

INSTITUTO DE PESQUISAS ESPACIAIS

PROJETO COLETIVO

Planejamento de um Curso de Treinamento de Avaliadores Educacionais

Submetido por

Glória Cardozo Bertti

Maria de Lourdes Neves de Oliveira

Maristela Arantes Marteleto

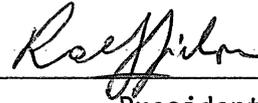
Suely Costa

Em cumprimento parcial dos requisitos exigidos para a obtenção
do título de Mestre em Tecnologia Educacional

- 1977 -

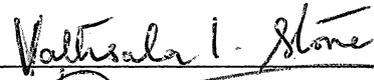
Aprovado pela Banca Examinadora
em cumprimento dos requisitos exigidos
para a obtenção do Título de Mestre em
Tecnologia da Educação

Dr. Ralf Gielow



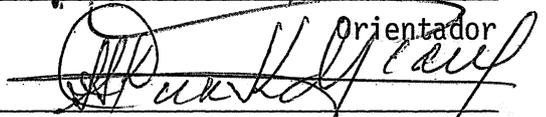
Presidente da Banca

Dra. Vathsala I. Stone



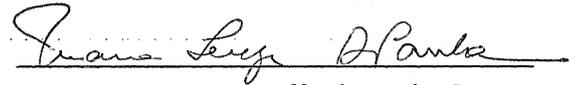
Orientador

Eco. Sérgio Suren Kurkdjian, MSc.



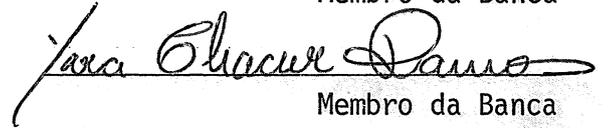
Co-Orientador

Sra. Maria Tereza Dejuste de Paula, MSc.



Membro da Banca

Sra. Yara Chacur Ramos, MSc.



Membro da Banca

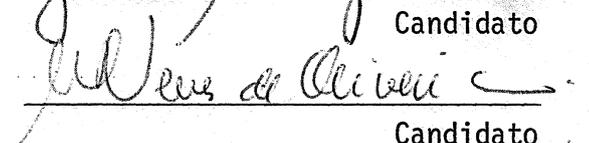
-convidada-

Glória Cardozo Bertti



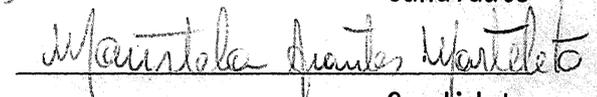
Candidato

Maria de Lourdes Neves de Oliveira



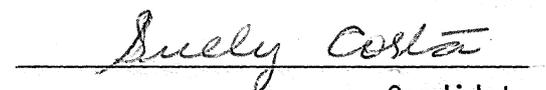
Candidato

Maristela Arantes Marteleto



Candidato

Suely Costa



Candidato

São José dos Campos, 10 de junho de 1977

Aos nossos pais

ÍNDICE

VOLUME I

SUMÁRIO	viii
ABSTRACT	ix
LISTA DE FIGURAS	x
LISTA DE TABELAS	xv
CAPÍTULO I - VISÃO GERAL DO TRABALHO	1
1.1 - Colocação do Problema	1
1.2 - Objetivo do Trabalho	5
1.3 - Clientela	6
1.4 - Revisão de Literatura - Modelos Instrucionais e de Treina mento	7
1.5 - Metodologia do Trabalho	26
1.6 - Revisão de Literatura Complementar	29
1.6.1 - Teoria de Avaliação Educacional	29
1.6.2 - Análise Ocupacional	38
1.6.3 - Estratégias Instrucionais	43
1.6.4 - Seleção de Meios Instrucionais	47
CAPÍTULO II - PLANEJAMENTO DO CURSO: DEFINIÇÃO DO UNIVERSO A SER TREINADO E PLANEJAMENTO DAS UNIDADES INSTRUCIONAIS.	63
2.1 - Definição do Universo de Atividades que Serão Treina das	63
2.1.1 - Identificação do Universo Teórico de Atividades que Carac terizam a Função do Avaliador	63
2.1.2 - Identificação do Universo Real de Atividades que Caracte rizam a Função do Avaliador	68

2.1.3 - Comparação das Atividades Constantes do Universo Real com Atividades do Universo Teórico	75
2.1.4 - Identificação de um Universo de Atividades do Trabalho ..	76
2.1.5 - Identificação de um Universo de Atividades que Serão Treinadas	76
2.2 - Planejamento das Unidades Instrucionais	85
2.2.1 - Estrutura e Sequência do Curso	85
2.2.2 - Objetivos Específicos das Unidades do Curso	87
2.2.3 - Pré-Requisitos do Curso	100
2.2.4 - Seleção de Meios	101
2.2.5 - Seleção de Estratégias Instrucionais, Meios e Distribuição de Horas de Aula para a Consecução de Cada Objetivo Instrucional	106
2.2.6 - Duração do Curso	136
2.2.7 - Número de Alunos por Professor	136
CAPÍTULO III - PLANEJAMENTO DO TESTE PILOTO E INSTALAÇÃO DO CURSO DE TREINAMENTO	137
3.1 - Ensaio do Curso	139
3.1.1 - Obtenção do Ensaio	139
3.1.2 - Operação do Ensaio	152
3.2 - Instalação do Curso	167
3.2.1 - Obtenção da Instalação	167
3.2.2 - Operação da Instalação	173
CAPÍTULO IV - CONSIDERAÇÕES FINAIS	185
AGRADECIMENTOS	189
BIBLIOGRAFIA	191

APÊNDICE A - LISTAGEM DE ENTIDADES EDUCACIONAIS FORNECIDAS PELA ABT.

APÊNDICE B - FORMULÁRIO.

APÊNDICE C - INSTRUMENTAL DE COLETA DE DADOS PARA A ANÁLISE OCUPACIONAL DO AVALIADOR EDUCACIONAL.

APÊNDICE D - INSTRUMENTAL DE COLETA DE DADOS, JUNTO AOS DECISORES, PARA ANÁLISE OCUPACIONAL DO AVALIADOR.

APÊNDICE E - FREQUÊNCIA DE EXECUÇÃO DAS ATIVIDADES.

VOLUME II

APÊNDICE F - ESPECIFICAÇÃO DO CONTEÚDO PARA O CURSO DE TREINAMENTO DE AVALIADORES EDUCACIONAIS	F.1
1.0 - Unidade 1: Contexto do Processo de Avaliação - Planejamento Sistêmico e Processo Decisório	F.2
1.1 - Sub unidade 1: Planejamento, Teoria Geral de Sistemas, Abordagem de Sistemas - Visão Geral	F.2
1.2 - Sub unidade 2: O Processo de Engenharia de Sistemas	F.22
1.3 - Sub unidade 3: A Avaliação Educacional dentro de uma Abordagem Sistêmica	F.41
1.4 - Sub unidade 4: Algumas Abordagens de Avaliação Educacional	F.54
1.5 - Sub unidade 5: Pesquisa X Avaliação	F.70
2.0 - Unidade 2 - Uma Abordagem Avaliativa - O Modelo de Stufflebeam et al	F.79
2.1 - Sub unidade 1: A Tomada de Decisão Educacional	F.79
2.2 - Sub unidade 2: Tipos de Avaliação	F.92
2.3 - Sub unidade 3: O Modelo de Avaliação segundo Stufflebeam et al	F.113
3.0 - Unidade 3 - Obtenção de Informações: Instrumentais de Medida	F.120
3.1 - Sub unidade 1 - Questionário e Entrevista	F.122
3.2 - Sub unidade 2 - Escalas de Atitude	F.132
3.3 - Testes de Aprendizagem	F.145
4.0 - Unidade 4: Obtenção de Informações: Avaliação Formativa de Materiais Instrucionais Impressos e de Alunos	F.214
4.1 - Sub unidade 1: Papel e Fontes de Informações de Avaliação Formativa	F.215

4.2 - Sub unidade 2: Avaliação Formativa de Materiais Instrucionais Impressos	F.220
4.3 - Sub unidade 3: Avaliação Formativa de Alunos	F.237
4.4 - Algumas Concepções Errôneas em Estudos de Avaliação Formativa	F.250
5.0 - Unidade 5 - Obtenção de Informações: Avaliação Somativa dos Produtos da Mudança em Termos de Aprendizagem	F.252
5.1 - Introdução: Dois Esquemas para Avaliação Somativa dos Produtos da Mudança	F.252
5.2 - Sub unidade 1: Um Esquema de Avaliação Somativa dos Produtos de Mudança (no caso Aprendizagem) baseado nos Procedimentos da Pesquisa Experimental	F.256
5.3 - Sub unidade 2: Um Esquema Descritivo da Avaliação Somativa dos Produtos da Mudança	F.352
6.0 - Unidade 6 - A Estrutura de um Plano de Avaliação	F.363

SUMÁRIO

Este trabalho refere-se ao Planejamento de um curso de treinamento de Avaliadores Educacionais a partir da identificação de um universo de atividades que caracteriza a função do Avaliador. Para tal identificação foi feito um levantamento de atividades constantes na literatura como características da função do Avaliador e também um levantamento das atividades exercidas pelo mesmo, na prática, em entidades educacionais. Assim, buscando o equilíbrio entre teoria e prática, o universo de atividades a ser treinado foi definido através do confronto entre os dois levantamentos. Partindo de tais atividades, foram gerados objetivos instrucionais e, para o seu alcance, foi planejada a instrução seguindo um esquema instrucional proposto pelas autoras. Utilizando técnicas de Engenharia de Sistemas, foram planejados o ensaio do curso para sua validação e a instalação posterior.

ABSTRACT

The planning of a training course for Educational Evaluators is presented, based on an universe of activities identified with the function of evaluation. Said universe was defined by the integration of a literature and a field survey, thus combining activities theoretically expected of the evaluator with those he would normally execute. This merging of the two surveys resulted in an universe which balances the theory and practice. Based on the identified activities, instructional objectives were generated for the course, and instruction was planned within a scheme proposed by the authors, for the attainment of the objectives. Using Systems Engineering Techniques, a trial of the pilot course was planned for the purposes of validation and subsequent installation.

LISTA DE FIGURAS

Figura I.1 - Modelo Serviço de Informação do Governo	9
Figura I.2 - "Design" de Sistemas Instrucionais	11
Figura I.3 - Modelo de Sistema de Treinamento e Desenvolvimento ...	13
Figura I.4 - Modelo de Desenvolvimento de Instrução Vocacional	15
Figura I.5 - Modelo para o Desenvolvimento de Sistemas Instru cionais	17
Figura I.6 - Modelo Operacional de Sistema Instrucional	19
Figura I.7 - Esquema para o Planejamento do Curso	20
Figura I.8 - Diagrama dos Pontos de Decisões Necessários para a Seleção de uma Simples Classe de Meios Apropria dos para uma Necessidade Instrucional	53
Figura I.9 - Critérios para Distinção entre Conteúdo Concreto e Abstrato	54
Figura I.10 - Decisão entre Apresentação em Movimento ou Estáti- ca	55
Figura I.11 - Escolha dos Meios de Narração	56
Figura I.12 - Visualização da Narração Audio	57
Figura I.13 - Visualização da Narração Impressa	58
Figura I.14 - Alternativas para Expressar o Conteúdo Concreto	59
Figura I.15 - Alternativas para Expressar o Conteúdo Abstrato	60
Figura II.1 - Estrutura e Sequência do Curso	85
Figura II.2 - Critério para Distinção entre Conteúdo Concreto e Abstrato	103

Figura II.3	- Escolha dos Meios de Narração	104
Figura II.4	- Visualização da Narração Impressa	105
Figura III.1	- Estrutura de Divisão de Atividades para a Obtenção do Ensaio do Curso	140
Figura III.2	- Estrutura de Divisão de Recursos para a Obtenção do Ensaio do Curso	141
Figura III.3	- Diagrama de Fluxo de Trabalho para a Obtenção do Ensaio do Curso	144
Figura III.4	- Cronograma para Obtenção do Ensaio do Curso	153
Figura III.5	- Estrutura de Divisão de Atividades para a Operação do Ensaio do Curso	154
Figura III.6	- Estrutura de Divisão de Recursos para a Operação do Ensaio do Curso	155
Figura III.7	- Diagrama de Fluxo de Trabalho para a Operação do Ensaio do Curso	159
Figura III.8	- Cronograma para Operação do Ensaio do Curso	168
Figura III.9	- Estrutura de Divisão de Atividades para a Obtenção da Instalação do Curso	170
Figura III.10	- Estrutura de Divisão de Recursos para a Obtenção da Instalação do Curso	171
Figura III.11	- Diagrama de Fluxo de Trabalho para a Obtenção da Instalação do Curso	172
Figura III.12	- Cronograma para Obtenção da Instalação do Curso	174
Figura III.13	- Estrutura de Divisão de Atividades para a Operação da Instalação do Curso	175

Figura III.14 - Estrutura de Divisão de Recursos para a Operação da Instalação do Curso	176
Figura III.15 - Diagrama de Fluxo de Trabalho para a Operação da Instalação do Curso	177
Figura III.16 - Cronograma para Operação da Instalação do Curso	183
Figura F.1 - Caracterização de um Sistema	F.15
Figura F.2 - Sequência de Passos no Planejamento e Execução de um Sistema	F.24
Figura F.3 - Fases Sequenciais do Planejamento Preliminar	F.27
Figura F.4 - Atividades da Fase de Operação do Sistema	F.31
Figura F.5 - Atividades da Fase de Obtenção do Sistema	F.33
Figura F.6 - Passos Gerais do Processo de Análise de Sistemas ...	F.34
Figura F.7 - Primeiro Conjunto de Etapas do Processo de Engenharia de Sistemas: Planejamento	F.36
Figura F.8 - Segundo Conjunto de Etapas do Processo de Engenharia de Sistemas: Execução	F.39
Figura F.9 - Sequência de Etapas no Processo de Engenharia de Sistemas	F.40
Figura F.10 - Domínio da Investigação	F.70
Figura F.11 - Domínio da Investigação: Avaliação e Pesquisa Educacional	F.73
Figura F.12 - Relação entre a Avaliação e a Tomada de Decisão	F.79
Figura F.13 - Um Modelo de Avaliação	F.116
Figura F.14 - Relação entre o Escore-Universo e o Escore Observado	F.192

Figura F.15 - Gráfico com as Percentagens e Escores Brutos para a Determinação do Percentil	F.203
Figura F.16 - Área sobre a Curva Normal	F.205
Figura F.17 - Perfil de um Aluno nos Exames Finais	F.207
Figura F.18 - Modelo de Avaliação Formativa de Materiais Instrucionais Impressos	F.235
Figura F.19 - Estrutura para Avaliação	F.250
Figura F.20 - Efeito de Interação da Variável Independente Primária e da Variável Moderadora	F.264
Figura F.21 - Relação entre os 5 Tipos de Variáveis	F.266
Figura F.22 - Efeito de Interação das Variáveis Método de Ensino e Zona	F.303
Figura F.23 - Ausência de Efeito de Interação das Variáveis Método de Ensino e Zona	F.303
Figura F.24 - Representação Gráfica de Séries Temporais	F.313
Figura F.25 - Região Crítica para Hipótese Nula com o Sinal de Igualdade (=)	F.328
Figura F.26 - Região Crítica para Hipótese Nula com Sinal Menor ou Igual (\leq)	F.329
Figura F.27 - Região Crítica para Hipótese Nula com Sinal Maior ou Igual (\geq)	F.330
Figura F.28 - Comparação Gráfica dos Resultados Teórico e Prático de uma Aplicação do Teste t de Student, para $H_0: \mu_1 = \mu_2$	F.347
Figura F.29 - Comparação Gráfica dos Resultados Teórico e Prático de uma Aplicação do Teste t de Student, para $H_0: \mu \leq K$	F.349

LISTA DE TABELAS

Tabela I.1 - Distinção entre Avaliação Intrínseca e Extrínseca e Avaliação Formativa e Somativa	34
Tabela I.2 - Comparação entre Duas Formas de Análise Ocupacional	40
Tabela I.3 - Os Meios de Comunicação - Uma Taxionomia	51
Tabela II.1 - Atividades que Constituem o Processo de Avaliação ...	64
Tabela II.2 - Sumário de Procedimentos de Coleta de Dados	72
Tabela II.3 - Formação Acadêmica	73
Tabela II.4 - Treinamento Recebido em Avaliação	74
Tabela II.5 - Aplicação dos Critérios para a Seleção das Atividades do Universo a ser Treinado	78
Tabela II.6 - Objetivos Específicos das Unidades do Curso	87
Tabela II.7 - Objetivos, Estratégias Instrucionais, Meios e Horas/Aula do Curso de Treinamento de Avaliação	107
Tabela F.1 - Distinção entre Avaliação Intrínseca e Extrínseca e Avaliação Formativa e Somativa	F.60
Tabela F.2 - Situações de Decisão	F.81
Tabela F.3 - Tipos de Decisão	F.89
Tabela F.4 - Avaliação de Contexto	F.98
Tabela F.5 - Avaliação de Input	F.103
Tabela F.6 - Avaliação de Processo	F.112
Tabela F.7 - Avaliação de Produto	F.114
Tabela F.8 - Vantagens e Desvantagens de Questionário e Entrevista	F.130

Tabela	F.9 - Tabela de Respostas	F.168
Tabela	F.10 - Registro dos Escores em Cada Item	F.169
Tabela	F.11 - Escores Ordenados	F.170
Tabela	F.12 - Número de Acertos em Cada Item	F.172
Tabela	F.13 - Índices de Discriminação, Dificuldade e Correlação Item-Total	F.173
Tabela	F.14 - Questões, Procedimentos, Aplicação e Exemplos dos Tipos de Validade	F.187
Tabela	F.15 - Acerto ou Erro nos Itens	F.210
Tabela	F.16 - Alcance de Objetivos	F.210
Tabela	F.17 - Resultados da Testagem Um a Um: Prē e Pōs-Teste	F.229
Tabela	F.18 - Ganho por Objetivo, Obtido pelo Grupo	F.230
Tabela	F.19 - Percentagem de Alunos por Objetivo	F.231
Tabela	F.20 - Percentagem de Atingimento por Objetivo e por Aluno	F.243
Tabela	F.21 - Um Exemplo de Ficha de Avaliação Formativa de Aluno, a Ser Preenchida pelo Professor	F.248
Tabela	F.22 - Um Exemplo de Ficha de Avaliação Formativa de Aluno, a Ser Preenchida pelo Professor e pelo Aluno ...	F.249
Tabela	F.23 - Tabela de Números Randômicos	F.274
Tabela	F.24 - Um Esquema de Coleta de Dados para "Designs" de um Simple Fator	F.299
Tabela	F.25 - Combinação de Níveis num Design Fatorial Completo ..	F.301

Tabela	F.26 - Combinação de Níveis num Design Fatorial Completo do Tipo 2 X 2	F.304
Tabela	F.27 - Um Esquema de Coleta de Dados de um "Nested Design"...	F.306
Tabela	F.28 - Um Esquema Alternativo de Coleta de Dados de um "Nested Design"	F.307
Tabela	F.29 - Um Esquema de Coleta de Dados de um "Nested Fatorial Design"	F.308
Tabela	F.30 - Controle de Fontes de Validade, e Utilidade para a Avaliação, de cada Design de Coleta de Dados	F.317
Tabela	F.31 - Quatro Situações Possíveis de Ocorrer quando é tomada uma Decisão acerca da Hipótese Nula	F.326
Tabela	F.32 - Testes Estatísticos Apropriados para Cada "Design" de Coleta de Dados	F.342
Tabela	F.33 - Resultados de uma Aplicação da Análise de Variância "Two-Way"	F.345
Tabela	F.34 - Resultados de uma Aplicação da ANAVA "One-Way"	F.350
Tabela	F.35 - Um Esquema de um Plano de Avaliação	F.365

CAPÍTULO I

VISÃO GERAL DO TRABALHO

1.1 - COLOCAÇÃO DO PROBLEMA

Estamos vivendo uma época de constantes mudanças educacionais decorrentes do avanço tecnológico. Há necessidade de estudos no sentido de que nossas estruturas educacionais fiquem em consonância com esse desenvolvimento. As diretrizes governamentais brasileiras estão preocupadas com a integração educação/desenvolvimento científico-tecnológico. Tais preocupações se evidenciam:

- Em Metas e Bases para Ação de Governo 70 é apresentado como uma das maneiras de renovar o sistema educacional brasileiro a "divisão do País em regiões educacionais, a fim de que seja possível a implantação de um sistema de Planejamento, Coordenação e Avaliação de todos os programas e projetos educacionais, diretamente ligados à Secretaria Geral do MEC". Ainda, neste documento, um de seus objetivos é a implantação progressiva de um Sistema de Tecnologia Avançada para a Educação, instituindo-se, no Brasil, os métodos de educação de massa: "há necessidade de preparar o sistema educacional brasileiro para realizar a institucionalização da mudança ... dotar o sistema de instituições mais flexíveis e torná-lo propenso à inovação: amplitude e flexibilidade para atingir a maior parcela possível da população, adaptação constante de currículos ... enriquecer o currículo e aumentar a eficiência do ensino".
- Na implantação da reforma do ensino superior e da lei 5.692/71 com novos currículos e novos procedimentos.
- No Plano Setorial de Educação 72/74: "estimular a expansão dos sistemas de treinamento e retreinamento de pessoal", e no seu projeto 23 - Reformulação de Currículos de ensino fundamental - que tem a finalidade de dotar os sistemas de ensino de currícu

los realistas, flexíveis e integrados, que sejam instrumentos eficientes de consecução dos objetivos dentro da reforma de ensino de 1ª e 2ª graus", pretendo-se, após o ano de 1974, a atividade contínua de Avaliação e de reajustamento permanente dos currículos.

