tempo e recursos” (Brasil, 2003e, p. 29), portanto, os recursos foram parcialmente suficientes.

Os procedimentos de coleta de informação não foram descritos com clareza (Questão 16); não se encontram nos anexos, sendo apenas citados no relatório de planejamento como questionário, e apesar da realização de grupo focal, não foram citados entrevista ou roteiro de questões. Diante dessa limitação, não há como inferir a validade dos instrumentos utilizados (Questão 17).

Não foram tomadas medidas para evitar erros na coleta de dados (Questão 18), embora isso tenha sido reconhecido pela auditoria na fase de planejamento, “[...] referente às respostas dos Prefeitos ao questionário postal, pois como não haverá triangulação, as respostas poderão conter um grave viés, já que os prefeitos tenderiam a superestimar os resultados obtidos e ocultar problemas graves” (Brasil, 2003e, p. 29). Embora no mesmo texto tenha sido citado que para contornar esse problema seriam realizadas entrevistas em campo com os Prefeitos, esse procedimento por si só não assegura a confiabilidade das informações.

A informação quantitativa foi adequadamente analisada (Questão 19) e demonstrada em gráficos, mapas e tabelas, estabelecendo-se comparação entre os usos e parâmetros atuais do programa com outras possibilidades de aplicação e uso dos recursos de fonte de energia. Também foi proposta e explicitada a adoção de indicadores para o Programa.

Não transparece análise de informações qualitativas no relatório final (Questão 20); foram utilizadas fontes de dados secundários (estudos específicos na área de energia), porém em forma de análise quantitativa. Não houve clareza do relato dos grupos focais nos resultados.

## **7. Programa Morar Melhor**

Os públicos do Programa e participantes da auditoria foram identificados satisfatoriamente (Questão 1). Os interessados no programa foram descritos por meio da análise *Stakeholder*, a saber: beneficiários finais; Secretaria Nacional de Habitação/Ministério das Cidades; parlamentares e tomadores de recursos (Estados, Municípios e Distrito Federal). O público do programa é formado por famílias de baixa renda, submetidas às condições precárias de habitabilidade.

A equipe de auditoria possuía treinamento adequado para realizar o trabalho de auditoria (Questão 2). Utilizou-se de forma parcial (fase de planejamento) um Painel de Referência, com auxílio de técnicos e especialistas do IPEA, UNB, SFCI e apoio de pessoal especializado do Ministério das Cidades e Caixa Econômica Federal.

Consultores externos participaram somente da fase de planejamento da auditoria (Questão 3), e não na execução, análise e conclusão dos trabalhos, portanto, houve participação parcial de consultores na auditoria.

As informações coletadas possuem alcance suficiente para entender o funcionamento do Programa (Questão 4). Verificou-se que a amostra foi representativa do universo pesquisado, contemplando observação direta em 8 estados (17 municípios que foram visitados pela equipe de auditoria); entrevistas com técnicas sociais em sete Gerências Regionais da Caixa; questionários enviados a 27 Gerências Regionais da Caixa e a 437 municípios (com taxa de retorno de 46,5%) e também foram enviados questionários a beneficiários.

As informações coletadas foram suficientes para refletir os objetivos da auditoria (Questão 5), uma vez que: “[...] foi verificar em que medida a concepção, execução e operacionalização do programa vêm contribuindo para a equidade de atendimento do público-alvo, efetividade dos projetos contratados e sustentabilidade dos empreendimentos” (Brasil, 2004ef, p. 18) e os resultados apresentados referem-se a estas questões especificamente.

As técnicas de análise de dados foram explicitadas (Questão 6), a saber: mapa de produtos e impactos; diagrama de avaliação de riscos; análise *Stakeholder*, análise SWOT, matriz de planejamento e encontram-se nos apêndices do relatório de planejamento. Foram citadas ainda, na matriz de planejamento: análise documental; análise de dados secundários; análise de conteúdo das entrevistas e análise estatística dos questionários.

O relatório descreveu claramente o programa (Questão 7). Foram apresentados os seus *Stakeholders*; legislação; principais produtos; aspectos orçamentários e financeiros; sistema de controle e processo de tomada de decisão (gestão).

O contexto do Programa foi claramente descrito (Questão 8), sua contextualização (ambiente de atuação do Programa); objetivo, público-alvo e tipos de intervenção.

O relatório descreveu claramente os objetivos do Programa (Questão 9), conforme Brasil (2003f, p. 3): “viabilizar o acesso à moradia, bem como elevar os padrões de habitabilidade e de qualidade de vida das famílias de baixa renda, assim consideradas aquelas

com rendimento mensal de até três salários mínimos, que vivem em localidades urbanas e rurais”.

Os procedimentos metodológicos foram descritos parcialmente (Questão 10). Em Brasil (2003f, p. 4), lê-se:

*A estratégia metodológica [...] contemplará consulta à documentação e dados secundários; estudo de caso em 8 estados brasileiros (contemplando 18 municípios e 5 gerências regionais da Caixa); pesquisa postal com 440 municípios; e pesquisa amostral com famílias beneficiadas em projetos implementados em agregados de municípios do Rio Grande do Norte (Acari, Campo Redondo, Lagoa Nova, Pedro Avelino, Grossos, Paranimirim, e São José e Mipibu), e nas cidades de Campo Grande/MS, Palmas/TO e Santo André/SP.*

Encontrou-se imprecisão metodológica na concepção de estudo de caso; na verdade, tratou-se de um estudo comparativo nos Estados (Alagoas, Bahia, Goiás, Paraíba, São Paulo, Rio Grande do Norte, Rio Grande do Sul e Tocantins) e de avaliação de desempenho do Programa. Foi realizada uma análise da série histórica da aplicação de recursos do Programa, com base em dados secundários (CAIXA), para o período de 2000 a 2003 por regiões do país além, do histórico orçamentário e financeiro do programa no referido período; este último não foi relatado no delineamento da auditoria.

O relatório descreveu claramente os resultados do Programa (Questão 11). Conforme pode ser resumido:

*Foi verificado que a contribuição do Programa para a redução do déficit global é diminuta, dado o pequeno volume de recursos alocados frente à demanda por moradia e inadequação dos domicílios que afeta parcela da população brasileira. Os resultados dos estudos de caso mostram que os poucos recursos alocados ao programa têm um impacto maior, especialmente no Nordeste, na diminuição do número de domicílios rústicos, bem como naqueles sem unidade sanitária. Essas duas situações têm implicações diretas na melhoria da saúde dos moradores (Brasil, 2004e, p. 13).*

Foi constatado também que há desequilíbrio na distribuição dos recursos de programas habitacionais operacionalizados pela Caixa, que não considerou as necessidades habitacionais das diversas faixas de renda. Verificou-se que há pouca influência do programa para a

regularização jurídica da propriedade e que há iniquidade no atendimento da população-alvo em São Bernardo do Campo/SP. Porém, no Rio Grande do Norte, são atendidas, na maioria, famílias pobres, com até um salário mínimo. Quanto ao atendimento de pessoas extremamente pobres e presença de pessoas portadoras de necessidade especiais, os resultados indicaram maior atendimento no RN, e situação intermediária em SP e TO.

Um resultado favorável apresentado pela auditoria foi a satisfação dos beneficiários com as novas moradias e “o alto índice de intenção de permanência na nova casa” (Brasil, 2004e, p. 13). Entre os resultados do Morar Melhor foi encontrado achado semelhante ao do Prodetur/NE, que foi a ausência de estudo ambiental antecedendo a instalação do Programa.

O relatório justifica as conclusões da auditoria (Questão 12) e reconhece que:

*Ante a flagrante insuficiência dos recursos para atender à demanda, já se poderia afirmar da ineficácia do programa para a solução do déficit habitacional brasileiro [...] mas, dentro do universo do que foi gasto, quanto dos recursos aplicados atingiu o objetivo, seja de produção de moradias ou de urbanização de áreas, ou se houve perdas significativas por paralisação de obras, por exemplo (Brasil, 2003f, p. 22).*

Nesse sentido, a auditoria concluiu que o Programa conseguiu elevar os padrões de habitabilidade e de qualidade de vida dos beneficiários, com critérios de equidade no seu atendimento.

As recomendações da auditoria foram justificadas claramente (Questão 13). As recomendações baseadas nos resultados encontrados foram dirigidas à Secretaria Nacional de Habitação do Ministério das Cidades e Caixa Econômica Federal.

A auditoria considerou parcialmente a postura de diferentes grupos de interesse no programa (Questão 14), pois identificou o interesse de movimentos sociais e ONGs no programa, entretanto, não os considerou no levantamento de dados.

Os recursos (tempo, dinheiro e funcionários) foram suficientes para a realização da auditoria (Questão 15), uma vez que não foram registradas informações contrárias nos documentos analisados.

Os procedimentos de coleta de dados foram descritos com clareza (Questão 16); utilizou-se grupo focal e nos anexos do relatório de planejamento, consta o roteiro com as questões aplicadas. Questionários foram enviados pelo correio às Prefeituras e foi realizada a aplicação de entrevistas pessoais e domiciliares através de questionário estruturado.

Diferentemente dos outros relatórios de planejamento, o instrumental utilizado foi descrito: “O instrumento de pesquisa postal a ser utilizado junto aos municípios contempla um total de 16 questões objetivas, agrupadas em quatro categorias que visam avaliar os seguintes aspectos: [...]” (Brasil, 2003f, p.17).

Os instrumentos para a coleta de dados foram válidos (Questão 17). É interessante notar que foi esclarecida a base conceitual para medir algumas variáveis, como a melhoria da *habitabilidade*; “a partir do conceito de déficit habitacional quantitativo da Fundação João Pinheiro (coabitação familiar, ônus excessivo com aluguel, diminuição de domicílios rústicos e improvisados)” (Brasil, 2003f, p. 22). Nesse caso, também foi apresentada, no relatório de planejamento, a base conceitual para a construção dos indicadores de desempenho. Entre as bases consideradas, além dos padrões mínimos de arquitetura (direito ao sol, dimensão do imóvel, e existência de janelas), citou-se a Hierarquia de Necessidades de Maslow:

*Tomando como referência a pirâmide de Maslow, a habitação supriria as necessidades primárias (fisiológicas e de segurança), bem como as secundárias (sociais e de estima) dos indivíduos. Talvez fosse importante distinguir as ações que suprem as necessidades primárias dos indivíduos e as ações que atenderiam as necessidades mais elevadas em termos de hierarquia. A obtenção da moradia por si só representa um avanço enorme de qualidade de vida para quem não a possuía. A metodologia proposta para aferição das necessidades habitacionais não hierarquiza as benfeitorias, colocando em um mesmo patamar, por exemplo, a situação de coabitação familiar com a sobrevivência em domicílios improvisados (Brasil, 2003f, p. 25).*

Nesse sentido, mesmo considerando que há teorias contemporâneas que poderiam explicar melhor a motivação humana, foi válida a busca de sustentação teórica para a construção do modelo de indicadores e a transposição de conceitos realizada.

Foi realizado um estudo-piloto em Campo Grande/MS (Questão18), considerou-se o estudo como forma de minimizar erros na coleta de dados. Além disso, “para obter as estimativas do universo da pesquisa com 95% de confiança e erro de amostragem de 5%, será considerada uma amostra aleatória simples de 440 municípios a serem entrevistados por questionário postal” (Brasil, 2003f, p. 18).

A informação quantitativa foi analisada adequadamente (Questão 19) conforme se verificou extensamente no relatório final, por meio das análises estatísticas. Por exemplo, foi

demonstrada a estimativa de *déficit* habitacional urbano por faixas de renda familiar (por regiões); o histórico financeiro e orçamentário do programa e a melhoria das condições de habitabilidade.

A análise qualitativa dos dados foi realizada parcialmente (Questão 20); destacaram-se as definições das variáveis medidas em relação a *déficit* qualitativo e *déficit* quantitativo; no entanto, não transparecem no texto indícios da análise de conteúdo das entrevistas e das questões abertas dos questionários e grupo focal.

### **8. Programa Nacional do Livro Didático – PNL D**

O público da auditoria do Programa Nacional do Livro Didático foi identificado, (Questão 1), conforme análise *Stakeholder*:

*Professores, alunos, donos de editoras, pessoal que trabalha nas duas Gerências do Livro Didático, professores responsáveis pela análise didático-pedagógica dos livros (Secretaria de Ensino Fundamental – SEF/MEC), funcionários das SECEXs estaduais (Brasil, 1999b, p. 8).*

Verificou-se que a equipe de auditoria possuía treinamento adequado para realizar a auditoria (Questão 2). Nesse sentido, a auditoria teve como principal foco a análise custo e efetividade da utilização dos livros didáticos, que segue a linha clássica de auditoria contábil, um dos principais focos de atuação da organização.

Houve participação de consultores na auditoria (Questão 3), conforme relatório de planejamento “[...] foram agendadas, no período, reuniões com os consultores ingleses, com integrantes dos outros grupos de auditoria e com assessores do Ministro Relator [...]” (Brasil, 1999b, p. 7). Entretanto, não se verificou nas demais partes do referido relatório, alusão à participação dos consultores, nem ao tipo de consultoria ou orientação prestada.

Quanto à Questão 4, verificou-se que as informações coletadas não possuíram alcance suficiente para entender o funcionamento do programa, uma vez que a auditoria identificou o público-alvo na fase de planejamento da auditoria, porém, na fase de coleta de dados inquiriu somente gerentes do FNDE, da SEF/MEC, integrantes da SEDUC de Belo Horizonte/MG e São Paulo/SP. Entende-se que seria necessária a consulta aos professores que trabalham com o livro didático, obtendo-se maiores informações sobre o processo de escolha e avaliação dos livros didáticos, assim como a opinião dos alunos que são beneficiários diretos do Programa.

A suficiência das informações coletadas para refletir os objetivos da auditoria (Questão 5) foi parcial. Os objetivos da auditoria foram formulados em forma de perguntas a serem investigadas:

- a) *Em que medida os atuais custos diretos e indiretos envolvidos no PNLD podem ser reduzidos?*
- b) *Como o rol de escolhas de livros pode ser aprimorado/reduzido?*
- c) *Em que medida cada etapa do processo pode ser agilizada?*
- d) *Em que medida a reserva técnica é necessária/suficiente para atender à demanda?*
- e) *Até que ponto é necessária a repetição do processo de escolha dos livros de 1ª série todos os anos?*
- f) *Em que medida os instrumentos de controle são efetivos?*
- g) *Como pode ser mensurado o desempenho do programa?*
- h) *Quais instrumentos poderiam ser implementados para reduzir a não utilização dos livros?*
- i) *Em que medida é viável a extensão do prazo de utilização dos livros de 2 para 3 anos, no caso da 2ª a 8ª séries?*
- j) *Em que medida o modelo de atendimento descentralizado é mais eficiente do que o centralizado?* (Brasil, 1999a, p. 6).

Para a análise de custo-benefício, as informações foram suficientes para ilustrar o quantitativo de livros utilizados, custos diretos e indiretos da aquisição de livros, quantidade de alunos beneficiados, pois foram estabelecidos alguns indicadores de desempenho para medi-los. Entretanto, algumas questões como a repetição do processo de escolha dos livros de 1ª série todos os anos e a mensuração do desempenho do programa poderiam obter uma visão mais ampla se contasse com a percepção dos professores e alunos que recebem e utilizam periodicamente os livros didáticos. Na questão pedagógica, foram ouvidos somente o Fundo Nacional de Desenvolvimento da Educação – FNDE e a Secretaria de Ensino Fundamental – SEF/MEC (Brasil, 1999a). Sugere-se que os beneficiários diretos e indiretos sejam ouvidos para avaliar esse tipo de questão.

As técnicas de análise de dados foram explicitadas parcialmente (Questão 6), na metodologia proposta foram citadas: mapeamento de produtos e processos, análise SWOT,

Análise *Stakeholder*, *Benchmarking* e análise de custo. Tais técnicas foram utilizadas na fase de planejamento do programa.

A descrição do Programa (Questão 7) não foi realizada com clareza. Tanto no relatório de planejamento, quanto no relatório de execução da auditoria, foi pouco enfatizada a descrição do programa, que contemplou, sobretudo, os aspectos orçamentários financeiros, foco da auditoria, conforme ilustrado a seguir:

*O PNLD é o segundo maior programa do FNDE, depois do Programa Nacional de Alimentação Escolar (Merenda). No Relatório das Contas do Governo do Exercício de 1997, está registrado que naquele ano foram gastos R\$ 232,9 milhões com a aquisição de livros e mais R\$55,9 milhões com os demais gastos na área, tendo sido distribuídos 64,7 milhões de livros para todas as unidades da federação, contra 41 milhões de livros didáticos distribuídos em 1996. Já em 1998, o PNLD consumiu 369,6 milhões de reais e, para 1999, a previsão é que sejam gastos 225 milhões de reais, o que reflete a magnitude do Programa (Brasil, 1999a, p.5).*

Nesse sentido, nem o programa, nem seu contexto foram descritos claramente (Questão 7 e 8). Descreveram-se dados que subsidiaram a proposição da redução de custos na aquisição dos livros, redução do custo na distribuição dos livros, redução no número de títulos ofertados para escolha e diminuição da não-utilização dos livros.

O objetivo do programa foi descrito parcialmente (Questão 9). A descrição sucinta não favorece a compreensão dos objetivos e do Programa como um todo. Por exemplo,

*O PNLD tem como missão contribuir para a universalização do ensino e para a melhoria de sua qualidade por meio da aquisição e distribuição de livros didáticos para todos os alunos matriculados nas escolas públicas do ensino fundamental de 1ª a 8ª séries. Seu principal objetivo é distribuir, gratuitamente, livros didáticos aos alunos do ensino público fundamental no tempo certo, na quantidade certa e com a especificação solicitada (Brasil, 1999a, p. 6).*

Os procedimentos metodológicos (Questão 10) foram citados e descritos parcialmente. Na fase de execução da auditoria, foram utilizados a pesquisa e o estudo de caso. O estudo de caso realizado em São Paulo e Minas Gerais, quanto à “pesquisa”, verificou-se tratar da aplicação de questionários e entrevistas pelas regionais da Organização. Porém, não há

descrição clara da amostra dos participantes da auditoria e nem dos instrumentos utilizados, que não constam no relatório de planejamento.

Diante dos objetivos da auditoria expostos na Questão 5, verificou-se que os resultados (Questão 11) foram expostos claramente, o que se verificou por meio da leitura do relatório final e apresentação de questões de auditoria e respostas.

Em relação às conclusões (Questão 12) e recomendações (Questão 13), foram descritas com clareza. Entretanto, verificaram-se no Comentário do Gestor em relação à matéria, acréscimos às respostas da auditoria e modificação de algumas delas. Por exemplo, que se preserve o aspecto qualitativo da avaliação pedagógica e que não seja apenas quantitativo. Que a classificação de livros pertencentes à linhas/procedimentos teóricos metodológicos distintos não se presta à quantificação, por sua própria natureza pedagógica e ética, não sendo suficiente a adoção de apenas uma medida isolada como a classificação fundamentada em critérios quantitativamente mensuráveis.

Apesar da identificação do público do programa, por meio da análise *Stakeholder* (vide Questão 1), a auditoria não considerou a opinião desses diferentes atores (Questão 14).

Uma vez que não foram elaborados comentários pela equipe de auditoria a respeito do tempo, dinheiro e quantidade de pessoas necessárias para a realização da avaliação (Questão 15) entende-se que teria sido satisfatória.

Os procedimentos para coleta de informações foram descritos parcialmente (Questão 16) não se esclareceu o procedimento utilizado junto às SEDUCs, para a realização dos estudos de caso, os quais geraram indicadores de desempenho. Não constou no relatório de planejamento, nem no relatório final, a descrição dos questionários enviados a SECEXs Estaduais que coletariam as informações junto às SEDUCs estaduais.

Em relação à validade dos instrumentos de coleta (Questão 17), não houve como verificar a adequação ou validade, uma vez que não constavam e não foram descritos na documentação analisada.

Não se verificou uso de medidas para assegurar o mínimo de erros na coleta de dados (Questão 18).

Verificou-se que a informação quantitativa foi analisada adequadamente por meio da análise custo-benefício (Questão 19), já a informação qualitativa não foi devidamente tratada (Questão 20), pois não transparece nos textos indícios de tal análise.

### **9. Monitoramento e Controle da Mortalidade Materna**

Os públicos do programa e participantes da auditoria foram identificados parcialmente (Questão 1). O público do programa não foi identificado já o público da auditoria foi: Secretarias de Saúde de todos os Estados e Distrito Federal, Comitês Estaduais e Municipais de Mortalidade Materna (das capitais); Diretores Clínicos de amostra selecionada a partir de hospitais considerados referência na área do atendimento à Saúde da Mulher; Secretarias Municipais de 542 municípios; Hospitais considerados como referência no atendimento à saúde da mulher no Rio Grande do Sul e Paraná.

A equipe de auditoria possuía treinamento adequado para a realização da auditoria (Questão 2). A auditoria contou com a ajuda de especialistas nas fases de planejamento e execução da auditoria, sendo entrevistados especialistas da área de mortalidade materna no Brasil (Questão 3).

As informações coletadas possuíam alcance parcial para o entendimento do funcionamento do Programa (Questão 4). A descrição não esclareceu o funcionamento do programa.

As informações coletadas foram parcialmente suficientes para refletir os objetivos da auditoria (Questão 5) que estavam relacionados ao monitoramento e prevenção da mortalidade materna coordenados pelo Ministério da Saúde – MS. Foi relatado “[...] a indisponibilidade de dados fidedignos e as divergências entre os especialistas entrevistados impossibilitaram à equipe responder a quinta questão de auditoria: existe correlação entre os elevados percentuais de cesarianas registrados no Brasil e os coeficientes de mortalidade materna?” (Brasil, 2001d, p. 24).

As técnicas de análise de dados foram explicitadas parcialmente (Questão 6). Citou-se: *Benchmarking* e análise RECI; entretanto, verificaram-se outros tipos de análises estatísticas na apresentação dos resultados.

O relatório não descreveu claramente o Programa (Questão 7). A descrição do Programa foi demasiado breve, por exemplo, não houve caracterização clara do público-alvo do Programa; foram apresentados dados de mortalidade materna; porém, não do funcionamento do Programa, houve pouco detalhamento da legislação e atribuições das esferas legais envolvidos. O enfoque em relação à descrição do contexto do Programa foi claro (Questão 8). Foram referidos no relatório final, dados internacionais sobre médias de

mortalidade materna no mundo; e dados de países da América Latina. Entre eles, o Brasil apresenta alto índice de mortalidade materna:

*A Organização Mundial de Saúde e o Banco Mundial divulgam, para o Brasil, um coeficiente de 114 óbitos por 100.000 nascidos vivos [...] estima um valor entre 100 a 150 óbitos, reconhecendo, no entanto, que os valores oficiais (da ordem de 50 mortos por 100.000 nascidos vivos) seriam inaceitáveis. Conforme já mencionado, a OMS e a OPAS consideram altos os coeficientes de mortalidade materna entre 50 e 149, e muito alto a partir de 150, estando o Brasil, portanto, numa situação limítrofe entre essas duas categorias, (Brasil, 2001d, p. 12).*

De acordo com dados apresentados em Brasil (2001d), os índices de mortalidade infantil no Brasil são altos, conforme ressaltado na conclusão do relatório “os níveis de mortalidade materna no Brasil são nove vezes maiores do que os admitidos pela Organização Mundial de Saúde” (Brasil, 2001d, p. 86).

O relatório não descreveu claramente os objetivos do Programa (Questão 9). Essa falha ocorreu em relação à descrição do Programa como um todo. Foram descritas, somente as diretrizes gerais dos Comitês de Mortalidade Materna. As atribuições dos Comitês dizem respeito à divulgação de diretrizes gerais aos Estados e Municípios; desenvolvimento de recursos humanos na área; informação, educação e comunicação; assessoria técnica; apoio à realização de estudos ou pesquisas sobre a morte materna e monitoramento de estudos sobre a morte materna.

Os procedimentos metodológicos foram descritos parcialmente (Questão 10) e referidos na fase de planejamento somente algumas etapas da auditoria, como identificação de pessoas e instituições a serem contatadas na fase de execução; entrevistas com pessoal do Ministério da Saúde; contatos com a CPI da mortalidade materna; pesquisas na internet; elaboração de questionários e definição da abordagem estatística. No projeto de auditoria, foi detalhada a forma de coleta dos dados através do questionário a ser enviado pelo correio e a descrição do cálculo do tamanho da amostra.

O relatório descreveu parcialmente os resultados da auditoria (Questão 11). As informações foram apresentadas de forma condensada e nos textos há poucos gráficos ou tabelas resumindo informações dos casos apresentados. Houve, entretanto, descrição e sugestão de indicadores de desempenho ao Programa e resposta a quase todas as questões de auditoria, exceto a quarta questão conforme verificado a seguir

*1ª questão - as taxas de mortalidade materna são efetivamente monitoradas e avaliadas pelas áreas técnicas do Ministério da Saúde?*

*[...] Assim sendo, conclui-se que as taxas de mortalidade materna são monitoradas de forma incompleta, devendo o monitoramento ser aprimorado a partir da integração entre os dados do sistema de vigilância epidemiológica e dos comitês de mortalidade materna (Brasil, 2001 d, p. 86).*

*2ª questão – os Comitês de Mortalidade Materna são efetivos em suas competências?*

*[...] Não há evidências que a atuação dos comitês tenha contribuído decisivamente para a queda daqueles coeficientes [...] (Brasil, 2001d, p. 87).*

*3ª questão – a existência de uma rede de postos de saúde vinculados ao GHC é um fator preponderante na busca da equidade?*

*[...] Sim. Apesar de limitada geograficamente, pode-se afirmar que a existência da rede de unidades de saúde vinculadas ao Grupo Hospitalar Conceição, em Porto Alegre, contribui de forma significativa para a equidade no acesso a serviços de saúde de qualidade (Brasil, 2001d, p. 88).*

*4ª questão – existe correlação entre os elevados percentuais de cesáreas e as taxas de mortalidade materna?*

*[...] em função da indisponibilidade de dados fidedignos para a avaliação estatística e das divergências entre os especialistas entrevistados, optou por não desenvolver este tema (Brasil, 2001d, p. 88).*

As conclusões da auditoria (Questão 12) e as recomendações (Questão 13) foram descritas claramente. As recomendações foram baseadas nos achados de auditoria e explicitadas ao final do trabalho.

A auditoria considerou de forma parcial as posturas dos diferentes grupos de interesse no programa (Questão 14), uma vez que as usuárias do programa não foram abordadas, porém, buscaram-se opiniões de organizações, por exemplo, “organizações femininas, em especial a Rede Feminista de Saúde e Direitos Reprodutivos, em São Paulo [...] essa organização congrega diversas entidades femininas e tem assento na Comissão Nacional de Mortalidade Materna” (Brasil, 2001d, p.22). Entende-se que uma visão mais ampla do Programa seria obtida por meio do levantamento de informações junto ao público-alvo e não

somente com instituições de atendimento, especialistas, organizações e órgãos do governo, conforme comentado na primeira questão.

Os recursos da auditoria (tempo, dinheiro e funcionários) (Questão 15) foram suficientes para a realização da auditoria, uma vez que não há indícios em sentido contrário nos documentos consultados.

Os procedimentos de coleta de informações não foram descritos com clareza (Questão 16). Não foram apresentados nos anexos dos documentos consultados e não houve indicação das bases sob as quais foram criados os questionários e as entrevistas.

Não houve como verificar a validade dos instrumentos de coleta de dados, uma vez que não se encontravam nos anexos na documentação consultada (Questão 17). Pressupõe-se, entretanto, que os resultados obtidos pela auditoria, mediante o auxílio dos especialistas no assunto, na fase de planejamento e execução tenha auxiliado na construção e qualidade dos mesmos.

Foram tomadas medidas de forma parcial para evitar erros na coleta de dados (Questão 18). Houve cálculo da amostra estatística de municípios (Secretarias Municipais) e escolha aleatória dos municípios após cálculo da amostra. Entretanto, o delineamento não experimental escolhido, ausência de grupo de comparação, não consulta a participantes do programa e mortalidade instrumental, limitaram a generalização dos resultados obtidos.

A informação quantitativa foi parcialmente analisada (Questão 19), as comparações (*Benchmarking*) entre os casos analisados não foram claras, pois há muita informação em textos e poucos gráficos e tabelas que ajudariam na compreensão dos resultados.

Não foram explicitadas e apresentadas as bases analíticas qualitativas das informações coletadas (Questão 20).

### ***10. Programa Novo Emprego e Seguro-Desemprego***

Os públicos do programa e participantes da auditoria foram identificados (Questão 1). De acordo com a análise *Stakeholder*, os interessados no Programa eram: desempregados do setor formal; trabalhadores desempregados ou com necessidade de qualificação; empresas; Ministro de Estado do Trabalho e Emprego; gestores do Seguro-Desemprego; entidades que promovem a qualificação de trabalhadores e famílias dos desempregados. Apesar dessa identificação na fase de planejamento, consta no relatório final à realização de entrevistas com outros interessados como representantes da Central Única dos Trabalhadores - CUT, além

desse, foi ouvido principalmente o pessoal dos postos de atendimento ao segurado e os responsáveis pelo programa e alguns beneficiários do SD foram entrevistados.

A equipe de auditoria possuía treinamento adequado parcial para a realização da auditoria (Questão 2). A especialidade da equipe de auditoria – análise de finanças e controle - poderia ter sido complementada pela participação de consultores ou especialistas em pesquisas, evitando os problemas do delineamento da auditoria identificados na meta-avaliação.

Não houve participação de consultores externos na auditoria (Questão 3), o que poderia ter facilitado o trabalho da equipe frente à amplitude do programa e ao pouco tempo para a realização da mesma.

As informações coletadas foram parcialmente suficientes para entender o funcionamento do programa (Questão 4).

As informações coletadas foram parcialmente suficientes para refletir os objetivos da auditoria (Questão 5). Considerando que o objetivo principal da auditoria foi identificar as causas que impedem ou dificultam o acesso dos trabalhadores ao Seguro-Desemprego, a auditoria procurou responder às seguintes questões:

- a) *Quais são os fatores que influenciam o acesso dos segurados ao benefício?*
- b) *Como o Seguro-Desemprego pode ampliar o atendimento aos desempregados sem justa causa que atendem aos requisitos do Programa?*
- c) *O prazo para pagamento da primeira parcela do Seguro-Desemprego é adequado?*
- d) *Como o Seguro-Desemprego pode atender aos empregados sem anotação do contrato de trabalho? (Brasil, 2002f, p. 6).*

Conforme se verificou, as respostas a estas questões envolveriam a participação de parcela significativa de segurados do Programa, o que não ocorreu devido às restrições de tempo e pessoal, para a realização da auditoria.

As técnicas de análise de dados foram explicitadas parcialmente (Questão 6). Utilizou-se na fase de planejamento: matriz de planejamento, análise SWOT, RECI, análise de riscos e análise *Stakeholder* e para a análise dos dados coletados: análise de conteúdo, técnicas de triangulação, interpretações alternativas e exames de casos negativos. Entretanto, não se esclareceu como seriam realizadas as *interpretações alternativas*.

O relatório descreveu claramente o Programa (Questão 7), apresentou-se o histórico do SD, as leis que o embasam; fontes de recurso e destinação; alíquotas; contribuintes; condições para habilitação ao benefício; funcionamento; controle de segurança e benefícios especiais.

O relatório descreveu claramente o contexto do Programa (Questão 8) a partir dos antecedentes, histórico de atendimentos do programa, modalidades e quadro atual.

O relatório descreveu claramente os objetivos do programa (Questão 9), que “[...] tem por objetivo, além de prover assistência temporária ao trabalhador desempregado em virtude de dispensa sem justa causa, auxiliá-lo na manutenção e busca de emprego, promovendo, para tanto, ações integradas de orientação, recolocação e qualificação profissional” (Brasil, 2002f, p. 2).

Os procedimentos metodológicos foram descritos parcialmente (Questão 10). Quanto ao delineamento, há “referência a estudo de caso, pesquisa documental e dados sobre amostras não probabilísticas e delineamentos não experimentais, definindo, nos casos cabíveis, os elementos que consubstanciam as conclusões ou achados da auditoria” (Brasil, 2002g, p. 5). Os instrumentos de coleta de dados (entrevista e questionário) não foram descritos e não se encontram nos anexos dos relatórios; também não foram explicitadas as bases sob os quais foram criados.

O relatório descreveu parcialmente os resultados (Questão 11), embora, a apresentação de dados quantitativos tenha sido adequada em forma de gráficos, tabelas, análises estatísticas.

As conclusões foram descritas claramente no relatório (Questão 12), destacando-as com base nos resultados obtidos, os pontos fortes do Programa e as oportunidades de melhoria de desempenho.

As recomendações foram apresentadas claramente no relatório (Questão 13), sendo que a maioria (13 recomendações) destinou-se à Secretaria de Políticas Públicas de Emprego do MTE, uma diretamente ao Ministro de Estado do Trabalho e do Emprego e uma ao departamento de Emprego e Salário, da Secretaria de Políticas Públicas do MTE.

A opinião de diferentes grupos de interesse foi considerada parcialmente (Questão 14), visto que foi citada no relatório final a realização de entrevistas com membros da CUT e os resultados dessas entrevistas não foram mencionados nos dados e conclusões finais do trabalho de auditoria.

Os recursos não foram suficientes para a realização da auditoria (Questão 15), em relação ao tamanho da equipe, três auditores (na fase de planejamento) e cinco pessoas (na fase de execução). O prazo também foi curto para a realização. Conforme Brasil (2002g, p. 26), “devido à limitação de tempo, não será suficiente para que seja feita a identificação da maior parte dos problemas existentes na execução do programa simplesmente [...]”. Além disso, (mais de 4 milhões de beneficiários), demandariam exaustivo trabalho da equipe que contou apenas com 54 dias úteis para a realização das atividades, considerando as atividades da fase de planejamento, execução e elaboração do trabalho de auditoria (Brasil, 2002f).

Os procedimentos de coleta de informações não foram descritos com clareza (Questão 16). Foram citados como procedimentos de coleta de dados as fontes secundárias, ressalta-se que dados secundários, são fontes de coleta de informação e não procedimento de coleta de dados. Além disso, os instrumentos (entrevistas) não se encontram nos anexos e não foram descritas suas bases conceituais.

Não houve como verificar a validade dos instrumentos de coleta de dados (Questão 17), pois não foram descritos nos materiais consultados e não se encontravam nos anexos.

Foram adotadas medidas de forma parcial para evitar erros na coleta de dados (questão 18). Houve cálculo da amostra estatística do número de atendentes que fariam parte da amostra. Ressalta-se que para atender ao escopo da auditoria seria necessário coletar informações junto aos beneficiários do Programa. Entretanto, a auditoria justificou a coleta de dados junto a número maior de atendentes do que de beneficiários, da seguinte forma:

*Foi escolhida a consulta aos atendentes pela absoluta impossibilidade de se realizar uma pesquisa com um número significativo de segurados, seja pela dificuldade de lidar com um universo muito maior, seja pela possível dificuldade de localizar e obter resposta de trabalhadores que estejam desempregados dentro do tempo requerido (Brasil, 2002g, p. 28).*

Além do problema de enfoque em relação ao respondente o estudo de caso realizado no Distrito Federal pode ter enviesado os resultados dessa etapa, na Análise de Risco da auditoria, reconheceu-se que: “[...] pelo fato da situação do DF ser inteiramente diferente da média encontrada no país. Devido a sua alta renda *per capita*, o DF é uma Unidade da Federação que possui mais recursos para investir em seus serviços públicos do que a maioria [...]” (Brasil, 2002g, p. 26).

Os dados quantitativos foram analisados adequadamente (Questão 19), embora não tenha sido esclarecido acerca do tratamento dos dados no tópico do relatório que descreve a metodologia da auditoria. Conforme apresentação dos resultados no relatório final, as estatísticas descritivas (percentuais) apresentadas basearam-se em fontes de dados secundários (Cadastro Geral de Empregados e Desempregados – CAGED e Sistema de Informações Gerenciais sobre Ações de Emprego – SIGAE).

Quanto à análise de dados qualitativos, foi parcialmente adequada (Questão 20). Foram citados: análise de conteúdo, “técnicas de triangulação, interpretações alternativas e exames de casos negativos (exceções que confirmam a regra e que não seguem a tendência central)” (Brasil, 2002g, p. 5). Entretanto, não se esclareceu acerca do significado de *interpretações alternativas* e não há exemplificações ou indícios da análise de conteúdo nos resultados apresentados.

#### ***4.4 Análise dos Conteúdos***

A seguir foram apresentadas as categorias decorrentes da análise de conteúdo dos relatórios de planejamento e relatórios finais de auditoria. Definem-se as categorias, os temas, e cita-se para o tema, a quantidade de recorrências nos documentos analisados e a fonte na qual foi obtido. Após as definições de cada categoria foi apresentado um quadro resumo da categoria, com os temas, frequências, exemplos e fonte do exemplo (programa auditado).

Para compreensão dos exemplos, em alguns casos, escreveu-se entre colchetes uma explicação ou sinônimo do termo de acordo com o significado pertinente ao programa do qual a expressão foi extraída.

Ao final foram analisados os conteúdos dos objetivos dos programas e classificados como implantação, intermediários e finais (Posavac & Carey, 2003). Essa classificação *a priori* foi defendida neste trabalho como etapa indispensável ao conhecimento do programa e delineamento da avaliação (Stufflebeam & Shinkfield, 1987; Posavac & Carey, 2003; Aguilar & Ander-Egg, 1994). Assim, por meio da classificação, verificou-se a interpretação e análise das ANOPs em relação aos objetivos dos programas auditados.

A partir do resultado de cada categoria, foram analisados qualitativamente os significados em relação ao escopo deste trabalho: meta-avaliação das auditorias do Tribunal de Contas da União.

## ***1. Avaliação de Processo***

### ***Definição***

A categoria avaliação de processo define-se pelo planejamento e avaliação de processo do programa, processo organizativo e a capacitação da equipe do programa para fazê-lo. O tema planejamento e avaliação de processos refere-se às atividades organizativas do programa, definição de objetivos, metas e produtos; envolve o planejamento estratégico e os estudos de viabilidade técnica para o programa. O tema processo organizativo significa a gestão do programa em suas unidades de funcionamento nos níveis municipal, estadual, federal e as condições institucionais que viabilizam essa execução. O tema execução do programa representa a implantação das tarefas necessárias para a prestação do serviço e o funcionamento do programa, as etapas da execução, sejam ações ou processos utilizados e mecanismos de atuação do programa. O tema capacitação da equipe do programa refere-se às atividades de treinamento para a atuação dos agentes envolvidos nos programas. A seguir, apresentou-se o quadro resumo da categoria.

**Quadro 6. Categoria Avaliação de Processo**

Categoria: Avaliação de processo			
Temas	Freq.	Exemplos	Programa auditado
Planejamento e avaliação dos processos	19	Como é realizado o planejamento da execução do programa a curto, médio e longo prazo? (Brasil, 2002g, p. 23).	Novo Emprego e Seguro Desemprego
		Capacidade de diagnóstico e planejamento das prefeituras como suporte a uma atuação mais efetiva na redução do déficit habitacional local (Brasil, 2003f, p. 34)	Morar Melhor
		[...] eventual necessidade de que a UGP-G, em conjunto com as demais unidades, defina claramente e internalize seus objetivos, metas e produtos (Brasil, 2000e, p. 16)	Pró-Água/Semi-Árido
Processo organizativo	38	Com vistas a se conhecer como se estabelece o processo organizativo nos assentamentos [...] (Brasil, 2003h, p.23)	Novo Mundo Rural
		As informações [...] sugerem que as condições institucionais para a execução das ações constituem-se em risco para o processo de implantação dos assentamentos rurais (Brasil, 2003h, p. 21)	Novo Mundo Rural
		Essa situação reflete condições institucionais decorrentes da sobreposição de interesses políticos na gestão da instituição, da alta rotatividade dos dirigentes e da cultura institucional [...] (Brasil, 2003h, p. 19)	Novo Mundo Rural

## Quadro 6. Categoria Avaliação de Processo

Categoria: Avaliação de processo			
Temas	Freq.	Exemplos	Programa auditado
Execução do programa	59	Os municípios estão conseguindo gerenciar adequadamente o programa com relação ao cumprimento das exigências técnicas e legais, processo decisório de seleção das famílias, acompanhamento dos projetos e realização de trabalho social junto às famílias? (Brasil, 2004e, p.18) [pergunta da auditoria]	Morar Melhor
Capacitação da equipe do programa	18	Capacitação dos agentes envolvidos com as medidas socioeducativas está contribuindo para a efetividade das ações do Programa (Brasil, 2003g, p. 32)	Reinserção Social do Adolescente em Conflito com a Lei
		Desenvolver formas de promover treinamento e capacitação de responsáveis pela manutenção, acompanhamento e assistência técnica dos equipamentos; (Brasil, 2003c, p.s/n)	Energia das Pequenas Comunidades
		Considerar a possibilidade de promover ações de sensibilização, mobilização e capacitação das comunidades, em parceria com as universidades (Brasil, 2000d, p.s/n)	Pró-Água/Semi-Árido
Total	<b>134</b>		

Legenda: Freq = frequência

### *Análise*

Na categoria Avaliação de processo, destaca-se fortemente o enfoque em relação à execução do programa e ao processo organizativo, planejamento e avaliação de processos e capacitação da equipe do programa. Com base na meta-avaliação, observou-se a necessidade de conhecimentos técnicos especializados sobre os programas sociais e políticas públicas para realizar a análise dos processos em áreas complexas de domínios específicos como, por exemplo, na área de energia e saúde.

### **2. Avaliação de Necessidades**

#### *Definição*

Refere-se ao diagnóstico de necessidades internas do programa, dos beneficiários e das localidades onde está ou será instalado o programa. O tema necessidade do programa representa as necessidades internas do programa originadas do funcionamento interno como necessidades relacionadas à gestão do programa em nível local. Necessidade dos beneficiários refere-se ao levantamento das necessidades dos beneficiários.

## Quadro 7. Avaliação de Necessidades

Categoria: Avaliação de necessidades			
Temas	Freq.	Exemplos	Programa auditado
Necessidades do programa	13	Os controles existentes são suficientes para as necessidades atuais e em caso de ampliação do percentual de segurados atendidos? (Brasil, 2002g, p. 24)	Novo Emprego e Seguro Desemprego
		Em que medida a reserva técnica [reserva de livros didáticos] é necessária/suficiente para atender a demanda? (Brasil, 1999a, p.6)	Programa Nacional do Livro Didático
		Os municípios estão conseguindo gerenciar adequadamente o programa com relação ao diagnóstico das necessidades habitacionais locais? (Brasil, 2003f, p. 34)	Morar Melhor
Necessidades do beneficiário	3	O quanto às políticas relativas à qualificação; atendem às necessidades do trabalhador? [emprego] (Brasil, 2002g, p. 23)	Novo Emprego e Seguro Desemprego
		Como pode o seguro desemprego ampliar o atendimento aos desempregados sem justa causa? (Brasil, 2002g, p. 23)	Novo Emprego e Seguro Desemprego
		Também não existem levantamentos das necessidades das comunidades a serem beneficiadas [...] (Brasil, 2003c, p. 3)	Energia das Pequenas Comunidades
Total	16		

Legenda: Freq = frequência

### Análise

Na avaliação de necessidades, a avaliação do atendimento é uma variável de difícil mensuração, pois geralmente envolve a coleta de dados junto aos beneficiários, e no caso das auditorias de programas sociais de governo meta-avaliadas, trata-se de milhares de beneficiados. Em alguns casos, esses beneficiários possuem baixa escolaridade, isso por sua vez, dificulta a leitura de questionários ou formulários. Além disso, alguns programas não possuem dados cadastrais atualizados para que se possa escolher a amostra da auditoria e envio de questionários.

O impacto do atendimento é verificado pelas ANOPs somente depois dos programas, não há comparações da ação do Programa antes e depois, ou uso de grupos de controle. Dentre os programas meta-avaliados, verificou-se uma exceção, o caso do Programa de Desenvolvimento do Turismo do Nordeste Prodetur/NE (Brasil, 2003b), auditoria na qual foi utilizado o grupo de controle. Entretanto, no referido caso, os resultados não se referem ao uso desse expediente de controle na pesquisa.

### 3. Avaliação de Impacto/Resultado

#### Definição

A categoria avaliação de impacto/resultado expressa as conseqüências diretas e indiretas dos serviços e produtos oferecidos para os beneficiários ou comunidade envolvida e indicadores de desempenho dos resultados do programa. O tema impacto direto representa aquele impacto originado do programa para o público-alvo. O tema impacto indireto significa o impacto indireto não previsto ou planejado que afeta pessoas ou localidades não envolvidas diretamente com o programa. O tema indicador de desempenho refere-se a índices numéricos que expressam o impacto do programa, comporta expressões numéricas que mensuram resultados e quantitativos ou estatísticas descritivas dos atendimentos realizados.

#### Quadro 8. Avaliação de Impacto/Resultados

Categoria: Avaliação de impacto/resultado			
Temas	Freq.	Exemplos	Programa auditado
Impacto direto	38	Em que medida as ações contribuem para a melhoria da qualidade de vida dos beneficiários, fortalecendo a sua auto-estima e convivência familiar comunitária? [a convivência familiar e comunitária é objetivo do programa] (Brasil, 2004h, p. 31).	Sentinela
		[...] uma vez que ambas as ações são complementares e voltadas a dotar os beneficiários de instrumentos técnicos e gerenciais necessários à melhoria da qualidade de vida nos assentamentos (Brasil, 2003f, p.8)	Morar Melhor
		Em que medida o processo de execução das ações relativas à organização social e produtiva nos assentamentos rurais contribui para a integração dos beneficiários na agricultura familiar de forma sustentável? (Brasil, 2003h, p.36)	Novo Mundo Rural
Impacto indireto	5	Existe correlação entre os elevados percentuais de cesarianas registradas no Brasil e os coeficientes de mortalidade materna? (Brasil, 2002f, p. 19). [o objetivo do programa é a diminuição da mortalidade materna]	Monitoramento da Mortalidade Materna
		Como o programa pode contribuir para o desenvolvimento de tecnologia e indústria nacionais e a capacitação de recursos humanos? (Brasil, 2003e, p. 20)	Energia das Pequenas Comunidades
		A existência de uma rede de postos de saúde – Unidade de Saúde Comunitária – vinculados ao Grupo Hospitalar Conceição, em Porto Alegre, é um fator preponderante na busca da equidade no acesso das populações menos favorecidas aos serviços de saúde? (Brasil, 2000f, p. 19)	Controle da Mortalidade Materna

### Quadro 8. Avaliação de Impacto/Resultados

Categoria: Avaliação de impacto/resultados			
Temas	Freq.	Exemplos	Programa auditado
Indicador de desempenho	93	<p>Institua os seguintes indicadores de desempenho [...] (Brasil, 2004d, p.96)</p> <p>Índice de investimento privado no turismo = investimento privado acumulado em turismo/investimento privado previsto pelo Prodetur/NE II em turismo (Brasil, 2004a, p.93)</p> <p>Percentual de trabalhadores desligados sem justa causa com renda até dois salários-mínimos que teriam direito de requerer o benefício, mas não o fizeram (Brasil, 2002f, p.s/n)</p>	<p>Sentinela</p> <p>Prodetur/NE</p> <p>Novo Emprego e Seguro Desemprego</p>
Total	<b>136</b>		

*Legenda: Freq = frequência*

#### *Análise*

A categoria Avaliação de Impacto expressa o objetivo e o tipo de auditoria conduzida pela organização, sua natureza operacional, pois mede principalmente os aspectos da utilização dos recursos dos programas em relação aos seus resultados.

A execução do programa em si gera auditorias de resultados/impactos. Destacou-se fortemente na categoria avaliação de impacto/resultados a recorrência do tema indicador de desempenho. Entende-se que a implantação de indicadores de desempenho nos programas, demanda a criação de sistemas informatizados para consolidar as informações em banco de dados, sendo essa uma recomendação que pode ser mais enfatizada, caso seja recomendado ao Programa a criação e o uso de indicador de desempenho.

Por exemplo, o Programa de Enfrentamento a Exploração e Abuso Sexual de Crianças e Adolescentes - Sentinela (Brasil, 2004d) controla os estudos qualitativos e quantitativos enviados pelos municípios ao Ministério do Desenvolvimento Social e Combate à Fome, por meio de planilhas manuais criadas no Software Excell. Assim como o PETI, daquele Ministério, a despeito de atendimento nacional de cerca de um milhão de beneficiários, realizava seus controles de atendimentos, pagamentos aos executores do Programa e bolsas concedidas às famílias, por meio de planilhas manuais em Excell, atualizadas manualmente e mensalmente, até o final do ano de 2005. Durante o ano de 2006, parte dos controles desse Programa passou a ser realizado por meio de bancos e sistemas de dados do Programa Bolsa Família, com o qual integrou o processo de controle do repasse financeiro às famílias beneficiárias.

#### 4. Avaliação da Eficiência

##### Definição

A avaliação da eficiência mede ao custo-benefício, a materialidade e a eficiência dos programas. O tema custo-benefício representa a avaliação da relação financeira entre o uso dos recursos e a sua finalidade, visando à otimização dos recursos em face do uso. O tema materialidade é definido pelo enfoque na utilização dos recursos no programa, sejam recursos orçamentários, materiais ou informações, exemplifica-se pelos investimentos e valores orçamentários do programa. O tema eficiência refere-se à mensuração da adequação dos meios utilizados aos fins que se deseja alcançar no programa.

#### Quadro 9. Avaliação da eficiência

Categoria: Avaliação da eficiência			
Temas	Freq.	Exemplos	Programa auditado
Custo-benefício	22	Em que medida a assistência prestada [salário] pelo Seguro Desemprego atende às necessidades do segurado? (Brasil, 2002g, p. 24)	Novo Emprego e Seguro Desemprego
		Adote providências para aperfeiçoar a análise da viabilidade socioeconômica dos projetos, em especial aqueles cuja operação dependa de intervenção de outros entes - públicos, privados ou não governamentais (Brasil, 2004a, p.93)	Energia das Pequenas Comunidades
Materialidade	27	Os recursos materiais, orçamentárias, financeiros, humanos e de informações da Secretaria Especial de Direitos Humanos - SEDH, dos estados e das entidades de atendimento são suficientes e compatíveis com os objetivos alcançados? (Brasil, 2003g, p. 19).	Reinserção Social do Adolescente em Conflito com a Lei
		[...] A escolha se deu em razão da sua elevada materialidade, que corresponde a cerca de 75% dos recursos do Programa (Brasil, 2004d, p. 17)	Sentinela
Eficiência	5	Estabeleça procedimentos e critérios quantitativamente mensuráveis que permitam a classificação e posterior redução de títulos a constar no Guia do Livro Didático sem, contudo, abrir mão do aspecto qualitativo da avaliação pedagógica (Brasil, 1999a, p.25).	Programa Nacional do Livro Didático
		Em que medida é viável a extensão do prazo de utilização dos livros de 2 para 3 anos, no caso da 2ª a 8ª séries? (Brasil, 1999ª, p. 18)	Programa Nacional do Livro Didático
<b>Total</b>	<b>54</b>		

Legenda: Freq = frequência

##### Análise

Os programas sociais auditados recebem altas dotações orçamentárias do governo em cada área de atuação federal. A sistemática de uso dos recursos públicos é balizada por leis e normas específicas que regulamentam seu uso. Assim, a estimação de custos e sua adequação

é um elemento de controle necessário a todo e qualquer programa social. Entretanto, essa estimativa de valor e custo do atendimento, serviços ou outros tipos de recurso materiais do programa é uma condição, em alguns casos, não encontrada nos programas, sendo apontada tal lacuna pelas ANOPs. Alguns exemplos dessa lacuna foram encontrados no Prodetur/NE I, como a não avaliação do impacto ambiental; já no Programa de Reinserção Social do Adolescente em Conflito com a Lei e também no Programa Sentinela não foi calculado o custo dos atendimentos.

Não há incompatibilidade entre a análise custo-benefício e programas sociais, entende-se que essa é uma análise necessária, entretanto, se auditoria atua somente com esse enfoque, caracteriza-se como financeira e contábil. Nesse caso, torna-se limitada a análise mais abrangente a que o programa poderia ser submetido mediante procedimentos de AP.

Os critérios utilizados pela organização em relação às auditorias foram: eficiência, eficácia, efetividade e equidade. Verificou-se nas questões de auditoria o enfoque principal em aspectos de eficiência e equidade e não se encontraram questões sobre efetividade. Também não se encontrou o conceito efetividade nos aspectos analisados da matriz de planejamento, resultados, recomendações e acórdãos dos relatórios finais.

Nas perguntas de auditoria, destacou-se fortemente o tema Materialidade, com ênfase no uso de recursos financeiros, isso representa a ênfase na busca de informações sobre os custos dos programas e sua utilização.

Verificou-se que a organização considera o volume de recursos recebidos como um dos critérios de seleção dos programas a serem auditados. Assim, na ANOP enfatiza-se sobremaneira a aplicação de recursos, realizando análise custo-benefício, o que é constatado por meio da busca do critério de eficiência nas avaliações. Além disso, é característica dos programas de governo auditados o recebimento de grande volume de recursos para a sua implantação, o que torna esse um critério relevante devido à natureza dos Programas auditados.

## ***5. Boas Práticas do Programa***

### *Definição*

A categoria Boas Práticas do Programa comporta as experiências bem sucedidas, programas análogos, critérios de equidade ou justiça social e estratégias de participação adotadas pelo programa. As experiências bem sucedidas são do próprio programa, ocorrem em outra localidade do o programa e apresentam diferenciais positivos nas ações. O tema

programas análogos refere-se a outros programas semelhantes que atuam sobre a mesma problemática social e podem oferecer “lições” ou exemplos de atuação. O tema critério de equidade social representa os critérios de justiça nas práticas do programa (gênero, etnia, origem social e diferenças regionais), no sentido de promover um tratamento justo ao público-alvo do programa. Por exemplo, através de formas diferenciadas de acesso ao programa ou o acesso às políticas públicas de atendimento a populações carentes. O tema estratégias de participação representa as formas de participação dos beneficiários e organizações parceiras, é tida como boa prática, pois visa a melhor prestação dos serviços do programa.

### Quadro 10. Boas Práticas do Programa

Categoria: Boas Práticas do Programa			
Temas	Freq.	Exemplos	Programa auditado
Experiências bem sucedidas	20	Como a UGP – G e as UEGPs identificam e aproveitam essas boas práticas? (Brasil, 2000e, p. 20)	Pró-água/Semi-árido
		Identificação de fatores relevantes e boas práticas a serem disseminadas [...] (Brasil, 2000e, p. 17)	Pró-água/Semi-árido
		Quais os fatores que constituem diferenciais de sucesso entre as Unidades Estaduais de Gestão do Programa - UEGPs Secretarias Estaduais e que poderiam ser disseminados? (Brasil, 2000e, p. 17).	Pró-água/Semi-árido
Programas análogos	5	[...] como as lições de programas análogos podem contribuir para atender às necessidades de energia elétrica das localidades isoladas? (Brasil, 2003e, p. 18).	Energia das Pequenas Comunidades
		Foram escolhidas [...] a presença de iniciativas com objetivos semelhantes ao do Programa [...] (Brasil, 2003e, p. 26)	Energia das Pequenas Comunidades
Critério de Equidade Social	27	Se o programa está enfatizando critérios de gênero, raça/etnia origem social e diferenças regionais no sentido de promover a equidade de acesso das famílias aos bens ofertados (Brasil, 2003f, p.32)	Morar Melhor
		Quais as ações das Unidades Estaduais de Gestão do Programa - UEGPs e das Secretarias estaduais - SEs que buscam garantir a participação equitativa dos mais carentes na gestão dos recursos hídricos? (Brasil, 2000e, p. 18).	Pró-água/Semi-árido
Estratégias de participação	25	Participação na escolha da empresa de assistência técnica e controle dos serviços prestados (Brasil, 2003h, p. 38)	Novo Mundo Rural
		Estimule as comunidades a apresentar idéias e propostas de ações, a serem financiadas pelo Programa, que envolvam educação e/ou conservação ambiental (Brasil, 2000d, s/n).	Pró-água/Semi-árido
		Deve também ser assegurada, na implantação do PDA, a efetiva participação das famílias beneficiárias [...] (Brasil, 2003h, p. 21)	Novo Mundo Rural
<b>Total</b>	<b>77</b>		

Legenda: Freq = frequência

### *Análise*

Destacou-se nessa categoria a recorrência dos temas: critérios de equidade e justiça social e estratégias de participação; seguido de experiências bem sucedidas.

A identificação de “boas práticas” é uma premissa adotada pela organização para conduzir a ANOP. Entretanto, chama-se a atenção para o fato de que características extremas de sucesso de um programa estão associadas à gestão, localização, características regionais e culturais específicas de um local. Tais características podem provocar viés nos resultados da ANOP quando adotadas como critério principal na seleção da amostra de programas a serem auditados, pois os resultados não expressariam a realidade na qual a maioria dos programas em funcionamento estaria operando. Um aspecto positivo de mapear boas práticas, tomando-se o devido cuidado em considerar os diferentes contexto do programa, é a possibilidade de aplicá-las nas demais unidades do programas e em programas semelhantes.

O tema equidade ou justiça social esteve presente fortemente nas questões de auditoria dos Programas Morar Melhor; Novo Emprego e Seguro Desemprego; Programa de Enfrentamento a Exploração e Abuso Sexual de Crianças e Adolescentes – Sentinela; Energia das Pequenas Comunidades e Controle da Mortalidade Materna.

As recomendações de adoção de novos critérios para o funcionamento do programa, sejam eles de acesso ao programa ou para o atendimento, trazem em si algumas implicações. Por exemplo, modificações de critérios de acesso ao programa precisam estar contempladas na política pública na qual se insere o programa. Esses critérios são definidos e pactuados em instâncias federais, estaduais e municipais, necessitando de discussão e aprovação. Devido à amplitude da maioria dos programas meta-avaliados e ao processo participativo de decisão, quando há esse tipo de recomendação, provavelmente seu cumprimento demandaria tempo superior a 60 ou 90 dias para ser colocada em prática e pode ainda, por decisão dos *stakeholders*, não ser aceita. Outro fato importante a ser considerado, quando da sugestão de modificação de critérios do programa, é que se está propondo uma mudança organizacional, o que traz em si a questão da resistência natural à mudança.

## **6. Monitoramento do Programa**

### *Definição*

A categoria Monitoramento do Programa refere-se ao acompanhamento dos participantes das ações e ao controle e pesquisas desenvolvidos sobre o programa. O tema acompanhamento dos participantes significa o acompanhamento dos beneficiários, egressos e

projetos. O tema acompanhamento das ações representa a atuação das instâncias de supervisão, na geração de informações sobre o desempenho do Programa, por meio de instrumentos e procedimentos que permitam a avaliação do desempenho do programa. O tema controle refere-se à notificação das ocorrências do fenômeno visado pelo programa; bem como ao fornecimento de quadros de acompanhamento e índices sobre checagem de informações de dados do programa. O tema pesquisas representa estudos e diagnósticos contratados ou realizados sobre o fenômeno no qual o programa intervém.