- No Plano Nacional de Desenvolvimento 75/79, no capítulo XIV - Política Científica e Tecnológica - : "Em educação, atualizar o conhecimento científico e técnico ministrado nos diferentes níveis de ensino, pela revisão dos textos didáticos e técnicos".
- No Plano Setorial de Educação 75/79 está declarado que: "a Educação se torna cada vez mais dispendiosa pela necessidade de se aliar a expansão quantitativa ao maior rendimento e à melhor qualidade. Nesse sentido, o sistema, para ampliar-se ... deverá recorrer a estruturas para e peri-escolares, bem como aos diferentes tipos de inovações, incorporando a Tecnologia Educacional impõe-se a renovação profunda das estruturas e dos métodos do sistema educativo ... impõe-se a elaboração dos novos métodos e novas técnicas de ensino que aumentam a eficácia do sistema".

Estas diretrizes, que implicam em mudanças educacionais exigem, a todo momento, a tomada de decisão em instalar, modificar ou abandonar programas educativos. A adequação destas decisões está na dependência direta de um sistema de Avaliação que forneça as informações certas nos momentos exatos. Para isso, a Avaliação deve ser considerada como um processo contínuo, aparecendo mesmo antes da ação de instalar um programa, acompanhando todo o seu processo de desenvolvimento e continuando após o respectivo término. A Avaliação, antes da ação de instalar um novo programa, consiste de estudos sobre o programa atual e a viabilidade de instalá-lo; durante o desenvolvimento do processo, consiste em verificar procedimentos e medir seus resultados parciais, podendo melhorá-los; após o término, consiste em medir os resultados finais alcançados.

Portanto, a Avaliação deve incluir a tarefa constante de fornecer informações relevantes sobre a natureza de programas, no que se refere a seus efeitos no desempenho do aluno e professor; nas atitudes de alunos, pais, comunidade; nos efeitos da metodologia e outros, permitindo ao sistema educacional, manter-se em equilíbrio com relações às mudanças frequentes que ocorrem.

Porém, a situação da Avaliação, no Brasil, de modo geral, encontra-se muito aquém das necessidades provocadas pelas mudanças contínuas do Sistema. E isto se deve à falta de Avaliadores especializados para o sistema de Avaliação exigido visto que, em Entidades Educacionais de vários Estados brasileiros, as atividades avaliativas, quando existentes, raramente são desempenhadas por pessoas com especialização no assunto. Na maioria dos casos, são exercidas por elementos com formação universitária, nas mais diversas áreas. Como consequência, a Avaliação apresenta, quase sempre, um enfoque predominantemente terminal e é centrada apenas no desempenho do aluno.

Afirmativas como as que seguem*, obtidas através de contatos nossos com pessoas-chave em Entidades Educacionais brasileiras, confirmam a precária situação da Avaliação:

- . "não desenvolvemos a atividade de Avaliação".
- . "nossos mecanismos de Avaliação são bastante simples, sendo que a ênfase maior é dada ao final do programa, não dispondo ainda de pessoal especializado para a utilização de técnicas e estratégias mais modernas e atualizadas para acompanhar o processo de desenvolvimento dos programas" (O grifo é nosso).

* Estas afirmativas foram coletadas com o envio de um formulário, no nosso 1º contato, às Entidades Educacionais listadas no Apêndice A.

- . "A atividade de Avaliação, em fase de iniciação, vem se constituindo de proposta de projeto de Avaliação do desempenho do aluno". (O grifo é nosso)
- . "As atividades de Avaliação foram iniciadas, de fato, em 1975, aplicando-se à implantação da Reforma de Ensino ... a partir dessa primeira experiência espera-se operacionalizar uma metodologia de Avaliação que deverá ser implantada gradativamente... porém, falta-nos recursos humanos". (O grifo é nosso).

Sabbag (1974) destaca ainda que as Entidades de televisão educativa, em particular, não têm dado à Avaliação a atenção merecida, pois não existem dados a respeito da efetividade de seus cursos.

Considerando a situação da Avaliação, no Brasil, vale acrescentar as colocações que seguem:

- . "Como é possível tomar decisões sem que a função da Avaliação tenha sido preenchida! Isso explica, em parte, as discrepâncias, dificuldades e deficiências das decisões que vêm sendo tomadas no campo educacional" (Oliveira e Oliveira, 1973).
- . "Talvez antes que possamos construir novos programas efetivos, devemos estabelecer modos criativos de monitorar e eventualmente julgar a efetividade de tais programas" Provus (1973)
- . "Os programas educacionais tendem a multiplicar mas podem ficar sem direção, por falta de realimentação constante, vinda de avaliadores habilitados" (Barry, 1974).

Existe uma relação direta entre eficácia do sistema e recursos humanos; e a demanda de avaliadores é muito maior do que a oferta! Nosso sistema de ensino não está em condições ainda de formar elementos para exercer atividades específicas de Avaliação. Consta no currículo de algumas Faculdades de Educação a disciplina "Testes e Medidas" e, em cursos de pós-graduação (CAPES, 1975) existem disciplinas como: Avaliação Educacional, Medida e Avaliação, Medidas Educacionais,

Avaliação da Aprendizagem, Testes e Medidas, Seleção e Avaliação, Avaliação de Currículo. Porém, o enfoque dado a esses conteúdos faz com que se conceba a Avaliação como sendo predominantemente medida de desempenho do aluno, que não é um enfoque restrito da Avaliação. O único curso que tem uma área de especialização em Avaliação Educacional é o mestrado em Tecnologia Educacional do Instituto de Pesquisas Espaciais (INPE, 1975/76).

Em resumo: as constantes mudanças educacionais necessitam de um sistema de Avaliação para sustentá-las, há falta de Avaliadores, e nosso sistema não oferece cursos suficientes para atender a demanda.

Como resolver o problema?

Devido à complexidade da área de Avaliação, propomos, como solução para minimizar o problema, um curso de treinamento formal de avaliadores educacionais. Este trabalho consiste no planejamento deste curso de treinamento, embasado em um universo de atividades que caracteriza a função do avaliador. Tais atividades são enfocadas dentro de uma abordagem atual de avaliação de tal modo que seja imprimido, ao curso, um caráter de agente catalizador de mudanças no próprio contexto educacional. Deste universo de atividades, foram deduzidos os objetivos instrucionais e, para alcançá-los, foi planejada a instrução.

O presente trabalho envolve também o planejamento do ensino e posterior instalação deste curso específico. Tal planejamento foi orientado para entidades já estruturadas para prestação de serviços na área de treinamento de recursos humanos.

1.2 - OBJETIVO DO TRABALHO

O objetivo deste trabalho é fornecer o planejamento de um curso de treinamento de avaliadores educacionais para entidades edu

cionais interessadas em treinar pessoal nesta área. O curso pretende atender a três propósitos:

- formar
- aperfeiçoar e/ou
- atualizar

indivíduos para desempenhar atividades pertencentes ao Universo identificado por nós.

Estes três termos, em nosso contexto, têm o significado que segue:

- . Formar - consiste em levar o indivíduo a adquirir conhecimentos teóricos, e a processar e manipular estes conhecimentos.
- . Aperfeiçoar - consiste em levar o indivíduo a adquirir maior grau de habilidade ou aptidão para executar uma atividade, desenvolvendo e ampliando seus conhecimentos teóricos e práticos.
- . Atualizar - consiste em levar o indivíduo a renovar seus conhecimentos em função de novas abordagens.

1.3 - CLIENTELA

Sendo nosso trabalho o Planejamento de um Curso de Treinamento, podemos localizar dois tipos de clientela do mesmo. O 1º tipo é aquele que vai executar nosso planejamento. Esta clientela deve ser uma Entidade Educacional que conte com recursos materiais e humanos necessários para a obtenção do sistema. Ela deixará, então, o curso pronto para funcionamento, ou seja, para a fase de instalação.

A Entidade Educacional pode ser uma daquelas que se preocupa com treinamento - Área de Desenvolvimento de Recursos Humanos - ou mesmo qualquer Entidade com interesses empresariais que pretenda oferecer treinamento na área de Avaliação.

O 2º tipo de clientela é aquela que vai receber o treinamento, isto é, indivíduos com nível universitário que trabalham em Entidades Educacionais que desenvolvam ou pretendam desenvolver atividades de Avaliação constantes no Universo de Atividades identificado.

Estas Entidades Educacionais podem ser: Departamentos do Ministério da Educação e Cultura; Secretarias de Educação e Cultura; Universidades; Instituições que desenvolvem Projetos educacionais utilizando ou não multimeios.

1.4 - REVISÃO DE LITERATURA: MODELOS INSTRUCCIONAIS E DE TREINAMENTO

Considerando a eficácia o objetivo central de um sistema de treinamento, abordamos o problema de modo sistêmico e integrado. A base de uma abordagem sistêmica ao treinamento está - 1) no conhecimento de quais são os objetivos comportamentais desejados que o aluno de verá demonstrar ao final do treinamento e 2) no desenvolvimento de um curso que forneça o que é necessário para que os alunos todos alcancem todos os objetivos desejados.

Para nos orientar na identificação desses elementos, partimos para uma revisão de modelos de desenvolvimento de Sistemas Instrucionais e de Treinamento. Revisamos os modelos:

- . Serviço de Informação do Governo (U.S. Civil Service Commission 1969).
- . "Design" de Sistemas Instrucionais (Banathy, 1968)
- . Modelo de Sistema de Treinamento e Desenvolvimento (Tracey, 1971).
- . Modelo de Desenvolvimento de Instrução Vocacional (Mager e Beach, 1967).
- . Modelo para o Desenvolvimento de Sistemas Instrucionais (Branson, et al, 1973).

- . Modelo Operacional de Sistema Instrucional Naval (Dederick, Sturge, 1975).

A escolha de um modelo a ser seguido, teve como base duas restrições:

- a limitação de tempo disponível das autoras para a execução do Planejamento do curso de Treinamento;
- economia de recursos humanos e materiais necessários a qualquer entidade para o Desenvolvimento, Validação e Instalação do curso de treinamento planejado.

Isto nos levou a analisar os modelos citados em busca de um que fosse, ao mesmo tempo, simples, operacionalizado e prático, nesta situação, e que considerasse os preceitos básicos da Tecnologia Educacional:

- . abordagem sistêmica do processo da aprendizagem e da instrução;
- . utilização de objetivos comportamentais;
- . embasamento em teorias de Aprendizagem e Comunicação.

MODELO DO SERVIÇO DE INFORMAÇÃO DO GOVERNO

O modelo ISG (Figura I.1) - Serviço de Informação do Governo (1969), foi criado em 1968 e é o modelo básico de Análise de Sistemas, para a utilização em problemas que requerem uma ação governamental. Foi aplicado por Lawrence Carter em 1969.

O modelo é geral, sendo aplicável a uma grande variedade de situações. Porém, como o objetivo de nosso trabalho está definido: - planejamento de um curso de treinamento - consideramos mais adequado procurar um modelo mais específico ao caso.

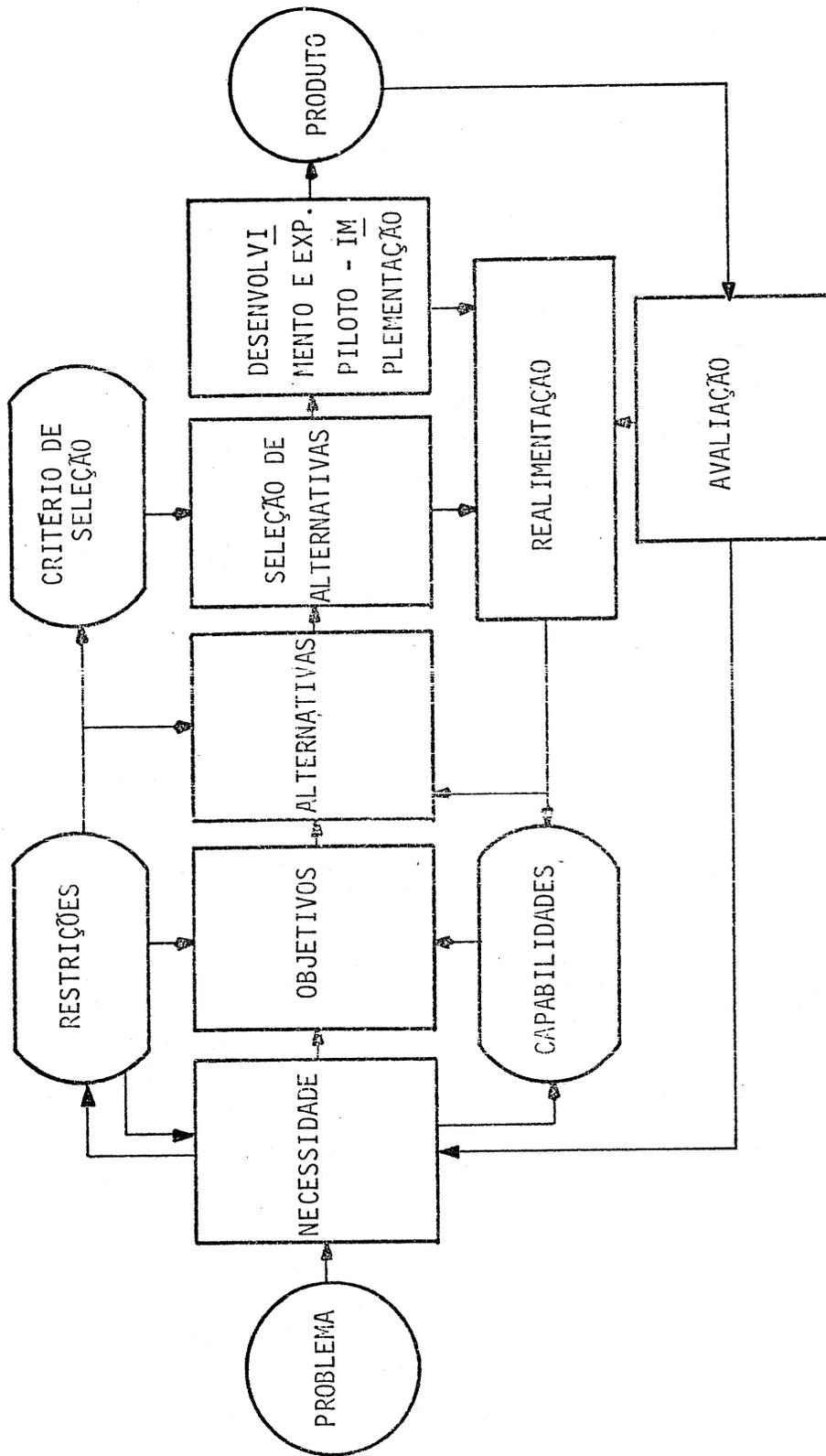


Fig. I.1 - Modelo Serviço de Informação do Governo

"DESIGN" DE SISTEMAS INSTRUCIONAIS

Banathy (1968) elaborou seu modelo "Design" de Sistemas Instrucionais em 1968 com o propósito de explicar e demonstrar o uso de abordagem de sistemas no desenvolvimento de Sistemas Instrucionais.

O modelo (Figura I.2) é apresentado em forma de diagrama através de 4 blocos interrelacionados e iterativos. O modelo é envolvido por uma linha de realimentação indicando retorno aos blocos.

Apesar do modelo não ser específico de treinamento, ele poderia ser adaptado a este propósito, se fosse realizada uma Análise Ocupacional antes da fase de Análise e Formulação de Objetivos e inserido, na fase de Implementação e Controle de Qualidade, um "follow up" dos treinados, no exercício de suas funções.

A utilização do modelo fica, também, dificultada, pelo fato de a distribuição de funções aos componentes, no bloco "design" do sistema, não estar totalmente operacionalizada de modo a permitir seu uso imediato. Esta operacionalização implicaria em um gasto de tempo não disponível pelas autoras.

O modelo não envolve, explicitamente, o desenvolvimento e validação de materiais instrucionais.

O modelo, portanto, não é inteiramente adequado ao nosso propósito.

MODELO DE SISTEMA DE TREINAMENTO E DESENVOLVIMENTO

Tracey (1971) formalizou seu modelo em 1968. Dirigiu-o inicialmente aos administradores de treinamento, com o objetivo de fornecer-lhes um procedimento sistemático para o "design" de sistemas de treinamento e desenvolvimento.

O modelo é apresentado na forma de um diagrama (Figura I.3). Representa um sistema em malha-fechada, numa sequência contínua de passos em um ciclo constante de avaliação, retroalimentação e melhoria.

O modelo é específico para treinamento. Suas fases constituem um esquema completo: requisitos, desenvolvimento e validação de um sistema.

Porém, na descrição dos componentes, a qualidade do modelo fica prejudicada:

- 1 - é proposta a definição de objetivos comportamentais sem explicitar o tipo de aprendizagem (habilidade), o que interfere nos componentes subsequentes, ficando a teoria de aprendizagem desligada do modelo.
- 2 - não trata explicitamente do desenvolvimento e validação dos materiais instrucionais.

A ênfase de seu modelo está na seleção de instrutores sendo que os meios são considerados auxílios de treinamento.

Uma contribuição importante dele é a orientação para a construção de instrumentais de medida para avaliar conteúdo, sua sequência, a estratégia utilizada, pessoal, facilidades, equipamentos e materiais.

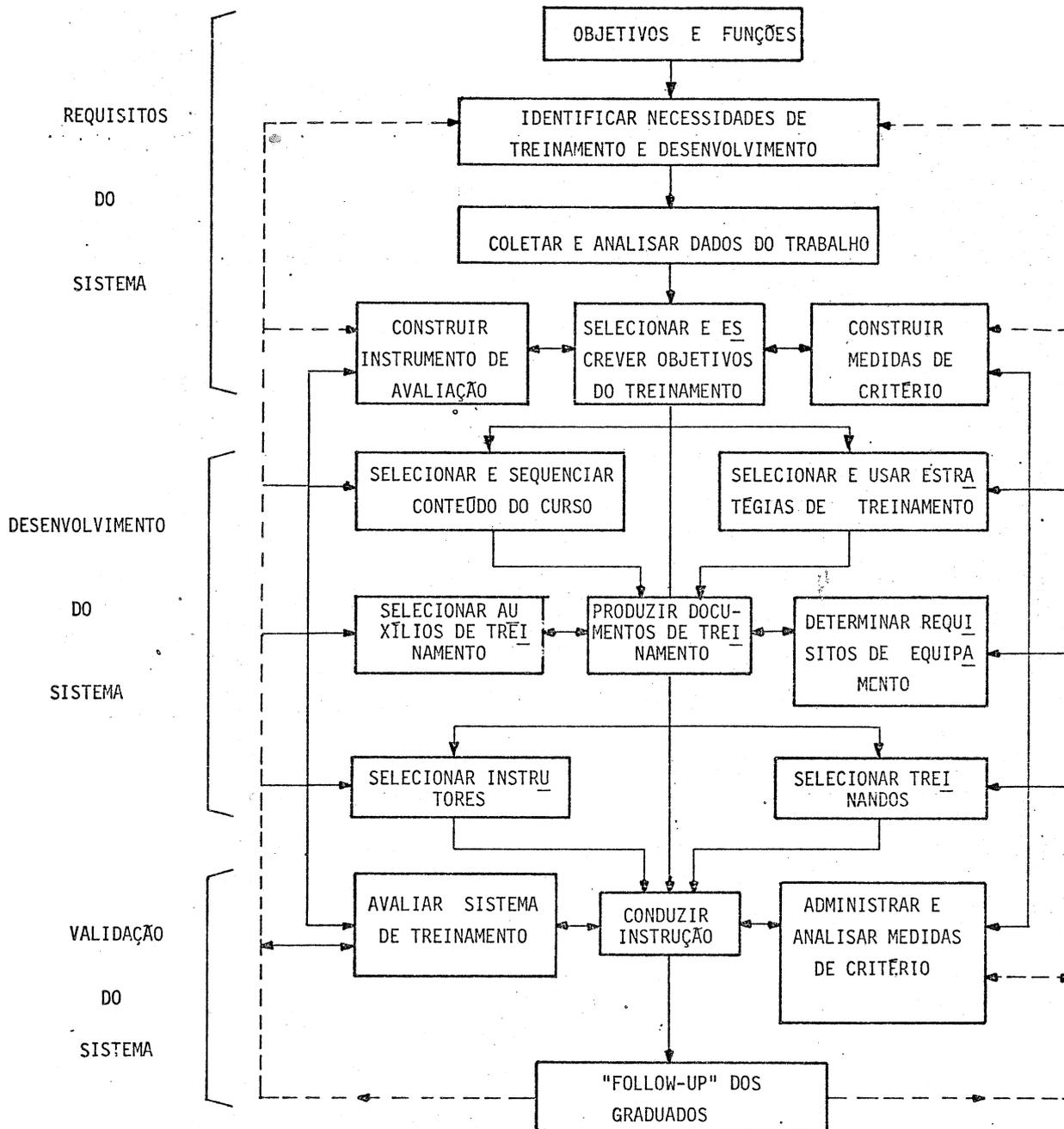


Fig. I.3 - Modelo de Sistema de Treinamento e Desenvolvimento

O modelo, no seu todo, não se adequa ao nosso propósito, sendo que alguns componentes prestaram ao mesmo tal como se apresentam e outros com reformulações, e outros ainda deveriam ser acrescentados.

MODELO DE DESENVOLVIMENTO DE INSTRUÇÃO VOCACIONAL

Uma versão mais condensada deste modelo foi preparada em consequência de um contrato dos autores Mager e Beach (1967) com o "United States Department of Health, Education, and Welfare, Office of Education". A presente versão data de 1967, e é resultado de revisões e ampliações da primeira.

A Figura I.4, representa um diagrama do modelo.

O modelo constitui um esquema simples que vincula a teoria de aprendizagem ao sistema de treinamento. Porém, não envolve, explicitamente, o desenvolvimento e validação de material.

O modelo, no seu todo, não é adequado ao nosso propósito, porém, alguns componentes são úteis tal como se apresentam.

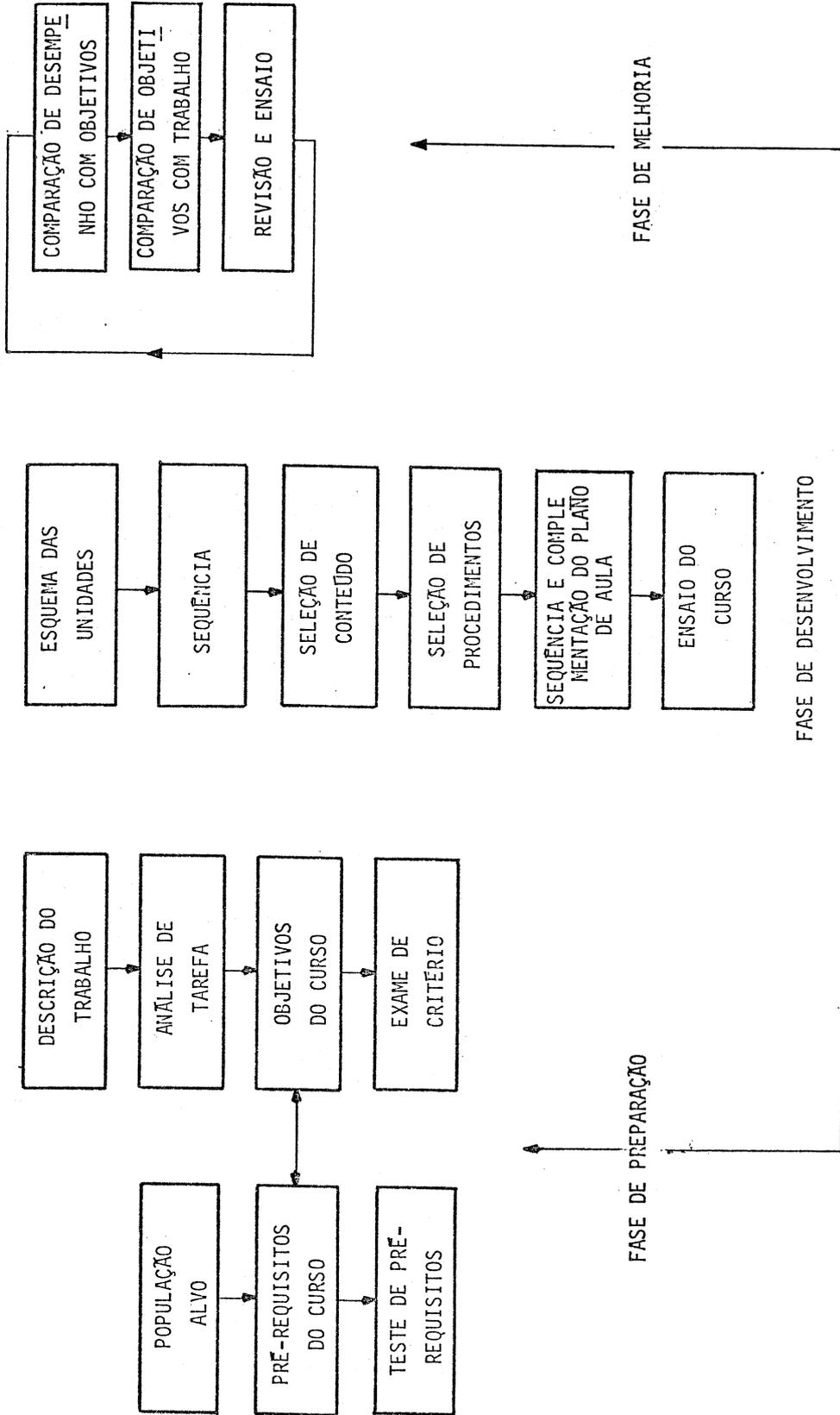


Fig. I.4 - Modelo de Desenvolvimento de Instrução Vocacional

MODELO PARA O DESENVOLVIMENTO DE SISTEMAS INSTRUCCIONAIS

O modelo de Branson e outros - ISD (1973) - Modelo para o Desenvolvimento de Sistemas Instrucionais (Figura I.5) foi elaborado em 1973 com o propósito de desenvolver Sistemas Instrucionais.

O modelo ISD é específico para o desenvolvimento de sistemas instrucionais. Suas fases constituem um esquema completo: planejamento, desenvolvimento instrucional e entrega.

Dentre os modelos revistos é o mais completo, pois apresenta todos os passos necessários para o desenvolvimento de um Sistema Instrucional. É o mais operacionalizado, pois especifica detalhadamente procedimentos para cada um dos passos.

Seria o mais adequado para o planejamento de nosso sistema de treinamento, se não estivéssemos limitadas às duas restrições básicas já citadas na página 8.

Sua utilização envolve o dispêndio de muito tempo, recursos humanos e materiais, o que o tornaria impraticável, em nossa situação, uma vez que nós temos uma restrição de tempo e que provavelmente haveria poucos recursos humanos e materiais disponíveis para a execução das fases de desenvolvimento e validação do curso.

I - Fase de Planejamento

II - Fase de Desenvolvimento/Instrucional

III - Fase de Entrega

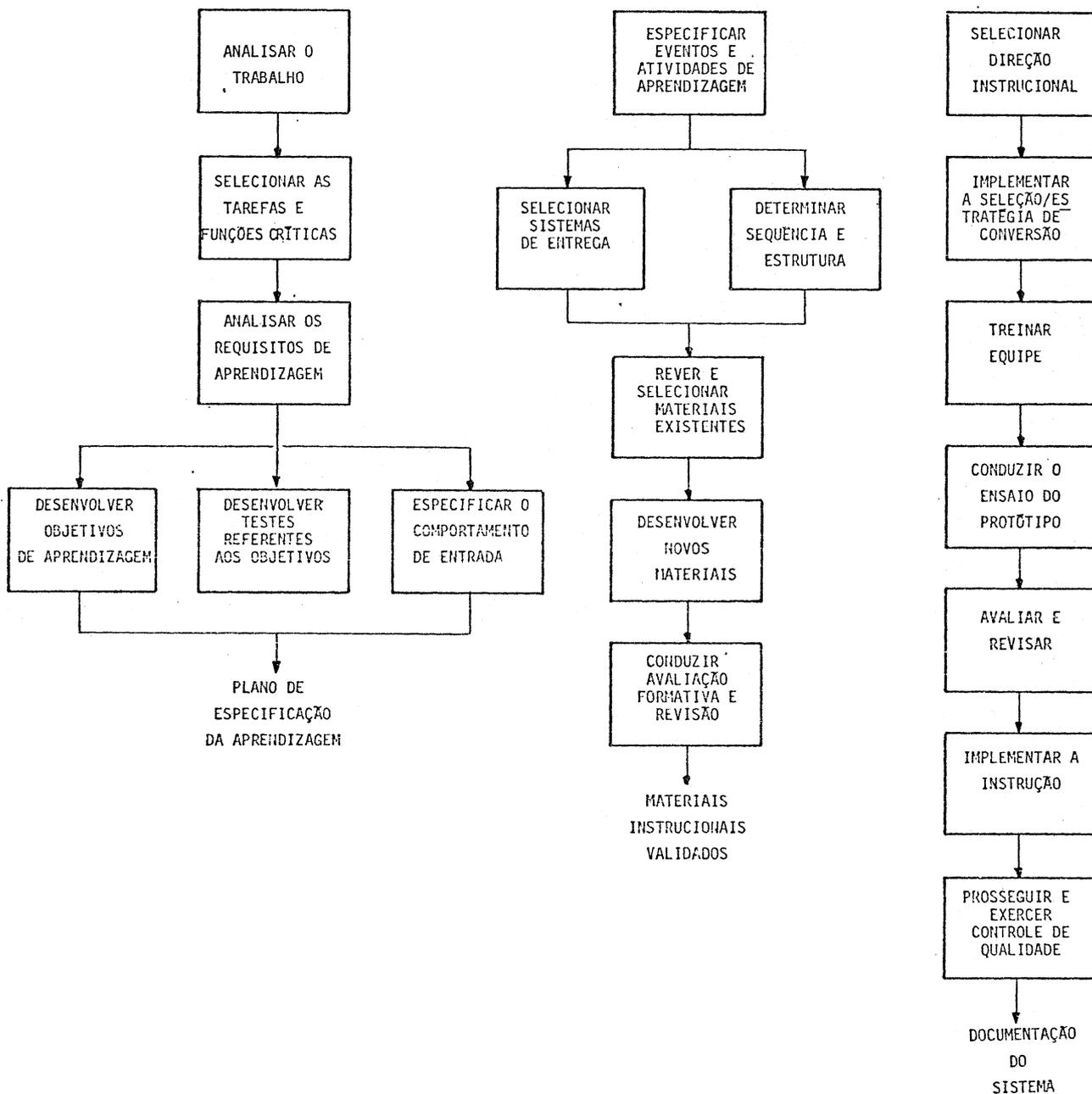


Fig. I.5 - Modelo para o Desenvolvimento de Sistemas Instrucionais

MODELO OPERACIONAL DE SISTEMAS INSTRUCCIONAIS

O modelo (Figura I.6) surgiu a partir das Diretrizes e Política Educacional Naval e do conhecimento na área de Análise de Sistemas. Os seus autores acreditam que ele pode ser generalizado para o desenvolvimento de programas educacionais em muitas outras áreas.

É um modelo específico para treinamento.

Um componente importante de seu modelo é a análise estrutural que sequencia os objetivos comportamentais. Esta análise permite uma reformulação de objetivos, bem como a identificação de objetivos necessários, omitidos no componente anterior.

Uma colocação importante deste modelo, é que a validação externa só deverá ocorrer após a comprovação da validade interna do sistema.

O modelo parece adequado ao nosso propósito, porém a falta de bibliografia relativa ao mesmo, impediu sua utilização.

Pela revisão dos modelos de treinamento em termos de adequação ao nosso propósito resolvemos pela não adoção de nenhum deles em sua íntegra.

Retiramos, porém, de cada um aquilo que consideramos importante ao nosso objetivo, e acrescentamos um passo necessário para a identificação do Universo de Atividades, que caracteriza o trabalho de Avaliação. O resultado disto se encontra, em forma do esquema apresentado na Figura 1.7, que nos orientou no planejamento do curso de treinamento.

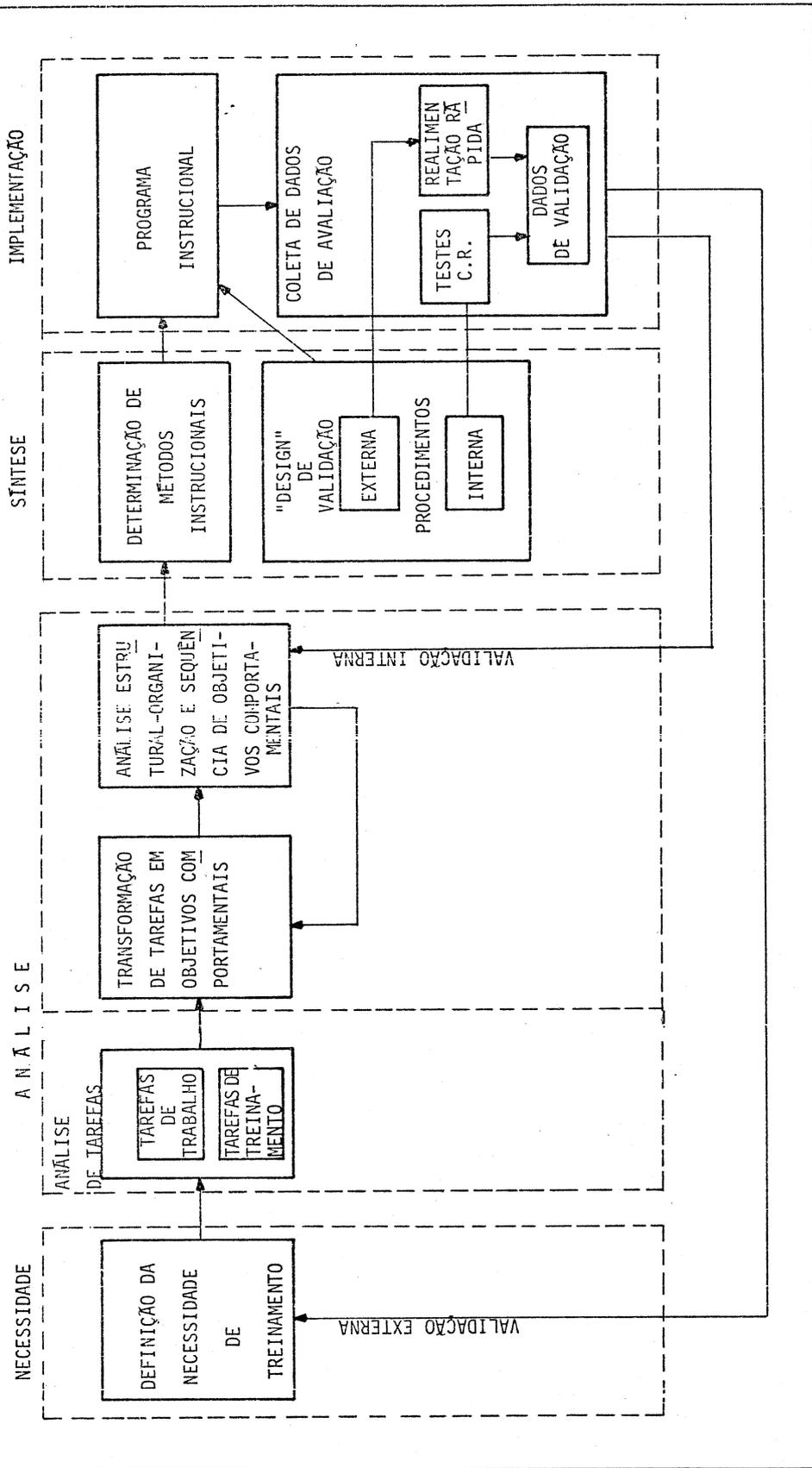


Fig. I.6 - Modelo Operacional de Sistema Instrucional

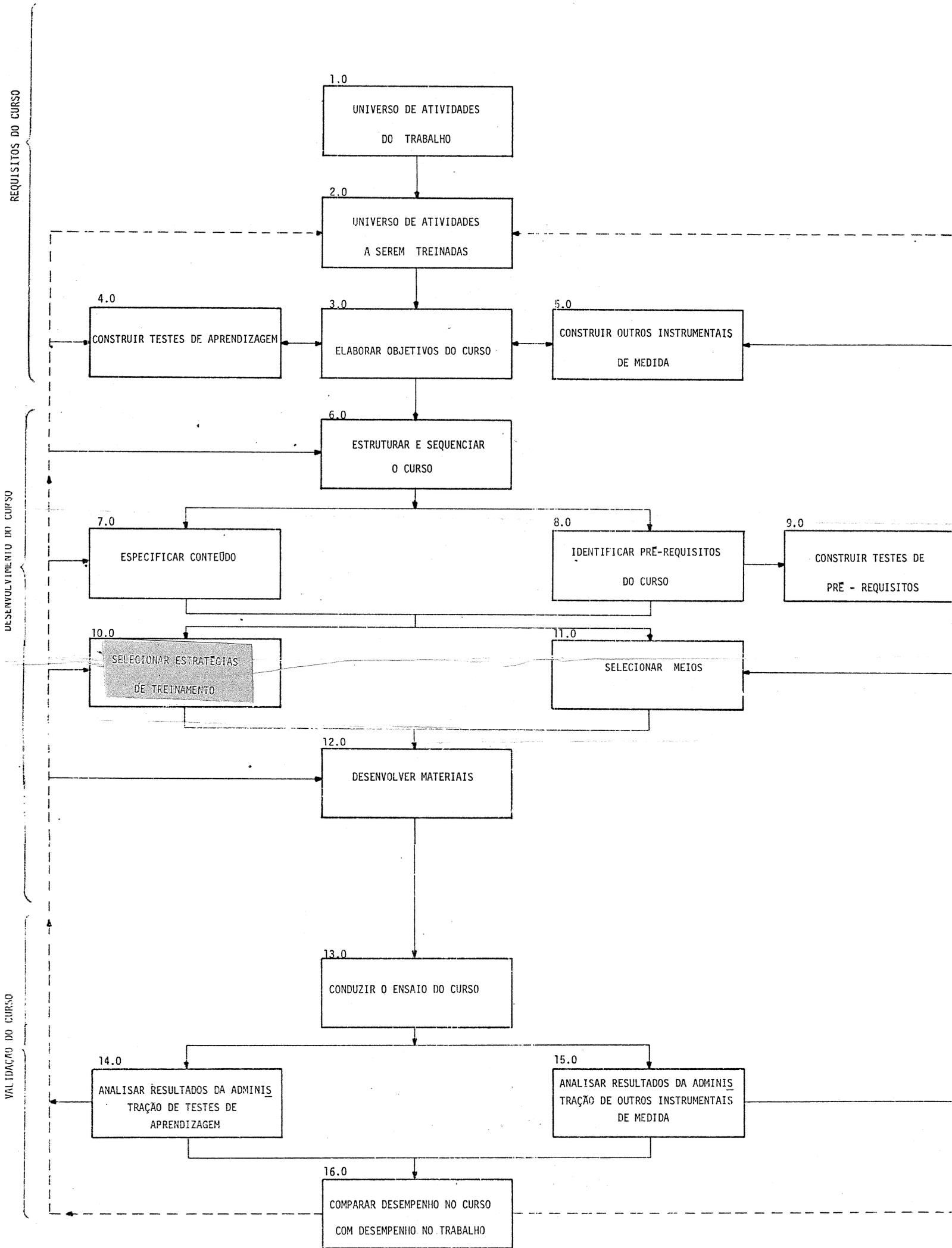


Fig. I.7 - Esquema para o Planejamento do Curso

Descrição sucinta dos componentes do esquema para o planejamento do curso de treinamento.