### Quadro 11. Monitoramento do Programa

Categoria: Monitoramento do Programa			
Temas	Freq.	Exemplos	Programa auditado
Acompanhamento dos participantes	10	Qual o grau de implantação das atividades de apoio e acompanhamento dos egressos? (Brasil, 2003g, p. 19)	Reinserção Social do Adolescente em Conflito com a Lei
		Inexistência de informações sobre os egressos (Brasil 2003g, p.34)	Idem
		Forma de acompanhamento e supervisão pelo INCRA (Brasil, 2003h, p. 37)	Novo Mundo Rural
Acompanhamento das ações	34	[...] de forma que sejam realizados serviços de acompanhamento do desenvolvimento dos assentamentos rurais [...] (Brasil, 2004 d, p.30)	Morar Melhor
		[...] acompanhar, fiscalizar e avaliar a execução dos serviços a serem realizados pela empresa contratada (Brasil, 2003h, p. 18)	Novo Mundo Rural
		Se as atividades de apoio e acompanhamento estão sendo implementadas (Brasil, 2003g, p.34)	Reinserção Social do Adolescente em Conflito com a Lei
Controle	10	Instrumentos, procedimentos e informações suficientes que permitam uma adequada avaliação do desempenho do programa. (Brasil, 2003f, p. 36)	Morar Melhor
		Em que medida os instrumentos de controle são efetivos? (Brasil, 1999a, p. 15)	Programa Nacional do Livro Didático
Pesquisas	13	[...] e elabore estudo com vistas a propor a sua inclusão como uma Ação do Programa Combate ao Abuso e à Exploração Sexual de Crianças e Adolescentes, na oportunidade da próxima revisão do PPA (Brasil, 2004d, p. 94)	Sentinela
		Elabore mapeamento com a identificação das dificuldades enfrentadas pelos municípios no desenvolvimento de ações na área de identificação e abordagem dos casos de exploração sexual; (Brasil, 2004d, p. 94)	Sentinela
Total	67		

Legenda: Freq = frequência

### *Análise*

O monitoramento é somente uma lacuna dos programas sociais avaliados, conforme resultados demonstrados pelas ANOPs nas dez auditorias meta-avaliadas. Também se configura como lacuna da própria organização no acompanhamento da implantação de suas recomendações, devido ao grande número de programas avaliados, e pequeno tamanho da equipe para acompanhar a implantação das recomendações nos programas. Para o acompanhamento da implantação das recomendações seria necessário criar uma nova sistemática em relação aos programas e que tornassem efetivos os grupos de contato de auditoria.

De acordo com análise dos resultados das ANOPs os temas controle, pesquisas e acompanhamento das ações pelos programas demonstram significativas deficiências nesses quesitos. Um olhar mais detido da organização em relação às pesquisas e diagnósticos realizados em alguns programas como o PETI e o Sentinela, do MDS pode demonstrar que diagnósticos e estudos são pouco utilizados pelos programas na sua gestão e não são usados como parâmetros para modificações nas ações.

## **7. Política do Programa**

### *Definição*

A categoria Política do Programa representa as ações governamentais em determinada área de atuação na qual o programa se insere e as articulações com parceiros governamentais e não governamentais. O tema política nacional representa as políticas extensivas à população por meio dos programas, ações e planos. São exemplos, as políticas de qualificação, de recursos humanos, para a criança e adolescente e política em relação ao turismo. O tema parceiros governamentais refere-se às organizações governamentais que articulam as ações do programa para a prestação de serviços e atendimento. O tema parceiros não-governamentais significa organizações não-governamentais que articulam as ações do programa para a prestação de serviços e atendimento. Exemplos: associações, organizações civis, sindicatos, cooperativas, conselhos, Sistema “S” Sesc, Sesi, Senai, Sebrae, Institutos de pesquisa, entre outros.

## Quadro 12. Política do Programa

Categoria: Política do Programa			
Temas	Freq.	Exemplos	Programa auditado
Política Nacional	21	Em que medida normas, diretrizes e ações governamentais constituem uma política nacional integrada de atendimentos ao adolescente em conflito com a lei? (Brasil, 2003g, p. 19)  Em 2000, o Governo Federal [...] implementou modificações na política agrária nacional, alterando o desenho institucional dos órgãos com atuação na área [...] (Brasil, 2003h, p. 9)	Reinserção Social do Adolescente em Conflito com Lei  Programa Novo Mundo Rural
Parceiros Governamentais	39	Estude, em conjunto com a SEDH, formas apropriadas de implantar os meios necessários para custear o deslocamento dos usuários até o centro de referência [...] (Brasil, 2004d, p. 72)  Em conjunto com o MDS, articule com os demais ministérios envolvidos no Programa o desenvolvimento de ações de prevenção, utilizando, inclusive, os recursos da Ação Orçamentária 0746 - Apoio a Projetos de Prevenção do Abuso e da exploração Sexual de Crianças e Adolescentes; (Brasil, 2004d, p.94)	Pró-Água/Semi-Árido  Sentinela
Parceiros Não-Governamentais	24	Representantes das associações de usuários [...] (Brasil, 2000e, p. 19)  Representantes [...] dos comitês de bacia (Brasil, 2000e, p. 19)  [...] se entidades representativas da população diretamente interessada estão participando do processo decisório local sobre a seleção das famílias e controle dos projetos (Brasil, 2003f, p. 35)	Pró-Água/Semi-Árido  Pró-Água/Semi-Árido  Novo Mundo Rural
Total	<b>84</b>		

Legenda: Freq = frequência

### Análise

Nessa categoria destacou-se o tema Parceiros Governamentais, o que se verificou na execução dos programas nos três níveis: federal, estadual e municipal.

Em relação à Política Nacional, as auditorias percebem e buscam como aspecto positivo dos programas a inter-relação e participação de outras entidades governamentais ou não-governamentais, o que pode levar a auto-sustentabilidade dos programas.

É característica dos programas de governo auditados uma estrutura na qual nos níveis federal, estadual e municipal exerçam papéis diferentes na gestão e execução do programa. Nesse sentido, a verificação desses aspectos é realizada por meio das comunicações formais entre os diferentes níveis de gestão do programa. As ANOPs verificam como são as estruturas de comunicação formais em relação ao programa e a sua eficiência; se, por exemplo, Estados

e municípios recebem e enviam informações entre si e o nível de disseminação e troca dessas informações. A articulação pode ocorrer em qualquer nível de gestão dos programas, seja na realização de fóruns para a discussão e construção de parâmetros para os programas de acordo com a política na qual se inserem.

Destaca-se, entretanto, que os programas sociais são concebidos em uma arena política, realizando suas ações através dos órgãos federais responsáveis por áreas da política e são dirigidos por diferentes partidos políticos. Assim, nem sempre a articulação sugerida pelas ANOPs aos programas é factível. Em alguns casos, o desconhecimento de peculiaridades dos programas impossibilita o cumprimento de uma recomendação. Por exemplo, ao recomendar que o Programa Sentinela institua indicador de proporção de encaminhamentos para o PETI (Brasil, 2004d). Esses dois programas, Sentinela e PETI, localizados na mesma gerência do MDS, não articularam ou estabeleceram até o presente momento em regulamentação ou na prática, o indicador designado. Ainda nesse exemplo algumas considerações: 1) o atendimento da clientela do Programa Sentinela só seria coerente com o PETI se configurasse caso de exploração sexual comercial, que é considerada como uma prática de trabalho infantil; 2) para compreensão da concepção de trabalho infantil e tipos de atendimento do Programa Sentinela, exploração sexual comercial de crianças e adolescentes, é necessário o entendimento da política de assistência social e direitos humanos, assim como aprofundamento do conhecimento das práticas desses programas. Assim, as recomendações sobre a articulação desses ou de outros programas auditados merece considerar a política na qual se inserem e os conhecimentos específicos sobre a problemática nas quais os programas intervêm.

## ***8. Fontes de dados da auditoria***

### *Definição*

A categoria fonte de dados da auditoria representa as fontes usadas no planejamento e execução da auditoria sejam bases de dados, fontes de dados secundários e bibliografia especializada. O tema base de dados representa dados de sistemas de informação. O tema fonte de dados secundários significa os materiais bibliográficos e outras fontes secundárias relacionadas ao tema. O tema bibliografia especializada contempla bibliografias técnicas específicas, artigos e documentos internos do programa como planos e relatórios. O tema documentos do programa refere-se a documentos específicos do programa produzidos internamente em qualquer nível de gestão.

**Quadro 13. Fonte de dados de auditoria**

Categoria: Fonte de dados de auditoria			
Temas	Freq.	Exemplos	Programa auditado
Base de dados	53	[...] os dados foram obtidos junto ao Ministério do Trabalho e do Emprego – MTE e solicitados bancos de dados adicionais a Datamec, empresa responsável pelo processamento do seguro desemprego [...] (Brasil, 2002f, p.s/n)	Novo Emprego e Seguro Desemprego
		Informações disponibilizadas pela base de dados do programa (Brasil, 2003f, p. 36)	Morar Melhor
		Dados sobre déficit habitacional por Unidade da Federação (Brasil, 2003f, p. 32)	Morar Melhor
Fonte de dados secundários	15	Inventários dos casos de interação do programa com outras iniciativas que contribuíram para reforçar o desempenho do Prodetur/NE [...] (Brasil, 2003b, p. 30)	Prodetur/NE
		Cadastros dos municípios beneficiários (Brasil, 2004h, p. 27)	Sentinela
		Cadastros dos adolescentes (Brasil, 2003g, p.34)	Reinserção Social do Adolescente em Conflito com a Lei
Bibliografia especializada	34	As fontes de dados para o estudo serão os PDAs, as bibliografias especializadas [...] (Brasil, 2003h, p. 25)	Novo Mundo Rural
		Estudos desenvolvidos por servidores do INCRA em 1971 já apontavam que o principal papel da instituição [...] (Brasil, 2003h, p. 19)	Novo Mundo Rural
		Segundo estudos do IPEA, entre 1995 e 1998 foram assentados [...] (Brasil, 2003h, p. 9)	Novo Mundo Rural
		Exame de documentos (Brasil, 2000e, p. 16)	Pró-Água/Semi-Árido
		Diretrizes e normas sobre o tema (Brasil, 2003g, p. 30)	Reinserção Social do Adolescente em Conflito com a Lei
<b>Total</b>	<b>102</b>		

Legenda: Freq = frequência

### *Análise*

Destacou-se fortemente na categoria Fontes de Dados da Auditoria a recorrência dos temas Base de dados e Bibliografia Especializada. As fontes de dados citadas nesta categoria são fonte de dados secundários, entretanto, para fins de caracterização e conhecimento do processo de ANOP, mantiveram-se as separações do temas, também devido à grande quantidade de recorrências de cada tema.

Na análise dos relatórios de planejamento e final das auditorias, observou-se apenas uma recorrência da citação “levantamento de dados primários [...]” (Brasil, 2003h, p. 37).

Devido à recorrência de apenas uma citação, não se considerou como um tema. Porém, é comum a utilização de dados primários nas ANOPs, pois são coletados por meio de questionários e entrevistas. Provavelmente, apenas uma citação literal de fonte de informação de dados primários ocorreu devido à imprecisão conceitual expressa nos materiais consultados.

Conforme informações obtidas nos documentos, a dificuldade de obtenção de dados dos programas auditados ocorre devido à indisponibilidade ou baixa qualidade e confiabilidade dos dados disponíveis. As equipes de auditoria apontam que não há deficiência de informações precisas nos programas avaliados como sistemas de processamento de dados que cubram as informações dos programas. Diante do exposto, entende-se que o Tribunal de Contas da União poderia enfatizar mais essa recomendação por ser uma deficiência constatada em vários Programas Sociais através de diversas ANOPs. Considerando-se também o forte “peso” das *recomendações* da organização em relação às normas do programa, esse aspecto poderia ser objeto de uma orientação especial a todos os programas.

### ***9. Respondente da auditoria***

#### ***Definição***

A categoria Respondente da Auditoria representa as pessoas pesquisadas para prestar informações acerca do programa sejam beneficiários, equipe técnica do programa, especialistas, instituições e parceiros do programa. O tema beneficiário é representado por egressos, usuários, alunos e beneficiários em geral. O tema equipe técnica designa atendentes dos postos, servidores e técnicos e gestores nos três níveis de atuação do programa. São pessoas que desenvolvem atividades operacionais nos programas, inclusive servidores e terceirizados. O tema especialista representa profissionais de áreas específicas de atuação, por exemplo, especialistas em educação, assistentes sociais, defensores públicos, delegados, juizes, professores, profissionais da área e promotores. O tema instituições significa associações, comitês, instituições, organizações modelos, programas relacionados e sindicatos. O tema parceiros do programa designa pessoas e entidades com participação direta no programa como promotores, juizes e conselhos tutelares que prestam atendimentos aos beneficiários em outras instâncias, exemplo: comitês e conselhos.

### Quadro 14. Respondentes da auditoria

Categoria: Respondentes da auditoria			
Temas	Freq.	Exemplos	Programa auditado
Beneficiários	17	Beneficiários (crianças, adolescentes e familiares) (Brasil, 2004h, p. 31)	Sentinela
		Beneficiários (Brasil, 2003f, p. 33).	Morar Melhor
Equipe técnica	100	Técnicos (Brasil, 2003h, p. 35).	Novo Mundo Rural
		[...] equipe técnica [...] (Brasil, 2004h, p. 31) Gerentes (Brasil, 2003h, p. 35)	Sentinela Novo Mundo Rural
		[...] gestores federais, estaduais, municipais, equipe técnica e responsáveis pelos centros de referência [...] (Brasil, 2004h, p.30)	Sentinela
Especialistas	100	[...] especialmente com relação às questões pedagógicas, foi consultada a Secretaria de Ensino Fundamental – SEF/MEC responsável pela avaliação pedagógica dos livros didáticos (Brasil, 1999 <sup>a</sup> , p. 7)	Programa Nacional do Livro Didático
		[...] entrevistas com especialistas [...] (Brasil, 2001d, p. 20) [...] assistentes sociais da CAIXA (Brasil, 2003f, p. 34)	Controle da Mortalidade Materna Morar Melhor
		[...] profissionais envolvidos na área – educação, saúde, conselho (Brasil, 2004h, p. 32)	Sentinela
Instituições	30	Ministério do Turismo e Banco do Brasil, (Brasil, 2003b, p. 31)	Prodetur/NE
		ONGs (Brasil, 2000e, p. 17)	Pro-Água/Semi-Árido
Parceiros do programa	7	Conselhos Estaduais, Municipais e Tutelares (Brasil, 2004h, p. 30).	Sentinela
		[...] entidades de atendimento para egressos (Brasil, 2003g, p. 34)	Reinserção Social do Adolescente em Conflito com a Lei
<b>Total</b>	<b>254</b>		

Legenda: Freq = frequência

#### Análise

Pela quantidade de frequências, destacou-se nesta categoria o tema especialistas e equipe técnica, que expressa o ponto de vista de quem participa do cotidiano ou atua direta ou indiretamente no programa. Diversas fontes de informação e profissionais são buscados nas ANOPs, como exemplo assistentes sociais. Por fim, instituições e parceiros do programa são pesquisados no levantamento de informações dos programas.

#### 10. Planejamento da Auditoria

##### Definição

A categoria Planejamento de Auditoria define-se pelas técnicas utilizadas pela organização para o planejamento da auditoria. São elas: análise RECI, análise *stakeholder*,

análise SWOT, avaliação de riscos, *benchmarking*, técnica de elaboração de indicadores de desempenho, mapa de processos e produtos e matriz de planejamento. Tais técnicas foram apresentadas e definidas no presente trabalho no item “3.8 Descrição das técnicas utilizadas pela organização”.

### Quadro 15. Planejamento da auditoria

Categoria: Planejamento da auditoria			
Temas	Freq.	Exemplos	Programa auditado
Análise RECI	6	Análise RECI (Brasil, 2003g, p.38)	Reinserção Social do Adolesc. Em Conflito com a Lei
		Análise RECI (Brasil, 2000f, p. 7)	Controle da Mortalidade Materna
Análise <i>Stakeholder</i>	12	Análise <i>Stakeholder</i> (Brasil, 2003h, p. 45)	Novo Mundo Rural
		Pessoas-Chave para os Trabalhos – <i>stakeholders</i> (Brasil, 1999b, p. 8)	Programa Nacional do Livro Didático
Análise SWOT	13	Análise de possíveis influências (SWOT Analysis) (Brasil, 1999b, p. 7)	Programa Nacional do Livro Didático
		Análise SWOT (Brasil, 2003f, p. 37)	Morar Melhor
Avaliação de riscos	9	Análise de risco aplicada à ação assistência técnica e capacitação (Brasil, 2003h, p. 43)	Novo Mundo Rural
		Diagrama de avaliação de riscos (Brasil, 2003b, 43)	Prodetur/NE
		Diagrama de avaliação de risco (Brasil, 2003b, 19)	Prodetur/NE
Benchmarking	21	<i>Benchmarking</i> (Brasil, 2000e, p. 17)	roa-Água/Semi-Árido
		<i>benchmarking</i> das estruturas organizacionais das UEE’s, considerando a estrutura da BA como referência (Brasil, 2003b, p. 30)	Prodetur/NE
		[...] <i>benchmarking</i> [...](Brasil, 2003e, p. 3)	Energia das Pequenas Comunidades
Técnica de Indicadores de desempenho	11	Matriz de Indicadores (Brasil, 2004h, p. 33)	Sentinela
		A partir dos conhecimentos na fase de execução da auditoria, pode-se revisar os indicadores de desempenho inicialmente propostos [...] (Brasil, 1999a, p. 16)	Programa Nacional do Livro Didático
Mapa de processos e produtos	14	Mapa de produtos (Brasil, 2000e, p. 16)	Pró-Água/Semi-Árido
		O mapa de produtos das ações relativas à organização social e produtiva dos assentamentos rurais indica os seguintes efeitos [...] (Brasil, 2003h, p. 16)	Novo Mundo Rural
Matriz de planejamento	10	Matriz de planejamento (Brasil, 2004h, p. 27)	Sentinela
		Matriz de planejamento (Brasil, 2003h, p. 34)	Novo Mundo Rural
Total	96		

Legenda: Freq. = frequência

### *Análise*

Verificou-se imprecisão conceitual no uso dos termos e informações agrupadas sob esses títulos na matriz de planejamento da organização. Por exemplo, na nomeação das técnicas de diferentes formas: “análise” de risco, “avaliação” de risco e “diagrama” de avaliação de risco; mapa de “processos”, mapa de “processos e produtos” e mapa de “produtos e impactos”; indicadores de desempenho e “matriz” de indicadores de desempenho. Nesse sentido, sugerem-se cuidados no uso preciso dos termos técnicos para clareza do planejamento e execução da auditoria e dos que dele fazem uso.

## ***11. Critério de Seleção da Amostra da Auditoria***

### *Definição*

A categoria Critério de Seleção da Amostra refere-se ao critério utilizado para a escolha da amostra da auditoria, por exemplo, uma amostra aleatória ou selecionada intencionalmente através de critérios como volume de recursos, demográficos, quantitativo de atendimentos e programas com diferenciais. O tema amostra aleatória representa o conceito usualmente utilizado em pesquisa, em que todos os participantes possuem a mesma probabilidade de ser escolhidos para formar a amostra. O tema volume de recursos envolve o porte dos recursos repassados pelo governo no nível federal ao programa e às contrapartidas financeiras estaduais e municipais ao programa. O tema diferenciais de atendimento representa a escolha da amostra com base em características que diferenciam o Programa positiva ou negativamente das outras unidades onde se encontra implantado. O tema distância geográfica refere-se à escolha da amostra mediante localização geográfica da unidade do programa. O tema quantidade de atendimentos significa a quantidade de atendimentos realizados, nº de famílias atendidas e sistemas do programa em funcionamento usados como critério para a seleção da amostra.

### Quadro 16. Amostra da auditoria

Categoria: Amostra da auditoria			
Temas	Freq.	Exemplos	Programa auditado
Amostra aleatória	3	<p>[...] os destinatários deste questionário foram escolhidos por meio de amostra aleatória simples (Brasil, 2002f, p. 13)</p> <p>Serão enviados questionários à 230 prefeituras, escolhidas por meio de amostra aleatória a partir e listagem encaminhada pela SFCI (Brasil, 2003e, p. 29)</p>	<p>Novo Emprego e Seguro Desemprego</p> <p>Energia das Pequenas comunidades</p>
Volume de recursos	5	<p>Distribuição dos recursos e existência de financiamento das esferas governamentais e não governamentais (Brasil, 2004h, p. 27)</p> <p>Para a determinação dos empreendimentos que irão compor a pesquisa, adotou-se como critério o volume de recursos alocados pelo Programa nas unidades da federação [...] (Brasil, 2003f, p. 19)</p>	<p>Sentinela</p> <p>Morar Melhor</p>
Diferenciais de atendimento	5	<p>[...] de uma amostra selecionada de hospitais considerados referência na área de atendimento à saúde da mulher [...] (Brasil, 2001d, p. 20).</p> <p>A auditoria concentrou seu foco de ação nos municípios das regiões Norte, Centro-Oeste e Nordeste Essa escolha deu-se em virtude dos seguintes fatores: a) o Programa faz parte do Projeto Alvorada, que concentra suas ações nos municípios com Índices de Desenvolvimento Humano – IDH abaixo de 0,5, [...] (Brasil, 2003c, p. 8)</p> <p>Projetos que sejam considerados “boas” e “más” práticas (Brasil, 2003f, p. 18)</p>	<p>Controle da Mortalidade Materna</p> <p>Sentinela</p> <p>Morar Melhor</p>
Distância geográfica	3	<p>Em cada estado, os trabalhos foram realizados em um município, além da capital, escolhido com base nos seguintes critérios: facilidade de acesso em relação à capital [...] (Brasil, 2004d, p. 18).</p> <p>O município foi selecionado em razão da pouca distância do DF [...] (Brasil, 2003h, p. 27)</p> <p>[...] Para a escolha dos municípios, ponderou-se, sobretudo, a localização geográfica próxima às capitais (Brasil, 2004h, p. 17)</p>	<p>Sentinela</p> <p>Novo Mundo Rural</p> <p>Sentinela</p>
Quantidade de atendimentos	4	<p>A auditoria centrou seu foco [...] nos municípios das regiões norte, centro-oeste e nordeste. Esta escolha deu-se em virtude dos seguintes fatores: [...] dos 5.912 sistemas adquiridos pelo Programa até sua Fase IV 2.254 sistemas destinaram-se à região Nordeste [...] (Brasil, 2003c, p. 8)</p> <p>[...] os assentamentos em função do número de famílias. (Brasil, 2003h, p. 27)</p>	<p>Energia das Pequenas Comunidades</p> <p>Novo Mundo Rural</p>
Total	20		

Legenda: Freq = frequência

### *Análise*

As amostras não estatísticas comprometem a generalização dos resultados de um estudo, pois podem não representar as características da população e não oferecerem a oportunidade que todas as pessoas da população participem da pesquisa. Nesse sentido, uma amostra não estatística pode servir a estudos qualitativos ou comparativos. Deve-se ressaltar que nos documentos consultados encontrou-se apenas uma recorrência de *amostra não-estatística*, embora esse tipo de amostra tenha sido utilizado em algumas ANOPs.

Chama-se atenção para as seguintes características dos programas sociais, a abrangência de participantes e amplitude do funcionamento em quase todas as unidades da federação. Essas características demandam amostras aleatórias para assegurar a validade dos resultados das ANOPs para o programa como um todo.

## **12. Técnicas de Coleta de Dados da auditoria**

### *Definição*

As Técnicas de Coleta de Dados da auditoria definem-se pelos instrumentos utilizados para a coleta de informações sobre o programa, sejam questionários, entrevistas, observação direta e grupo focal. As técnicas que formam os temas desta categoria foram apresentadas e definidas no item “3.8 Descrição das técnicas utilizadas pela organização”.

### **Quadro 17. Técnica de coleta de dados**

Categoria: Técnica de coleta de dados			
Temas	Freq.	Exemplos	Programa auditado
Questionário	27	Questionário semi-estruturado para equipes técnicas das Superintendências Regionais do INCRA (Brasil, 2003h, p. 36)	Novo Mundo Rural
		Questionários (Brasil, 2003f, p. 32)	Morar Melhor
Entrevista	40	Entrevista estruturada (Brasil, 2003f, p. 32)	Morar Melhor
		Entrevista semi-estruturada com membros das associações; com grupos de beneficiários; técnicos de assistência técnica e de empresas contratadas (Brasil, 2003h, p. 37)	Novo Mundo Rural
		Entrevistas com as mães e responsáveis pelos beneficiários (Brasil, 2004h, p. 32)	Sentinela
		Entrevista não-estruturada com gestores (Brasil, 2003g, p. 30)	Reinserção Social do Adolescente em Conflito com a Lei

(cont.) Quadro 17. Técnica de coleta de dados

Categoria: Técnica de coleta de dados			
Temas	Freq.	Exemplos	Programa auditado
Observação direta	12	Observação local direta (Brasil, 2003h, p. 37)	Novo Mundo Rural
		Observação direta (Brasil, 2000e, p. 17)	Pró-Água/Semi-Árido
		Observação direta (Brasil, 2003g, p. 30)	Reinserção Social do Adolescente em Conflito com a Lei
Grupo focal	23	Grupo focal (Brasil, 2003g, p. 33).	Reinserção Social do Adolescente em Conflito com a Lei
		O levantamento <i>in loco</i> será realizado por meio de grupos focais (técnicos do INCRA) [...] (Brasil, 2003h, p. 26)	Novo Mundo Rural
Total	<b>102</b>		

Legenda: Freq = frequência

### Análise

Verificou-se na análise dos documentos consultados que a diferenciação nos tipos de técnicas não está clara nos relatórios de planejamento e final. Encontrou-se a ocorrência de imprecisão conceitual na denominação e conhecimento das técnicas utilizadas nas ANOPs. Os instrumentos em anexo aos relatórios de auditoria não se configuram como entrevistas. Por exemplo, semi-estruturadas, são estruturadas; não apresentam questionários semi-estruturados, trata-se de questionários estruturados. Há diferenciação de citação de questionário semi-estruturado e questionário; entretanto, não se encontrou modelo de questionário semi-estruturado. Há também diferenciação de entrevista semi-estruturada, entrevista estruturada e entrevista, entretanto, no material coletado as entrevistas são estruturadas. Na citação de observação local direta, não se esclarece acerca dos aspectos que seriam observados *in loco* e o tipo de instrumental de apoio para a coleta dos dados.

O tema observação direta refere-se a visitas *in loco* realizadas para a coleta de dados e estava descrito sob o título delineamento na matriz de planejamento, em vez de ser exemplificado como um procedimento de coleta de dados. As cinco recorrências de *transcrição, narração e descrição*, citadas na matriz de planejamento do Programa Novo Mundo Rural, estão sob o título método de análise de dados, porém, trata-se de procedimento de tabulação de dados.

Por fim, conforme leitura dos documentos de planejamento da auditoria, o grupo de discussão, com duas recorrências, refere-se ao grupo focal.

### 13. Procedimento de Coleta e Tabulação dos Dados

#### Definição

A categoria Procedimento de Coleta e Tabulação de Dados é formada por pesquisa documental, pesquisa postal, requisição de documentos e tabulação de dados. O tema pesquisa documental representa a análise de documentos do programa, leitura de bibliografia sobre o tema, pesquisas sobre o programa e fenômeno a ele relacionado na internet. O tema pesquisa postal configura-se como o envio de questionários (tipo lápis e papel), pelo correio convencional, aos participantes da pesquisa. O tema requisição de documentos contempla a solicitação formal dos documentos do programa aos gestores. O tema tabulação de dados significa a digitação dos dados em planilhas eletrônicas para análise estatística.

#### Quadro 18. Procedimento de coleta e tabulação de dados

Categoria: Procedimento de coleta e tabulação de dados			
Temas	Freq.	Exemplos	Programa auditado
Pesquisa documental	24	Os dados foram extraídos de bibliografia sobre o tema, pesquisa na Internet, dados do DNDE/MME, Aneel, Eletronorte, Cepel e Chesf. [...] (Brasil, 2003c, p. 9) Pesquisa bibliográfica e documental (Brasil, 2003h, p. 37).	Energia das Pequenas Comunidades Novo Mundo Rural
		Pesquisa documental (Brasil, 2004h, p. 32)	Sentinela
Pesquisa postal	15	Pesquisa postal para todos os centros de referência (Brasil, 2004h, p. 31)	Sentinela
		Questionário postal para todos os gestores municipais e estaduais e todos os centros de referência (Brasil, 2004h, p.30)	Sentinela
		Aplicação de questionários às Secretarias de Saúde de todos os Estados [...] (Brasil, 2001d, p. 19)	Controle da Mortalidade Materna
Requisição de documentos	10	Requisição formal de documentos (Brasil, 2003f, p. 33)	Prodetur
		Requisição formal de documentos (Brasil, 2004h, p. 27)	Sentinela
Tabulação de dados	5	[...] transcrição, narração, descrição (Brasil, 2003h, p. 39)	Novo Mundo Rural
<b>Total</b>	<b>54</b>		

Legenda: Freq = frequência

#### Análise

Verificou-se imprecisão conceitual nas ANOPs em relação às terminologias adotadas para coleta e análise de dados, havendo também ambigüidade na descrição da aplicação das mesmas. Dentre as terminologias descritas, houve uma recorrência de “oficina de trabalho”

(Brasil, 2000e, p. 16), entretanto, nos documentos consultados não se esclarece o tipo de atividade realizada na oficina. Houve também duas recorrências de pesquisas via internet, narração e observação direta, sem maiores esclarecimentos sobre a aplicação das mesmas e o tipo de informação buscada.

#### ***14. Análise dos Dados***

##### *Definição*

A categoria Análise dos Dados refere-se às análises realizadas pela auditoria em relação aos dados coletados dos programas sociais, a saber: análise de conteúdo; análise de dados secundários; análise de riscos; análise qualitativa; análise quantitativa, análise orçamentária e financeira. O tema análise de conteúdo refere-se à análise qualitativa originalmente baseada em Bardin (1977), porém, nos documentos consultados não há descrições que fundamentem essa definição. O tema análise de dados secundários representa a análise de dados de terceiros ou não coletados diretamente pelos auditores. O tema análise de risco está definido no item “3.8 Descrição das técnicas utilizadas pela organização”, no presente trabalho. O tema análise qualitativa significa uso de análises subjetivas e o tema análise quantitativa refere-se a análises numéricas. Houve recorrências desses temas como técnicas de análise de dados nos documentos consultados, mas não se encontrou elementos para a definição dos mesmos. O tema análise orçamentária e financeira representa a análise do uso dos recursos em relação aos valores autorizados, empenhados e liquidados pelo programa. O tema série temporal com grupo de controle é aquele empregado usualmente em pesquisa com comparação de variáveis através do tempo.