O esquema contém 3 etapas:

- requisitos
 - desenvolvimento
 - validação
- } do curso

REQUISITOS DO CURSO

1.0 - UNIVERSO DE ATIVIDADES DO TRABALHO

Consiste na identificação de atividades que poderiam ser treinadas. Esta identificação é embasada no confronto entre um universo teórico e um universo real de atividades. O universo teórico de atividades, ou seja, as atividades que o trabalhador deveria desempenhar, deve ser identificado através de uma revisão de literatura do trabalho em questão. O universo real de atividades, isto é, o das atividades, que o trabalhador vem desempenhando na realidade, deve ser identificado através de levantamentos em Entidades que possuem as atividades em pauta. Estes levantamentos têm como objetivo analisar o trabalho, através da obtenção de informações, utilizando métodos como observação, entrevista e questionário.

O confronto entre os universos pode mostrar que são equivalentes, que o universo real é maior que o teórico ou vice-versa.

2.0 - UNIVERSO DE ATIVIDADES A SEREM TREINADAS

Trata-se da seleção, dentre as atividades que podem ser treinadas, daquelas que serão incluídas no treinamento. Esta seleção deve ser feita com base nos dados do trabalho e critérios como:

- . dificuldade
 - . universalidade
 - . crucialidade
 - . frequência
 - . necessidade
- Dificuldade: de acordo com este critério as atividades cuja execução exige habilidades de complexidade média, devem ser incluídas na lista de objetivos do treinamento (Tracey, 1968);
- Universalidade: segundo este critério, devem ser incluídas, no treinamento, as atividades que são requeridas, não importa onde ou quando o trabalho de avaliação seja realizado (Tracey, 1968);
- Crucialidade: segundo este critério, uma habilidade específica, embora não frequente, mas sendo crítica para um desempenho adequado no trabalho, deve ser enfatizada num treinamento (Tracey, 1968);
- Frequência: as atividades executadas frequentemente, no trabalho, devem ser consideradas no treinamento;
- Necessidade: aquelas atividades que, não sendo realizadas na prática, são apontadas pelos decisores como desejáveis, devem ser consideradas no treinamento.

A aplicação desses critérios deve ser sustentada por julgamento profissional de especialistas.

3.0 - ELABORAR OBJETIVOS DO CURSO

A partir das atividades selecionadas para o curso de treinamento são elaborados os seus objetivos instrucionais. Os objetivos descrevem o que se espera do aluno ao final da instrução. Representam afirmativas claras e intenções instrucionais. Os objetivos devem conter um verbo no infinito expressando a habilidade (Gagné, 1973).

Os objetivos são continuamente modificados se revelarem falhas ou expectativas irrealísticas.

4.0 - CONTRUIR TESTES DE APRENDIZAGEM

Estes são construídos com base nos objetivos instrucionais. O propósito é determinar se, ao final de um segmento da instrução, os alunos têm desempenho coincidente com o desempenho pedido no objetivo. Tais testes devem ser construídos dentro da abordagem com referência a critério (CR). Nesta etapa, inicia-se o processo de avaliação formativa dos mesmos, através da revisão de especialista e testagem com indivíduos da clientela alvo. Estes testes são importantes para a validação e avaliação do sistema instrucional e seus componentes.

5.0 - CONSTRUIR OUTROS INSTRUMENTAIS DE MEDIDA

Trata-se da elaboração de outros instrumentais de medida (questionários, entrevistas, escalas) que são utilizados na validação e avaliação do sistema e seus componentes, naqueles aspectos que não podem ser avaliados por testes de aprendizagem. Nesta etapa inicia-se o processo de avaliação formativa destes instrumentais, através da revisão de especialista e testagem face a face com indivíduos da clientela alvo.

DESENVOLVIMENTO DO CURSO

6.0 - ESTRUTURAR E SEQUENCIAR O CURSO

Refere-se à organização dos objetivos das unidades e sub-unidades com base na lógica da matéria e na hierarquia de habilidades expressas nos objetivos (Gagné, 1973).

7.0 - ESPECIFICAR CONTEÚDO

Com base nos objetivos instrucionais é especificado o conteúdo, isto é, as informações necessárias para o domínio dos objetivos.

8.0 - IDENTIFICAR PRÉ-REQUISITOS DO CURSO

À medida que se especifica o conteúdo, deve-se identificar os pré-requisitos do curso. A determinação dos pré-requisitos é feita levando em conta, também, a população-alvo. Se os pré-requisitos referem-se às habilidades de uma população de estudantes muito diferente da pretendida, algum ajustamento precisa ser feito. Talvez alguns pré-requisitos devam ser incluídos nos objetivos do curso.

9.0 - CONSTRUIR TESTE DE PRÉ-REQUISITOS

São construídos com base nos pré-requisitos estabelecidos e têm o propósito de verificar se o indivíduo tem a habilidade que foi suposta. O teste deve ser aplicado aos alunos antes do início do curso. O resultado desta aplicação pode levar ou à incorporação de outros objetivos (antes supostos como pré-requisitos) ao curso, ou à necessidade de instrução remedial, antes de se permitir que os alunos frequentem o curso.

Nesta fase inicia-se o processo de avaliação formativa do teste de pré-requisito.

10.0 - SELECIONAR ESTRATÉGIAS DE TREINAMENTO

Trata-se da seleção de técnicas de ensino, organização de instrutores e treinandos, com o propósito de alcançar os objetivos instrucionais. A escolha da melhor estratégia deve ser de acordo com os objetivos da instrução, levando-se em consideração a natureza do conteúdo.

11.0 - SELECIONAR MEIOS

A partir dos objetivos da instrução e considerando as estratégias de treinamento, são selecionados os meios mais adequados a cada um deles. A seleção dos meios deve ser orientada pelas características e disponibilidades dos mesmos e pela natureza do conteúdo.

12.0 - DESENVOLVER MATERIAIS

Nesta etapa, inicia-se o desenvolvimento dos materiais instrucionais que contêm a informação necessária à consecução dos objetivos do curso. Aqui, estes materiais são preparados, revisados por especialista e testados face a face com indivíduos da clientela alvo.

VALIDAÇÃO DO CURSO

13.0 - CONDUZIR ENSAIO DO CURSO

No ensaio conduz-se a instrução, em pequena escala, para testar o curso, coletando-se dados para o desenvolvimento das atividades 14.0 e 15.0.

Sõ apõs o ensaio o curso pode ser aceito como um curso regular.

14.0 - ANALISAR RESULTADOS DA ADMINISTRAÇÃO DE TESTES DE APRENDIZAGEM

Consiste em responder a questão:

"Quão bem os estudantes alcançaram os objetivos instrucionais especificados?", através da análise dos resultados dos testes de aprendizagem.

15.0 - ANALISAR RESULTADOS DA ADMINISTRAÇÃO DE OUTROS INSTRUMENTAIS DE MEDIDA

Consiste na análise dos resultados da aplicação dos ins

instrumentais de medida obtidos dos treinandos instrutores, e administradores com o objetivo de responder a questão. "Quão bem o sistema está operando?" para permitir retroalimentação sistemática no mesmo.

As atividades desenvolvidas em 14.0 e 15.0 consistem no processo de validação interna do sistema. Seu foco é o planejamento do curso em si, verificando se o conteúdo, sua sequência, procedimentos de instrução e materiais instrucionais são consistentes com os objetivos comportamentais e apropriados aos treinandos, portanto, nestas etapas, conclui-se a avaliação formativa do teste de pré-requisitos, testes de aprendizagem, outros instrumentais de medida, materiais instrucionais e realiza-se a avaliação somativa do curso.

16.0 - COMPARAR DESEMPENHO NO CURSO COM DESEMPENHO NO TRABALHO

Trata-se da validação externa do sistema, que indica se os objetivos estão de acordo com as necessidades reais do trabalho. Esta informação pode ser obtida através de entrevistas ou questionários aplicados aos treinados quando no exercício do trabalho para o qual foram preparados. Pode-se também obter informações junto aos supervisores desses trabalhadores.

O sistema terá validade externa se as informações coletadas, na realidade do trabalho, indicarem que o curso auxiliou na melhoria do desempenho dos treinados no trabalho. Caso contrário, deve-se corrigir a definição original dos objetivos.

De acordo com os resultados dos componentes 14.0, 15.0 e 16.0 o sistema deve ser revisado e novamente executado, num ciclo constante de avaliação, retroalimentação e melhoria.

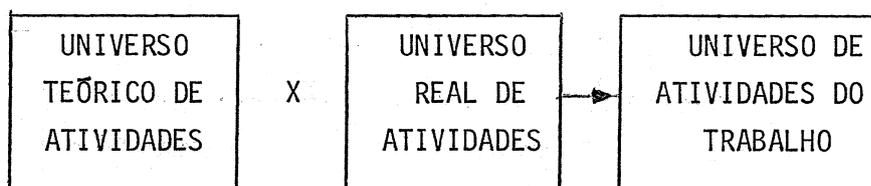
1.5 - METODOLOGIA DO TRABALHO

A partir do esquema adotado (Figura I.7), foi determinada a metodologia seguida na consecução deste trabalho.

Para que os programas de treinamento forneçam um retorno adequado do investimento, eles precisam ser relevantes ao trabalho, isto é, as experiências de aprendizagem precisam ser diretamente e explicitamente relacionadas às atividades que os indivíduos devem realizar, quando no exercício da função.

No caso da Avaliação, apesar deste não ser um conceito novo, sendo, inclusive, atualmente bastante explorado na teoria, não existe ainda um consenso de quais atividades caracterizam tal função.

Conscientes desta lacuna sentimos que seria essencial, para o planejamento do curso de treinamento, iniciar com a identificação de um Universo de Atividades que caracterize a função do Avaliador. Este universo foi obtido através da identificação de um Universo Teórico e de um Universo Real de atividades em termos de Avaliação:



O Universo Teórico de atividades, ou seja, as atividades que o avaliador deveria desempenhar, foi identificado através de uma revisão de literatura de Avaliação (página 29). O Universo Real de atividades, isto é, as atividades que o avaliador vem desempenhando na nossa realidade, foi identificado através de levantamentos em Entidades Educacionais que possuem atividades avaliativas. Para a realização deste levantamento fizemos uma revisão da literatura de análise ocupacional (página 38) buscando técnicas para a descrição da função do avaliador.

A comparação entre o teórico e o real pretende conduzir ao planejamento de um curso de treinamento em que os sujeitos serão

treinados para o exercício de atividades ditadas por uma combinação, visando o equilíbrio entre atividades definidas na teoria e aquelas exercidas na realidade.

Identificado o Universo de Atividades do trabalho foi necessário limitá-lo, para que a duração do curso não tornasse este último impraticável. Isto foi feito através da aplicação dos critérios de dificuldade, universalidade, crucialidade, frequência e necessidade (página 76) que conduziu à identificação de um Universo de Atividades que serão treinadas.

A partir destas atividades foram elaborados os objetivos instrucionais das unidades e sub-unidades. Em todos estes houve a preocupação de incluir um verbo no infinito que representa o tipo de aprendizagem envolvido na habilidade a ser adquirida Gagné (1973). Na medida do necessário foram incluídos, nos objetivos, outros elementos propostos por Gagné e Briggs (1974) para formulação de objetivos instrucionais, ou seja, situação e verbo de ação.

A partir dos objetivos das unidades foi definida a estrutura e sequência do curso segundo uma ordem lógica dos mesmos.

Com base nos objetivos instrucionais e na sua estruturação foi relacionado, sequenciado e especificado o conteúdo do curso.

À medida que foi especificado o conteúdo, foram identificados os pré-requisitos, ou seja, as habilidades que, necessariamente, já deveriam ser dominadas pela clientela alvo.

Os passos seguintes foram a seleção de estratégias de treinamento e de meios instrucionais a serem utilizadas para a consecução dos objetivos propostos. Para tal, foi feita uma revisão da literatura pertinente (página 43 e 47, respectivamente). Para a seleção de meios adotamos o modelo de Bretz, (1971).

Utilizando técnicas de engenharia de sistemas foi planejado o ensaio e a instalação do curso.

1.6 - REVISÃO DE LITERATURA COMPLEMENTAR

Conforme citado nas páginas 27 e 28, da revisão de literatura sobre modelos de treinamento, foi necessário revisar a literatura acerca da teoria de Avaliação Educacional, teoria de Análise Ocupacional, Estratégias Instrucionais e Modelos de Seleção de Meios Instrucionais.

1.6.1 - TEORIA DE AVALIAÇÃO EDUCACIONAL

O propósito desta revisão de literatura foi a identificação de um universo teórico de atividades que caracterizasse a função do avaliador. Considerando que o tipo de atividade desempenhado pelo avaliador está diretamente relacionado com o conceito de Avaliação adotado, iniciamos com um estudo daqueles mais representativos no passado e mais relevantes na literatura atual.

Inicialmente revisamos as definições de avaliação como:

- 1a.) medida
- 2a.) congruência
- 3a.) julgamento

A 1a. definição surgiu na década de 1920-1930, (Stufflebeam et al, 1971) identificando avaliação como medida. Esta Avaliação era vista apenas como um processo de desenvolvimento e interpretação de instrumentos. Assim, no processo instrucional, somente aquelas variáveis passíveis de mensuração, podiam ser avaliadas. Outras variáveis, que não podiam ser medidas, não eram consideradas.

A Avaliação dentro desta definição tem, portanto, um enfoque limitado e uma abordagem mecânica.

A definição de congruência foi proposta por Tyler por volta de 1950 (Stufflebeam et al, 1971). Esta definição considera a Avaliação como um processo de determinar o grau de consecução dos objetivos educacionais que um curso ou programa procura alcançar. Em outras palavras, a Avaliação procura determinar a congruência entre desempenho do aluno e objetivos comportamentais estabelecidos.

A Avaliação, nesta definição, tem sido entendida, comumente, como uma técnica terminal.

A 3a. definição enfatiza a Avaliação como um processo de julgamento de profissionais. Um grupo de especialistas, em diferentes áreas, emite julgamentos de valor a respeito de um currículo, um estudante, uma escola, etc.

Dentro desta abordagem, na maioria das vezes, não existem critérios bem definidos para o julgamento.

Dentre as abordagens mais atuais, revisamos as de Stake (1973), Scriven (1973), Provus (1973) e Stufflebeam et al (1971). Tais autores foram revisados, por serem os mais difundidos entre nós e serem representativos do que existe atualmente em termos de Avaliação.

Dentre estas abordagens, a Avaliação passou a envolver um julgamento do valor dos fenômenos educacionais.

Stake (1973) define a Avaliação como um processo de descrever e julgar programas educacionais através de investigação formal. Nesta abordagem, descrição e julgamento de dados relativos ao programa educacional estudado, são consideradas atividades básicas do avaliador.

Estes dois termos chaves, nesta definição, são apresentados a seguir:

19) Descrição

Na descrição de dados, o avaliador deve identificar três conjuntos de informações: antecedentes, transações e resultados.

Tais elementos são assim definidos:

Antecedentes são todas as condições que existem antes do processo de ensino, tais como estudantes, equipes, materiais, facilidades, que podem relacionar-se com os resultados.

Transações são todas as atividades realizadas para o alcance dos objetivos. Incluem relações e interações entre estudantes, professor e estudante, equipe do programa e de todo o pessoal envolvido no processo educacional.

Resultados referem-se aos produtos da aprendizagem que podem ser identificados como habilidades, realizações, atitudes e aspirações de estudantes. Incluem não só aqueles que são evidentes, como fins de aprendizagem, mas também aplicação, transferência e efeitos de reaprendizagem. Os resultados podem ser consequências, imediatas ou a longo prazo, da educação.

Identificados estes três conjuntos de informações (antecedentes, transações e resultados) eles são descritos em termos de "Intenções" e "Observações".

O termo intenção é aqui considerado como sinônimo daquilo que é pretendido, seja em relação a antecedentes, transações ou resultados. A descrição das intenções é feita através de afirmativas explícitas das mesmas.

O termo observação refere-se ao que ocorre realmente em termos de antecedentes, transações e resultados.

Os dados descritos são processados de duas maneiras:

- identificando contingência entre antecedentes, transações e resultados, pretendidos e observados.
- procurando congruência entre "intenções" e "observações".

Contingência - refere-se a lógica do programa. Há contingência, quando existe uma conexão lógica entre antecedentes, transações e resultados.

Congruência - existe congruência, quando o que estava sendo pretendido, ocorre realmente, ou seja, quando não há discrepância entre antecedentes, transações e resultados pretendidos e observados.

29) Julgamento

Julgamento é aqui definido como o ato de decidir sobre que padrões considerar, para a avaliação de programas educacionais.

Para julgar as características de um programa, o avaliador pode considerar duas bases: (1) padrões relativos, que se referem às características de programas alternativos; (2) padrões absolutos, que se referem a julgamentos pessoais.

O julgamento final, feito com base em dados relativos ou absolutos, deve prover uma resposta imediata para a tomada de decisões educacionais.

Na abordagem de Stake a Avaliação é orientada mais para os programas educacionais em si mesmo, do que para seus produtos.

Para Scriven (1973) a Avaliação consiste:

- na coleta sistemática de dados de desempenho, na combinação des

tes dados com um conjunto de metas valoradas, e na análise desta informação com o propósito de julgar o valor ou mérito de um programa educacional.

- em justificar:

- . instrumentos para a coleta de dados
- . atribuição de pesos aos objetivos
- . seleção de critérios e padrões.

A Avaliação está diretamente relacionada com os objetivos do programa. O avaliador identifica estes objetivos, julga o seu valor e verifica o seu alcance.

Segundo Scriven, dois tipos de padrões podem ser considerados na avaliação de programas educacionais: padrões absolutos e relativos.

Em relação ao seu papel, a Avaliação pode ser formativa ou somativa, dependendo se é feita para conduzir à melhoria, ou à determinação do mérito de um programa educacional.

Scriven relaciona, ainda, a estes papéis formativo e somativo, dois tipos de Avaliação: intrínseca, relativa ao conteúdo do material ou currículo, e extrínseca, relativa aos seus efeitos.

A Tabela I.1, apresentada por Sanders (1973) visualiza estas relações.

DISTINÇÃO ENTRE AVALIAÇÃO INTRÍNSECA E EXTRÍNSECA E
AVALIAÇÃO FORMATIVA E SOMATIVA

	INTRÍNSECA	EXTRÍNSECA
FORMATIVA	Julgamento da integridade (ex. estrutura e sequência de conteúdo) do conteúdo.	Julgamento dos efeitos intermediários para fornecer realimentação aos elaboradores de materiais.
SOMATIVA	Julgamento final dos materiais	Julgamento final dos efeitos dos materiais, em alunos, professores, pais, administradores, etc.

Para Provus (1973) - "Avaliação é o processo de:

- a) levar a um acordo em relação aos padrões do programa.
- b) determinar se existe discrepância, entre algum aspecto do programa e os padrões que governam aquele aspecto,
- c) usar a informação de discrepância para identificar os "pontos fracos do programa", com o propósito de monitorar e eventualmente julgar a efetividade de programas, determinando se devem ser melhorados, conservados como estão ou eliminados".

O modelo de Avaliação, proposto por Provus, aplica-se aos programas em desenvolvimento e com pessoal já alocado. Tais programas contêm quatro estágios maiores de desenvolvimento, e o processo de Avaliação se desloca através de cada um deles: 1º) definição, 2º) instalação, 3º) processo e 4º) produto.

No primeiro estágio a tarefa do avaliador é obter uma definição do programa, com base numa taxionomia de conteúdo (um modelo ge

ral que auxilia a definição do programa em termos de seus "inputs", processos e "outputs"). A definição obtida torna-se a informação do desempenho do programa, a ser comparado com a taxionomia que constitui o padrão de desempenho deste estágio.

No segundo estágio, a informação do desempenho do programa é obtida através de observações de campo, relativas à instalação dos componentes do programa. Neste estágio o padrão para comparação é a definição do programa, obtida no primeiro estágio.

No terceiro estágio, em que o programa é lançado à população alvo, a informação, sobre o desempenho do programa, é obtida em termos de seus processos e efeitos iniciais do tratamento. Neste estágio, o padrão de comparação é aquela parte da definição que descreve relações entre processos do programa e objetivos de capacitação.

No quarto estágio, as informações de desempenho do programa são relativas às medidas de alcance de objetivos terminais do projeto, sendo seu padrão de comparação aquela parte da definição do programa referente aos objetivos terminais.

A informação de discrepância é utilizada em todos os estágios de desenvolvimento e leva sempre a uma decisão que pode ser de:

- ir ao estágio seguinte,
- reciclar o estágio depois de mudanças nos padrões ou operações do programa,
- reciclar para o primeiro estágio,
- interromper o programa.

Assim, a Avaliação é um processo contínuo de fornecimento de informações de discrepância, pertinentes e confiáveis, que servem como base de decisões administrativas.

Stufflebeam et al (1971) define a Avaliação como "um processo de delinear, obter e prover informações úteis para o julgamento de alternativas de decisão".

A Avaliação é, então, vista como um processo contínuo, que fornece informações relevantes em todas as fases e momentos do desenvolvimento de um programa. Quando, porém, acontece de a Avaliação não ter sido conduzida desde o início do programa, ela pode atuar em qualquer um dos estágios.

A Avaliação tem, essencialmente, uma função de assistência aos decisores e administradores pelo fornecimento de informações úteis a serem utilizadas na tomada de decisão educacional. Nesta decisão está implícito um julgamento de valor. O decisor possui um sistema de valores, e com base neste, especifica critérios de decisão. Cada alternativa é então pesada, em função destes critérios. E, para a determinação dos pesos de cada uma das alternativas, em função dos critérios especificados, o avaliador delinea, obtém e provê informações úteis, ao decisor, para que ele possa julgar cada uma das ações em relação aos seus valores.

Nesta definição, um grande destaque é dado ao processo de julgamento, e ao objetivo com que é feito este julgamento, isto é, a tomada de decisão.

Stufflebeam et al (1971) identifica 4 tipos de decisão dentro de um programa educacional:

- . decisões de planejamento para determinar os objetivos.
- . decisões de estruturação para planejar procedimentos, que conduzam ao alcance dos objetivos.
- . decisões de implementação para utilizar, controlar e refinar procedimentos.
- . decisões de reciclagem para julgar e reatuar os produtos.

Relacionado a cada um destes tipos de decisão, ele pro
põe respectivamente quatro tipos de Avaliação: Contexto, "Input", Pro
ceso e Produto.

A Avaliação de Contexto provê uma lógica para determi
nação dos objetivos de um programa.

A Avaliação de "Input" provê informações para determinar
como utilizar recursos, para encontrar os objetivos do programa. Ela
identifica e avalia as estratégias relevantes, as capacidades e pro
cedimentos da implementação.

A Avaliação de Processo provê realimentação peri
odica, às pessoas responsáveis pelos procedimentos e plano de implementação, a
fim de antecipar e superar dificuldades relativas a procedimentos.

A Avaliação de Produto verifica a extensão na qual os
fins estão sendo alcançados, através da medida e interpretação dos efei
tos da mudança.

Os dados obtidos através destes quatro tipos de Avalia
ção, desde que obedeam aos critérios científicos, práticos e de efi
ciência, resultam em informações úteis que podem ser consideradas essen
ciais em toda situação que envolva mudanças educacionais.

Considerando que:

- . o estudo destes conceitos de Avaliação foi feito com o prop
osito de chegarmos a uma listagem de atividades contidas, na lite
ratura, como atividades do avaliador;
- . os conceitos atuais de Avaliação concebem-na como um processo
contínuo e sistemático;
- . os conceitos atuais de Avaliação possuem uma característica cô
mum, ou seja, utilizam uma abordagem sistêmica, tratando dos
"inputs", processos e produtos dos programas educacionais, seria

indiferente a adoção de qualquer um desses conceitos para embasamento da listagem de atividades.

Porém, considerando que:

- . dentre estas abordagens, as de Provus e de Stufflebeam et al são as mais específicas para Macro Avaliação sendo, portanto, mais abrangentes, permitindo a geração de uma listagem de atividades mais ampla; prestando-se, ambas, pois, ao nosso propósito;
- . dentre ambos, o modelo de Stufflebeam et al é o mais difundido e está sendo aplicado em projetos educacionais brasileiros, inclusive a nível nacional (MEC/DSU, Cenafor e outros)*, e que existe mais bibliografia disponível a respeito.

Pareceu-nos que sua abordagem seria a mais adequada para embasar a listagem das atividades do avaliador.

Assim, de um estudo mais aprofundado da teoria de Stufflebeam et al, chegamos a um universo teórico de atividades exercidas pelo Avaliador Educacional. Este universo se encontra listado na página 65.

1.6.2 - ANÁLISE OCUPACIONAL

O propósito desta revisão de literatura foi a identificação de técnicas para determinar o universo real de atividades exercidas pelo Avaliador Educacional.

A Análise Ocupacional é uma técnica utilizada, comumente, em indústrias, e possui duas classificações gerais (Otis e Leukart, 1954) em função do propósito para o qual é realizada:

* Esta informação foi coletada através de interface com estes sistemas.

- 1 - Análise de um Trabalho do ponto de vista do trabalhador.
- 2 - Estudo de Tempo e Movimento, e Análise de Métodos, do ponto de vista do trabalho.

Estes dois tipos de Análise Ocupacional são bem distintos tanto quanto à abordagem, como às técnicas utilizadas e ao conteúdo da matéria.

A Tabela I.2 apresenta uma comparação destas duas formas de Análise Ocupacional (Otis e Leukart, 1954).

Tendo em vista o nosso propósito de planejar um Curso de Treinamento, julgamos mais adequado abordar o trabalho do ponto de vista do trabalhador.

Otis e Leukart (1954) definem a Análise Ocupacional como "o processo de, pela observação e estudo, determinar e relatar informações pertinentes, relacionada à natureza de um trabalho específico. Consiste na determinação das tarefas que compreendem o trabalho e a prática, conhecimentos, habilidades e responsabilidades, requeridas de um trabalhador, para um desempenho bem sucedido".

Flippo (1971), define a Análise Ocupacional como o "processo de estudo e coleta de informações relativas a operações e responsabilidades de um trabalho específico". A Análise Ocupacional é, portanto, o processo ou técnica de obter informações sobre um trabalho para determinar desempenhos que o caracterizam.

Para coletar informações adequadas para descrição das tarefas do trabalho e para indicação das qualificações necessárias ao trabalhador para seu desempenho, foi esquematizado no Guide for Analyzing Jobs of the War Manpower Commission a seguinte fórmula de Análise Ocupacional apresentada por Otis e Leukart (1954):

TABELA I.2

COMPARAÇÃO ENTRE DUAS FORMAS DE ANÁLISE OCUPACIONAL

Análise de um Trabalho como um Papel	Estudo de Tempo e Movimento, e Análise de Métodos
Abordagem: Do ponto de vista do trabalhador	Abordagem: Do ponto de Vista das operações do Trabalho
Principais Técnicas 1. Questionário 2. Entrevista com o trabalhador 3. Entrevista com supervisores 4. Observação	Principais Técnicas 1. Observação 2. Fotografia, Estática e Dinâmica 3. Cronometragem
Conteúdo da Matéria: relativo aos trabalhadores 1. Deveres 2. Habilidades 3. Conhecimentos 4. Responsabilidades 5. Esforços 6. Condições de Trabalho	Conteúdo da Matéria: relativo a tarefas, deveres, <u>mo</u> vimentos 1. Sequência 2. Tempo e Movimento 3. Movimentos envolvidos 4. Métodos, máquinas e equipa mentos
Propósitos e Usos 1. Promoção, transferência, alocação e seleção de pessoal. 2. Avaliação do Trabalho: administração de salários e remuneração 3. Treinamento 4. Mudanças na estrutura da organização 5. Mudanças no conteúdo do trabalho 6. Melhoria na segurança e condições de trabalho 7. Padronização de títulos e deveres do trabalho para todo o pessoal de operação 8. Base para Avaliação do empregado. 9. Levantamento e Orçamento para pessoal	Propósitos e Usos 1. Melhoria de sequência 2. Estabelecimento de padrões 3. Simplificação de Trabalho 4. Pagamento - Incentivo 5. Economia de Esforços, Tempo e Movimento 6. Melhoria de equipamento e maquinaria.

O que o trabalho faz

Como ele faz

Por Que ele faz

Habilidades Envolvidas, que incluem:

- Responsabilidade;
- Conhecimento do Trabalho;
- Aplicação Mental;
- Destreza e Acurácia;

O Que o trabalhador faz envolve atividade física como: cortar, moer, montar, regular, e também atividade mental na qual ele pode planejar, computar, dirigir, etc.

Como o trabalhador executa sua tarefa envolve fatores físicos: maquinaria, instrumentos, dispositivos de medida, rotinas e também atividades mentais: cálculo, uso de fórmulas e aplicação de julgamentos e decisão.

O Por Que, da fórmula da Análise Ocupacional, abrange o propósito ou a razão pela qual cada tarefa é desempenhada, em relação às outras tarefas, no mesmo trabalho, e a outros trabalhos no mesmo departamento ou companhia.

A habilidade exigida para a execução de um trabalho determina o grau de dificuldade do mesmo. Os 4 fatores básicos listados sob esta fórmula são responsabilidade, conhecimento do trabalho, aplicação mental, destreza e acurácia.

No "Training and Reference Manual Jobs Analyses of the Occupational Analyses Section of the United States Employment Service", citado por Otis e Leukart (1954), tais termos são detalhados como segue:

- Responsabilidade - sobre material ou produto, segurança dos outros, equipamentos ou processo, cooperação com outros, instruções de outros, contatos públicos e similares;
- Conhecimento do Trabalho - considera empregos anteriores e conhecimento do equipamento em serviço, materiais, procedimentos de trabalho, técnicas e processos;
- Aplicação Mental - considera iniciativa, adaptabilidade, julgamento independente, prontidão mental;
- Destreza e Acurácia - considera a velocidade e grau de precisão, coordenação, perícia, prudência e destreza de manipulação, operação ou processamento de materiais, ferramentas, instrumentos.

Como o produto da Análise Ocupacional resulta na descrição de uma função específica para uma Organização específica, envolvendo, portanto, detalhes que não são de nosso interesse, uma vez que pretendemos definir o Universo Real de Atividades do Avaliador a nível nacional, não consideramos todos os elementos da fórmula da Análise Ocupacional. Os elementos utilizados se encontram nas páginas 68 e 69.

Há três métodos através dos quais pode ser obtida informação sobre trabalhos - : observação, entrevista e questionário.

Para obtenção de informação sobre um trabalho, através da observação e entrevista, são sugeridos vários esquemas ou guias para registro dos desempenhos observados ou descritos pelo trabalhador. A utilização destes esquemas ou guias fornece um relatório melhor e mais completo da Análise Ocupacional, dando todos os dados fonte. Estes dados podem ser, posteriormente, rearranjados e adaptados de várias formas para preparar a informação sobre o trabalho. Tais formas incluem descrição do trabalho, especificação salarial, esquema de treinamento ou qualquer projeto similar que pode precisar certos desempenhos básicos sobre trabalhos.

Para analisar os trabalhos, o procedimento que em geral se adota é o seguinte:

- 1) Obter toda a informação pertinente disponível
- 2) Contatar o supervisor do departamento, seguindo qualquer formalidade necessária.

- 3) Discutir, com o chefe do departamento, o plano geral de coleta de dados para a análise ocupacional;
- 4) Com a ajuda do chefe, contatar outras pessoas chave dentro do departamento para enriquecer as informações;
- 5) Analisar os trabalhos chave do departamento principal, para encobrir, em maiores detalhes, os processos desenvolvidos no departamento, e para estabelecer uma estrutura básica para comparar a informação sobre outros trabalhos que dependem ou suplementem estes trabalho chave;
- 6) Verificar e obter aprovação, dos supervisores e outras pessoas no departamento, de todos os relatórios de informações sobre o trabalho.

O registro dos desempenhos detetados pela Análise Ocupacional é encontrado na Descrição e na Especificação do Trabalho. A descrição do trabalho se constitui, usualmente, em uma descrição escrita de operações, deveres, métodos, condições de trabalho, equipamento e material usado, linhas de autoridade e outros fatos essenciais sobre uma posição ou trabalho. A Especificação do trabalho é um relatório das qualidades ou habilidades que um trabalhador deve possuir para desempenhar um trabalho de uma maneira satisfatória.

Quando o propósito da Análise Ocupacional é planejar um curso de Treinamento, o registro das informações acerca dos desempenhos do trabalho é feito através do próprio esquema de treinamento.

Esta revisão de literatura embasou os procedimentos seguidos para a identificação do universo real de atividades do avaliador educacional. Tais procedimentos se encontram na página 68.

1.6.3 - ESTRATÉGIAS INSTRUCCIONAIS

Em busca de uma maneira de selecionar as estratégias instrucionais, retomamos os modelos revisados na página 7, no que se refere a esta seleção.

O modelo Serviço de Informação do Governo (1968), no componente seleção de alternativas, refere-se à necessidade de um plano para a implementação e previsão dos efeitos das alternativas sob várias condições futuras.

Banathy (1967) refere-se ao elemento análise dos componentes, em termos do "que" e "quem" deve ser empregado, para conduzir funções específicas identificadas como resultado de seu componente anterior-análise de funções (o que deve ser feito pelo sistema). As funções são:

- seleção e organização do conteúdo
- seleção e organização de experiências de aprendizagem
- direção aos alunos
- avaliação da aprendizagem e da operação do sistema

Os componentes são:

- humanos - professor, alunos e pessoal envolvido no sistema
- materiais - "software" e "hardware", como livros, materiais de instrução programada, "tapes", filmes, máquina de ensinar e outros.

A seleção de um componente deve basear-se em critérios como potencial para cumprir uma função particular, habilidade para integrar-se com outros componentes, relevância para o aprendiz, praticabilidade e economia.

Tracey (1971), no componente seleção e uso de estratégias de treinamento, trata das estratégias como uma combinação de métodos e técnicas de ensino, organização de treinandos e instrutores com o propósito de alcançar um objetivo instrucional. A seleção da melhor estratégia deve ser determinada pelos objetivos da instrução, natureza do conteúdo, população dos treinandos, instrutores, tempo, facilidades, equipamentos e materiais disponíveis e exigidos.

Mager (1967) refere-se ao elemento seleção de procedimentos relacionado a cada um dos objetivos de desempenho. Deve-se selecionar a técnica que:

- mais se aproxima do objetivo
- levará o aluno a realizar, de maneira mais aproximada, o desempenho exigido pelo trabalho
- permitirá ao aluno executar o maior número de respostas relevantes, por unidade de tempo.

Em seguida, selecionam-se, dentre aqueles relacionados aos objetivos de desempenho, os procedimentos mais práticos, com base em critérios administrativos.

Branson et al (1973) tratam da seleção e direção instrucional, significando o esquema particular selecionado para organizar os eventos de aprendizagem e alunos, para otimizar os resultados do sistema. Tentam, com isto, acomodar as necessidades de treinamento aos recursos disponíveis, levando em conta restrições reais e utilizando múltiplas abordagens em partes diferentes do mesmo curso.

Dederick e Sturge (1975) referem-se à determinação de métodos instrucionais. A partir da definição dos objetivos comportamentais do curso, elabora-se um programa de instrução que deve considerar restrições como limitação de orçamento, pessoal, equipamento, tempo e outros.

Conforme se pode notar, todos os modelos referem-se, de uma maneira ou de outra, ao elemento seleção de estratégias instrucionais, porém, nenhum deles apresenta suas próprias estratégias e sua adequação ou não às diversas situações.

Além disso, em alguns modelos, as estratégias instrucionais são tratadas a nível de sistema. Nosso enfoque, no momento, se refere a estratégias instrucionais, a nível de ensino – aprendizagem,

ou seja, em termos de técnicas de ensino e de organização de instrutores e treinandos.

Assim sendo, recorreremos à literatura, à procura de técnicas relacionadas a cada um dos objetivos de desempenho. Dentre as técnicas encontradas, baseadas na terminologia de Nérici (1969), foram selecionadas as mais adequadas, em termos de suas potencialidades para cumprir um propósito particular, ou seja, para atender a natureza do conteúdo dos nossos objetivos. São elas:

- técnica expositiva (AE) - Consiste na exposição oral, por parte do professor, do assunto da aula, devendo este desencadear a participação dos alunos;
- técnica de estudo dirigido individual (EDI) - Consiste na ligeira apresentação, pelo professor, do assunto a ser trabalhado, com finalidade de motivação e orientação. Depois desta exposição, os alunos devem, individualmente, ler os textos preparados. Durante a leitura, o professor esclarece as possíveis dúvidas e responde as perguntas levantadas pelos alunos. Após estes esclarecimentos, o professor distribui os objetivos ou questões sobre as quais os alunos deverão trabalhar, lendo novamente o texto para encontrarem elementos que resolvam as proposições. Em seguida, o professor solicita a exposição do trabalho por um ou mais alunos. Esta exposição deve ser acompanhada de discussão. No final desta, são tiradas conclusões sobre o tema em estudo, devendo o professor, depois, fazer uma conclusão final do trabalho (retroalimentação);
- técnica de estudo dirigido em grupo (EDG) - Consiste nos mesmos procedimentos da técnica anterior, exceto que após a distribuição dos objetivos ou questões, os alunos deverão se reunir em grupo para encontrarem elementos que resolvam as proposições;
- técnica de tarefa dirigida (TD) - Trata-se da execução, a partir de orientações dadas pelo professor, de exercícios práticos sobre objetivos dominados anteriormente. Após a execução, o professor solicita a exposição da tarefa por um ou mais alunos.

Esta exposição deve ser acompanhada de discussão. No final desta, são tiradas conclusões sobre a tarefa, devendo o professor fazer, depois, uma conclusão final (retroalimentação).

A tarefa dirigida pode ser individual (TDI) ou em grupo (TDG).

A indicação das técnicas selecionadas, neste trabalho, em relação a cada um dos objetivos instrucionais, é apresentada na página 107.

1.6.4 - SELEÇÃO DE MEIOS INSTRUCIONAIS

O propósito desta revisão foi a identificação de procedimentos para a seleção de meios instrucionais.

Os modelos de treinamento revistos na página 7, para auxiliar no planejamento do curso não apresentam, também, procedimentos de seleção de meios, adequados aos propósitos deste trabalho.

O modelo Serviço de Informação do Governo (1968), modelo básico de Análise de Sistemas e, portanto, bastante geral, não trata do problema de seleção de meios.

O modelo de Banathy (1967), no bloco relativo a "Design" do Sistema, trata das funções que devem ser executadas pelo sistema e dos componentes que devem ser utilizados para conduzir cada função específica. Contudo, não operacionaliza a distribuição das funções aos componentes, não discriminando procedimentos de seleção de meios que conduziram algumas destas funções.

O modelo de Tracey (1971), trata dos meios como auxiliares de treinamento e coloca a seleção deles como uma questão de julgamento, não apresentando procedimentos específicos para esta seleção.

O modelo de Mager e Beach (1967) não operacionaliza procedimentos para seleção de meios, fornecendo apenas critérios gerais relativos a esta atividade.

O modelo de Branson et al. (1973) trata da seleção do Sistema de Entrega, que contém e apresenta a informação ao aluno, considerando as características e restrições das alternativas de meios. Porém, sua abordagem é bastante complexa, partindo da análise de objetivo instrucional identificado, e da especificação dos eventos para cada um destes objetivos. Estes procedimentos extremamente operacionalizados não poderiam ser, aqui, utilizados, considerando que o tempo disponível para a execução deste trabalho, isto é, para o planejamento do curso de treinamento, é limitado.

O modelo de Dederick e Sturge (1975) não inclui como um de seus passos a seleção de meios.

Foi considerado, ainda, o Modelo de Briggs (1970). Neste Modelo os procedimentos propostos para a seleção de meios, bastante complexos, não poderiam, também, ser utilizados aqui, pelos mesmos motivos da não utilização do Modelo de Branson e outros (1973). Briggs (1970) propõe a seleção de meios, em função de suas capacidades para transmitir os estímulos específicos, de cada evento instrucional dos componentes de cada objetivo. Neste trabalho não se pretendeu chegar a este nível de detalhamento.