#### **Quadro 19. Análise dos dados**

Categoria: Análise de dados			
Temas	Freq.	Exemplos	Programa auditado
Análise de Conteúdo	10	Análise de conteúdo do grupo focal (Brasil, 2003b, p. 26)	Prodetur/NE
		Análise de conteúdo das entrevistas e grupos focais (Brasil, 2003f, p. 34)	Morar Melhor
Análise de dados secundários	4	Análise de dados secundários (Brasil, 2003e, p.3)	Energia das Pequenas Comunidades
		Análise documental (Brasil, 2003f, p. 32)	Morar Melhor

(cont.) Quadro 19. Análise dos dados

Categoria: Análise de dados			
Temas	Freq.	Exemplos	Programa auditado
Análise de riscos	3	A apuração dos custos de execução do PNLD [...] (Brasil, 1999b, p. 11)	Programa Nacional do Livro Didático
		Fontes de financiamento; destinação e aplicação de recursos [...] (Brasil 2002g, p. 3)	Novo-Emprego e Seguro Desemprego
		Análise orçamentária (Brasil, 2003e, p.10)	Energia das Pequenas Comunidades
Análise qualitativa	12	Análise qualitativa dos dados coletados (Brasil, 2003b, p. 26)	Prodetur/NE
Análise quantitativa	17	Tratamento estatístico (Brasil, 2000e, p. 17)	Pró-Água/Semi-árido
		Tratamento estatístico de questionários e entrevistas estruturadas (Brasil, 2003b, p. 26)	Prodetur/NE
Análise orçamentária e financeira	6	Execução orçamentária e financeira dos recursos autorizados para o Programa Assentamento de Trabalhadores Rurais no exercício de 2000, 79,6% foram utilizados para aquisição de terras [...] (Brasil, 2003h, p. 11)	Novo Mundo Rural
		A execução orçamentária efetiva (valor liquidado no exercício (X) menos o resto a pagar cancelado no exercício seguinte (X + 1)) em 2002 foi de 77,5% dos recursos (Brasil, 2003h, p. 16)	Novo Mundo Rural
		A Tabela 1 demonstra a execução orçamentária e financeira do Programa Sentinela no período de 2000/2003 (Brasil, 2004h, p. 14)	Sentinela
Total	<b>52</b>		

Legenda: Freq = frequência

### Análise

Verificou-se que mesmo sendo citado nas ANOPs o uso da análise de conteúdo, não foram mencionadas nos relatórios finais as referências as falas dos entrevistados às equipes das ANOPs. Também praticamente não houve casos de citações de falas no corpo dos documentos e não se encontraram análises categoriais temáticas.

A expressão *análise de dados secundários* não esclarece o tipo de análise, ocorreu o mesmo na referência a análise estatística, que por si só, não revela o procedimento adotado. Nas ANOPs encontrou-se bem descritas estatísticas descritivas como médias, percentuais e totais dos dados, em alguns casos foram usadas correlações.

A menção a análise qualitativa e análise quantitativa não especificam os tipos de análises realizadas, sendo possível encontrá-las somente através da leitura e manuseio dos documentos analisados nesta tese.

Foram mencionadas como análise dos dados as análises RECI, *Stakeholder* e SWOT. Entretanto, essas são técnicas utilizadas para o planejamento das auditorias e não na análise dos dados coletados posteriormente ao planejamento. Verificou-se esse equívoco no preenchimento da matriz de planejamento dos programas, no item que prevê a técnica de análise de dados.

Devido a grande quantidade de informações qualitativas e quantitativas obtidas nas ANOPs, sugere-se que a análise dos dados seja realizada com auxílio de softwares estatísticos, como o *Statistical Package of Social Science* – SPSS e softwares para análise de dados qualitativos como o Alceste, Evoque ou outros.

## **15. Métodos da Auditoria**

### *Definição*

A categoria Métodos da Auditoria foi formada pelos temas: estudo de caso; estudo-piloto, exemplos extremos e série temporal com grupo de controle. De acordo com o TCU o tema estudo de caso, representa um ou mais casos analisados, a partir dos quais se deduz o funcionamento dos demais programas (casos), embora, metodologicamente essa definição se refira ao estudo de um caso. O tema estudo-piloto representa a execução da auditoria em uma unidade do programa para testar os instrumentais e os demais procedimentos conforme o planejamento da auditoria. O tema exemplos extremos refere-se a delineamento com base em casos de desempenho excepcional do programa, seja positivo ou negativo, representando realidades extremas. O tema série temporal com grupo de controle possui a definição semelhante em pesquisa avaliativa e foi definida no item 2.4.1 Delineamento de Pesquisa.

## Quadro 20. Delineamento da Auditoria

Categoria: Métodos da auditoria			
Temas	Freq.	Exemplos	Programa auditado
Estudo de caso	16	Estudo de caso (Brasil, 2000e, p. 16)  A efetividade dos instrumentos de controle do PNLD foi aferida por intermédio [...] da realização de estudos de casos junto as SEDUCs (Brasil, 1999a, p. 7)	Proáguia  Programa Nacional do Livro Didático
Estudo-piloto	3	Foi realizado teste-piloto no Distrito Federal [...] (Brasil, 2003d, p. 18)  Em função do grande número de informações colhidas durante a realização do teste-piloto [...] (Brasil, 2003h, p. 26)	Reinserção Social do Adolescente em Conflito com a Lei  Novo Mundo Rural
Exemplos extremos	3	Em cada grupo, foram identificados dois Estados com atuação considerada boa dos comitês de mortalidade materna, e dois em que essa atuação é deficiente ou inexistente. (Brasil, 2001d, p. 21)	Controle da Mortalidade Materna
Série temporal com grupo de controle	3	Série temporal com grupo de controle dos dados sobre emprego, renda e n°. de leitos estabelecidos de hospedagem (Brasil 2003b, p. 26)  Análise de série temporal (Brasil, 2004a, p.17)	Prodetur/NE  Prodetur/NE
Total	<b>25</b>		

Legenda: Freq = frequência

### Análise

Verificou-se imprecisão conceitual nas ANOPs em relação ao significado das estratégias metodológicas, inclusive sob esse título são listados nas matrizes de planejamento das auditorias: procedimentos de coleta de dados como consulta a dados secundários, pesquisa postal, pesquisa documental, escolha da amostra, visitas exploratórias, visitas *in loco*, critérios de escolha dos estados. A denominação de estudo de caso foi utilizada de forma incorreta, na verdade as ANOPs realizaram estudos comparativos ou longitudinais, com coleta de dados em diversas unidades, como exemplo em 71 comunidades conforme ocorreu em uma auditoria de programa. As denominações observação direta, observação *in loco*, visita e visita exploratória foram usadas para descrever o que é na verdade refere-se à pesquisa de campo. Conforme se verificou nos documentos consultados as equipes de auditoria necessitam de aperfeiçoamento sobre os tipos de delineamentos, suas vantagens e desvantagens em relação à validade interna e externa para a pesquisa avaliativa e a auditoria.

Foi citado em Brasil (2003b, p. 26) o delineamento “Série temporal com grupo de controle de dados sobre emprego, renda e n°. de leitos em estabelecimentos de hospedagem”, o que indica conhecimento de delineamento adequado a estudos longitudinais para programas que possuem longa duração, como foi o caso do Prodetur/NE.

Encontraram-se duas recorrências do termo pesquisa amostral nas ANOPs, embora não seja um tema da categoria Delineamento da Auditoria, o que é outro exemplo de imprecisão na descrição sobre o tipo de pesquisa realizado.

## ***16. Dificuldades da Auditoria***

### *Definição*

A Categoria Dificuldades da Auditoria refere-se às dificuldades na execução do processo e está relacionada à mortalidade instrumental, prazos curtos para cumprimento das recomendações pelo programa; dificuldades técnicas da auditoria e do programa; criação do grupo de contato de auditoria e acesso geográfico aos participantes do programa. O tema mortalidade instrumental representa a baixa taxa de resposta aos questionários. O tema prazo para a auditoria significa a limitação de tempo para a realização da ANOP. O tema Prazos para Cumprimento das Recomendações refere-se ao período concedido aos responsáveis pelo programa para cumprimento e envio ao TCU, do plano de ação, contemplando prazos de implantação, cronograma de adoção das medidas ao cumprimento das recomendações e a indicação dos responsáveis pelas ações. Tais prazos de 60 ou 90 dias são curtos para os programas, que em geral pedem prorrogação para a resposta ao TCU e para a implantação das recomendações. O tema dificuldades técnicas da Auditoria refere-se a dificuldades internas da equipe de auditoria para a realização do processo de ANOP, por exemplo, a necessidade de treinamento em análises qualitativas como a análise de conteúdo para respostas de grupos focais. O tema dificuldades técnicas do programa são aquelas de ordem de funcionamento do programa. O tema Grupo de Contato de Auditoria refere-se à criação de um grupo de pessoas do programa com o objetivo de atuar como canal de comunicação entre o Programa e o Tribunal para implantação e acompanhamento das recomendações. O tema Acesso geográfico significa a dificuldade de localização de endereço e acesso à moradia dos beneficiários.

## Quadro 21. Dificuldades da Auditoria

Categoria: Dificuldades da Auditoria			
Temas	Freq.	Exemplos	Programa auditado
Mortalidade instrumental	7	Baixa taxa de retorno dos questionários (Brasil, 2003g, p. 32)  [...] questionários por via postal aos 411 conselheiros de turismo [...] obtendo resposta a 218 questionários (Brasil 2004a, p. 17)	Reinserção Social do Adolescente em Conflito com a Lei  Prodetur/NE
Prazo para a auditoria	3	Dada a limitação de tempo para a realização dos trabalhos, que é uma característica das auditorias-piloto do Projeto [...], o planejamento indicou a impossibilidade de se realizar uma avaliação completa do Programa. [...] (Brasil, 2000d, p. 33)	Pró-Água/Semi-Árido
Prazos para cumprimento das recomendações		Remeta ao Tribunal, no prazo de 60 dias, plano de ação contendo o conjunto de metas correspondentes aos indicadores de desempenho recomendados na alínea "9.3.11 ", contemplando prazo para o atingimento dessas metas, e o cronograma de adoção das medidas necessárias à implantação das recomendações [...] com o nome dos responsáveis pela implantação dessas medidas (Brasil, 2003d, p. 98)  Remeta ao Tribunal, no prazo de 90 dias, plano de ação, elaborado em conjunto com as Superintendências Regionais, contendo o conjunto de metas correspondentes aos indicadores de desempenho recomendados, contemplando prazo para o atingimento dessas metas, e o cronograma de adoção das medidas necessárias à implantação das recomendações [...] com a identificação dos responsáveis pela implantação dessas medidas (Brasil, 2004c, p.58).	Pró-Água/Semi-Árido  Novo Mundo Rural
Prazos para cumprimento das recomendações	5	Elabore plano de implantação de versões do Sistema de Informação para a Infância e a Adolescência - Sipiá, com cronograma que contemple informações a respeito do custo de execução da medida socioeducativa por adolescente e das atividades de profissionalização, escolarização e lazer desenvolvidas durante o cumprimento de medida (Brasil, 2003d, p. 97)	Reinserção Social do Adolescente em Conflito com a Lei
Dificuldades técnicas da auditoria	3	Insuficiência e falta de clareza dos atos normativos disciplinando a matéria (Brasil, 2003g, p. 30)  risco de não contar com assessoria ou software de análise de conteúdo para grupos focais	Reinserção Social do Adolescente em Conflito com a Lei

## (cont.) Quadro 21. Dificuldades da Auditoria

Categoria: Dificuldades da Auditoria			
Temas	Freq.	Exemplos	Programa auditado
Dificuldades técnicas do programa	8	Quanto ao Sistema Sipra, além de problemas técnicos, de confiabilidade e de segurança dos dados, constatamos ainda que este não é alimentado regularmente pelas Superintendências Regionais (Brasil, 2003h, p. 19)	Novo Mundo Rural
		Quanto aos fatores que impõem limites à realização da auditoria, consideramos como principais pontos: baixa execução das ações; baixa confiabilidade dos dados disponíveis [...] (Brasil, 2003h, p. 26)	Novo Mundo Rural
Grupo de contato de auditoria	7	Estabeleçam um grupo de contato de auditoria [...], para atuar como canal de comunicação com este Tribunal e para acompanhar a implantação das recomendações desta Corte de Contas e a evolução dos indicadores de desempenho e o atingimento das respectivas metas (Brasil, 2003d, p.79)	Reinserção Social do Adolescente em Conflito com a Lei
Acesso geográfico	7	A despeito da equipe de auditoria ter visitado localidades de difícil acesso, só possível por meio de horas de viagens de barco ou trator, não houve condição operacional de visitar aquelas ainda mais isoladas [...] (Brasil, 2003c, p. 10)	Energia das Pequenas Comunidades
		Há também dificuldades de acesso às associações, o que poderá dificultar a pesquisa (Brasil, 2003h, p. 26)	Novo Mundo Rural
Total	<b>40</b>		

*Legenda: Freq = frequência*

### *Análise*

Na categoria Dificuldades Internas da Auditoria foram mencionados problemas como mortalidade instrumental, dificuldades em relação a prazo para execução da auditoria e dificuldades técnicas, tais problemas também são comuns à pesquisa avaliativa em geral. Entretanto, o conhecimento dessas limitações precisa ser considerado na adoção de providências pelo TCU, como forma de apoiar o trabalho das equipes que realizam as ANOPs e minimizar seus impactos no resultado das ANOPs. Considerando que uma das principais metas do TCU é a verificação da prestação de contas dos governos em relação ao uso de recursos, ao fazê-lo por meio da auditoria, torna-se necessário assegurar sua viabilidade.

Diante das limitações de tempo para a realização das ANOPs constatadas na presente pesquisa, sugere-se à organização considerar a necessidade de ampliação dos prazos para a execução do processo de ANOP. Assim como, a necessidade de continuar capacitando sua equipe técnica no aparato metodológico necessário à “boa” realização do trabalho de auditoria, conforme constatado na meta-avaliação. Entre as possibilidades de melhoria da

capacitação técnica metodológica e suporte organizacional, a estratégia de contratação de consultores, adotada em algumas ANOPs, mostrou-se eficiente e poderia ser considerada como praxe para auditorias em áreas específicas e de difícil domínio público dos programas públicos.

Conforme verificado por meio do tema Prazos para Cumprimento das Recomendações a organização, existe a necessidade de rever os prazos para que os programas implantem as recomendações. Pois, os prazos de 60 e 90 dias não parecem suficientes à amplitude e complexidade dos programas e os esforços necessários para fazê-lo. Nesse sentido, identificou-se uma diferenciação nos prazos concedidos a dois programas, o Reinserção Social do Adolescente em Conflito com a Lei e o Novo Mundo Rural, para cumprir as mesmas recomendações. As recomendações foram: a elaboração do plano de ação com prazos para o alcance das metas e cronograma das recomendações, com responsáveis pela implantação das medidas. Não se identificou nos documentos consultados fatores que levassem à diferenciação de prazos para as mesmas recomendações. À medida que os prazos de 60 ou 90 dias tornam-se curtos para os gestores dos programas, tornam-se curtos também para o próprio TCU acompanhá-los.

Apesar das dificuldades apontadas para a realização das ANOPs, o TCU possui alternativas para superá-las. Devido ao conhecimento de seu corpo técnico e evolução da ANOP, os ajustes metodológicos necessários poderiam ser alvo de discussão e aprimoramento. Uma alternativa à avaliação e auditoria de grandes programas seria trabalhar com amostras aleatórias que garantam a representatividade dos resultados e utilizar delineamentos de pesquisa que minimizem as ameaças à validade interna e externa.

## ***17. Responsabilização***

### ***Definição***

A categoria Responsabilização representa o comportamento dos gestores do programa sua consistência com leis e normas, regulamentação interna do programa e modificação de procedimentos internos. O tema consistência com leis e normas significa a adequação do programa a leis, decretos, normas operacionais, normativos, diretrizes e manuais. O tema papel desempenhado representa a execução das ações do programa com base em propostas pré-estabelecidas, atribuições das instâncias envolvidas no Programa, definição de competência na execução das ações e da atuação institucional e de seu papel institucional. O tema regulamentação interna do programa refere-se à regulamentação do Programa, por meio

de documentos, convênios e proposição de modelos a serem adotados no programa. Exemplo, a elaboração de cartilhas e manuais de procedimentos do programa, detalhando metodologias de atendimento ao beneficiário. O tema modificação de procedimentos internos do programa refere-se a novas diretrizes e critérios a serem incorporados ao programa em suas documentações legais e o estabelecimento ou modificação de prazos das atividades do programa.

## Quadro 22. Responsabilização

Categoria: Responsabilização			
Temas	Freq.	Exemplos	Programa auditado
Consistência com leis e normas	55	Os profissionais que atuam no programa Sentinela deverão, obrigatoriamente, de acordo com o ECA, seguir os mandamentos estabelecidos para a Política de atendimento (arts. 86 e 87, incisos) [...] (Brasil, 2004h, p. 16)	Sentinela
		[...] se as atribuições das instâncias envolvidas estão claras (Brasil, 2003g, p. 30)	Reinserção Social do Adolescente em Conflito com a Lei
		Estabeleça, nos contratos de subempréstimo, o compromisso do submutuário (Brasil, 2004a, p. 93)	Prodetur/NE
Papel desempenhado	21	A Superintendência Nacional de Desenvolvimento agrário, no exercício de suas atribuições [...] (Brasil, 2003h, p. 17)	Programa Novo Mundo Rural
		[...] uma vez que não cabe ao Incra, como órgão executor da política fundiária, a execução dos programas de fomento e assistência técnica (Brasil, 2003h, p. 15)	Programa Novo Mundo Rural
Regulamentação interna do programa	18	Disponibilize as UEGPs documentos considerados modelos, desenvolvidos na própria Unidade de Gerenciamento do Programa -Gestão - UGPG ou em alguma das unidades estaduais, como, por exemplo, termos de referência para o desenvolvimento de planos diretores de recursos hídricos, para a elaboração de estudos concernentes à cobrança pelo direito de uso da água, assim como documentos sobre diretrizes gerais e específicas relativas ao Pró-Água/Semi-Árido (Brasil, 2000d, s/n).	Pró-Água/Semi-Árido
		O regimento interno do INCRA [...] prevê três instâncias [...] (Brasil, 2003h, p. 17)	Programa Novo Mundo Rural
Modificação de procedimentos internos do programa	12	Promova alterações nos regulamentos dos conselhos de turismo, com vigência a partir do segundo mandatos desses colegiados [...] (Brasil, 2004a, p.94)	Prodetur/NE
		Estabeleça cláusula nos convênios/termos de responsabilidade [...] (Brasil, 2004d, p.95)	Sentinela
<b>Total</b>	<b>106</b>		

Legenda: Freq = frequência

### *Análise*

Na categoria Responsabilização tanto pelos temas nomeados nessa categoria e pela expressiva quantidade de fontes de dados secundários utilizados nas ANOPs, verificou-se o aspecto sancionador das avaliações, pois há forte ênfase na Normatização para os programas. Nesse sentido, também se destacou recorrência do tema Consistência com Leis e Normas e Papel Desempenhado.

### **18. Comunicação Institucional**

#### *Definição*

A categoria Comunicação Institucional define-se pelos sistemas de informação dos programas, comunicação com beneficiários, parceiros e publicização. Os sistemas de informação são os sistemas informatizados como bancos de dados para armazenamento do perfil dos beneficiários e controle de atividades relacionadas ao programa. O tema comunicação com os parceiros significa a divulgação do programa por meio de documentos e orientações sobre procedimentos e ocorre nos níveis federal, estadual e municipal em relação aos parceiros e beneficiários. O tema publicização representa a divulgação das ações do Programa para a sociedade, ações de sensibilização, mobilização das comunidades e divulgação da política nacional na qual o programa se insere.

### **Quadro 23. Comunicação Institucional**

Categoria: Comunicação Institucional			
Temas	Freq.	Exemplos	Programa auditado
Sistemas de informação		Elabore plano de implantação de versões do Sistema de Informação para a Infância e a Adolescência - Sipi, com cronograma que contemple informações a respeito do custo de execução da medida socioeducativa por adolescente e das atividades de profissionalização, escolarização e lazer desenvolvidas durante o cumprimento de medida (Brasil, 2003d, p. 78)	Reinserção Social do Adolescente em Conflito com a Lei
	5	Implemente ou amplie, a exemplo do que já acontece no Programa Habitar Brasil/BID, ações voltadas ao Desenvolvimento Institucional dos agentes executores das políticas municipais de habitação (Brasil, 2004e, p. 108)	Morar Melhor

## (cont.) Quadro 23. Comunicação Institucional

Categoria: Comunicação Institucional			
Temas	Freq.	Exemplos	Programa auditado
Comunicação com parceiros	16	<p>Desenvolva canais de comunicação regulares com estados e municípios por meio de, por exemplo, fóruns de discussão na internet e eventos regulares para troca de experiências entre os executores das ações, incluindo a disseminação de boas práticas e de informações registradas em textos, cartilhas ou outros meios sobre a importância das atividades do trabalho social para as famílias atendidas (Brasil, 2004e, p.132)</p> <p>Institua canal formal de comunicação permanente, entre o Ministério e os demais gestores do Programa, como por exemplo, por meio da criação de um sítio específico para o Programa Sentinela, com lista de discussão, correio eletrônico, concurso de boas práticas, espaço para divulgação de informações e encaminhamentos de interesse do Programa; (Brasil, 2004d, p. 95)</p> <p>[...] falta de comunicação e coordenação entre a sede e as superintendências regionais [...] (Brasil, 2003h, p. 34)</p>	<p>Prodetur</p> <p>Sentinela</p> <p>Novo Mundo Rural</p>
Publicização	6	<p>Dissemine as boas práticas desenvolvidas por municípios e estados apresentadas neste Relatório. (Brasil, 2004d, p.96)</p> <p>Divulgar mais amplamente a existência de um protocolo de pré-natal elaborado pelo Ministério da Saúde, certificando-se de que está atingindo seu público alvo e ressaltando a necessidade de sua adoção (Brasil, 2001d, p. 92)</p>	<p>Sentinela</p> <p>Controle da Mortalidade Materna</p>
Total	27		

Legenda: Freq = frequência

### Análise

A Comunicação Institucional nos níveis federal, estadual e municipal e a publicização parecem ser um problema comum aos programas avaliados, que possuem sistemas de informação e comunicação deficientes, conforme apontam as auditorias. A busca de publicização pode ser verificada no canal de comunicação aberto que a organização busca manter com a sociedade civil para divulgar seus resultados, a exemplo do Evento realizado periodicamente *Diálogo Público*. Esse evento caracteriza-se pela apresentação de resultados das ANOPs a membros internos e externos da organização. Assim, possui participação aberta, são convidados ONGs, representantes do Governo dos programas Auditados e outros.

### 19. Objetivos dos programas auditados

A seguir, a definição dos objetivos de implantação, intermediários e finais de acordo com Posavac e Carey (2003). O tema objetivos de implantação representa os níveis de atividade do programa necessários para a sua instalação, funcionamento e alcance dos resultados e refere-se às atividades necessárias para iniciar o funcionamento integral do programa. O tema Objetivos intermediários define o que se espera como resultado da implantação do programa e representam o programa em funcionamento. O tema Objetivos finais descreve os resultados finais que se espera alcançar com o programa, refletem a modificação da linha de base, antes da instalação do programa, ao momento da avaliação.

#### Quadro 24. Frequência da Classificação dos Objetivos dos Programas Avaliados

Auditorias de Programas	Objetivo			Total
	Implan- tação	Interme- diários	finais	
Programa Controle da Mortalidade Materna	9	10	0	19
Programa Novo Mundo Rural	2	13	4	19
Programa Reinserção Social do Adolescente em Conflito com a Lei	2	10	5	17
Programa Energia das Pequenas Comunidades	4	9	2	15
Programa de Desenvolvimento Sustentável de Recursos Hídricos para o Semi-árido Brasileiro – Pró-Água/Semi-Árido	4	4	5	13
Programa Morar Melhor	1	4	4	9
Programa de Desenvolvimento do Turismo do Nordeste – Prodetur/NE	1	4	2	7
Programa Nacional do Livro Didático – PNLD	0	1	0	1
Programa de Enfrentamento a Exploração e Abuso Sexual de Crianças e Adolescentes – Sentinela.	2	2	2	6
Programa Novo Emprego e Seguro Desemprego	0	1	1	2
Total	25	58	25	108

#### *Análise*

Conforme destacado no quadro acima de acordo com a distribuição da classificação dos objetivos baseada na recorrência dos mesmos dentro dos documentos consultados, verificou-se que não constam nos documentos pesquisados os objetivos de implantação do Programa Nacional do Livro Didático e Novo Emprego e Seguro Desemprego. Isso não significa que o programa não possua tal tipo de objetivo, e sim que na realização da ANOP não se descreveu o objetivo. Isso pode ter ocorrido devido a dificuldade que os formadores de política e programas possuem em definir claramente os objetivos, ou devido a falhas no levantamento da auditoria por não ter apreendido de forma suficiente o desenho do programa.

Verificou-se também que não constam os objetivos finais dos Programas Controle da Mortalidade Materna e Programa Nacional do Livro Didático. Acredita-se que qualquer

programa possua objetivo final. Entretanto, o fato de não constarem nos relatórios analisados também pode ter ocorrido devido em função da dificuldade dos formuladores dos programas em definir os objetivos, ou pode ter havido a falha no levantamento da auditoria, por não ter apreendido de forma suficiente o desenho do programa.

No planejamento e criação de objetivos de um programa, em geral a cada objetivo final, correspondem objetivos de implantação e objetivos intermediários (Posavac & Carey, 2003). Há grande variedade nas frequências constatadas na classificação dos objetivos, embora avaliações de diferentes programas. Por exemplo, no Programa Novo Mundo Rural (Brasil, 2003b, p. 10) o objetivo “prover os assentamentos de infra-estrutura social e de produção, de modo a garantir sua sustentabilidade” (p.10) inclui na primeira parte da definição um objetivo intermediário; e na segunda parte, onde se lê “garantir a sustentabilidade” dos assentamentos, um objetivo final. Entretanto, os resultados sugerem maior concentração da ANOP na compreensão e avaliação de objetivos intermediários dos programas. É importante que numa avaliação se considerem todos os três tipos de objetivo, nesse sentido, o enfoque nos objetivos intermediários dos programas possui característica de avaliação de processo (ERS, 1998). Para avaliar um programa de forma integral, é importante considerar todas as fases do programa e verificar se realmente ele possui as condições de avaliabilidade necessárias para fazê-lo (Cotta, 1998; Posavac & Carey, 2003). Ou ainda, planejar uma avaliação que supra tal lacuna.

Concluída a apresentação das categorias, apresentou-se a seguir, um quadro resumo das 19 categorias encontradas na presente pesquisa.

**Quadro 25. Resumo das Categorias da Análise de Conteúdo**

Categoria	Freq. Cat.	Temas	Freq. Temas
1. Avaliação de processo	134	Planejamento e avaliação dos processos	19
		Processo organizativo	38
		Execução do programa	59
		Capacitação da equipe do programa	18
2. Avaliação de necessidades	16	Necessidades do programa	13
		Necessidades do beneficiário	3
3. Avaliação de Impacto/Resultado	136	Impacto direto	38
		Impacto indireto	5
		Indicador de desempenho	93
4. Avaliação da eficiência	54	Custo-benefício	22
		Materialidade	27
		Eficiência	5
5. Boas práticas do programa	77	Experiências bem sucedidas	20
		Programas análogos	5
		Critério de equidade social	27
		Estratégias de participação	25
6. Monitoramento do programa	67	Acompanhamento dos participantes	10
		Acompanhamento das ações	34
		Controle	10
		Pesquisas	13
7. Política do programa	84	Política nacional	21
		Parceiros governamentais	39
		Parceiros não governamentais	24
8. Fontes de dados da auditoria	102	Base de dados	53
		Fontes de dados secundários	15
		Bibliografia especializada	34
9. Respondentes da auditoria	254	Beneficiários	17
		Equipe técnica	100
		Especialistas	100
		Instituições	30
		Parceiros do programa	7
10. Planejamento da auditoria	96	Análise RECI	6
		Análise Stakeholder	12
		Análise SWOT	13
		Avaliação de riscos	9
		Benchmarking	21
		Técnica de indicadores de desempenho	11
		Mapa de produtos e processos	14
		Matriz de planejamento	10
11. Critério de seleção da amostra da auditoria	20	Amostra aleatória	3
		Volume de recursos	5
		Diferenciais de atendimento	5
		Distância geográfica	3
		Quantidade de atendimentos	4
12. Técnicas de coleta de dados da auditoria	102	Questionário	27
		Entrevista	40
		Observação direta	12
		Grupo focal	23

(cont.) Quadro 25. Resumo das Categorias da Análise de Conteúdo

Categoria	Freq. Cat.	Temas	Freq. Temas
13. Procedimento de coleta e tabulação dos dados	54	Pesquisa documental Pesquisa postal Requisição de documentos Tabulação de dados	24 15 10 5
14. Análise dos dados	52	Análise de conteúdo Análise de dados secundários Análise de riscos Análise qualitativa Análise quantitativa Análise orçamentária e financeira	10 4 3 12 17 6
15. Métodos da auditoria	25	Estudo de caso Estudo-piloto Exemplos extremos Série temporal com grupo de controle	16 3 3 3
16. Dificuldades internas da auditoria	40	Mortalidade instrumental Prazo para a auditoria Prazo para cumprimento das recomendações Dificuldades técnicas da auditoria Dificuldades técnicas do programa Grupo de contato da auditoria Acesso geográfico	7 3 5 3 8 7 7
17. Responsabilização	106	Consistência com leis e normas Papel desempenhado Regulamentação interna do programa Modificação de procedimentos internos do programa	55 21 18 12
18. Comunicação Institucional	27	Sistemas de informação Comunicação com parceiros Publicização	5 16 6
19. Objetivos do programa	108	Implantação Intermediários Finais	25 58 25

Legenda: freq. (frequência); Cat. (Categoria)

#### ***4.5 Síntese de Resultados das Categorias da Análise de Conteúdo***

Assim, as 19 categorias resultantes da análise de conteúdo foram sintetizadas em cinco agrupamentos: a) Características dos programas; b) Enfoque da auditoria; c) Contexto Político; d) Métodos e Técnicas e; e) Dificuldades da auditoria. Os agrupamentos seguiram o critério de semelhança entre os temas das categorias e por consequência das categorias. A seguir são listadas as categorias dos agrupamentos, seguindo-se a numeração da análise de conteúdo.