Nenhum dos modelos instrucionais revistos era adequado ao propósito deste trabalho, quanto aos procedimentos de seleção de meio instrucionais. Recorremos então a um modelo específico de seleção de meios: Bretz (1971).

Nesta abordagem, "um meio de comunicação é um sistema completo para a transmissão de mensagens, geralmente envolvendo um processo para a produção de programas, um processo para sua entrega e um modo de apresentação dos mesmos. As entradas primárias, deste sistema,

são as mensagens que ele processa. As entradas secundárias consistem em recursos tais como: equipamentos, trabalho, materiais utilizados no planejamento e produção do programa, administração e outros. As saídas do sistema de comunicação são essencialmente as mensagens como apresentadas nos programas". Mensagem, neste contexto, significa o conteúdo da comunicação, os significados, idéias, conceitos e dados que são comunicados.

O sistema de comunicação serve sempre a sistemas mais amplos. Em se tratando do sistema instrucional, o sistema de comunicação se propõe a ajudá-lo a desempenhar suas funções no processo de ensino/aprendizagem. Faz isto conduzindo os programas instrucionais ou lições. Portanto, as funções de um meio de comunicação são: expressar a mensagem como um programa, registrá-la e/ou transmiti-la e exibi-la para os alunos.

Observamos que Bretz (1971) considera o professor como um elemento do sistema instrucional distinto do sub-sistema de comunicação, com funções específicas decorrentes da estratégia instrucional utilizada.

Bretz (1971) aborda o problema da seleção de meios do ponto de vista do uso dos mesmos. A escolha é baseada nos objetivos de aprendizagem, previamente determinados, na natureza do conteúdo a ser ensinado, em certas características da população de aprendizes e na estratégia instrucional selecionada.

O autor coloca sua abordagem como sendo semelhante à abordagem de Briggs, na medida em que ambas consideram a seleção de meios a partir da identificação e especificação dos requisitos instrucionais. Entretanto, a abordagem de Briggs (1970) apresenta um nível de análise mais detalhado, com procedimentos mais complexos. Além disso, Briggs trata de todos os meios de instrução, incluindo professores e objetos (pedras, folhas, produtos manufaturados, etc...), enquanto a abordagem de Bretz (1971) limita-se aos meios de comunicação como definidos na página 48.

Dentro desta abordagem são considerados propriedade e praticabilidade dos meios. Estes conceitos são assim entendidos:

Propriedade: capacidade do meio de expressar a informação decorrente do objetivo instrucional. Por exemplo, um meio não visual não seria apropriado nos casos em que os objetivos da aprendizagem envolvem reconhecimento visual.

Praticabilidade: envolve várias considerações tais como: conveniência, disponibilidade, intercambialidade de "software", requisitos de treinamento para seu uso efetivo, custo relativo em comparação com o custo de outros meios igualmente apropriados. Depende, portanto, em grande extensão, das condições locais de uma aplicação específica.

De sua abordagem, utilizamos os procedimentos especificados para a identificação dos meios mais apropriados (propriedade) para cada objetivo instrucional. A metodologia geral desta tarefa, consiste em:

- analisar os objetivos instrucionais de cada aula ou segmento de aula, para identificar a classe de meios mais apropriada a cada um deles. São propostas 8 classes, e dentro de cada uma delas, todos os meios são igualmente apropriados para apresentar uma mesma informação;
- determinar o tempo instrucional para o qual cada classe de meio é apropriada;
- identificar classes alternativas ou combinações multi-classes igualmente apropriadas para apresentar uma mesma informação.

As 8 classes identificadas por Bretz (1971) são apresentadas na Tabela I.3.

TABELA I.3

OS MEIOS DE COMUNICAÇÃO - UMA TAXIONOMIA

	MODOS DE APRESENTAR A INFORMAÇÃO					Gravação
	SOM	IMAGEM	DESENHOS	IMPRESSOS	ANIMAÇÃO	
Telecomunicação						
CLASSE I: AUDIO-VISUAL E ANIMAÇÃO						
	x	x	x	x	x	Filme sonoro
Televisão	x	x	x	x	x	{ "Video-tape" Gravação de filmes em-TV
	x	x	x	x	x	Gravação Holográfica
Telefone com imagem	x	x	x	x	x	
CLASSE II: AUDIO-VISUAL ESTÁTICO						
TV de varredura lenta TV compartilhada	x	x	x	x		Gravação estática de vídeo
	x	x	x	x		"Filmstrip" sonoro
	x	x	x	x		Jogo de "slide" com som
	x	x	x	x		Som gravado no "slide"
	x	x	x	x		Página sonora
	x	x	x	x		Livro falado
CLASSE III: AUDIO-SEMI-MOVIMENTO						
Teleimpressão	x		x	x	x	Teleimpressão gravada
Áudio para sincronização	x		x	x	x	
CLASSE IV: VISUAL COM MOVIMENTO						
		x	x	x	x	Filme mudo
CLASSE V: VISUAL ESTÁTICO						
"Fac-símile"		x	x	x		Páginas impressas
		x	x	x		"Filmstrip" mudo
		x	x	x		Conjunto de imagens
		x	x	x		Microformas
		x	x	x		Arquivo de vídeo
CLASSE VI: SEMI-MOVIMENTO						
Teleautógrafo			x	x	x	
CLASSE VII: AUDIO						
Telefone } Rádio }	x					{ Disco Fita
CLASSE VIII: IMPRESSÃO						
Teletipo				x		Fita perfurada

Os meios identificados nesta taxionomia subdividem-se em:

- meios de telecomunicação, capazes de transmitir programas através das distâncias, em tempo real;
- meios de gravação, capazes de registrar programas, estocá-los e repeti-los posteriormente (inclusive via meios de telecomunicação).

e são classificados nas oito classes, segundo seu modo de apresentar a informação.

O procedimento proposto para a identificação da classe de meios apropriada para cada objetivo instrucional consiste em tomar uma série de decisões sequenciais, identificadas nos esquemas que seguem. O primeiro destes esquemas (Figura I.8) apresenta uma visão geral do processo, e os esquemas seguintes (Figuras I.9 a I.13) são detalhes deste esquema genérico.

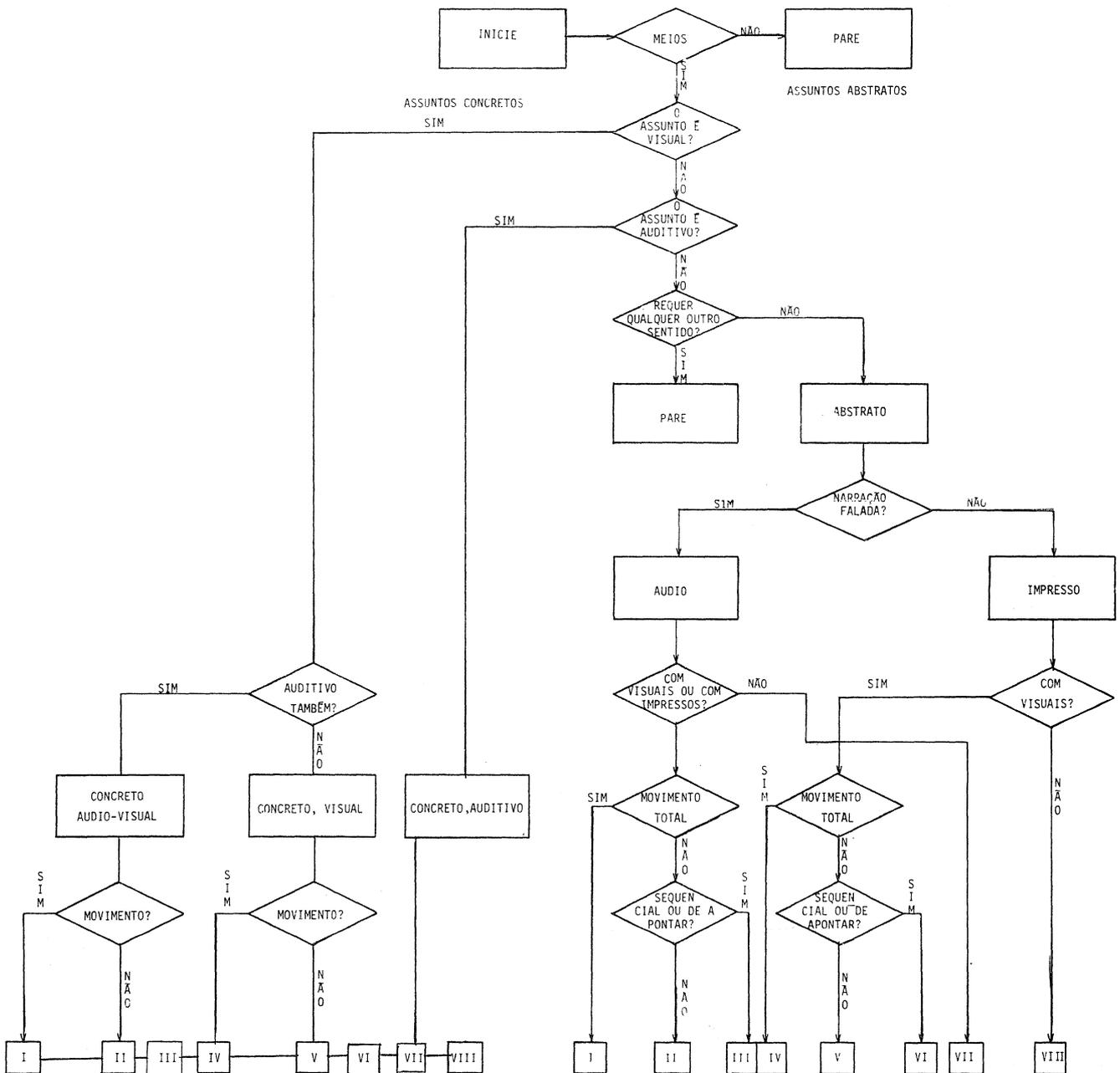


Fig. I.8 - Diagrama dos Pontos de Decisões Necessários para a Seleção de uma Simples Classe de meios Apropriados para uma Necessidade Instrucional (Bretz, 1971)

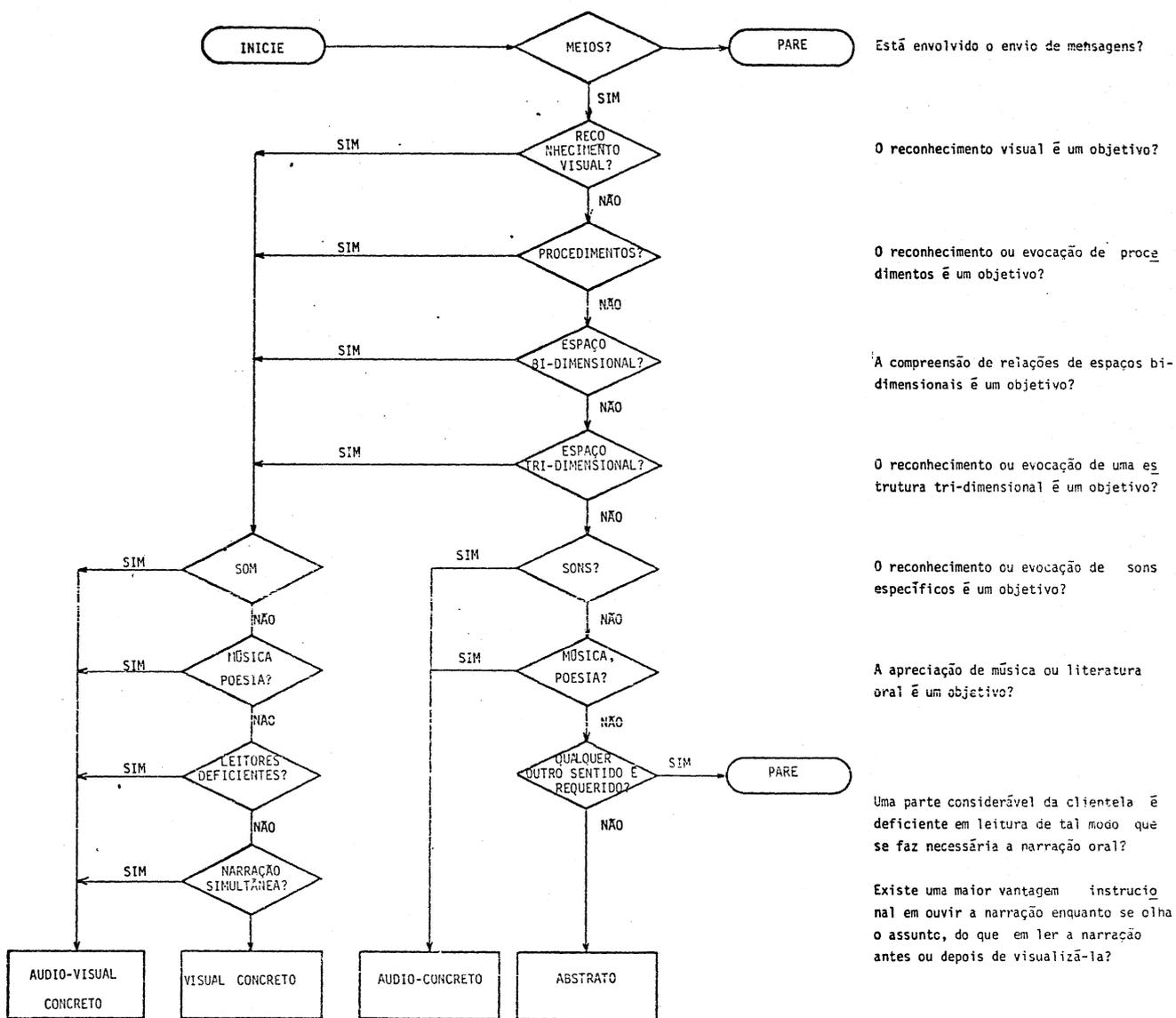


Fig. I.9 - Critério para Distinção entre Conteúdo Concreto e Abstrato (Bretz, 1971)

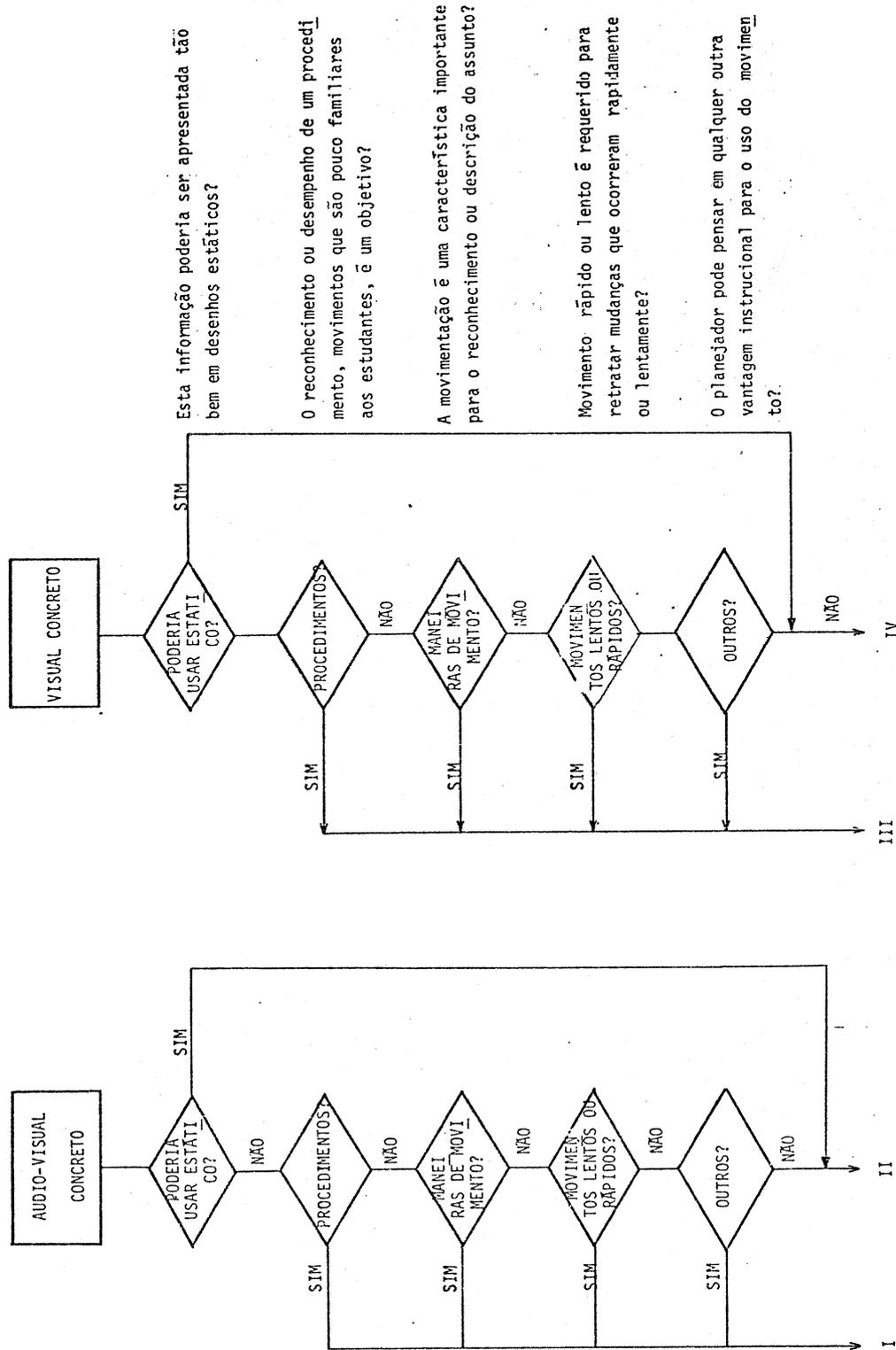


Fig. I.10 - Decisão entre Apresentação em Movimento ou Estática (Bretz, 1971)

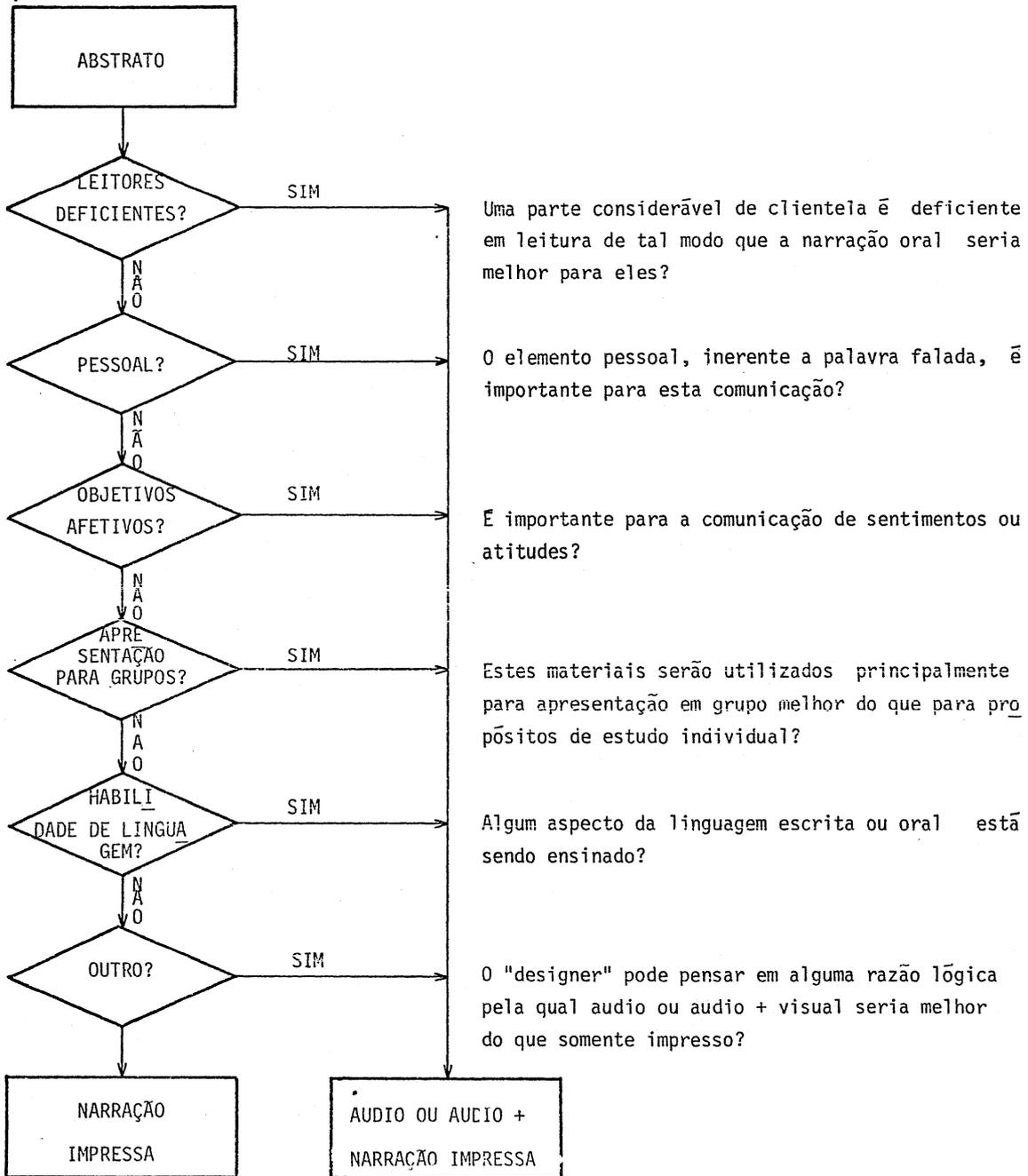


Fig. I.11 - Escolha dos Meios de Narração (Bretz, 1971)