**Quadro 26. Síntese de Categorias: Características dos Programas**

<b>Características dos Programas</b>		
<b>Categorias</b>	<b>Temas</b>	<b>Frequência do agrupamento</b>
5. Boas práticas do Programa	Experiências bem sucedidas Programas análogos Critério de equidade social Estratégias de participação	77
6. Monitoramento do Programa	Acompanhamento dos participantes Acompanhamento das ações Controle Pesquisas	67
18. Comunicação Institucional	Sistemas de informação Comunicação com parceiros Publicização	27
19. Objetivos do Programa	Implantação Intermediários Finais	108
<b>TOTAL</b>		<b>279</b>

As características dos Programas são descritas nas ANOPs e envolvem as Boas Práticas, Monitoramento dos Programa, Comunicação Institucional e os Objetivos do Programa. Com 279 recorrências, indica as especificidades do funcionamento dos Programas Sociais.

Os programas sociais de governo possuem elementos que os destacam, **as boas práticas**, que não abrangem o Programa como um todo; porém, são diferenciais de funcionamento cuja disseminação é estimulada pelas ANOPs e apresentaram 77 recorrências.

O **monitoramento dos programas** é uma lacuna dos programas sociais apontado pelas ANOPs, 67 recorrências, e precisa ser aperfeiçoado e acompanhado efetivamente na prática cotidiana dos programas sociais.

As ANOPs estimulam o aperfeiçoamento da **comunicação institucional** dos programas, dos seus sistemas de informação, comunicação com parceiros e melhor divulgação das ações do Programa.

A descrição dos **objetivos dos programas** apresenta deficiências e nem sempre contempla as fases de implantação, intermediária e final dos Programas. Algumas descrições são imprecisas, dificultando a compreensão do Programa e o planejamento de sua avaliação. A insuficiente descrição de objetivos que se origina dos programas afeta a compreensão e descrição dos objetivos no planejamento das ANOPs.

### Quadro 27. Síntese de Categorias: Enfoque da Auditoria

Enfoque da Auditoria		
Categorias	Temas	Frequência do agrupamento
1. Avaliação de processo	Planejamento e avaliação dos processos Processo organizativo Execução do programa Capacitação da equipe do programa	134
2. Avaliação de necessidade	Necessidades do programa Necessidades do beneficiário	16
3. Avaliação de impacto/resultado	Impacto direto Impacto indireto Indicador de desempenho	136
4. Avaliação da eficiência	Custo-benefício Materialidade Eficiência	54
17. Responsabilização	Consistência com leis e normas Papel desempenhado Comprometimento Regulamentação interna do programa	106
TOTAL		<b>446</b>

O **enfoque** da ANOP é definido a partir do estudo de viabilidade dos Programas e planejamento da auditoria, os enfoques podem ser nos processos do programa, necessidades, resultados, eficiência e na responsabilização dos gestores e executores. O Enfoque da Auditoria foi um dos cinco agrupamentos com maior quantidade de recorrências (446) e indica a importância do enfoque das ANOPs em relação aos programas. Verificou-se que uma ANOP pode possuir mais de um enfoque, sendo que os três enfoques mais recorrentes foram nos resultados (136), processo (134) e responsabilização (106).

ANOPs que enfatizam o **processo** desde o planejamento dos Programas à gestão e qualificação da equipe técnica que executa as ações.

O enfoque menos recorrente (16) foi o de **avaliação de necessidades** que enfatiza as necessidades do Programa e dos beneficiários. Este resultado é coerente com a natureza da ANOP, visto que o seu enfoque principal, conforme definição é o processo, impacto e responsabilização.

As ANOPs enfocam **resultados** dos programas, quer sejam diretos ou indiretos, e a partir deles busca o estabelecimento de indicadores de desempenho. Esse foi o enfoque mais recorrente (136).

ANOPs que enfatizam a **eficiência** se atêm ao custo-benefício, materialidade e eficiência dos Programas, características próprias de auditoria contábil e adequados aos objetivos gerais da organização, seja o controle dos gastos e uso dos recursos públicos.

Uma característica marcante e a terceira mais recorrente (106) foi ANOPs que enfatizam a cobrança da **responsabilização** dos gestores em relação às ações do Programa, o que se observa no cumprimento do seu papel em relação às Leis, normas e regulamentações que sustentam os Programas.

#### Quadro 28. Síntese de Categorias: Contexto Político

Contexto Político		
Categoria	Temas	Frequência do agrupamento
7. Política do Programa	Política nacional	84
	Parceiros governamentais	
	Parceiros não governamentais	
	TOTAL	<b>84</b>

Justifica-se a denominação Síntese de Categorias para apenas uma categoria, no caso, o contexto político, pela sua relevância para a compreensão da ANOP. Além disso, adota-se o pressuposto de Bardin (1977) que uma análise de conteúdo deve considerar, entre outros critérios, a representatividade e pertinência da categoria em relação ao conteúdo analisado.

Os Programas Sociais auditados estão inseridos num contexto político amplo, que é a **política nacional** dos governos, característica relevante com frequência de 84 que incide sobre as ações dos Programas e das ANOPs. A política dos programas envolve as **parcerias** com instituições governamentais e não-governamentais e a ação governamental nas suas diferentes áreas de atuação.

**Quadro 29: Síntese de Categorias: Métodos e Técnicas**

<b>Métodos e Técnicas</b>		
<b>Categorias</b>	<b>Temas</b>	<b>Frequência do agrupamento</b>
8. Fonte de dados da auditoria	Base de dados Fontes de dados secundários Bibliografia especializada	102
9. Respondentes da auditoria	Beneficiários Equipe técnica Especialistas Instituições Parceiros do programa	254
10. Planejamento da auditoria	Análise RECI Análise Stakeholder Análise SWOT Avaliação de riscos Benchmarking Técnica de indicadores de desempenho Mapa de produtos e processos Matriz de planejamento	96
11. Critério de Seleção da Amostra da auditoria	Amostra aleatória Volume de recursos Diferenciais de atendimento Distância geográfica Quantidade de atendimentos	20
12. Técnicas de coleta de dados da auditoria	Questionário Entrevista Observação direta Grupo focal	102
13. Procedimento de coleta e tabulação de dados	Pesquisa documental Pesquisa postal Requisição de documentos Tabulação de dados	54
14. Análise dos dados	Análise de conteúdo Análise de conteúdo Análise de dados secundários Análise de riscos Análise qualitativa Análise quantitativa Análise orçamentária e financeira	52
15. Método da auditoria	Estudo de caso Estudo-piloto Exemplos extremos Série temporal com grupo de controle	25
<b>TOTAL</b>		<b>705</b>

O maior agrupamento de categorias e com maior número de recorrências (705) foi o de Métodos e Técnicas da auditoria. Esse agrupamento abrange fonte de dados, respondentes, planejamento da auditoria, critérios de seleção, técnicas e instrumentos de coleta, procedimentos de coleta, análise de dos dados e o método da auditoria.

Os métodos e técnicas adotados nas ANOPs possuem semelhança aos métodos e técnicas da pesquisa avaliativa. Entretanto, em alguns aspectos poderiam obter maior precisão metodológica no seu uso e aplicação aos Programas, por sua vez, contribuiriam mais na produção de resultados para os gestores dos programas.

As ANOPs utilizam fortemente fontes de dados secundários como: bases de dados dos Programas e outras bibliografias especializadas (total de 102 recorrências). Entretanto, as bases de dados dos Programas apresentam-se em geral desatualizadas e não capazes de proporcionarem informação confiável. Essa limitação dos programas Sociais poderia ser superada nas ANOPs com ênfase maior na coleta de dados primários pela auditoria. Isso demandaria maior tempo e recursos humanos, necessitando de apoio institucional para fazê-lo.

A auditoria busca dados primários, conforme demonstra a categoria de respondentes, com o maior número de recorrências (254) no agrupamento. Participam da auditoria beneficiários, equipe técnica, especialistas, instituições e parceiros dos Programas. Conforme análise de conteúdo, a ANOPs poderia enfatizar mais os beneficiários diretos dos Programas Sociais, que pelo seu tamanho e abrangência demandaria estratégias de apoio institucional à equipe de auditores para realizá-lo.

A fase de planejamento da auditoria demanda tempo considerável no processo de ANOP, essa categoria teve 96 recorrências. O planejamento da auditoria é realizado com base em um conjunto de técnicas utilizadas de forma alternada pelas equipes de auditoria em relação a cada Programa auditado. Nesse sentido, transparece nos relatos o domínio do uso das técnicas pelas equipes de auditoria.

Os critérios de seleção da amostra das auditorias são geralmente intencionais (volume de recursos recebido pelo programa, diferenciais de atendimento, distância geográfica), as ANOPs buscam amostras representativas através de amostras estatísticas.

As técnicas e os instrumentos de coleta de dados utilizados nas ANOPs, com 102 recorrências, foram questionários, entrevistas, observação direta e grupo focal. Verificou-se que não é uma praxe da ANOP a descrição do processo de construção do instrumental de coleta de dados. Porém, a descrição dos mesmos e a explicitação de sua base teórica reforçariam a validade do processo de criação, aplicação e resultados produzidos. Em alguns casos foram utilizados consultores da área do Programa para a criação dos instrumentos, considerando-se as especificidades dos Programas Sociais. Sugere-se que essa prática seja

mantida para todas as ANOPs, para assegurar a produção de dados confiáveis que denotem o fenômeno no qual o Programa intervêm.

Os procedimentos de coleta e tabulação de dados usuais das ANOPs foram: pesquisa documental, postal, requisição de documentos e tabulação de dados, com um total de 54 recorrências nessa categoria.

Com 52 recorrências a análise de dados é um elemento importante para o trabalho da ANOP, destacando-se a capacidade das equipes de auditoria em realizar análises quantitativas. As ANOPs desenvolvem análises qualitativas e quantitativas de dados dos Programas. A equipe de auditoria referiu necessidade de treinamento em análise qualitativa e a importância da participação de consultores nessa área, principalmente em análise de conteúdo.

Os delineamentos das ANOPs foram não experimentais, verificou-se que as auditorias carecem de aperfeiçoamento conceitual no uso e compreensão das terminologias e tipos delineamentos de pesquisa avaliativa, já que se propõe a realizar um trabalho de auditoria semelhante a uma avaliação de programas. Poderiam ser mais bem utilizados o *survey* e estudo de caso, as estratégias para controlar ameaças internas e externas aos resultados produzidos.

### Quadro 30. Síntese de Categorias: Dificuldades para realização da auditoria

Dificuldades para realização das auditorias		
Categorias	Temas	Frequência do agrupamento
16. Dificuldades internas da auditoria	Mortalidade instrumental Prazo para a auditoria Prazo para cumprimento das recomendações Dificuldades técnicas da auditoria Dificuldades técnicas do programa Grupo de contato da auditoria	
TOTAL		40

Justifica-se a denominação Síntese de Categorias para apenas uma categoria, no caso, dificuldades para a realização da auditoria pela sua relevância para a compreensão da ANOP.

As dificuldades para a realização das ANOPs que são internas da organização, com 40 recorrências. As dificuldades estão relacionadas a processos internos de realização da ANOP, como mortalidade instrumental, dificuldades técnicas e o estabelecimento de grupo de contato de auditoria. Aspectos da organização que intervêm no processo de ANOPs são os prazos concedidos para a realização da auditoria e prazo que a auditoria concede aos programas para

cumprir as recomendações. O grupo de contato de auditoria pode ser um elemento facilitador do monitoramento da adoção das recomendações, nesse sentido, precisa de um acompanhamento após auditoria para auxiliar o intercâmbio com a organização auditada.

As restrições de mortalidade instrumental e dificuldades técnicas que afetam as ANOPs, poderiam ser melhoradas mediante capacitação em pesquisa avaliativa, que possui estratégias para diminuição da mortalidade instrumental e oferece referencial teórico e metodológico para a realização de APs.

Os prazos para a realização das auditorias foram em geral exíguos. Nesse sentido, necessitam de apoio institucional para que sejam repensados ou seja, criadas estratégias alternativas para que não afetem negativamente a qualidade do trabalho das ANOPs. Prazos exíguos podem enfraquecer os resultados das auditorias, pois não permitem o cuidado e tempo necessários para a realização dos processos de planejamento, execução e divulgação dos resultados da auditoria.

#### ***4.6 Síntese de Estudos Qualitativos: Critérios do Joint Committee e Síntese de Categorias da Análise de Conteúdo***

As análises qualitativas empregadas nesta tese apresentaram resultados comuns, o que tornou possível a elaboração da Síntese dos Estudos Qualitativos, a partir da qual foram resumidos os resultados da meta-avaliação, considerando-se as pontuações das auditorias conforme critérios do *Joint Committee*, e as frequências da Síntese Categorias da Análise de Conteúdo. Para melhor visualização dos resultados, apresentou-se primeiro o quadro que possibilitou a comparação conceitual e em seguida o quadro comparativo dos resultados das análises. Após os quadros os resultados foram comentados, seguindo-se os cinco agrupamentos temáticos das sínteses das análises: Contexto Político; Características dos Programas; Enfoque da Auditoria; Métodos e Técnicas e Dificuldades para a realização da Auditoria.

Assim, pode-se visualizar na próxima página dois quadros um ao lado do outro, possibilitando as comparações entre as questões orientadoras da meta-avaliação. Foram apresentados na primeira coluna os critérios do *Joint Committee*, demonstrando o que foi avaliado; na segunda coluna, as pontuações de cada critério, na terceira coluna, a Síntese de Categorias da Análise de Conteúdo e na última coluna, a frequência dos Agrupamentos das Categorias.

As pontuações para os critérios foram ponderadas considerando-se:

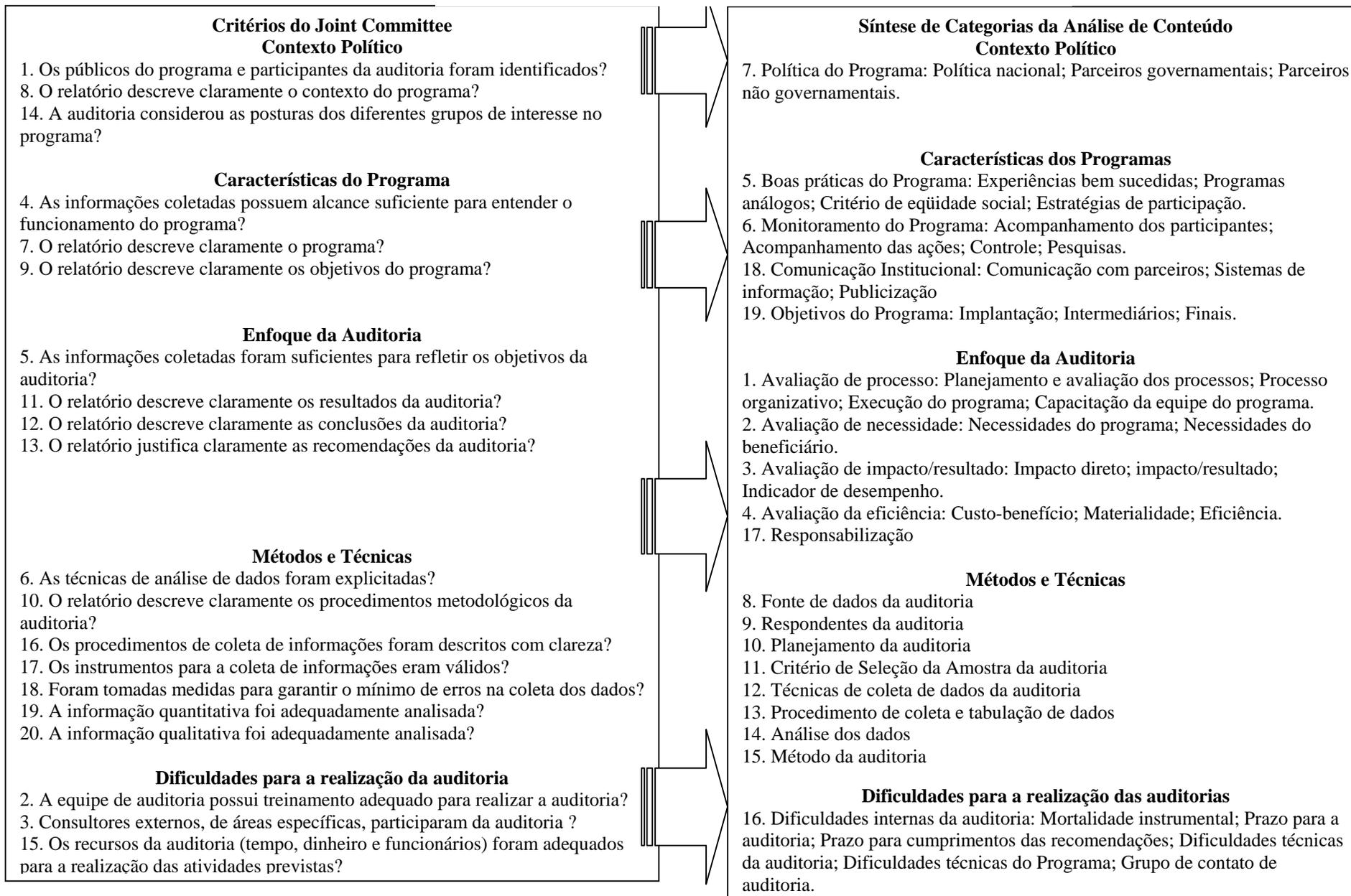
EXCELENTE  $\geq 90 \leq 100$  de pontos atribuídos como SIM (3), ou seja, entre 9 e 10 pontos.

BOM  $\geq 70\% \leq 80\%$  de pontos atribuídos como SIM (3), ou seja, entre 7 e 8 pontos.

PARCIAL  $\geq 50\% \leq 60\%$  de pontos atribuídos como SIM (3), ou seja, entre 5 e 6 pontos.

INSUFICIENTE  $\leq 50\%$  de pontos atribuídos como SIM (3), ou seja, até 4 pontos.

Essa categorização de excelente, bom, parcial e insuficiente foi criada para possibilitar e facilitar a comparação dos resultados. Os pontos de corte foram escolhidos dedutivamente em relação à cognição de pontuações atribuídas em avaliações em geral.



**Figura 9. Síntese de Estudos Qualitativos**

<b>Crítérios do Joint Committee</b>	<b>Avaliação</b>	<b>Síntese de Categorias da Análise de Conteúdo</b>	<b>Freq.</b>
<p align="center"><b>Contexto Político</b></p> 1. Os públicos do programa e participantes da auditoria foram identificados? 8. O relatório descreve claramente o contexto do programa? 14. A auditoria considerou as posturas dos diferentes grupos de interesse no programa?	EXCELENTE BOM INSUFICIENTE	<p align="center"><b>Contexto Político</b></p> 7. Política do Programa	<p align="center"><b>84</b></p> 84
<p align="center"><b>Características do Programa</b></p> 4. As informações coletadas possuem alcance suficiente para entender o funcionamento do programa? 7. O relatório descreve claramente o programa? 9. O relatório descreve claramente os objetivos do programa?	INSUFICIENTE  BOM BOM	<p align="center"><b>Características do Programa</b></p> 5. Boas práticas do Programa  6. Monitoramento do Programa 18. Comunicação Institucional 19. Objetivos do Programa	<p align="center"><b>279</b></p> 77  67 27 108
<p align="center"><b>Enfoque da Auditoria</b></p> 5. As informações coletadas foram suficientes para refletir os objetivos da auditoria? 11. O relatório descreve claramente os resultados da auditoria? 12. O relatório descreve claramente as conclusões da auditoria? 13. O relatório justifica claramente as recomendações da auditoria?	INSUFICIENTE  PARCIAL BOM EXCELENTE	<p align="center"><b>Enfoque da Auditoria</b></p> 1. Avaliação de processo  2. Avaliação de necessidade 3. Avaliação de impacto/resultado 4. Avaliação da eficiência 17. Responsabilização	<p align="center"><b>446</b></p> 134  16 136 54 106
<p align="center"><b>Métodos e Técnicas</b></p> 6. As técnicas de análise de dados foram explicitadas? 10. O relatório descreve claramente os procedimentos metodológicos da auditoria? 16. Os procedimentos de coleta de informações foram descritos com clareza? 17. Os instrumentos para a coleta de informações eram válidos? 18. Foram tomadas medidas para garantir o mínimo de erros na coleta dos dados? 19. A informação quantitativa foi adequadamente analisada? 20. A informação qualitativa foi adequadamente analisada?	INSUFICIENTE INSUFICIENTE  PARCIAL INSUFICIENTE INSUFICIENTE  PARCIAL INSUFICIENTE	<p align="center"><b>Métodos e Técnicas</b></p> 8. Fonte de dados da auditoria 9. Respondentes da auditoria  10. Planejamento da auditoria 11. Critério de Seleção da Amostra da auditoria 12. Técnicas de coleta de dados da auditoria  13. Procedimento de coleta e tabulação de dados 14. Análise dos dados 15. Método da auditoria	<p align="center"><b>705</b></p> 102 254  96 20 102  54 52 25
<p align="center"><b>Dificuldades para a realização da auditoria</b></p> 2. A equipe de auditoria possui treinamento adequado para realizar a auditoria? 3. Consultores externos, de áreas específicas, participaram da auditoria ? 15. Os recursos da auditoria (tempo, dinheiro e funcionários) foram adequados para a realização das atividades previstas?	INSUFICIENTE BOM BOM	<p align="center"><b>Dificuldades para a realização da auditoria</b></p> 16. Dificuldades internas da auditoria	<p align="center"><b>40</b></p> 40

Legenda: freq. (frequência)

### *Contexto Político*

A identificação dos participantes dos Programas nas ANOPs foi excelente (Questão 1 *Joint Committee* – JC). Das dez auditorias meta-avaliadas, apenas uma atingiu parcialmente esse critério. No entanto, foi insuficiente a consideração dos diferentes grupos de interesse nas ANOPs. Tal enfoque pode estar relacionado a fatores associados à **Política do Programa** (categoria 7) e à natureza das auditorias. No Brasil existem 27 Estados e o Distrito Federal, a cada período de quatro anos, os representantes políticos mudam, assim como os partidos políticos. A predominância de partidos de governo ou da oposição pode facilitar ou dificultar o acesso, de forma geral, à pesquisa em relação ao programa auditado, pois o Tribunal, mesmo sendo um órgão independente, é visto como um representante do governo federal. Já a natureza da ANOP recai em enfoque interno ao programa, nos gestores, na responsabilização e no cumprimento de leis e normas do programa. Nesse sentido, dentro do Contexto Político, os dados da presente pesquisa indicam que as informações coletadas pelas ANOPs, poderiam ser expandidas, explorando-se os parceiros não governamentais. Dessa forma permitiria-se transparecer na ANOP a posição desses atores, que são representantes da sociedade civil, sindicatos e outras entidades que possuem interesse ou são atendidos pelos programas sociais.

O Contexto Político influencia a Política do programa em relação aos parceiros. O contexto político afeta as prioridades de governo para determinadas áreas de atuação da Política Nacional através dos Planejamentos Pluri-Anuais e da votação do orçamento para as áreas prioritárias de atuação do governo. Isto, por sua vez, também influenciará o processo de escolha dos programas a serem auditados pelo Tribunal de Contas da União.

O contexto dos Programas (Questão 8) foi descrito adequadamente na maioria das auditorias (sete), descritos parcialmente em uma e inadequadamente descrito em duas auditorias, obtendo avaliação boa nesse quesito. Os contextos dos Programas foram descritos por meio da apresentação do histórico, dados de estudos sobre a temática envolvida no Programa, opiniões de especialistas e principalmente dos gestores. Contudo, conforme destacado na primeira questão, o enfoque interno restringe a descrição ampla do contexto político que envolve o interesse de diferentes grupos nos Programas.

As ANOPs identificaram claramente a existência de diferentes grupos de interesse (Questão 14), o que se observou em duas auditorias; em cinco identificaram parcialmente e em três não o consideraram, o que resultou em uma avaliação insuficiente nesse critério. Esse aspecto revela as diferentes posturas políticas em relação aos Programas do Governo. A

postura favorável ou desfavorável do gestor do programa em nível estadual e municipal pode resultar numa melhor ou pior execução do programa social. Dessa forma, estaria além do controle das ANOPs, visto que o jogo de interesses políticos exposto “foge” à governabilidade e controle do TCU. Essa parcialidade foi constatada na identificação dos atores através de análise *Stakeholder* e da não participação dos mesmos na auditoria do Programa.

Conforme comentado anteriormente o enfoque, devido à natureza da ANOP é interno (gestores e pessoas ligadas diretamente ao Programa). O enfoque interno associado ao contexto político dos programas sociais pode produzir situações de difícil penetração à equipe de auditoria. Principalmente se ocorrerem problemas técnicos internos à equipe, como pouco tempo para a realização da auditoria, carência de pessoal, ou insuficiência de dados cadastrais e ações dos Programas, dessa forma dificultaria o mapeamento preciso dos principais envolvidos com os Programas Sociais, os beneficiários diretos.

Comparando-se a avaliação dos critérios JC com o resultado da frequência da Síntese de Categorias da Análise de Conteúdo, 84 frequências representam 12% da frequência máxima (705) obtida nos agrupamentos. Isto representa a importância da consideração do contexto político para a ANOP. Já a síntese de categorias da análise de conteúdo obteve avaliação excelente e boa em duas questões. Nesse sentido, verifica-se o potencial da ANOP para explorar o contexto político, assim como autonomia e possibilidade de vencer limitações que possam advir desse contexto. Como sugestão de melhoria a apenas um item, que obteve avaliação insuficiente, sugere-se, mediante os motivos expostos anteriormente, que sempre que possível, sejam consultados e considerados a postura dos diferentes grupos de interesse no programa, mesmo que representem dificuldades à execução da ANOP e postura antagônica em relação ao programa. Se a cobertura desse aspecto for possível, pode representar mais um fator positivo que assegura a fidedignidades dos resultados das ANOPs.

### ***Características dos Programas***

As ANOPs buscam nos Programas as experiências bem sucedidas, critérios de equidade social e estratégias de participação, o que pode ser observado na *Categoria 5 – Boas Práticas do Programa*. Das informações coletadas nas ANOPs meta-avaliadas, quatro possuíam alcance suficiente para entender o funcionamento do Programa (Questão 4) e seis apresentaram alcance parcial, o que gerou um desempenho insuficiente nesse quesito. A

coleta de dados com amostras que não sejam representativas do Programa pode estar associada a essa insuficiência. Entretanto, receberam boa avaliação a descrição do programa, contexto e objetivos. Essas informações são colhidas pelas ANOPs a partir de fontes secundárias e entrevistas com pessoas do Programa.

Uma característica comum identificada pelas ANOPs, e são falhas comuns aos programas auditados, é demonstrada pela categoria **18 - Comunicação Institucional**, onde se verificou que os Programas necessitam aprimorar seus sistemas de informação; comunicação com seus parceiros e divulgação de suas ações. Isto significa que as ANOPs identificaram nos programas meta-avaliados sistemas manuais de controle de beneficiários, cadastro em duplicidades ou com ausência de dados dos beneficiários que produzem dificuldades gestão do programa, assim como para o seu acompanhamento, monitoramento e a realização de pesquisas externa como a da própria ANOP do TCU.

Assim, os Programas foram descritos adequadamente (Questão 7) na maioria das auditorias (sete), descritos parcialmente em uma e inadequadamente descritos em duas, o que gerou uma boa avaliação nesse quesito. As ANOPs descrevem com detalhamento a Legislação que sustenta os Programas, assim como a responsabilização dos gestores pelas ações.

Por fim, outras características comuns aos Programas Sociais Meta-avaliados foram algumas limitações em relação à definição e descrição dos objetivos (Questão 9). Embora os objetivos tenham sido descritos adequadamente na maioria das auditorias (sete), descritos parcialmente em duas e inadequadamente descrito em uma auditoria, resultando em boa avaliação nesse critério.

Nessa linha, na análise de conteúdo da *categoria 19 – Objetivos do Programa*, observaram-se imprecisões na descrição dos objetivos dos programas como descrição das fases de implantação, intermediária e final dos Programas. Tais descrições dificultam a compreensão do Programa e o planejamento de sua avaliação e parecem encerrar uma questão de silogismo, na qual os objetivos de auditoria possuem falhas na elaboração, pois foram baseados em objetivos de Programas que também possuem falhas, conseqüentemente, a descrição dos objetivos na ANOP dificulta seu alcance e avaliação.

Outra lacuna apontada pelas ANOPs em relação aos Programas foi demonstrada no exemplos da categoria **6 – Monitoramento do Programa** e são as falhas no acompanhamento

dos participantes e das ações do Programa, pois possuem pouco controle e fazem baixo investimento em pesquisas.