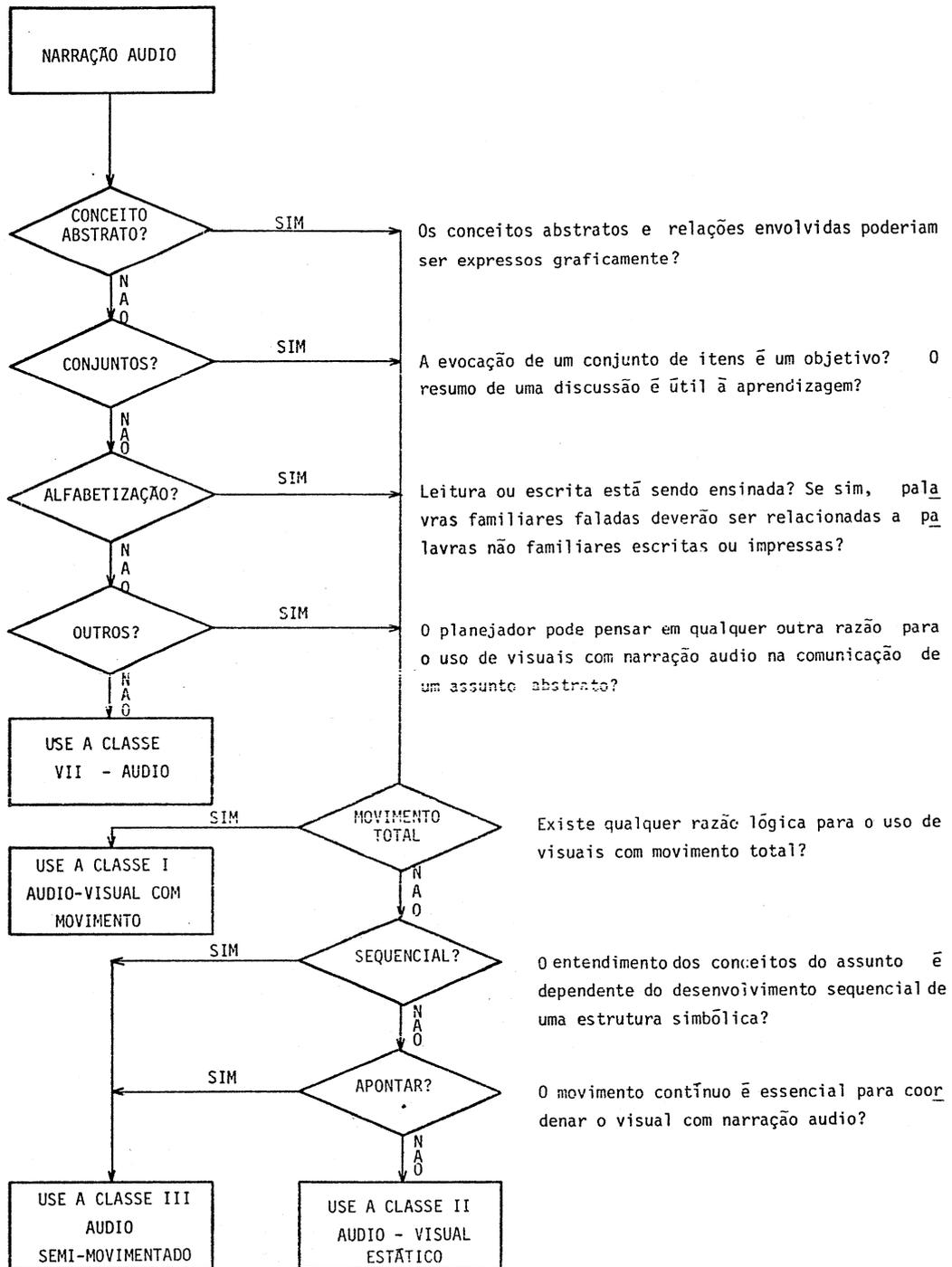


Fig. I.12 - Visualização da Narração Audio (Bretz, 1971)

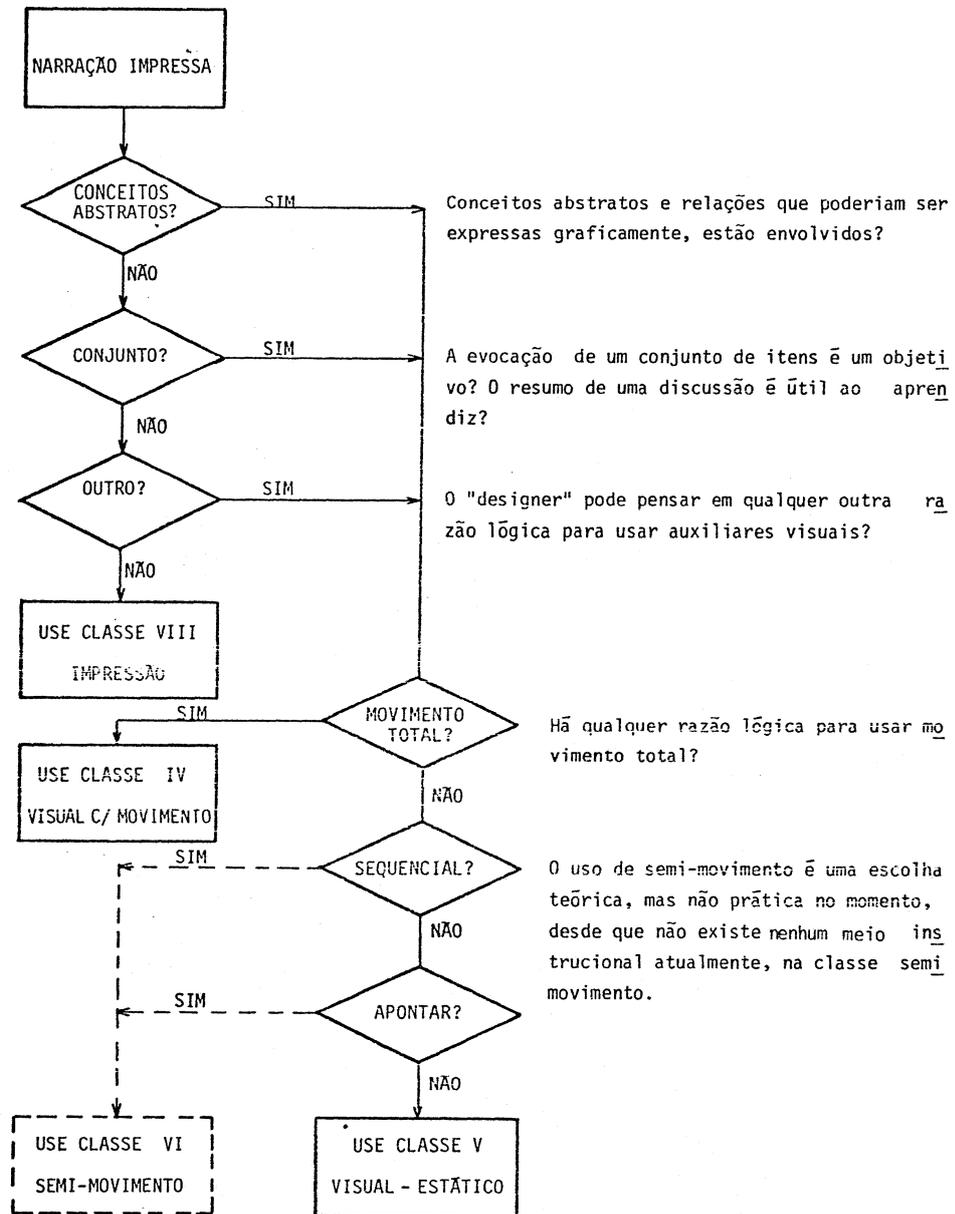


Fig. I.13 - Visualização da Narração Impressa (Bretz, 1972).

Através da utilização destes diagramas, chega-se à identificação da classe de meios mais apropriada para cada objetivo.

Bretz (1971) especifica, ainda, conforme as Figuras I.14 e I.15 para cada classe de meios, as classes ou combinações multi-classes que lhes são teoricamente equivalentes, ou seja, que podem substituí-las apropriadamente.

CLASSE I AUDIO-VISUAL COM MOVIMENTO	CLASSE II AUDIO-VISUAL ESTÁTICO	CLASSE III AUDIO SEMI MOVIMENTO	CLASSE IV VISUAL COM MOVIMENTO	CLASSE V VISUAL ESTÁTICO	CLASSE VI SEMI MOVIMENTO	CLASSE VII AUDIO	CLASSE VIII IMPRESSÃO
--	---------------------------------------	---------------------------------------	--------------------------------------	--------------------------------	--------------------------------	---------------------	--------------------------

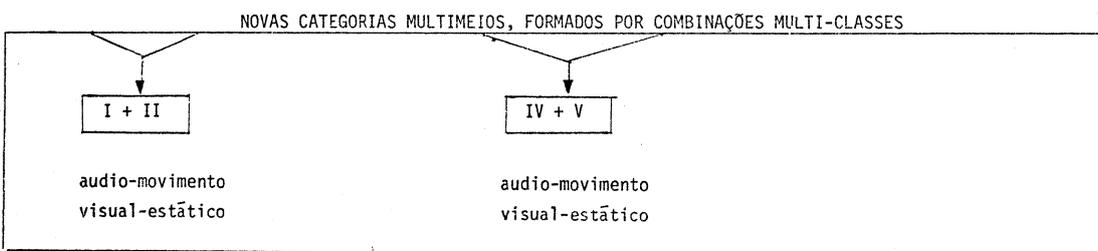
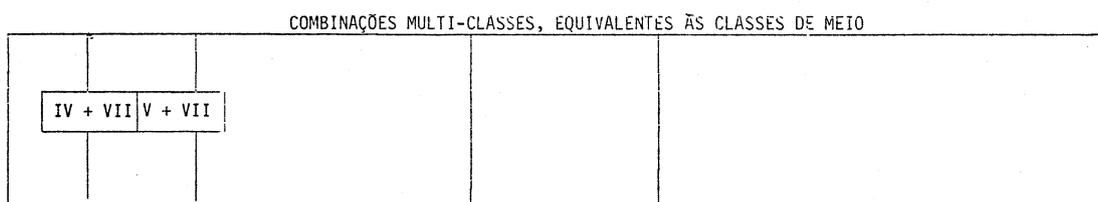
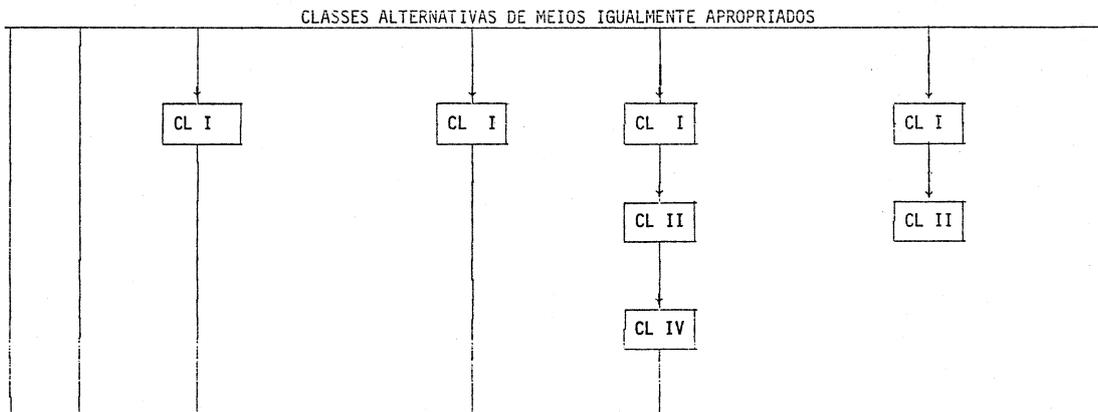
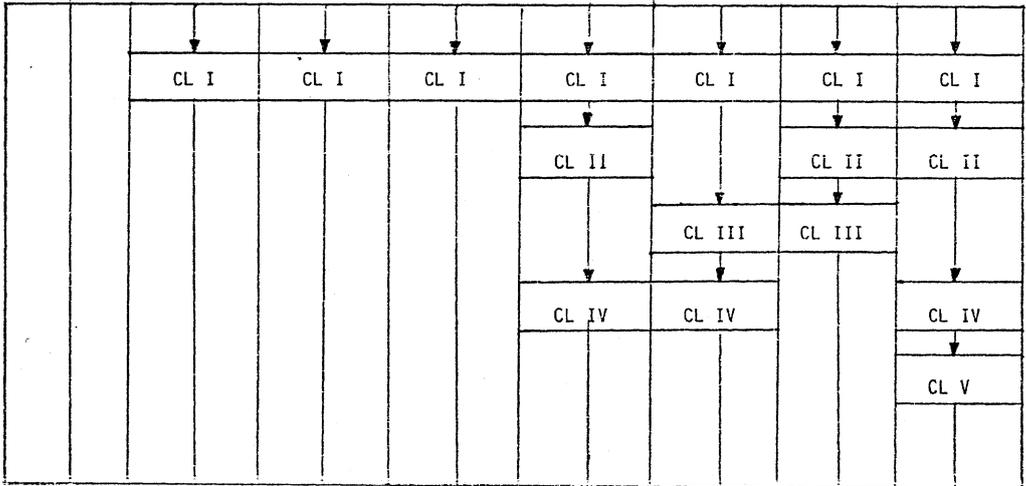
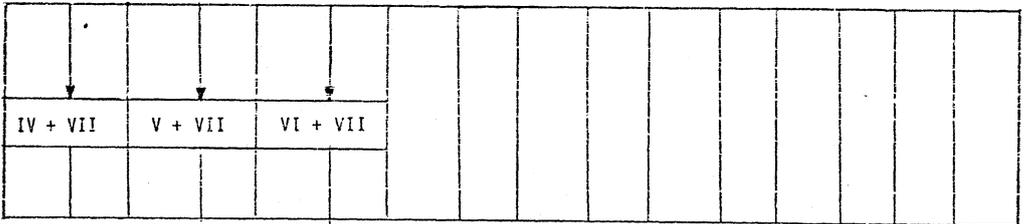


Fig. I.14 - Alternativas para Expressar o Conteúdo Concreto (Bretz, 1971).

CLASSE I AUDIO VISUAL C/ MOVIMENTO	CLASSE II AUDIO VISUAL ESTÁTICO	CLASSE III AUDIO SEMI- MOVIMENTO	CLASSE IV VISUAL E MOVIMENTO	CLASSE V VISUAL ESTÁTICO	CLASSE VI SEMI MOVIMENTO	CLASSE VII AUDIO	CLASSE VIII IMPRESSÃO
---	--	---	---------------------------------------	--------------------------------	--------------------------------	---------------------	--------------------------



Combinações multi-classes, equivalentes às classes de meios



Novas categorias multi-meios, formadas por combinações multi-classes:

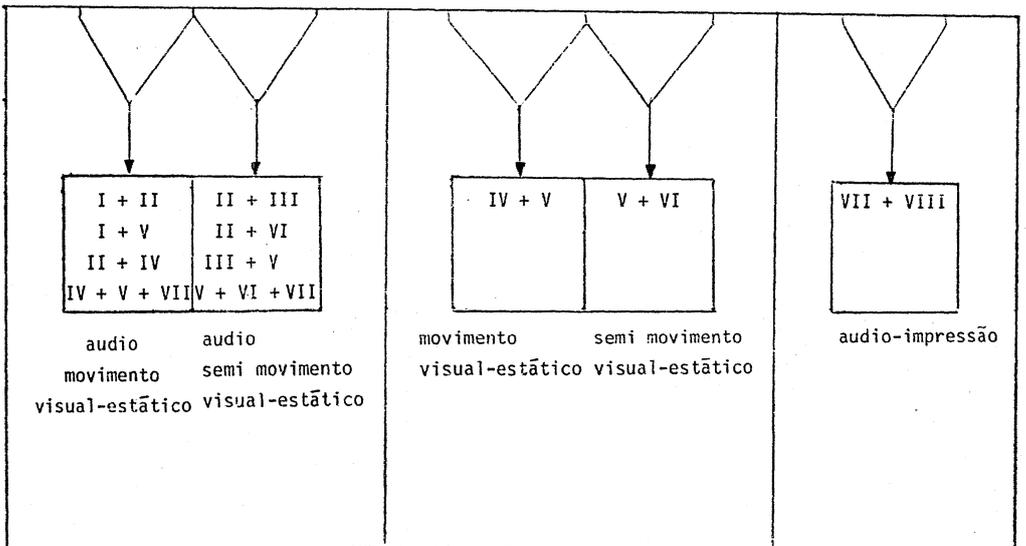


Fig. I.15 - Alternativas para Expressar o Conteúdo Abstrato

Após a identificação da classe de meios mais apropriada e das classes que lhe são alternativas, é considerada a praticabilidade de dos meios nela incluídos, para a seleção final dos mesmos.

O modelo de Bretz (1971) foi adotado para a seleção de meios instrucionais, a serem utilizados no curso de treinamento de Avaliadores Educacionais, por ser um modelo simples, prático e operacionalizado. Um exemplo da aplicação deste modelo, para a seleção de meios para a consecução de um objetivo instrucional do curso, é apresentado no Capítulo II. Os resultados da aplicação deste modelo para todos os objetivos do curso são apresentados na Tabela II.7, no citado capítulo.

O capítulo que segue apresenta o planejamento instrucional do curso de treinamento. A revisão de literatura, realizada neste capítulo, embasa tal planejamento.

CAPÍTULO II

PLANEJAMENTO DO CURSO: DEFINIÇÃO DO UNIVERSO A SER TREINADO E PLANEJAMENTO DAS UNIDADES INSTRUCCIONAIS

Este capítulo apresenta a identificação do universo de atividades que serão treinadas e o planejamento das unidades instrucionais do curso.

2.1 - DEFINIÇÃO DO UNIVERSO DE ATIVIDADES QUE SERÃO TREINADAS

A identificação do universo de atividades que serão treinadas foi resultado dos seguintes passos: identificação do universo teórico de atividades do avaliador, identificação do universo real de atividades, comparação entre o universo teórico e o real para determinação do universo de atividades que poderiam ser treinadas, e identificação do universo de atividades que serão treinadas.

2.1.1 - IDENTIFICAÇÃO DO UNIVERSO TEÓRICO DE ATIVIDADES QUE CARACTERIZAM A FUNÇÃO DO AVALIADOR

Pela revisão da literatura da teoria de Avaliação e posterior adoção da abordagem de Stufflebeam et al (1971), pág. 36, chegamos à Tabela II.1 que apresenta uma visão geral, abrangente, do processo de avaliação, em termos das atividades que o avaliador desempenha.