Quanto à Síntese de Categorias da Análise de Conteúdo, o agrupamento Características do Programa recebeu uma frequência de 279, o que representa 39,5% da frequência no maior agrupamento, isso sugere que as características poderiam ser mais exploradas nas ANOPs.

### ***Enfoque da Auditoria***

Na avaliação do critério 5, as informações coletadas pelas ANOPs foram insuficientes para alcançar os objetivos da auditoria, em três foram suficientes e em sete foram parcialmente suficientes. Os objetivos das ANOPs são expressados na forma de perguntas que direcionam o foco e os objetivos que se espera alcançar. Conforme verificado na síntese da análise de conteúdo, o *Enfoque da auditoria* foi demonstrado na *Categoria 1 - Avaliação de processo; Categoria 2 – Avaliação de necessidades; Categoria 3 – Impactos/resultado; categoria 4 - Eficiência e Categoria 17 - Responsabilização*.

Os casos em que as informações não foram suficientes para descrever os resultados do Programa estavam relacionados à busca de informações e objetivos sem instrumental adequado (questionários e entrevistas); como no Programa Sentinela, onde se tentou avaliar melhoria de auto-estima, sem questões válidas a este respeito. Conforme exemplo, não foi possível avaliar o impacto do Programa. Outro fator associado a essa limitação foi a necessidade de adoção de delineamentos que permitam comparações dos resultados do Programa. Isso foi reconhecido pela própria auditoria do Programa Prodetur/NE e Energia das Pequenas Comunidades. A equipe também reconheceu a necessidade de amostras generalizáveis, como no caso do Programa Nacional do Livro Didático e Programa Novo Emprego e Seguro-Desemprego.

Os resultados foram descritos adequadamente (Questão 11) em seis auditorias e parcialmente em quatro, isso produziu uma avaliação parcial nesse quesito. Em geral, as descrições estatísticas dos dados foram adequadas e parece ser uma especialidade das ANOPs, como no caso de descrições de análises orçamentárias, financeiras e apresentações de resultados em gráficos e tabelas. Os resultados de acordo com o Enfoque da Auditoria foram apresentados organizadamente em tópicos, seguindo as questões de auditoria, demonstrando qualidade nesse processo. Além disso, as ANOPs apresentam os resultados positivos dos Programas e não somente os negativos. Nesse sentido, os resultados positivos são

considerados boas práticas, havendo ênfase para a sua disseminação no programa como um todo. Já, as descrições qualitativas dos resultados, em especial resultados de análise de conteúdo de entrevistas, foram identificadas nas auditorias meta-avaliadas, embora não no formato de análise de conteúdo com indicado nos relatórios. Verificaram-se também resultados apresentados de forma excessivamente resumida, como no Programa de Monitoramento e Controle da Mortalidade Materna.

As conclusões (Questão 12) foram descritas claramente em oito auditorias e parcialmente em duas, recebendo uma avaliação boa nesse critério. Nos casos em que as conclusões foram inadequadas, como no Programa Sentinela, verificaram-se imprecisões conceituais na compreensão das premissas do Programa e falha na medição de aspectos das perguntas de auditoria, o que por sua vez gerou conclusões sem base empírica. Já no caso do Prodetur/NE, as conclusões foram dúbias quanto aos impactos ambientais produzidos pelo Programa na fase de implantação, produzindo conclusões confusas. Em relação aos acórdãos e recomendações das ANOPs verificou-se correspondência de 100% das recomendações da equipe de auditoria pelo relator da matéria, que as transforma em Acórdão.

As recomendações (Questão 13) foram apresentadas claramente em todas as auditorias. Apesar de deficiência na descrição de resultados qualitativos e algumas descrições parciais dos mesmos, as argumentações para as recomendações foram construídas com base nos resultados e conclusões, o que por sua vez as sustenta. As recomendações de praxe foram apresentadas, em fase preliminar, aos gestores dos Programas, que as respondem; sendo as suas considerações incluídas na versão do relatório final. Os verbos utilizados nas recomendações e acórdãos estão geralmente no imperativo, por exemplo: dote, promova, recomende, amplie, divulgue, incentive, estabeleça, realize, possibilite, libere, determine e possuem característica de lei a ser cumprida, pois são oriundas de uma corte de contas, conforme mencionado em todos os Acórdãos analisados.

O imperativo das recomendações está associado à categoria **17 – Responsabilização**; na qual foram enfatizados: a consistência com leis e normas; papel desempenhado pelos gestores; comprometimento e regulamentação interna do Programa.

### ***Métodos e Técnicas***

O maior agrupamento de critérios e de categorias, os Métodos e Técnicas da Auditoria, foi o que recebeu uma avaliação mais fraca em relação ao aspecto metodológico dos critérios

do *Joint Committee*. Considerando também que na Síntese de Categorias foi o agrupamento com maior número de frequências (705), isso indica a sua importância para a ANOP e a necessidade de aperfeiçoamento.

As técnicas de análise de dados (Questão 6) foram explicitadas adequadamente em três auditorias; em seis, parcialmente, e em uma não foram explicitadas, levando a uma avaliação insuficiente nesse critério. Nas descrições **dos métodos e técnicas das auditorias**, houve várias imprecisões conceituais em relação aos termos coleta de dados, técnica de análise e tabulação de dados. Verificaram-se casos onde foram citadas análises qualitativas, mas não foram encontradas nos relatos dos resultados das auditorias, como no caso do Programa Novo Mundo Rural. Observou-se que na maioria (nove) das auditorias foram utilizadas técnicas de análise quantitativa e qualitativa.

Os procedimentos metodológicos das ANOPs (Questão 10) foram descritos parcialmente em nove das dez auditorias meta-avaliadas, resultando em uma avaliação insuficiente nesse critério. Os **métodos e técnicas** das auditorias são contemplados, em especial no planejamento da auditoria, conforme *Categoria 10 – Planejamento da auditoria*. Entretanto, os problemas nas descrições dos procedimentos metodológicos estão relacionados a incorreções conceituais dos delineamentos. Por exemplo, a confusão entre estudos comparativos e estudo de caso. Outra lacuna refere-se à ausência de descrição dos procedimentos ou técnicas de análise de dados, como referências vagas à análise qualitativa, sem especificação da mesma. Os delineamentos carecem de rigor metodológico e adoção de procedimentos, como grupos de controle e medição de linha de base dos Programas. Conforme análise de conteúdo, a categoria *15 – Métodos da auditoria* indicou o uso de estudo de caso, estudo-piloto, exemplos extremos e série temporal com grupo e controle. Ou seja, delineamentos não-experimentais com restrições em relação à validade interna e externa. Para aperfeiçoar a metodologia é também necessário descrever em todos as ANOPs a amostra e os procedimentos de coleta de dados também não foram descritos claramente.

Os procedimentos de coleta de informações (Questão 16) foram descritos adequadamente em cinco auditorias, parcialmente em duas e não foram descritos em três, obtendo-se uma avaliação parcial nesse critério. Conforme categoria *12 – Técnicas de coleta de dados da auditoria*, os instrumentos usados foram questionários, entrevistas, observação direta e grupo focal. De acordo com o modelo do relatório de planejamento, não há descrição no corpo do texto dos instrumentos de coleta de dados; entretanto, esses instrumentos

encontram-se nos anexos dos relatórios. A breve descrição dos instrumentos no corpo do texto do relatório de planejamento favoreceria a sua compreensão e principalmente das bases teóricas e conceituais sob as quais são construídos. Tal base não foi descrita em nenhuma auditoria meta-avaliada. Além disso, as descrições realizadas com base em fontes de dados secundários foram encontradas na categoria **8 – Fonte de dados da auditoria** que contemplou: base de dados, fontes de dados secundários e bibliografia especializada. Apesar do uso de diversas fontes de dados, as falhas na coleta de dados (mortalidade instrumental, escolha da amostra e insuficiência de dados do próprio Programa), ocasionaram em alguns casos (programas A2, A6, A8, A9 e A10, ver Tabela 3, p. 148) a não compreensão total do funcionamento do Programa.

Acredita-se que os fatores internos, relacionados ao treinamento adequado das equipes, possam ser aprimorados. Já os fatores internos dos Programas, precisariam ser considerados na fase preliminar, na qual ainda se está planejando a auditoria. Portanto, a análise de riscos realizada na fase de planejamento precisaria aperfeiçoar a análise das condições de avaliabilidade dos Programas, e quando possível traçar estratégias de superação às dificuldades técnicas e de aplicação metodológicas antevistas.

Nas meta-avaliações, verificou-se enfoque na coleta de dados com público interno do Programa. Isso pode também ser verificado na Questão 14, pois, as ANOPs não demonstraram, em geral, a postura de diferentes grupos de interesse no Programa. Esse resultado é corroborado pela análise de conteúdo, na categoria **9 - Respondentes da Auditoria**, que apresenta beneficiários, equipe técnica, especialistas, instituições e parceiros, como os participantes das auditorias, conforme pode ser observado nos Programas: Energia das Pequenas Comunidades; Prodetur NE; Programa Nacional do Livro Didático; Monitoramento da Mortalidade Materna e Novo Emprego – Seguro Desemprego. Nesse último, os beneficiários diretos foram ouvidos, mas a amostra não foi representativa.

Os instrumentos para a coleta de dados foram válidos (Questão 17) em três auditorias; parcialmente válidos em três; não válidos em uma auditoria e não foi possível verificar a validade em três auditorias, uma vez que os instrumentos não constavam nos anexos dos relatórios de planejamento consultados, resultando em avaliação insuficiente nesse critério. Esse resultado indica necessidade de melhoria das técnicas de coleta de dados das auditorias.

Em uma auditoria foi tomada medida para evitar erros (questão 18) na coleta de dados; em quatro foram tomadas medidas que atenderam parcialmente a esse critério e em cinco

auditorias não foram tomadas medidas, resultando em avaliação insuficiente nesse critério. As medidas utilizadas foram a aplicação de teste-piloto para experimentação dos instrumentos e cálculo de amostras probabilísticas para possibilitar generalização de resultados. No entanto; apesar do teste-piloto para avaliação dos instrumentos, não foi relatado se houve ajustes nos mesmos.

Evitar erros nas ANOPs de Programas Sociais poderia ser reforçado através da adoção de grupos de controle, uma vez que há outras variáveis intervenientes que podem estar associadas aos resultados obtidos pelos Programas. Esse resultado, semelhante à Questão 17, indica necessidade de delineamentos que permitam controlar ameaças à validade interna e externa das auditorias.

As informações quantitativas (Questão 19) foram analisadas e relatadas adequadamente em seis auditorias e parcialmente em quatro, esse resultado produziu uma avaliação parcial. Conforme *Categoria 14 – análise dos dados*, as mais citadas foram: análise de conteúdo; de dados secundários; de riscos; qualitativa; quantitativa; orçamentária e financeira. A descrição quantitativa dos dados nos textos poderia ser aperfeiçoada através da adoção de uma linguagem mais precisa na descrição dos resultados, relatando-se a opinião de todos os interessados e calculando-se as diferenças entre as respostas seriam significativas.

As informações qualitativas (Questão 20) foram apresentadas de forma parcial em seis auditorias e de forma inadequada em quatro delas, resultando em avaliação insuficiente nesse critério. As inadequações referem-se principalmente à descrição de resultados de análise de conteúdo de entrevistas e grupos focais, pois nos resultados não há referências claras às mesmas. A auditoria que melhor relatou análises qualitativas foi a do Programa Reinserção Social do Adolescente em Conflito com a Lei. A falta de conhecimento das equipes de auditoria para trabalhar com análise qualitativa (análise de conteúdo) foi citada pelas equipes, como fator de restrição aos trabalhos das ANOPs. Esse resultado indica necessidade de aperfeiçoamento das técnicas de análise de dados nas auditorias.

### ***Dificuldades para a realização da auditoria***

Quanto à análise do treinamento adequado da equipe técnica das auditorias meta-avaliadas (Questão 2), considerou-se as pontuações totais dos 20 critérios do *Joint Committee*. Assim, verificou-se que quatro equipes possuíam treinamento adequado ANOP; duas equipes treinamento parcialmente suficiente e cinco não possuíam treinamento adequado para a

realização da auditoria do Programa, resultando em avaliação insuficiente nesse critério. O treinamento adequado das equipes de auditoria pode ser potencializada ou complementada através da participação de consultores externos, em áreas de domínios específicos dos Programas. Avalia-se que esse seja um procedimento favorável ao êxito das ANOPs que são realizadas em **contextos políticos e sociais** tão diferenciados, pertinentes à atuação dos Programas.

Foi comum a participação de consultores (Questão 3) nas ANOPs. Das dez auditorias meta-avaliadas, sete tiveram a participação de consultores na fase de planejamento e execução, uma na fase de planejamento e em duas auditorias não houve participação, resultando em boa avaliação no critério. Destaca-se que a contratação de consultorias especializadas para compor as equipes auditorias é fator positivo para contornar as dificuldades relacionadas à complexidade da área na qual se insere cada Programa. Além disso, a limitação de tempo para a realização das auditorias foi citada como fator limitador nas auditorias dos Programas Pró-Água-Semi-Árido e Energia das Pequenas Comunidades. Verificou-se que a limitação do tamanho da equipe e pouco tempo para a realização da auditoria afetou o resultado da auditoria do Programa Novo Emprego e Seguro Desemprego.

Os recursos para a realização das ANOPs (Questão 15) foram suficientes em sete auditorias, parcialmente suficientes em duas e insuficientes em uma, com avaliação boa nesse critério.

No último capítulo, serão retomadas as principais discussões dos resultados da meta-avaliação e indicadas as conclusões da pesquisa.

## Capítulo 5. Discussão da meta-avaliação

Foram discutidos os principais resultados deste trabalho; as categorias resultantes da análise de conteúdo; síntese das categorias da análise de conteúdo; a análise dos critérios do *Joint Committee*; seu agrupamento dedutivo e a comparação dos resultados da síntese de categorias com o agrupamento dedutivo. Esse conjunto de resultados de análises formou indutivamente um modelo para a realização de meta-avaliações, conforme apresentado na Figura 3. Ao final do capítulo foram retomadas as perguntas de pesquisa norteadoras do presente trabalho.

### 5.1 Síntese de estudos qualitativos

O ponto de partida da discussão do trabalho parece estar no conhecimento e identificação do processo de realização de ANOP.

Assim, foi produzido através da análise de conteúdo um total de 19 categorias que representam características e limitações desse processo de trabalho e retratam as características dos Programas Sociais, do Contexto Político, do Enfoque das ANOPs, seus Métodos e Técnicas, e sua influência sobre o trabalho de ANOP. As 19 categorias encontradas foram: 1) avaliação de processo; 2) avaliação de necessidades; 3) avaliação de impacto/resultados; 4) avaliação da eficiência; 5) boas práticas do programa; 6) monitoramento do programa; 7) política do programa; 8) fontes de dados da auditoria; 9) respondentes da auditoria; 10) planejamento da auditoria; 11) critérios de seleção da amostra da auditoria; 12) técnicas de coleta de dados da auditoria; 13) procedimentos de coleta e tabulação de dados; 14) análise de dados; 15) métodos da auditoria; 16) dificuldades internas da auditoria; 17) responsabilização; 18) comunicação institucional e, 19) objetivos do programa.

Essas 19 categorias se parecem com uma fotografia instantânea de como ocorreu o processo de ANOP nos dez programas meta-avaliados. Assim, ao se aproximar o foco dessa fotografia nas duas análises realizadas (Síntese de Categorias da Análise de Conteúdo e Critérios do *Joint Committee*), encontraram-se agrupamentos que se destacam pela aproximação de significados. Inovou-se no sentido de comparar os resultados dessas duas análises formando o que foi denominado nesta tese de Síntese de Estudos Qualitativos. A

seguir foram apresentados os comentários e conclusões sobre a Síntese de Estudos Qualitativos.

### **5.1.1 Contexto político**

Um aspecto importante encontrado na presente meta-avaliação refere-se ao Agrupamento de Categoria **Contexto Político**, no qual a auditoria e os programas se inserem. Assim, a Política do Programa é estabelecida diante de uma política maior, a política nacional e através das parcerias do programa com outros agentes e instituições. Em suma verificou-se a existência de conflitos entre o corpo técnico e político, como as ingerências sobre a condução e os resultados dos processos de trabalho desenvolvidos no TCU.

A interferência do contexto político foi identificada na avaliação insuficiente pelas ANOPs das posturas políticas dos diferentes grupos de interesse no programa (Questão 14). Considerando que o enfoque originário da ANOP é interno ao programa, no caso, nos responsáveis diretos, hipotetiza-se que a vinculação partidária do gestor possa ser um elemento facilitador ou dificultador da realização da auditoria, pela visão que o gestor possui do TCU como representante do governo.

Por outro lado, verificou-se claramente que a ANOP identifica o público e grupos de interesse do programa (Questão 1), além disso, descreve claramente o contexto do programa (Questão 8), uma vez que recebeu avaliação excelente e boa respectivamente nesses critérios, mas não usa essa especialidade de identificação (*análise Stakeholder*) e seu resultado como elemento para formar suas amostras dos programas.

Retomando a análise da Categoria Política do Programa, constata-se que os programas sociais são concebidos em uma arena política e, nesse sentido, suas ações atravessam áreas da política e recebem influência das coalizões e configurações de poder predominantes na organização (Mendes et al. 1999). Parece haver um conflito entre as decisões políticas e técnicas.

Entende-se que tanto a auditoria quanto a avaliação de programas precisam contemplar a política na qual os programas se inserem (Posavac & Carey, 2003; Cotta, 1998). Constatou-se, que o foco da auditoria, baseado nos métodos e técnicas específicos das ANOPs incidem sobre os programas avaliados, que possuem características próprias de Programas Sociais e são parte e afetados pelo Contexto Político no qual são criados. Assim, a metodologia da ANOP

precisaria ser aperfeiçoada para abranger de forma mais ampla as características do cenário político relacionado aos programas e que são comentadas no próximo tópico.

### ***5.1.2. Características dos programas sociais auditados***

Encontraram-se as seguintes características nos Programas Sociais: Boas Práticas; Monitoramento do Programa; Comunicação Institucional e, Objetivos dos Programas.

Em relação aos objetivos do programa, no que diz respeito à elaboração das perguntas da auditoria que checam esses objetivos, sugere-se expressar uma pergunta de cada vez para fins de clareza dos relatórios e compreensão dos resultados. O problema na elaboração de questões de auditoria foi verificado em algumas questões do Programa Reinserção Social do Adolescente em Conflito com a Lei. Dessa forma, direciona-se a investigação para aspectos específicos do problema, expressados sob enfoques diferentes em cada questão.

O conjunto de categorias formadoras do Agrupamento Características do Programa possui grande importância e influência sobre o processo de realização das ANOPs. Com 279 frequências, isso indica que as ANOPs valorizam e buscam essas características nos programas auditados.

Esse olhar sobre as características dos programas, por assim dizer, é um olhar direcionador que busca características especiais e de sucesso nos programas sociais, que são boas práticas. A comunicação institucional dos programas é um aspecto investigado pelas ANOPs. Esse aspecto ainda precisa ser aprimorado, pois conforme a Questão 4 dos critérios *Joint Committee* no agrupamento Características dos Programas foi avaliado como insuficiente nas ANOPs meta-avaliadas.

### ***5.1.3 Enfoques das ANOPs***

As ANOPs possuem diferentes **enfoques**, seja avaliação de processo; necessidades; impacto/resultado; eficiência ou responsabilização. O enfoque é escolhido em função das demandas do contexto político, das características do programa e por questões de ordem de funcionamento interno do departamento que realiza a ANOP, como, conhecimentos e habilidades dos auditores.

Os enfoques mais encontrados foram nos resultados, no processo e na responsabilização. Além disso, baseia-se em parâmetros de custo-benefício, efetividade e eficiência (Brasil, 2000c), analisando os programas dentro de uma concepção exata e quantitativa da realidade.

Enquanto que programas sociais têm suas ações pautadas em concepções de política de direitos e política pública (Cotta, 1998). Sendo assim, há certo distanciamento entre a forma de realizar a auditoria e a forma de acompanhar a execução do Programa.

Na análise da categoria Avaliação de Processos da presente meta-avaliação na qual se enfatiza a execução, gestão, planejamento do programa e capacitação da equipe do programa, verificou-se a necessidade de conhecimentos técnicos especializados sobre programas sociais e políticas públicas para a sua realização.

No enfoque de Avaliação de Necessidades, as ANOPs identificam necessidades internas dos Programas e dos Beneficiários. Porém, para fazê-lo necessitam vencer limitações de informações e dados dos programas e, sobretudo, ampliar sua forma de avaliar necessidades dos beneficiários dos programas com delineamentos metodológicos e instrumentais que o permitam, conforme demonstrado pela avaliação insuficiente da Questão 5 sobre o alcance dos objetivos da auditoria e descrição parcial dos resultados e Questões 11 e 14.

Já em relação aos acórdãos e recomendações, problematiza-se até que ponto o TCU pode estabelecer parâmetros de funcionamento (categoria responsabilização, tema: consistência com leis e normas) em nível de detalhamento nos guias e Manuais dos Programas. Isso envolve capacitação em nível nacional para os operadores dos programas; bem como a construção de manuais e guias pelos programas, no nível federal, de acordo com linguagem técnica específica da área, e por esse motivo, de difícil compreensão aos operadores da política no nível local. Por exemplo, algumas diretrizes de atendimento depois de criadas, precisam ser pactuadas em instância federal, estadual e municipal, o que torna o processo lento.

As auditorias possuem aspectos positivos e pontos que necessitam de melhorias na forma de sua condução, para que assegurem a validade dos seus resultados. A elevada frequência demonstra a importância da eficiência (recorrência 155 na análise de conteúdo) e responsabilização (recorrência 106 na análise de conteúdo) no enfoque das ANOPs. Quanto aos critérios do *Joint Committee*, as conclusões (Questão 12) foram boas e as recomendações (Questão 13) excelentes.

Os aspectos positivos apresentados sobre as ANOPs demonstram a evolução de uma prática – a auditoria - de um enfoque mais clássico – o de *accountability* – para o enfoque de avaliação de programas – com uma compreensão maior do funcionamento dos Programas, e

por sua vez, da forma de contribuir com os gestores, produzindo fatores associados à eficiência e continuidade dos Programas (Chelimsky, 1985).

Na mesma linha Aguilar e Ander-Egg (1995, p. 32) comentam que a avaliação de programas adota os procedimentos científicos da pesquisa social, e para realizar APs precisa-se “[...] estabelecer uma estratégia dentro de um processo que tem uma clara direcionalidade e para cuja consecução se estabelecem antecipadamente os cursos de ação mais racionais”. De acordo com essa compreensão, foram apontados neste trabalho, os cursos de ação das ANOPs, e dentre esses, os que fogem do padrão de cientificidade e necessitam de aprimoramento, conforme comentado no agrupamento Métodos e Técnicas, para que a lógica da racionalidade, eficiência e eficácia que justificam esse trabalho sejam mantidos. Dentre os tipos de APs apresentados nesta tese existem aqueles cuja direcionalidade produzirá uma compreensão mais completa e legitimização das ANOPs como procedimento de auditoria dos Programas Sociais. Nesse sentido, as APs são tipos avaliativos mais abrangentes que auditorias, devido as diferentes naturezas das duas modalidades (Chelimsky, 1985). Nesse sentido, verificou-se a potencialidade de ampliação da expertise das ANOPs, que podem manter suas características de auditoria de conformidade e de desempenho e agregar conhecimentos de AP. A seguir foi feita a apresentação das conclusões sobre os métodos e técnicas das ANOPs.

#### 5.1.4 Métodos e *técnicas*

A síntese que comportou o maior número de categorias foi a dos **métodos e técnicas**, sua frequência total foi de 705 recorrências, isso indica a importância desse aspecto para a ANOP e para a meta-avaliação, que se debruçou sobre critérios metodológicos reconhecidos internacionalmente como aqueles esperados em um processo avaliativo. As Categorias sob esse agrupamento referem-se a: fonte de dados; respondentes; planejamento da auditoria; critérios de seleção da amostra; técnicas e procedimentos de coleta de dados; análise de dados e métodos da auditoria.

Foram inúmeras as constatações da necessidade de aprimoramento de conhecimentos e aplicação metodológica de AP no processo de ANOP. Embora o processo de ANOP tenha suas características próprias, conforme demonstrado na sua descrição e de suas técnicas, o planejamento e a execução da ANOP é realizado com base em um conjunto de técnicas combinado de acordo com o Programa a ser auditado. Sendo que as dez ANOPs meta-

avaliadas apresentaram aspectos metodológicos comuns a serem melhorados. Esses aspectos dizem respeito a:

- ênfase na coleta de dados com os responsáveis diretos, conforme encontrado nas categorias Fontes de Dados da Auditoria e Respondentes da Auditoria e necessidade de ampliar suas amostras em relação aos beneficiários diretos e grupos de interesse nos programas sociais;

- necessidade de melhoria na construção, teste e aplicação dos instrumentos de coleta de dados, conforme categorias Planejamento da Auditoria, Técnica de Coleta de Dados e Procedimento de coleta e tabulação de dados e Questões 5 e 17. Em relação aos instrumentos de pesquisa, sugere-se que pelo menos as bases teóricas, conceituais ou práticas sob as quais foram elaborados os instrumentos sejam sucintamente descritas no relatório de planejamento de auditoria, aprimorando a validade e fidedignidade dos resultados (Questão 18).

- treinamento em análises qualitativa dos dados conforme categoria Análise de dados como análise de conteúdo. Em relação aos dados qualitativos obtidos através dos procedimentos de entrevista e grupo focal, devem ser realizadas referências diretas aos conteúdos dos mesmos.

- treinamento em delineamento e aplicação de métodos de pesquisa avaliativa, conforme descrito na categoria Método da Auditoria.

- aperfeiçoar os procedimentos de coleta de dados e sua descrição nos relatórios (Questão 16);

- aprimorar a análise de informação quantitativa (Questão 19) que foi avaliada como parcialmente adequada.

- capacitar a equipe no uso e compreensão das etapas da pesquisa avaliativa, e em conhecimentos sobre delineamentos de pesquisa, visto que nas ANOPs (relatório de planejamento) não se diferencia claramente *estratégia metodológica, a tipologia, métodos de seleção da amostra e método de coleta e dados*, essas distinções se perdem no trabalho realizado pela organização. Além disso, apresentou confusão entre estudo de caso, um delineamento e análise de dados, análise *benchmarking* - métodos de análise de dados e confundiu análise de dados com procedimento de coleta de dados.

- aprimorar a descrição dos relatos dos resultados de análise de conteúdo e análise documental nos relatórios finais, pois não se verificou categorias, temas, recorrências ou qualquer outra referência que fornecesse indicativo de resultado desse tipo de análise.

- reestruturar a matriz de planejamento, pois há redundância entre *fontes de informação e informação requerida*; estas poderiam ser unificadas sob um título comum como *Fontes de dados*, organizando-a de forma a se referirem às etapas de avaliação, descrevendo realmente quais procedimentos serão adotados na ANOP, tipos de análises a serem realizadas, estatísticas descritivas ou outras, e não citar por exemplo, somente análise estatística.

- considerar outros critérios de seleção da amostra da pesquisa, uma vez que, geralmente consideram as boas práticas e os casos deficientes, seguindo-se depois à escolha aleatória. Pois, a escolha dos participantes da pesquisa, entretanto, com base em casos extremos da população pode produzir viés nos resultados. Boas práticas e programas com problemas de execução podem ser utilizados para análises do tipo *Benchmarking*, mas quanto aos seus resultados, provavelmente não representam a situação da maioria das implementações.

Em suma, é necessário continuar o processo de aquisição de conhecimento interno e externo e socialização dos conhecimentos sobre as ANOPs. Segundo Freitas (2005), os treinamentos externos tiveram intensidade contínua entre os anos de 1993 a 2004 e variedade diversa.

Assim, o debate teórico e metodológico no campo da avaliação demonstra que a avaliação de programas sociais é abrangente e comporta o uso de diversas abordagens e procedimentos metodológicos (Patton, 2001; Günther, 2006). Não se pretendeu nesta pesquisa, mediante a análise das auditorias, sugerir a superioridade da AP em relação à auditoria e sim a possibilidade de sua ampliação.

Quanto ao objetivo principal desta pesquisa, utilizaram-se os resultados da meta-avaliação de forma conceitual, para aperfeiçoar o modelo avaliativo do TCU a partir da metodologia avaliativa existente na organização, a ANOP. Por fim, apresentam-se as conclusões em relação às dificuldades para a realização das ANOPs.

### ***5.1.5 Dificuldades para a realização de auditorias***

As dificuldades para a realização das ANOPs perpassam as características dos programas sociais, características das ANOPs e contexto político no qual é realizada a ANOP.

Identificaram-se dificuldades internas do TCU como: mortalidade instrumental; prazo para a auditoria; dificuldades técnicas da auditoria e formação do grupo de contato de auditoria, que também é uma questão de interna aos Programas Sociais. As questões externas que dificultam o processo de ANOP são os prazos para os programas cumprirem as

recomendações, dificuldades técnicas do programa e a criação do grupo de contato de auditoria, ou seja, grupo formado por pessoas do programa para tratar com a equipe do TCU sobre as recomendações da auditoria do Programa. Considerando-se que a “criação do grupo de contato de auditoria” foi recorrente em sete dos dez programas auditados, e provavelmente essa recomendação foi criada com o objetivo de facilitar do contato do TCU com o programa, sugere-se que esse estilo de aproximação do TCU ao programa seja intensificado e priorizado para favorecer o monitoramento e acompanhamento dos programas. Nessa linha sugere-se ainda, que em estudos futuros o monitoramento dos programas auditados pelo TCU sejam meta-avaliados para comparar a eficiência desse procedimento de monitoramento da organização.

Entende-se a importância e a dificuldade de estabelecer esse intercâmbio junto aos programas, considerando a arena política na qual se inserem. Entende-se arena política traçando um paralelo com a concepção de Mintzberg (1983) como uma configuração de poder. Nesse sentido, a política nacional do governo na qual o programa se insere, é permeada de elevada atividade política. Assim, o conflito entre diferentes coalizões de interesses internos e externos ao programa, afetaria o ciclo de vida dos programas sociais.

A dificuldade de implantar grupos e contato de auditoria pode estar associada ao conflito entre uma postura de fiscalização constante, com ênfase nos benefícios que trariam o processo de trabalhar junto com o cliente, ou atuar dentro de um modelo corretivo, de julgamento, caso as recomendações não sejam cumpridas dentro dos prazos estabelecidos. Devido a grande abrangência e os programas avaliados serem de grande porte, existem limitações de tempo para a realização das auditorias e outras como “dificuldade de acesso às famílias” (Brasil, 2003f, p. 33); “dificuldade de acesso às associações via Correio” (Brasil, 2003h, p. 35); “[...] limitação de tempo para a realização dos trabalhos, que é uma característica das auditorias-piloto do Projeto [...]” (Brasil, 2000e, p. 11). O TCU poderia superar essas dificuldades, modificando alguns aspectos do método de intervenção ANOP, para assegurar maior validade aos seus resultados. Constatou-se necessidade de ampliação do prazo de realização das auditorias, considerando-se o escopo e tamanho de cada auditoria, para não prejudicar a qualidade de seus resultados. Sugere-se que a ANOP cubra períodos menores dos programas e depois seja realizado o monitoramento das auditorias, conforme já ocorre no TCU, para que outras interferências externas não sejam confundidas com os impactos avaliados pelas ANOPs em relação aos programas.

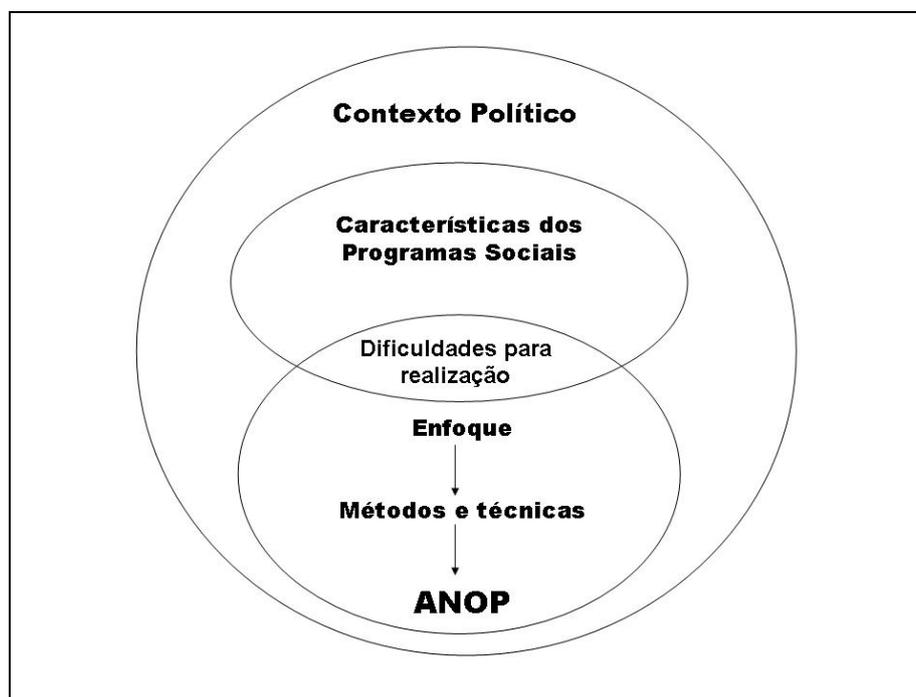
Apesar desse contexto político envolvendo o estabelecimento do grupo de contato, essas dificuldades técnicas, processuais e internas demandam estratégias de enfrentamento às ameaças de validade interna e externa, na realização do processo de ANOP.

Por fim, retratou-se a influência desses agrupamentos, compreendendo-os como variáveis associadas à realização do processo de ANOP. Assim, a variável Contexto Político perpassa todas as outras variáveis encontradas, sejam características dos programas sociais, dificuldades para a realização do trabalho, o enfoque, seus métodos e técnicas, por fim a ANOP resultante.

Por conseguinte, as características dos programas sociais são determinadas pelo contexto político e, por sua vez, influenciam a forma de condução da ANOP, inclusive a escolha do seu enfoque, métodos e técnicas. A relação das características dos programas apresenta uma dicotomia, uma vez que, a partir das características dos programas são feitas escolhas para auditá-lo de acordo com o conjunto de métodos e técnicas das ANOPs, e por outro lado o resultado da ANOP poderá produzir modificações nos Programas. Conforme foi discutido, as variáveis se relacionam de tal forma que o contexto social influencia as características dos programas sociais, o enfoque e a escolha dos métodos e técnicas para a realização da ANOP. Por sua vez, existem dificuldades para a realização das auditorias, que se originam tanto das características dos programas sociais, quanto do enfoque e métodos de realização das ANOPs.

A seguir, foi apresentada a ilustração do processo e relação entre as variáveis encontradas.

**Figura 10. Variáveis associadas ao processo de realização de ANOP**



### ***5.2 Respostas às perguntas de pesquisa***

Retomando as questões de pesquisa da presente meta-avaliação, seguem-se as constatações em relação às perguntas de pesquisa:

A *metodologia avaliativa* do Tribunal de Contas da União - a auditoria de natureza operacional - sustenta o formato dos programas de governo avaliados?

Conforme Síntese de Categorias, o Contexto Político e Características dos Programas envolvem a estrutura dos programas, abrangência, cobertura, quantidade de atendimentos; gestão dos programas nas esferas de governo federal, estadual e municipal. Além disso, os programas possuem em sua definição e concepção de objetivos uma estreita relação com leis e normas e recebem grandes dotações orçamentárias. Em relação a essas características, o arcabouço técnico das ANOPs possibilita avaliar ou auditar programas, entretanto, conforme análise em relação aos métodos, técnicas e procedimentos empregados é necessária maior qualificação técnica e conceitual para aplicar os métodos de AP em ANOPs. Pois a ANOP apresenta limitações na forma com que foi conduzida nos dez programas avaliados, apresentando fortes limitações metodológicas em relação à pesquisa avaliativa, conforme

análise dos critérios do *Joint Committee* e síntese de Categorias, referentes aos Métodos e Técnicas de auditoria.

A *finalidade* da avaliação configura-se como avaliação ou trata-se de auditoria?

Apesar do Tribunal de Contas da União caracterizar sua auditoria como avaliação de programas (Brasil, 2000c), entende-se que as suas atividades caracterizam-se, principalmente, como auditoria, visto que estão ligadas à própria natureza da organização, na modalidade de controle e fiscalização (Brasil, 2003a). Levanta-se a hipótese de que o uso do termo avaliação, em vez de auditoria, esteja associado à questão política da atuação do TCU enquanto tribunal que realiza a fiscalização e o controle dos gastos públicos, sua missão principal e realizada também a modalidade de auditoria de natureza operacional. Nesse sentido, o uso do termo avaliação de programas não anula ou diminui a efetividade do enfoque e a aplicação das técnicas que compõem o instrumental de ação da ANOP. Entretanto, o uso intercambiável desses termos pode alimentar uma ambigüidade na realização desse tipo de auditoria, pois existe diferença entre as duas metodologias. A auditoria é uma modalidade de controle que atende ao objetivo principal da organização que é “(...) julgar as contas dos administradores e demais responsáveis por dinheiros, bens e valores públicos federais, bem como as contas de qualquer pessoa que der causa a perda, extravio ou outra irregularidade de que resulte prejuízo ao erário” (Brasil, 2003a, p. 9).

Parece haver um conflito entre auditoria clássica (*accountability*) e auditoria de natureza operacional. Entretanto, a auditoria de natureza operacional possui características que se assemelham à modalidade de avaliação de programas, por isso a utilidade dos conceitos de AP para as ANOPs, a necessidade de treinamento de habilidades e as treinamento requerido para conduzir o processo, conforme verificado nas auditorias meta-avaliadas.

Em relação às técnicas e modelos das ANOPs descritos: Indicadores de desempenho, mapas de produtos, marco lógico e matriz de verificação de risco são técnicas claramente associadas a processo de auditoria e não de avaliações de programas, no sentido utilizado dentro da área da pesquisa avaliativa. Transparece nas técnicas e modelos a concepção positivista na avaliação dos programas, ligando o conceito de avaliação à idéia de controle e quantificação de resultados (Cotta, 1998).

Organização auditada é a forma com que a organização que recebe a auditoria de natureza operacional é denominada, termo presente em diversos relatórios de avaliação de

programas (Brasil, 2004f; Brasil, 2004e; Brasil, 2004d; Brasil, 2004a) entende-se que o termo auditada estaria relacionado à auditoria e não à avaliação. Além disso, nos textos que descrevem as técnicas utilizadas nas avaliações, o termo que surge é auditoria e não avaliação (Brasil, 2000a; Brasil, 2000b; Brasil, 2001a, Brasil, 2001b; Brasil, 2002a; Brasil, 2002b). Os termos são utilizados de forma intercambiável, como se possuíssem o mesmo significado, mas não o possuem, conforme discutido.

Entende-se que historicamente o processo de trabalho que envolve a auditoria de natureza operacional tem ampliado o seu foco, abrangendo aspectos que são comuns em AP. Porém, essa ampliação de foco encontra-se confusa no âmbito interno da organização, pois ela parece estar buscando realizar avaliações, entretanto realiza, sobretudo auditorias, dando oportunidade à responsabilização. De acordo com Lima (2003), atualmente as auditorias e avaliações realizadas pela organização inserem-se no contexto de performance *accountability*, ou seja, prestação de contas por desempenho e responsabilização por desempenho. Segundo o autor, a contribuição da organização será na responsabilização pública, atingindo os seguintes objetivos: a) garantir para o Poder Legislativo e para a sociedade; b) que as contas sejam fidedignas; c) as receitas e as despesas estejam de acordo com o determinado pela Lei; e d) que os recursos públicos estejam sendo usados com eficácia, eficiência e efetividade. Nesse contexto, entende-se que se trata de uma mudança cultural, auditoria *versus* avaliação de programas, que a organização busca por meio de sua prática avaliativa (Lima, 2003). Tal mudança pode estar ocorrendo por diversos fatores como “[...] busca de legitimidade e da aceitação social, as organizações procuram tornar suas ações e práticas mais próximas dos padrões tidos por corretos socialmente” (Freitas, 2005, p. 33).

Uma outra possibilidade seria a mudança no ambiente, no caso a sociedade, que historicamente busca se apropriar do seu estado de direitos, no sentido de conhecer e exigir esse estado. Verificou-se que as diferenças em termos do treinamento e orientação teórica estudada nos trabalhos intitulados *avaliações ou auditorias de natureza operacional* enfatizam o melhor uso dos recursos públicos, sua otimização, assemelhando-se mais a *auditoria de programa* e não *avaliação de programa*.

***Há regularidade*** na forma de conduzir as avaliações nas diferentes áreas, com um processo avaliativo consistente com a ***abordagem avaliativa*** da organização?

Existe regularidade na forma de conduzir as avaliações, os critérios e práticas são semelhantes (Bastos, 2004). Entretanto, apresentam as mesmas deficiências metodológicas para sustentar os resultados das ANOPs.

As avaliações de programa realizadas por governos, não possuem a mesma finalidade que a pesquisa científica e nem o rigor científico de pesquisas acadêmicas (Posavac & Carey, 2003). Constatou-se no caso do Tribunal de Contas da União que:

- A avaliação pauta-se na concepção de ferramenta para a verificação da racionalização da alocação de recursos orçamentários dos programas inseridos nas políticas públicas. As suas funções, conforme comentado na descrição da organização são: fiscalizadora, consultiva, informativa, judicante, sancionadora, corretiva, normativa e de auditoria; podendo também assumir caráter pedagógico (Brasil, 2003a); e na prática a abordagem avaliativa defronta-se com dificuldades de ordem interna e externa e do contexto da política.
- Houve predominância nos dez casos analisados da condução de um processo semelhante a *avaliação somativa e avaliação de impacto* conduzida para o estudo da eficácia e julgamento de valor dos programas (ERS, 1982). É utilizada para decidir sobre a alocação de recursos do programa.

Essas características foram encontradas nas ANOPs da presente meta-avaliação e são amostras do trabalho realizado pelo TCU no período entre 1999 à 2005. Nessa linha, Freitas (2005) encontrou em período anterior, de 1999 a 1998, casos de ANOPs que só apresentavam essa denominação – ANOP, sendo, entretanto, executadas de forma semelhante à auditoria clássica. Porém, no período de 1993 a 1998, verificou tendência à evolução das características-chaves das ANOPs. Comparando-se os achados da presente meta-avaliação com os de Freitas (2005) parece que as ANOPs ainda continuam evoluindo em suas práticas para um modelo mais próximo à avaliação de programas.

Por fim, o modelo de avaliação de ANOP pode obter incremento de melhoria a partir do conhecimento dos programas avaliados e dos seus objetivos, por isso se faz necessária uma melhor qualificação ou formação em pesquisa avaliativa de programas e políticas públicas. Assim, a partir desses objetivos é importante ter claro quais são as possibilidades de delineamentos a adotar para a auditoria, mediante a readaptação de procedimentos e técnicas para as ANOPs.

Quanto ao planejamento da avaliação, que é uma das etapas mais importantes de uma avaliação de programas e requer disponibilidade de tempo, para planejar e executar a ANOP, essa é uma limitação a ser vencida nas ANOPs, conforme análise da Questão 15 e encontrada nos programas Pró-Água/Semi-Árido; Energia das pequenas Comunidades e programa Morar Melhor.

Em relação ao aspecto metodológico, que envolve a escolha do delineamento, métodos e técnicas ou método precisa-se considerar fatores relativos ao programa conforme agrupamento Características do Programa que são boas práticas, monitoramento, comunicação institucional e, sobretudo objetivo dos programas.

Assim, escolher não somente com base no viés ou conhecimento do pesquisador, recaindo sempre sobre a escolha do mesmo método ou técnica independente da configuração do programa. Isso pode ser verificado no alcance parcial dos objetivos da ANOP sobre o programa auditado (Questão 5) em sete ANOPs e alcance total ou excelente somente em três (ver Tabela 3) nos Programas Novo Mundo Rural, Consolidação de Assentamento; Reinserção Social do Adolescente em Conflito com a Lei e Programa Morar Melhor.

Sugere-se que os conhecimentos dos auditores sejam reforçados mediante treinamento em métodos (tipos de delineamento de pesquisa) e as diversas técnicas (*survey*, análise de dados qualitativos e quantitativos), para depois decidir como aplicá-las em contextos diferentes.

A contribuição ao modelo avaliativo da ANOP proposto nesta tese considerou os pressupostos de Cook e Shadish, (1986); Krueger, Casey, Donner, Kirsch e Maack (2001), bem como a tipologia de objetivos de Posavac e Carey (2003). Conforme Cook e Shadish (1986), o planejamento da avaliação com base nos objetivos do programa não é nova, pois foi utilizada desde o início dos processos de avaliação. Para tentar suprir a fraqueza da forma com que algumas vezes são definidos os objetivos do programa agregou-se a contribuição de Krueger, Casey, Donner, Kirsch e Maack (2001). Assim, propõe-se a consideração das seguintes etapas ao modelo de ANOP do TCU.

Sugestões para a ANOP:

1. Realizar reuniões para levantar informações que possam fornecer entendimento sobre os grupos minoritários, a partir dos *stakeholders*, e identificar como o trabalho pode influenciar o contexto social;
  - 1.1 Identificar e compreender profundamente o programa, mediante entrevistas e reuniões com os *stakeholders* e coleta de dados secundários;
  - 1.2 Examinar as expectativas dos diversos atores envolvidos sobre o que os objetivos que o programa irá alcançar e sobre como esses fatores conduzirão ao sucesso ou fracasso;
2. Verificar e identificar modelos, a partir de fonte de dados secundários, para entender a lógica subjacente que os resultados do programa produzem.
  - 2.1 Caracterizar as necessidades de informação. Isso envolve levantar quais recursos, sejam humanos, materiais e informacionais estão disponíveis para dar suporte à ANOP.
  - 2.2. Avaliar qualquer modelo de sistema que tenha sido desenvolvido para explicar como o programa produz *inputs* relacionados aos processos de resultados.
  - 2.3 Planejar o delineamento da avaliação, a escolha da técnica de coleta e análise de dados, considerando: o estágio do programa e suas condições de avaliabilidade, o orçamento, o tempo disponível para a avaliação, a amplitude e a profundidade da análise, as vantagens e desvantagens dessas escolhas.
3. Escolher os critérios para avaliação do programa, considerar:
  - a. Se o programa incorpora valores da Política Pública na qual se insere;
  - b. Se o programa incorpora as necessidades das pessoas atendidas;
  - c. O grau com que foram satisfeitas as necessidades materiais dos clientes;
  - d. O grau de atendimento dos objetivos planejados;
  - e. Se os resultados obtidos estão de acordo com as metas e especificam os níveis dos resultados esperados;
  - f. Se os recursos estão sendo gastos apropriadamente pelo programa; observa-se que essa etapa já é perfeitamente realizada pelas ANOPs;

g. Se as recomendações afetarão a tomada de decisões dos criadores de política ou gerentes no programa. Isso poderia ser alcançado com maior publicidade das ANOPs junto aos beneficiários e sociedade em geral

4. Compreender e classificar os objetivos do programa em: a) objetivos de implantação; b) intermediários e, c) finais (Posavac & Carey, 2003) para verificar se são possíveis de avaliar.

Por fim, sugere-se que a aplicabilidade dessas questões sejam testadas nas ANOPs para verificar a sua incorporação ou não ao modelo do TCU.

### ***5.3 Limitações da presente meta-avaliação***

Considerando que a avaliação de programas e a ANOP são práticas que atravessam diferentes áreas de atuação profissional e de conhecimento científicos, reconhece-se a necessidade de maior intercâmbio com disciplinas de políticas sociais e políticas públicas para ampliar e apoiar o foco das análises dos programas sociais da meta-avaliação.

Entende-se que os resultados obtidos da meta-avaliação são aplicáveis à amostra em questão, limitando-se à generalização, dentro da possibilidade do que uma análise qualitativa pode oferecer.

A criação de categorias de forma indutiva, dentro na análise de conteúdo, poderia ter sido favorecida pela utilização do software Alceste devido à grande quantidade de material escrito analisado.

Poderia-se agregar fontes de documentos secundárias das organizações responsáveis pelos programas para facilitar a compreensão dos objetivos dos programas e de seu funcionamento.

Em estudos futuros, sugere-se que esse tipo de agrupamento de questões dedutivo do *Joint Committee* também seja submetido à análise de juízes, assim como as pontuações para cada item. Entretanto, tal procedimento não foi realizado no presente trabalho devido a restrições de tempo para a conclusão do trabalho e a grande quantidade de páginas dos documentos a serem analisadas pelos juízes para opinarem sobre a pertinência dos critérios para cada auditoria de programa, pois demandaria tempo maior para a leitura e análise de 1.206 páginas de documentos.

#### *5.4 Agenda de pesquisa*

A partir das categorias levantadas na presente pesquisa, sugere-se a construção e validação de instrumento a ser aplicado na organização para complementação da apreensão do seu processo de trabalho, visto que a aplicação do questionário poderia captar a percepção dos pontos de vistas da equipe que executa a auditoria.

Comparar os objetivos do programa a partir de documentos procedentes diretamente dos programas, com os objetivos descritos nos documentos produzidos pelas ANOPs, analisados nesta meta-avaliação.

Realizar uma meta-avaliação comparando uma ANOP de um programa social com o monitoramento da ANOP sobre o mesmo programa.

Em estudo futuros o monitoramento dos programas auditados pelo TCU poderiam ser meta-avaliados comparando-se a eficiência do procedimento de monitoramento do TCU.

Realizar a comparação entre a ANOP de um programa e uma AP do mesmo programa, baseando-se na classificação dos objetivos do programa e, assim, comparar a abrangência dos resultados obtidos a partir dos dois diferentes métodos.

Finalmente, poderia ser realizada também uma análise quantitativa, como uma meta-análise das ANOPs objeto de estudo desta tese.

## Apêndice

**Quadro 31. Classificação dos Objetivos dos Programas Auditados**

<b>1. Programa de Desenvolvimento Sustentável de Recursos Hídricos para o Semi-árido Brasileiro – Pró-Água/Semi-Árido (Brasil, 2000d)</b>
<b>Objetivos de Implantação</b>
Desenvolver estudos com vistas a identificar soluções eficazes para adequar as disponibilidades e as demandas de água nas diferentes bacias hidrográficas contempladas pelo Programa (p.4)
Elaborar Planos Diretores de bacias hidrográficas, considerando uma visão integrada e de múltiplos usos da água, além do reconhecimento do valor econômico deste recurso (p.4)
Detalhar em nível de projeto, as intervenções prioritárias selecionadas tanto nos estudos desenvolvidos quanto nos Planos Diretores já elaborados (p.4)
Estimular à construção e instalação de reservatórios e sistemas de distribuição que otimizem a capacidade de reserva e abastecimento de áreas rurais com alta concentração de famílias de baixa renda (p.17)
Prover acesso confiável e sustentável à água para uso doméstico, municipal e outros (p.4)
<b>Objetivos Intermediários</b>
Fortalecer o setor de recursos hídricos na região semi-árida brasileira, mediante a capacitação das instituições públicas e das entidades civis que serão envolvidas na implantação do Programa (p.4)
Implementar novas obras que, além de garantir o aumento da disponibilidade de água e de privilegiar soluções locais para operação e manutenção [...] (p.4)
Garantir a ampliação da oferta de água de boa qualidade para o semi-árido brasileiro com promoção do uso racional desse recurso, de tal modo que sua escassez relativa não continue a constituir impedimento ao desenvolvimento sustentável da região (p.17)
Estabelecer, de forma sustentável, de sistemas de administração, operação e manutenção da infra-estrutura de abastecimento de água, incluindo mecanismos seguros e eficientes de recuperação de custos (p.17)
<b>Objetivos finais</b>
Promover o uso racional e sustentável e o gerenciamento participativo dos recursos hídricos (p.4)
Concluir obras hidráulicas, já iniciadas, e que apresentem retorno sócio-econômico inquestionável (p.4) obras [...] que promovam tanto a melhoria da qualidade de vida da população local quanto a proteção do meio ambiente (p. 4)
Prover o acesso confiável e sustentável à água para consumo doméstico em âmbito municipal, com prioridade para as bacias hidrográficas do Nordeste (p. 17)
Promover o uso racional e sustentável dos recursos hídricos, com ênfase na gerência participativa, particularmente na região do semi-árido brasileiro (p.17)
<b>2. Programa Controle da Mortalidade Materna (Brasil, 2001d)</b>
<b>Objetivos de implantação</b>
Elaborar e distribuir, aos Comitês de Mortalidade Materna e Secretarias Estaduais da Saúde, do "Manual dos Comitês de Mortalidade Materna" (p.15)
Oferecer subsídios para o aperfeiçoamento da política nacional de saúde da mulher, estabelecendo correlações com os comitês estaduais, regionais e municipais (p.15)
Manter atualizado amplo diagnóstico da situação atual da mortalidade materna no Brasil, enfocando todos os seus múltiplos aspectos sociais, econômicos, políticos, jurídicos e outros que facultem ações específicas na sua solução (p.15)
Investigar a respeito da natureza do óbito, com triagem dos óbitos declaradamente maternos, dos não-maternos e dos presumíveis, preenchimento da ficha de investigação; busca ativa, investigando todos os óbitos de mulheres em idade fértil; e pesquisa sobre as condições em que ocorreu o óbito (p. 17)
Participar em reuniões com instituições governamentais e não governamentais para definição de estratégias de ação conjunta (p. 15)
Elaborar impressão e distribuição, às Secretarias Estaduais de Saúde e Comitês Estaduais de Mortalidade Materna, do "Plano de Ação para Redução da Morte Materna"; (p. 15)

## (cont.) Quadro 26. Classificação dos Objetivos dos Programas Auditados

<b>2. Programa Controle da Mortalidade Materna (Brasil, 2001d)</b>
<b>Objetivos de implantação</b>
Elaborar e distribuição, às Secretarias Estaduais de Saúde e Comitês Estaduais de Mortalidade Materna, da ficha de investigação da morte materna (p. 15)
Confeccionar e distribuir cartazes e <i>folders</i> sobre temas relacionados à morte materna; campanhas educativas; assessoria técnica
Propor normas de funcionamento dos serviços, visando à melhoria da qualidade da assistência; (p.17)
<b>Objetivos intermediários</b>
Mobilizar os diversos setores da sociedade afetos à questão, com a finalidade de melhorar a saúde da mulher (p.15)
Monitorar mortalidade materna. (p. 17)
Apoiar o técnico-financeiro às Secretarias Estaduais de Saúde e/ou aos Comitês Estaduais de Mortalidade Materna para criação de comitês regionais ou municipais de morte materna (p.15)
Acompanhar as ações do Ministério da Saúde no processo de articulação e integração das diferentes instituições e instâncias envolvidas na questão (p.15)
Discutir casos clínicos nos comitês hospitalares; casuística (acumulação de evidências epidemiológicas) e promoção de discussões sobre o problema da mortalidade materna (p. 17)
Corrigir estatísticas oficiais e divulgação, em informativos médicos ou outros meios, de relatórios para todas as instituições e órgãos competentes que possam intervir na redução da mortalidade materna (p. 17)
Avaliar a evitabilidade ou não da morte, sobre a responsabilidade pelo óbito e sobre as medidas de prevenção cabíveis (p. 17)
Integrar todas as instituições pertencentes a qualquer dos poderes públicos ou setores organizados da sociedade, com a finalidade de melhorar a saúde da mulher (p. 17)
Apoiar as Secretarias Estaduais de Saúde, Comitês Estaduais de Mortalidade Materna e sociedades científicas para a realização de eventos alusivos à mortalidade materna (p. 15)
<b>3. Programa Energia das Pequenas Comunidades (Brasil, 2003e)</b>
<b>Objetivos de Implantação</b>
Viabilizar a instalação de microssistemas energéticos de produção e uso locais em comunidades carentes isoladas não servidas por rede elétrica, destinados a apoiar o atendimento das demandas sociais básicas; (p.3)
Adotar uma estrutura mais eficaz de gestão que possibilite a transferência das atividades do Programa, em Brasília, para os agentes descentralizados do mercado e do setor privado (p.8)
Elaborar [...] um Projeto de Trabalho Técnico Social (p.22)
Desenvolver estudos necessários para medir o potencial do mercado das microrregiões e para definir o tipo de produção de energia mais adaptável às características locais e às necessidades dos consumidores (p.8)
Implementar [...] um Projeto de Trabalho Técnico Social (p.22)
Acompanhar o desenvolvimento do mercado-alvo e propor sugestões, às autoridades competentes, de política, legislação e regulamentação relevantes para o mercado; (p.8)
Expandir a capacidade de informação e disseminação do Programa e dos agentes do mercado, pela mídia; (p.8)
Introduzir as melhores tecnologias (p.8)
Promover a capacitação de recursos humanos e o desenvolvimento da tecnologia e da indústria nacionais, imprescindíveis à implantação e à continuidade operacional dos sistemas a serem implantados (p.3)
Demonstrar a viabilidade dos modelos de negócios que poderão servir como padrão para o desenvolvimento do mercado da energia renovável, otimizando a participação da sociedade civil, setor privado e das organizações não-governamentais; (p.8)
Promover programas de treinamento e capacitação para todos os tipos de participantes no mercado, inclusive agentes governamentais, em todo Brasil; (p.8)
<b>Objetivos intermediários</b>
Desenvolver mecanismos financeiros inovadores e/ou adaptados, destinados aos consumidores finais de energia renovável, bem como às empresas e provedores de serviços de energia renovável no meio rural (p.8)
Transferir, gradativamente, para entidades locais, as responsabilidades para a identificação, desenvolvimento, implantação, operação e manutenção dos projetos de energia renovável nas comunidades carentes; (p.8)

## (cont.) Quadro 26. Classificação dos Objetivos dos Programas Auditados

<b>3. Programa Energia das Pequenas Comunidades (Brasil, 2003e)</b>
<b>Objetivos finais</b>
Complementar a oferta de energia dos sistemas convencionais com a utilização de fontes de energia renováveis descentralizadas (p.3)
Promover o aproveitamento das fontes de energia descentralizadas no suprimento de energéticos aos pequenos produtores, aos núcleos de colonização e às populações (p.3)
<b>4. Programa Morar Melhor (Brasil, 2004e)</b>
<b>Objetivos de Implantação</b>
Prever diretrizes específicas, a título de orientação, para a elaboração dos projetos de intervenção pelos agentes executores (p.22)
<b>Objetivos intermediários</b>
Integrar outras intervenções ou programas da União, em particular com aqueles geridos pelo Ministério Extraordinário de Segurança Alimentar e Combate à Fome, ou de demais esferas de governo. (p.21)
Atender a população urbana e rural e prioridade às famílias de mais baixa renda e à mulher chefe de família (p.21)
Atender aos idosos ou portadores de deficiências físicas, previamente identificadas, pela adoção de projetos ou soluções técnicas que eliminem barreiras arquitetônicas ou urbanísticas [...] (p. 22)
Funcionar plenamente as obras e serviços propostos que devem reverter-se, ao seu final, em benefício imediato à população (p.22)
<b>Objetivos finais</b>
Melhorar as condições de habitabilidade das famílias de baixa renda, assim consideradas aquelas com rendimento mensal de até três salários mínimos atendendo, preferencialmente, às populações que habitam em áreas de risco. (p.21)
Promover a melhoria da qualidade de vida das famílias beneficiadas, agregando-se às obras e serviços a execução de trabalho social (p.21)
Executar equipamentos comunitários públicos voltados a esse segmento da população (p.22)
Beneficiar famílias localizadas em áreas sujeitas a fatores de risco, insalubridade ou degradação ambiental. (p.21)
<b>5. Programa Nacional do Livro Didático (Brasil, 1999a)</b>
<b>Objetivos intermediários</b>
Distribuir, gratuitamente, livros didáticos aos alunos do ensino público fundamental no tempo certo, na quantidade certa e com a especificação solicitada (p.5)
<b>Objetivos finais</b>
Contribuir para a universalização do ensino e para a melhoria de sua qualidade por meio da aquisição e distribuição de livros didáticos para todos os alunos matriculados nas escolas públicas do ensino fundamental de 1ª a 8ª séries. (p.5)
<b>6. Programa Novo Mundo Rural: Assentamento de Trabalhadores Rurais e Novo Mundo Rural: Consolidação de Assentamentos (Brasil, 2003h)</b>
<b>Objetivos de implantação</b>
Desenvolver Plano de Desenvolvimento do Assentamento Rural; (p.10)
Acompanhar a instalação de projetos de assentamentos rurais; (p.10)
<b>Objetivos intermediários</b>
Conceder crédito para aquisição de imóveis rurais – Banco da Terra (p.10)
Alimentar o Sistema de Informações de Projetos de Reforma Agrária –SIPRA; (p.10)
Discriminar Áreas Devolutas e Terras Irregularmente Ocupadas (p.10)
Prover Assistência Jurídica às Famílias Acampadas (p.10)
Prover Assistência Social às Famílias Acampadas (p.10)
Conceder crédito para implantação de infra-estrutura básica –Banco da Terra; (p.10)

## (cont.) Quadro 26. Classificação dos Objetivos dos Programas Auditados

Implantar infra-estrutura básica em assentamentos rurais [...] (p. 10)
<b>6. Programa Novo Mundo Rural: Assentamento de Trabalhadores Rurais e Novo Mundo Rural: Consolidação de Assentamentos</b> (Brasil, 2003h)
Investir em Infra-estrutura básica para assentamentos rurais; (p.10)
Prover os assentamentos de infra-estrutura social e de produção [...] (p. 10)
Identificar Imóveis Passíveis de Desapropriação e Aquisição; (p.10)
Realizar perícia Judicial em Ações de Desapropriação de Imóveis; (p.10)
Realizar Topografia em Áreas de Assentamento Rural; (p.10)
Prestar Assistência técnica e capacitação de assentados. (p.10)
<b>7. Programa de Desenvolvimento do Turismo no Nordeste – PRODETUR NE</b> (Brasil,2003b)
<b>Objetivos de implantação</b>
Realizar planejamento estratégico e preparação de projetos [...] (p. 2)
<b>Objetivos intermediários</b>
Conceder de crédito para implantação de infra-estrutura básica –Banco da Terra (p.10)
Realizar Planejamento estratégico: treinamento e infra-estrutura para o crescimento turístico [...] campanhas de conscientização, treinamento profissional e capacitação da população local, água potável e saneamento, e obras de infra-estrutura (p.2)
Prover de infra-estrutura básica e serviços públicos a áreas atualmente de expansão turística e onde a capacidade do Estado não acompanhou a demanda por tais serviços (p.2)
Promover investimentos do setor privado –financiamento de seminários e cursos de treinamento para pequenos e médios proprietários de agências de turismo e membros de organizações não governamentais do setor e, ainda, de serviços de consultoria para captação de investimentos privados (p.8)
<b>Objetivos finais</b>
Reforçar a capacidade da Região Nordeste em manter e expandir sua crescente indústria turística, contribuindo para o desenvolvimento sócio-econômico regional. (p.2)
Fortalecer a capacidade municipal para a gestão do turismo: gestão administrativa e fiscal dos municípios, gestão de resíduos sólidos, proteção e conservação de recursos naturais e do patrimônio cultural, e urbanização de áreas turísticas; (p.7)
<b>8. Programa Reinscrição Social do Adolescente em Conflito com a Lei</b> (Brasil, 2003d)
<b>Objetivos de Implantação</b>
Implantar serviços sociopsicopedagógicos destinados aos adolescentes em conflito com a lei e sua família (p.22)
Realizar estudos e pesquisas na área de reinscrição social do adolescente em conflito com a lei (p. 22)
<b>Objetivos intermediários</b>
Aplicação as seguintes medidas socioeducativas, estabelecidas nos artigos 112 e 115 a 123 do Estatuto da Criança e do Adolescente:
a) advertência – admoestação verbal; (p. 21)
b) internação – medida privativa de liberdade, sujeita aos princípios de brevidade, excepcionalidade e respeito à condição peculiar de pessoa em desenvolvimento [...] obrigatórias atividades pedagógicas. (p. 21)
c) prestação de serviços à comunidade – realização de tarefas gratuitas de interesse geral; (p. 21)
d) liberdade assistida – acompanhamento do adolescente, com a realização de: promoção social dele e de sua família; inserção em programa de auxílio e assistência social, se for o caso; supervisão da frequência e do aproveitamento escolar; diligências no sentido da profissionalização e da inserção no mercado de trabalho (p. 21)
semiliberdade – possibilita a realização de atividades externas e são obrigatórias a escolarização e a profissionalização (p. 21)
e) obrigação de reparar o dano -restituição, ressarcimento do dano ou compensação do prejuízo; (p. 21)
Realizar a produção de material esportivo por menores infratores Cidadania ao Menor. (p.22)
Prestar atendimento socioeducativo ao adolescente em conflito com a lei. (p. 22)
Promover eventos relativos à reinscrição social do adolescente em conflito com a lei. (p. 22)

## (cont.) Quadro 26. Classificação dos Objetivos dos Programas Auditados

Apoiar técnico e financeiro aos projetos de parceiros governamentais e não-governamentais comporta a realização de convênios e outras formas de cooperação. A ação de apoio ao atendimento de medidas não-privativas de liberdade constitui a prioridade fundamental
<b>8. Programa Reinserção Social do Adolescente em Conflito com a Lei</b> (Brasil, 2003d)
<b>Objetivos finais</b>
Articular e estimular os esforços do sistema socioeducativo instituído pelo Estatuto da Criança e do Adolescente de forma a possibilitar a inclusão do adolescente em conflito com a lei no meio social (p. 21)
Fortalecer os mecanismos de aplicação e execução das medidas socioeducativas, por meio de um conjunto de ações, incentivando experiências referenciais em cada unidade da federação
<b>9. Programa Novo Emprego Seguro-Desemprego</b> (Brasil, 2002f)
<b>Objetivos Intermediários</b>
Auxiliar a manutenção e busca de emprego, promovendo, para tanto, ações integradas de orientação, recolocação e qualificação profissional (p. 2)
<b>Objetivos Finais</b>
Prover assistência temporária ao trabalhador desempregado em virtude de dispensa sem justa causa (p.2)
<b>10. Programa Enfrentamento a Exploração e ao Abuso Sexual de Crianças e Adolescentes - Sentinela</b> (Brasil, 2004d)
<b>Objetivos de implantação</b>
Articular a implantação das políticas públicas ligadas ao combate da violência sexual infanto-juvenil (p. 22)
Deliberar sobre o desenho e a implantação do Programa Sentinela; [...] (p,22)
<b>Objetivos intermediários</b>
Garantir às vítimas e suas famílias o acesso a serviços de assistência social, saúde, educação, justiça, segurança, esporte, lazer e cultura (p. 22)
Coordenar a operacionalização, principalmente no que diz respeito às funções de monitoramento e avaliação; e atuar na coordenação política e na articulação com estados e municípios (p.22)
<b>Objetivos finais</b>
Atender, por meio de um conjunto de ações articuladas, de crianças e adolescentes vitimados pela violência com ênfase no abuso e na exploração sexual. (p.22)
Prevenir e combater a violência, o abuso e a exploração sexual comercial de crianças e adolescentes (p.22)

## Referências bibliográficas

- Aguilar, M. J., & Ander-Egg, E. (1995). *Avaliação de serviços e programas sociais*. Petrópolis: Vozes.
- Ala-Harja, M., & Sigurdur, H. (2000 Outubro-Dezembro). *Em direção às melhores práticas de avaliação*. In Revista do Serviço Público. Fundação Escola Nacional de Administração Pública, v.1, nº1. Ano 51. Brasília: ENAP.
- Ashworth, K. C. A., & Greenberg, D. W, Robert (2004). *Meta-evaluation: Discovering what works best in welfare provision*. Evaluation. [On-line] Vol. 10. Disponível: <http://evi.sagepub.com/cgi/content/abstract>.
- Bardin, L. (1977). *Análise de Conteúdo*. Lisboa: Edições 70
- Barreira, M. C. R. N. (2002). *Avaliação Participativa de Programas Sociais*. São Paulo: Veras Editora CPIHTS.
- Bastos, G. M. M. C. (2004). *Ciclo de trabalho de ANOP*. Tribunal de Contas da União. Secretaria de Fiscalização e Avaliação de Programas de Governo SEPROG. Material interno não publicado.
- Braga Lins, C. (2005). *Programa de erradicação do trabalho infantil: Desafios e potencialidades*. Ministérios do Desenvolvimento Social. Secretaria Nacional de Assistência Social. Documento Interno, não publicado. 22 p.
- Brasil (1992). Boletim do Tribunal de Contas da União. [On-line], v. 25, n ° 37. Brasília: TCU. Disponível:[http://www.tcu\\_sarq\\_prod/Sistemas/Btcu](http://www.tcu_sarq_prod/Sistemas/Btcu) (Sede) - [http://tcu\\_sarq\\_an/Sistemas/Btcu/tcu\\_tcu\\_UF/Sistemas/Btcu](http://tcu_sarq_an/Sistemas/Btcu/tcu_tcu_UF/Sistemas/Btcu) (Estaduais)
- Brasil (1998). Tribunal de Contas da União. *Técnicas de auditoria: Técnica de entrevista para auditorias*. Brasília: TCU. Coordenadoria de Fiscalização e Controle. 32p.
- Brasil (1999a). *Relatório de auditoria: Programa nacional do livro didático – PNLD*. Tribunal de Contas da União: Brasília. 11 p.
- Brasil (1999b). Relatório de planejamento de auditoria de desempenho. *Programa nacional do livro didático – PNLD*. TCU: Brasília. 13 p.
- Brasil (2000a). Tribunal de Contas da União. *Técnicas de auditoria: Indicadores de desempenho e mapa de produto*. Brasília: TCU. Coordenadoria de Fiscalização e Controle. 32p.
- Brasil (2000b). Tribunal de Contas da União. *Técnicas de auditoria: Mapa de processo*. Brasília: TCU. Coordenadoria de Fiscalização e Controle. 18p.

- Brasil (2000c). Tribunal de Contas da União. *Manual de auditoria de natureza operacional*. Brasília: TCU. Coordenadoria de Fiscalização e Controle. 114p.
- Brasil (2000d). Relatório de Auditoria de Natureza Operacional. *Programa de desenvolvimento sustentável de recursos hídricos para o semi-árido brasileiro – PROÁGUA/SEMI-ÁRIDO SR/MMA*. Brasília: TCU. Acordo de Cooperação TCU/Reino Unido. Brasília/DF.
- Brasil (2000e). Tribunal de Contas da União. Relatório de Levantamento de Auditoria. *Proágua/semi-árido SR/MMA*. Brasília: TCU. Acordo de Cooperação TCU/Reino Unido.
- Brasil (2000f). Relatório de levantamento de auditoria de natureza operacional. *Monitoramento da mortalidade materna*. Brasília: TCU: 14 p.
- Brasil (2001a). Tribunal de Contas da União. *Técnicas de auditoria: Análise RECI*. Brasília: TCU. Secretaria de Fiscalização e a Avaliação de Programas do Governo. 18p.
- Brasil (2001b). Tribunal de Contas da União. *Técnicas de auditoria: Marco lógico*. Brasília: TCU. Secretaria de Fiscalização e a Avaliação de Programas do Governo. 26p.
- Brasil (2001c). Tribunal de Contas da União. Relatório de levantamento de auditoria. *Programa de erradicação do trabalho infantil - PETI*. Brasília: TCU. Secretaria de Fiscalização e a Avaliação de Programas do Governo. 25p.
- Brasil (2001d). Relatório de auditoria de natureza operacional. *Monitoramento da mortalidade materna*. TCU: Brasília: TCU. 16 p.
- Brasil (2002a). Tribunal de Contas da União. *Técnica de auditoria: Análise stakeholder*. Brasília: TCU. Secretaria de Fiscalização e a Avaliação de Programas do Governo. 33p.
- Brasil (2002b). Tribunal de Contas da União. *Técnica de auditoria: Análise SWOT e matriz de verificação de risco*. Brasília: TCU Secretaria de Fiscalização e a Avaliação de Programas do Governo. 33p.
- Brasil (2002c). Tribunal de Contas da União. *Avaliação do TCU sobre o programa de erradicação do trabalho infantil*. Brasília: TCU Secretaria de Fiscalização e Avaliação de Programas de Governo. 20 p.
- Brasil (2002d). *Programa de erradicação do trabalho infantil – PETI*. Manual de Orientações. Brasília: Ministério da Previdência e Assistência Social. 44 p.
- Brasil (2002e). *Relatório de auditoria de natureza operacional: Programa de erradicação do trabalho infantil – PETI*. Brasília: Tribunal de Contas da União Secretaria Fiscalização e Avaliação de Programas de Governo. 45 p.

- Brasil (2002f). Relatório de auditoria de natureza operacional. *Programa novo emprego e seguro desemprego*. Brasília: TCU. 13 p.
- Brasil (2002g). Relatório de levantamento de auditoria de natureza operacional. *Programa Novo Emprego e seguro desemprego*. TCU: Brasília. 39 p.
- Brasil (2003a). *Conhecendo o Tribunal de Contas da União*, 2ª ed. Brasília: Gabinete da Presidência.
- Brasil (2003b). Relatório de planejamento de auditoria. Tribunal de Contas da União. Secretaria de Fiscalização e avaliação de Programas de Governo. *Programa de desenvolvimento do turismo do nordeste - PRODETUR/NE*. TCU: Brasília. 60p.
- Brasil (2003c). Relatório de auditoria. *Programa energia das pequenas comunidades*. TCU: Brasília, 60 p.
- Brasil (2003d). Relatório de avaliação de programa: *Programa reinserção social do adolescente em conflito com a lei*. Brasília: TCU. 104p.
- Brasil (2003e). Relatório de levantamento de auditoria. *Programa energia das pequenas comunidades*. TCU: Brasília, 35 p.
- Brasil (2003f). Relatório de planejamento de auditoria de natureza operacional: *Programa morar melhor*. TCU. Brasília. 77p.
- Brasil (2003g). Relatório de planejamento de auditoria de natureza operacional. *Programa reinserção social do adolescente em conflito com a lei*. Brasília: TCU. 87p.
- Brasil (2003h). Relatório de planejamento de auditoria de natureza operacional. *Programa novo mundo rural*. TCU: Brasília, 75 p.
- Brasil (2004a). Relatório de avaliação de programa: *Programa desenvolvimento do turismo no nordeste*. Brasília: TCU. Secretaria de Fiscalização e Avaliação de Programas de Governo.
- Brasil (2004b). *Cartilha do programa de erradicação do trabalho infantil*. Ministério do Desenvolvimento Social e Combate à Fome/MDS: Brasília.
- Brasil (2004c). Relatório de auditoria de natureza operacional. *Programa novo mundo rural*. TCU: Brasília, 65 p.
- Brasil (2004d). Relatório de avaliação de programa. *Programa combate ao abuso e à exploração sexual de crianças e adolescentes*. TCU: Brasília, 98p.
- Brasil (2004e). Relatório de avaliação de programa. *Programa desenvolvimento do turismo no nordeste*. Brasília: TCU. 96p.

- Brasil (2004f). Relatório de avaliação de programa: *Programa morar melhor*. Brasília: TCU. 138p.
- Brasil (2004h). Relatório de planejamento de auditoria de natureza operacional. *Programa combate ao abuso e à exploração sexual de crianças e adolescentes*. TCU: Brasília, 61p.
- Brasil (2005a). Portaria nº 666, de 28 de Dezembro de 2005. Ministério do Desenvolvimento Social e Combate à Fome. *Disciplina a integração entre o programa bolsa família e o programa de erradicação do trabalho infantil*. Edição nº 251 de 30/12/2005. Brasília: MDS.
- Brasil (2006). *Planilha consolidada Brasil Janeiro 2006*. Documento de circulação interna do Departamento de Proteção Social Especial – DPSE/MDS: Brasília.
- Campbell, D. T., & Stanley, J. C. (1979). *Delineamentos experimentais e quasi-experimentais de pesquisa*. São Paulo: Editora Pedagógica e Universitária.
- Cano, I. (2002). *Introdução à avaliação de programas sociais*. Editora. FGV.
- Chelimsky, E. (1985). *Comparing and contrasting auditing and evaluation: Some notes on their relationship*. Evaluation Review, [On-line], vol. 9, nº 4, p. 483-503. Disponível: <http://erx.sagepub.com/cgi/content/abstract/9/4/483>.
- Chianca, T., & Youker, B. (2004). *La evaluacion en America Latina y el Caribe: Visión general de los desarrollos recientes*. JMDE Journal of Multidisciplinary Evaluation. Number 01 [On-line], Disponível: [http://www.wmich.edu/evalctr/jmde/content/Articulo\\_Thomaz\\_Crianca\\_redes\\_regionales\\_esp.pdf](http://www.wmich.edu/evalctr/jmde/content/Articulo_Thomaz_Crianca_redes_regionales_esp.pdf).
- Cook, T. D., & Gruder, C. L. (1978). *Metaevaluation research*. Evaluation Review. [On-line], Sage Disponível: <http://erx.sagepub.com/cgi/content/abstract/2/1/5>.
- Cook, T. D., & Shadish, Jr., W. R. (1986). *Program evaluation: The worldly science*. *Annual Review Psychology*. 37: 193-232.
- Cotta, T. C. (1998 Abril - Junho). *Metodologias de avaliação de programas e projetos sociais: análise do resultado de impacto*. Revista do Serviço Público. Fundação Escola Nacional de Administração Pública - Ano 49, nº.2. Brasília: ENAP.
- Cozby, P. C. (2003). *Métodos de pesquisa em ciências do comportamento*. São Paulo: Ed. Atlas.

- Dershen, H. A. (1990). *Community crime prevention: A review essay on program evaluations and policy implications*. Criminal Justice Policy Review [On-line]. Sagi Publication. Disponível: <http://cjp.sagepub.com/cgi/content/abstracts/4/1/53>.
- Evaluation Research Society – ERS (1998). *Standards for program evaluation*. California: Jossey-Bass Inc.
- Faleiros, V. P. (2000). *Natureza e desenvolvimento das políticas sociais no Brasil*. In: Capacitação em serviço social e política social, módulo 3. Brasília: UnB. Centros de Educação Aberta, Continuada a Distância.
- Fernández-Ballesteros, R., Vedung, E., & Seyfried, E. (1998). *Psychology in program evaluation*. *European Psychologist*, 3, 143-154.
- Freitas, C. A. S. (2005). *A aprendizagem, isomorfismo e institucionalização: O caso da atividade de auditoria de natureza operacional no TCU*. Dissertação de Mestrado da Faculdade de Economia, Administração, Contabilidade e Ciência da Informação e Documentação, apresentada a Pós-graduação em Administração PPGA, da Universidade de Brasília – UnB.
- Günther, H. (2006). Pesquisa qualitativa versus pesquisa quantitativa: Esta é a questão? *Série: Textos de Psicologia Ambiental*, Nº 07. Brasília, DF: UnB, Laboratório de Psicologia Ambiental.
- Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: correcting error and biases in research findings*. USA. Sage Publications.
- Hunter, J. E., & Schmidt, F. L. (1996). *Measurement error in psychological research: lessons from 26 research scenarios*. *Psychological Methods*. Vol. 1, nº 2, 199-223
- Hunter, J. E., & Schmidt, F. L. (1999). *Comparison of three meta-analysis revisited: An analysis of Johnson, Lullen, and Salas (1995)*. *Journal of Applied Psychology*, vol. 84, nº 1, 144-148.
- Kanter, A. (2001). *After-School programs for adolescents*. NASSP Bulletin [On-line], Sagi Publications. Disponível: <http://www.sagepub.com/cgi/content/abstract/85/626/12>.
- Krueger, R. A., Casey, M. A., Donner, J. K., S., & Maack, J. N. Orgs. (2001 June). *Social analysis selected tools and techniques*. Social Development Department. The World Bank. Washington: Paper number 36.
- Lima, D. H. (2004). *O papel do Tribunal de Contas da União na responsabilização dos agentes públicos pelo desempenho da ação governamental*. Monografia não publicada

apresentada ao Curso de Especialização do Instituto Serzedello Corrêa, do Tribunal de Contas da União.

- Lima, S. M. V. (2003). *Mudança organizacional: Teoria e gestão*. Rio de Janeiro. Editora FGV, 348p.
- Mendes, A. M. B., Tamayo, A., Paz, M. G. T., Neiva, E. R., Tamayo, N., Silva, P. T., Souza, A. C., Martins, A. J., & David, R. G. (1999). *Análise da cultura organizacional do Tribunal de Contas da União – TCU*. Brasília: Relatório Final. O&T Consultoria/FINATEC/Unb.
- Mintzberg, H., Ahlstrand, B., & Lampel, J. (2002). *Safári de estratégia: Um roteiro pela selva do planejamento estratégico*. Tradução Nivaldo Montigelli Jr. Porto Alegre: Bookman.
- Mintzberg, H. (1983). *Power in and around organizations*. Englewood Cliffs: Prentice Hall.
- Oskamp, S. (1984). *Evaluation research – Sesame street*. Applied Social Psychology. New Jersey: Prentice Hall, Inc.
- Oskamp, S. (1984). *Quasi-experimental studies: Effects of television*. Applied Social Psychology. New Jersey: Prentice-Hall Inc.
- Pasqually, L. (1997). *Delineamentos de Pesquisa em Ciência*. Série: Métodos e Medidas em Ciências Psicossociais. Brasília. Material não publicado.
- Patton, M. Q. (2001). *Qualitative research e evaluation methods*. 3ª ed. EUA. Sage Publications, Inc.
- Popham, W. J. (1977). *Manual de avaliação: Regras práticas para o avaliador educacional*. Petrópolis: Vozes.
- Posavac, E. J., & Carey, R. G. (2003). *Program evaluation. Methods and case Studies*. New Jersey: Prentice Hall. 6ª edição.
- Ribeiro, N. G. F. (2004). *Trabalho e família: Um estudo sobre a dinâmica da congruência entre demandas múltiplas*. Tese de Doutorado não publicada apresentada ao Instituto de Psicologia da Universidade de Brasília. Brasília.
- Rossi, P. H., & Freeman, H., E. (1993). *Evaluation: A systematic approach*. 5 th. ed. United States of America: Sage Publications.
- Schwandt, T. A. (1989). *The politics of verifying trustworthiness in evaluation auditing*. American Journal of Evaluation, 10; 33.

- Shadish, W., Cook, T., & Levinton, L. (1995). *Foundations of program evaluation: Theories of practices*. California: Sage Publications.
- Silva, P. L. B. (2002a). *A avaliação de programas públicos: Reflexões sobre a experiência brasileira*. Relatório técnico. Brasília: IPEA.
- Silva, R. T. (2002b). *Eficiência e eficácia da ação governamental: Uma análise comparativa de sistemas de avaliação: relatório técnico*. Brasília: IPEA.
- Smith, P. B., & Bond, M. H. (1999). *Social psychology across cultures*. Allyn e Bacon: EUA.
- Stufflebeam, D. L., & Shinkfield, A. J. (1987). *Evaluación sistemática: Guia teórica y práctica*. Barcelona: Ediciones Paidós Ibérica.
- Weber, S. (2004 October). *Network evaluation as a complex learning process*. In: Journal of Multidisciplinary Evaluation. Number 01.
- Woodside, A. G., & Sakai, M. Y. (2001). *Meta-evaluation of performance audits of government tourism-marketing programs*. Journal of Travel Research; 39, 369. [On-line], Disponível: <http://jtr.saepub.com/cgi/content/abstract/39/4/369>.
- Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (2004). *Avaliação de programas: Concepções e práticas*. São Paulo: Editora Gente.

ABIH - Associação Brasileira da Indústria de Hotelaria  
ACE - Centro Americano de Avaliação  
AIS - Ações Integradas de Saúde  
Aneel – Agência Nacional de Energia Elétrica  
ANOP - Auditoria de Natureza Operacional  
AP - Avaliação de Programas  
ARCO – Agência Regional de Comercialização  
BB - Banco do Brasil  
BID - Banco Interamericano de Desenvolvimento  
BNDES - Banco Nacional de Desenvolvimento Econômico e Social  
CAIXA – Caixa Econômica Federal  
CBO - Classificação Brasileira de Ocupações  
CEME - Central de Medicamentos  
CEPEL – Centro de Pesquisa de Energia Elétrica  
CERDS – Aperfeiçoamento do Controle Externo com Foco na Redução da Desigualdade Social  
CET/UnB - Centro de Excelência em Turismo, da Universidade de Brasília  
CHESF – Companhia Hidro Elétrica do São Francisco  
CONANDA - Conselho Nacional dos Direitos da Criança e do Adolescente  
CPI - Comissão Parlamentar de Inquérito  
CREAS – Centro de Referência Especial de Assistência Social  
DATAPREV - Empresa de Processamento de Dados da Previdência Social  
DF – Distrito Federal  
ECA – Estatuto da Criança e do Adolescente  
EMATER - Empresa de Assistência Técnica e Extensão Rural  
EMBRAPA - Empresa Brasileira de Pesquisa e Agropecuária  
Embratur - Instituto Brasileiro de Turismo  
ERS - *Evaluation Research Society*  
FNDE – Fundo Nacional de Desenvolvimento da Educação  
FNPETI – Fórum Nacional de Prevenção e Erradicação do Trabalho Infantil  
FUNABEM - Fundação para o Bem-Estar do Menor  
GHC – Grupo hospitalar Conceição

IBAMA – Instituto Brasileiro de Meio Ambiente  
IBGE - Instituto Brasileiro de Geografia e Estatística  
IDH - Índices de Desenvolvimento Humano  
INAN - Instituto Nacional de Alimentação e Nutrição  
INAMPS - Instituto Nacional de Previdência Social  
INCRA – Instituto Nacional de Colonização e Reforma Agrária  
INPS - Instituto Nacional de Previdência Social  
INTOSAI - International Organization of Supreme Audit Institutions  
IPEA - Instituto de Pesquisa Econômica Aplicada  
LBA - Legião Brasileira de Assistência  
LDO - Lei de Diretrizes Orçamentárias  
LOAS - Lei Orgânica da Assistência Social  
LOPS - Lei Orgânica da Previdência Social  
MDA – Ministério do Desenvolvimento Agrário  
MDS - Ministério do Desenvolvimento Social e Combate à Fome  
MEC – Ministério da Educação e Cultura  
MI – Ministério da Integração Nacional  
MMA – Ministério do Meio Ambiente  
MPOG - Ministério do Planejamento, Orçamento e Gestão  
MS – Ministério da Saúde  
MTE - Ministério do Trabalho e do Emprego  
NOB/SUAS - Norma Operacional Básica/Sistema Único de Assistência Social  
OCDE - *Organization for Economic Cooperation and Development*  
OMS - Organização Mundial de Saúde  
ONG – Organização não Governamental  
OVE - *Office of Evaluation and Oversight*  
PAS - Programa de Avaliação Seriada  
PBF - Programa Bolsa Família  
PDA – Plano de Desenvolvimento do Assentamento  
PDITS – Plano de Desenvolvimento Integrado de Turismo Sustentável  
PETI - Programa de Erradicação do Trabalho Infantil  
PNAD - Pesquisa Nacional por Amostra de Domicílio

PNAS - Política Nacional de Assistência Social

PNLD - Programa Nacional do Livro Didático

PPA – Plano Pluri-Anual

Pró-Água/Semi-Árido - Programa de Desenvolvimento Sustentável de Recursos Hídricos para o Semi-Árido Brasileiro

Prodetur/NE - Programa Desenvolvimento do Turismo no Nordeste

PRONAF - Programa de Incentivo a Agricultura Familiar

RAIS - Relação Anual de Informações Sociais

RECI – Responsável (*responsible*), Executor (*accountable*), Consultado (*consulted*), Informado (*informed*)

RELAC - Rede de Seguimento, Avaliação e Sistematização da América Latina e Caribe

ROM - *Results Oriented Management*

SEAIN - Secretaria de Assuntos Internacionais

SECEX – Secretaria Estadual de Controle Externo

SEDH - Secretaria Especial de Direitos Humanos

SEF/MEC - Secretaria de Ensino Fundamental/Ministério da Educação e Cultura

Sentinela - Programa de Enfrentamento ao Abuso e Exploração Sexual de Crianças e Adolescentes

SEPDCA - Secretária Especial dos Direitos da Criança e do Adolescente

SFCI - Secretaria Federal de Controle Interno

SIAFAS - Sistema Integrado de acompanhamento de Finanças da Assistência Social –

SIPIA - Sistema de Informação para a Infância e Adolescência

SIPRA - Sistema de Informações de Projetos de Reforma Agrária

Sistema S - Sesc, Sesi, Senai, Sebrae

SPSS - *Statistical Package of Social Science*

SUAS – Sistema Único de Assistência Social

SUDS - Sistema Único de Desenvolvimento Social

SUS - Sistema Único de Saúde

SWOT – *Strengths, Weaknesses, Opportunities, Threats*

UEEs - Unidades Executoras Estaduais

UEGP - Unidade Estadual de Gestão do Programa

UEGP's - Unidades Estaduais de Gestão do Programa

UFCE - Universidade Federal do Ceara

UGP-G - Unidade de Gerenciamento do Programa – Gestão

UMA - Unidade de Monitoramento e Avaliação

UNB - Universidade de Brasília

USAID - Agência dos Estados Unidos para o Desenvolvimento Internacional