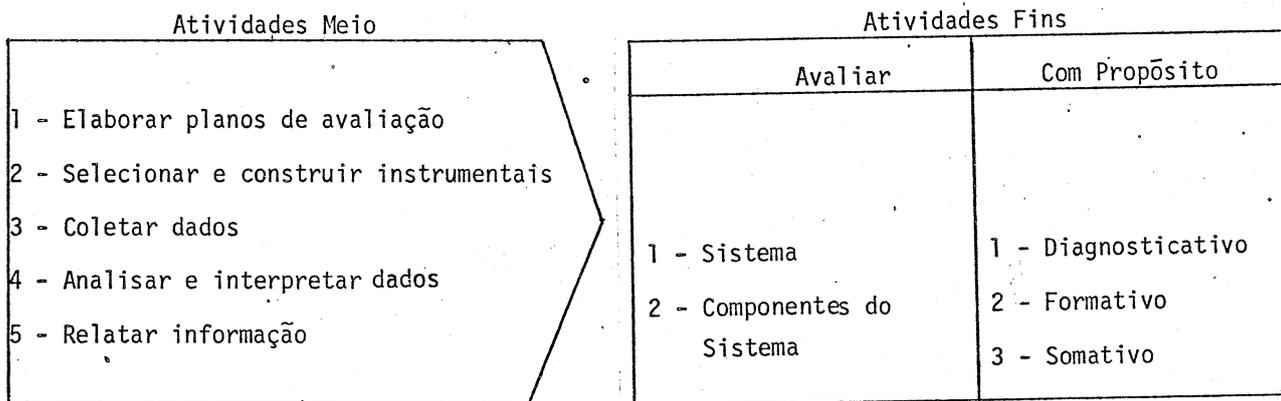
Classificamos as atividades do processo de Avaliação em atividades fins e meio.

As atividades fins deste processo são:

- Avaliar sistemas
- Avaliar componentes do sistema.

TABELA II.1

ATIVIDADES QUE CONSTITUEM O PROCESSO DE AVALIAÇÃO



Atividades meio são atividades gerais que suportam a realização de qualquer atividade fim, isto é, são imprescindíveis para qualquer Avaliação, independente do seu propósito.

As atividades fins são aquelas que caracterizam o propósito da Avaliação.

As atividades acima apresentadas foram desdobradas em atividades menores, que compõem a listagem final a que chegamos, através da revisão da literatura. Apenas a atividade ; avaliar componentes do sistema - não foi desdobrada em atividades menores, uma vez que estas estariam incluídas dentro daquelas relativas a avaliar sistemas.

A lista que segue, apresenta o universo teórico de atividades de avaliação identificadas por nós, a partir da teoria de Stufflebeam et al (1971), e de uma revisão da pesquisa de Worthen (1975).

ATIVIDADES FINIS

1.0 - Avaliar Sistema

1.1 - Propósito Diagnóstico

1.1.1 - Diagnosticar Problema

- definir o sistema
- identificar os objetivos do sistema
- determinar o valor e viabilidade dos objetivos
- descrever o sistema corrente em termos de variáveis populacionais, instrucionais e comportamentais
- comparar a situação atual do sistema com a desejada, identificando discrepâncias
- identificar, fora do sistema, oportunidades não usadas por este
- levantar questões que orientam o diagnóstico dos problemas que geram as discrepâncias (entre o real e o desejado), e impedem que as oportunidades, fora do sistema, sejam usadas
- explicar o problema.

1.1.2 - Identificar e Avaliar Alternativas de solução para o Problema

- identificar estratégias alternativas para solucionar o problema
- operacionalizar os objetivos da mudança a ser introduzida, para solucionar o problema, tornando-os precisos e mensuráveis
- especificar critérios para avaliar o alcance dos objetivos
- identificar recursos materiais e humanos disponíveis, para orientar a escolha da melhor estratégia
- identificar "designs" de procedimentos (para implementação da estratégia escolhida).

1.2 - Propósito Formativo e Somativo

1.2.1 - Controlar o Sistema e explicar resultados

- manter um registro constante dos procedimentos, tal como ocorrem;
- detectar ou prever defeitos no "design" de procedimentos;
- fornecer realimentação periódica sobre o desempenho do sistema;
- medir alcance dos objetivos;
- comparar a efetividade com padrões pré-estabelecidos.

2.0 - Avaliar Componentes.

ATIVIDADES MEIOS

1.0 - Elaborar Plano de Avaliação

2.0 - Selecionar e Construir Instrumentais

- estabelecer os propósitos do uso do instrumental de medida;
- selecionar instrumental adequado;
- adaptar instrumental;
- desenvolver instrumental.

3.0 - Coletar dados

- planejar coleta de dados, incluindo:
 - . fontes de informação;
 - . procedimentos de amostragem;
 - . instrumentais a serem utilizados;
 - . cronograma de coleta de dados.
- formular hipóteses testáveis;
- especificar dados para um teste de hipótese;
- utilizar "designs" pré-experimentais, quase-experimentais e experimentais, de acordo com Campbell e Stanley; *Stanley*
- utilizar meios informais de investigação, como:
 - . estudar e visitar outros sistemas;
 - . participar de seminários e conferências.
- utilizar procedimentos de investigação formal, como:

- . técnica Delphi de convergência e opinião
- . técnica de pesquisa bibliográfica
- utilizar técnicas de controle, como:
 - . diagrama de marcos
 - . cronogramas mestres e parciais
 - . técnicas de O & M
- aplicar instrumentais de medida, tais como:
 - . testes de aprendizagem
 - . questionários
 - . escalas de atitude
 - . fichas de observação
 - . entrevistas.
- atribuir pontos aos resultados obtidos com a aplicação de instrumentais de medida
- registrar dados.

4.0 - Analisar e Interpretar Dados

- processar dados
- escolher e empregar técnicas de análise estatística
 - . descritiva
 - . inferencial
- utilizar técnicas analíticas, tais como: gráficos, tabelas, perfis
- interpretar dados em termos de critérios especificados.

5.0 - Relatar Informação

- Identificar audiência
- Elaborar relatórios escritos e orais, com estilo e nível apropriado para a audiência especificada (decisores, seus subordinados, sua clientela, agências, etc...):

2.1.2 - IDENTIFICAÇÃO DO UNIVERSO REAL DE ATIVIDADES QUE CARACTERIZAM A FUNÇÃO DO AVALIADOR

Para a identificação do universo real de atividades que caracterizam a função do avaliador, utilizamos os procedimentos que seguem.

Para identificarmos as fontes de coleta de dados, enviamos, para diversas Entidades Educacionais*, um formulário que se constituiu em nosso primeiro contato com as mesmas. Este teve o propósito de obter toda informação pertinente disponível, de acordo com os procedimentos propostos pela análise ocupacional. Este formulário teve o objetivo de obter informações referentes a:

- existência de avaliadores
- qualificação dos mesmos
- natureza das atividades por eles desempenhadas.

O modelo deste formulário é apresentado no Apêndice B.

À medida que chegaram os primeiros formulários respondidos, tornou-se possível a seleção das Entidades que constituíram as fontes, propriamente ditas, de nossa coleta de dados. Assim, foram selecionadas apenas aquelas que possuíam equipes de Avaliação.

Nosso segundo contato teve o propósito de identificar um Universo Real de Atividades do Avaliador. Para isto, utilizamos, da tēcnica de Análise Ocupacional, aquilo que pudesse nos conduzir à determinação.

* A listagem das Entidades Educacionais foi fornecida pela Associação Brasileira de Teleducação (ABT) e se encontra no Apêndice A.

nação das tarefas que compreendem o trabalho em estudo. Assim, da fórmula O QUE, COMO, PORQUE e HABILIDADES ENVOLVIDAS, que orientam a descrição das tarefas do trabalho e a indicação das qualificações necessárias ao trabalhador, nos restringimos ao O QUE (Atividades Fins) o avaliador faz, e ao COMO (Atividades Meios) ele faz.

Assim, no segundo contato nossa intenção foi solicitar dos:

- avaliadores, uma descrição das tarefas específicas que realizam: O QUE o avaliador faz e COMO faz (Apêndice C).
- decisores, informações relativas aos objetivos de suas Entidades, a importância da Avaliação para a consecução dos mesmos, tarefas que a Avaliação realiza, quais deveria realizar e quais deveriam ser treinadas (Apêndice D)

Para estes contatos optamos pelo Método de Questionário, combinado, em alguns casos, com a Entrevista. O questionário foi aplicado após um processo de validação.

O questionário foi utilizado por nós, considerando que:

- foi o único método que possibilitou a inclusão, em nossa coleta de dados, de Entidades localizadas em vários Estados do País.
- pôde ser aplicado em um grande número de pessoas, ao mesmo tempo, uma vez que pôde ser enviado pelo correio.

A Entrevista foi utilizada em combinação com o Questionário, naquelas Entidades em que a nossa presença foi possível. Sua flexibilidade, isto é, a possibilidade de se repetir e de se apresentar de outro modo, as questões a serem feitas, permitiu uma verificação e suplementação das informações obtidas pelo questionário.

A observação não foi adotada por nós, devido à própria natureza do trabalho em análise, que envolve muito pouca atividade física. Além do que, a observação seria limitada pela duração dos acontecimentos.

O questionário utilizado junto aos avaliadores, foi o resultado da combinação da listagem de atividades de Avaliação, deduzidas da teoria, e dos modelos de instrumentais de coleta de dados, apresentados por Otis e Leukart (1954).

Parte das informações solicitadas se referiam à identificação do avaliador respondente, sua formação acadêmica, sua experiência profissional e ao recebimento ou não de treinamento em Avaliação. Isto com o propósito de verificar as afirmativas que seguem e que são originárias de informações verbais:

- 1 - No Brasil, as atividades de Avaliação são exercidas por pessoas de diversas áreas e, predominantemente, sem formação acadêmica específica de Avaliação.
- 2 - As pessoas que exercem atividades de Avaliação não têm uma "tradição" em termos de carreira. Apresentam as mais variadas experiências profissionais.
- 3 - As pessoas que exercem atividades de Avaliação, em geral, não recebem treinamento específico para tal.

Com relação às atividades de Avaliação, foi fornecida ao respondente uma listagem contendo aquelas já apresentadas na página 65. Entre estas, ele deveria assinalar as que realizava. Foi ainda pedido que descrevesse o seu trabalho, citando, com frases chave, as suas atividades principais. Com isto haveria possibilidade de serem registradas as atividades não listadas.

Foi solicitado, também, do respondente, que citasse as funções e respectivas atividades relacionadas àquelas que executasse.

Isto, com o propósito de verificar a interrelação ou não das atividades de outras funções, como por exemplo, as de Planejamento. Na associação ou dissociação, destas funções, transpareceria a abordagem recebida pela Avaliação dentro da Entidade.

Finalmente, foram pedidas informações referentes às suas atividades diárias, semanais, mensais, trimestrais, semestrais, anuais e ocasionais. Foi pedido, ainda, que especificasse as atividades mais difíceis ou complexas, com o objetivo de destacar as mais significativas para o treinamento, do ponto de vista da pessoa que realiza o trabalho.

O questionário, em sua forma final, encontra-se no Apêndice C.

A Tabela II.2 sumaria os procedimentos de coleta de dados.

A análise dos dados coletados, através da aplicação do instrumental para a identificação de um Universo Real de atividades foi realizada da seguinte maneira:

A partir dos dados obtidos com a aplicação dos questionários junto aos avaliadores, nas Entidades Educacionais, e das entrevistas realizadas com os mesmos, foram reforçadas as afirmativas já citadas na página 70. São as seguintes:

- a - No Brasil, as atividades de Avaliação são exercidas por pessoas de diversas áreas e, predominantemente, sem formação acadêmica específica de Avaliação.

TABELA II.2

SUMÁRIO DE PROCEDIMENTOS DE COLETA DE DADOS

PROCEDIMENTOS	INSTRUMENTAIS E FONTES DE INFORMAÇÃO	DADOS COLETADOS
- identificar Entidades Educacionais que serão contadas.	consulta à ABT.	lista com 62 Entidades Educacionais.
- identificar Entidades Educacionais que executam atividades de Avaliação.	formulário para 62 Entidades Educacionais.	respostas recebidas de 43 entidades.
- identificar atividades do Avaliador.	- questionário para 68 Avaliadores*. - questionário para 34 decisores. ou - questionário + entrevista com Avaliadores - questionário + entrevista com decisores.	- respostas em 26 questionários de avaliadores. - respostas em 12 questionários de decisores.

* Devido à impossibilidade prática de ser enviado um questionário para cada um dos Avaliadores de cada Entidade, foram fixados os limites de 1 a 3, em função da diversidade de atividades desenvolvidas nas mesmas. Esta informação foi coletada pelo formulário.

A Tabela II.3 apresenta, em percentagens, a quantidade de Avaliadores respondentes, por curso de graduação.

TABELA II.3

FORMAÇÃO ACADÊMICA

CURSO DE GRADUAÇÃO	PERCENTAGEM DE AVALIADORES RESPONDENTES
Pedagogia	37
Administração de Empresas	19
Ciências Sociais	13
Filosofia	13
Administração Escolar	6
Economia	6
Psicologia	6
Total	100

Dos avaliadores respondentes, 37% são formados em Pedagogia e 6% em Administração Escolar, que são as áreas mais afins com Avaliação Educacional. Os demais respondentes acham-se distribuídos, segundo sua formação acadêmica, em cursos de Administração de Empresas, Ciências Sociais, Filosofia, Economia, e Psicologia, desvinculados, portanto, da função que exercem.

Além disto, nenhum dos respondentes tem pós-graduação na área específica de Avaliação Educacional.

- b - As pessoas que exercem atividades de Avaliação não têm uma "tradição", em termos de carreira. Apresentam as mais variadas experiências profissionais.

Tradição foi aqui entendida como continuidade, através dos anos, na função de Avaliação. O indicador de continuidade foram as funções exercidas anteriormente. Assim, aqueles indivíduos que atualmente exercem a função de avaliador e que já a exerceram no trabalho anterior foram considerados como tendo tradição, em termos de carreira. Aqueles que estavam exercendo a função pela primeira vez foram considerados como não tendo tradição. Os dados coletados mostraram que 100% dos atuais avaliadores não tem tradição, em termos de carreira.

c - As pessoas que exercem atividades de avaliação, em geral, não recebem treinamento específico para tal.

TABELA II.4

TREINAMENTO RECEBIDO EM AVALIAÇÃO

ALTERNATIVAS	PERCENTAGEM
sim	33
não	50
em branco	17

O tipo de treinamento referia-se a curso e treinamento em serviço. A duração máxima dos cursos, para aqueles que os receberam, foi de 40 horas de aula.

Foi identificado, também, pelos resultados dos questionários e entrevistas junto aos Avaliadores e decisores, o universo de atividades que vem sendo executadas por eles, ou seja, o universo real de atividades do avaliador.

Conforme já foi exposto, o instrumental aplicado junto aos avaliadores apresentou uma listagem de atividades identificadas na literatura, onde o respondente deveria assinalar as desempenhadas por ele. Continha, ainda, uma questão aberta para que ele pudesse descrever o seu trabalho, incluindo atividades não listadas pelas autoras.

As atividades que constituíram a listagem são apresentadas no Apêndice E, com as respectivas percentagens com que são executadas pelos respondentes.

Com relação à questão aberta, seu preenchimento reforçou nossa suposição, relativa à falta de um consenso de quais atividades caracterizam a função de Avaliação. Encontramos listadas, como atividades de Avaliação, algumas que são caracteristicamente de outras funções tais como: atividades de planejamento, docência, orientação pedagógica, elaboração de material didático, elaboração de orçamento-programa, e outros. Este fato evidencia ainda a formação e o desempenho polivalente dos elementos envolvidos em Avaliação nas nossas Instituições.

Todas as atividades descritas pelos respondentes como atividades de Avaliação desempenhadas pelos mesmos, foram julgadas por nós para concluirmos a respeito da sua real caracterização, como atividades avaliativas. O Universo real constou das atividades que são executadas na prática.

2.1.3 - COMPARAÇÃO DAS ATIVIDADES CONSTANTES DO UNIVERSO REAL COM AS ATIVIDADES DO UNIVERSO TEÓRICO.

Identificadas as atividades do Universo Real, elas foram comparadas com aquelas do Universo Teórico. Esta comparação teve o propósito de verificar se existia uma correspondência entre as atividades constantes nos dois Universos. Foi constatado que todas as atividades identificadas no Universo Real se encontravam no Universo Teórico. Nenhuma nova atividade foi descrita. No entanto, esta comparação mostrou, também, que o Universo Real é menor que o Universo Teórico.

2.1.4 - IDENTIFICAÇÃO DE UM UNIVERSO DE ATIVIDADES DO TRABALHO

O Universo Teórico, tal como foi originado, constituiu-se no Universo de atividades do trabalho, isto é, as atividades que poderiam ser treinadas.

2.1.5 - IDENTIFICAÇÃO DE UM UNIVERSO DE ATIVIDADES QUE SERÃO TREINADAS

A partir do Universo de Atividades que poderiam ser treinadas, foi identificado o Universo de Atividades que serão treinadas. Na identificação deste, foi considerada a seguinte restrição proposta por nosso trabalho:

- Na identificação de um Universo de Atividades a serem treinadas deve ser levada em conta uma restrição de tempo, ou seja, uma orientação no sentido de selecionar aquelas estritamente adequadas para que a duração do curso não o torne impraticável.

Tal restrição embasou nos procedimentos na identificação das atividades que serão treinadas no curso. Assim sendo, foram selecionadas atividades com base nos critérios de:

- . dificuldade (D) - através de julgamento profissional nosso e de especialistas;
- . universalidade (U) - pelas informações obtidas pelo questionário dos avaliadores. Foram consideradas universais as atividades exercidas por 75% ou mais dos respondentes;
- . crucialidade (C) - julgamento profissional nosso e de especialistas;
- . frequência (F) - informações obtidas pelo questionário dos avaliadores. Foram consideradas frequentes as atividades indicadas pelos avaliadores como diárias, semanais e mensais;
- . necessidade (N) - informações obtidas pelo questionário dos avaliadores. Foram consideradas necessárias as atividades requisitadas por 25% ou mais dos respondentes.

A seleção final das atividades que realmente serão treinadas foi feita através de julgamento profissional nosso, que combinava os critérios citados.

A Tabela II.5 apresenta os resultados da aplicação destes critérios às atividades que serão treinadas. Nesta Tabela o (X) assinalado junto aos critérios indica que, de acordo com os mesmos, a atividade deveria ser incluída no treinamento. O (X) na última coluna indica que a respectiva atividade será incluída no treinamento. Assim, por exemplo, a atividade "elabora Plano de Avaliação para orientar a execução de suas atividades" atendeu aos critérios de:

- . dificuldade, porque é uma atividade viável de ser ensinada em um Curso de Treinamento;
- . universalidade, porque 100% dos avaliadores respondentes executam esta atividade;
- . crucialidade, porque é uma atividade crítica para o bom desempenho do trabalho.

Assim sendo, julgamos que esta atividade deve ser incluída no Curso de Treinamento.

As células em branco indicam que a atividade não atingiu o critério correspondente. Observe-se que algumas atividades não satisfizeram nenhum dos critérios especificados.

TABELA 11.3

APLICAÇÃO DOS CRITÉRIOS PARA A SELEÇÃO DAS ATIVIDADES DO UNIVERSO A SER TREINADO

ATIVIDADES	CRITÉRIOS					ATIVIDADES QUE SERÃO TREINADAS
	D	U	C	F	N	
- define o sistema - (Currículo, Programa, Curso, Projeto): seus limites, elementos e características destes elementos.	X	X	X			X
- identifica os objetivos de sistema.	X	X	X			X
- determina os padrões e sua aplicação sobre os objetivos do sistema para estimar o valor dos mesmos.		X	X			X
- determina a viabilidade dos objetivos do sistema em relação aos recursos disponíveis.		X	X			X
- estuda e descreve detalhadamente o sistema corrente em termos de variáveis populacionais.	X	X	X			X
- estuda e descreve detalhadamente o sistema corrente em termos de variáveis instrucionais.	X	X	X			X
- estuda e descreve detalhadamente o sistema corrente em termos de variáveis comportamentais.	X	X	X			X

ATIVIDADES	CRITÉRIOS					ATIVIDADES QUE SERÃO TREINADAS
	D	U	C	F	N	
- compara a situação atual do sistema com a desejada, identificando discrepâncias	X	X	X			X
- identifica, fora do sistema, oportunidades não usadas por este			X			X
- levanta e responde questões, através de um estudo de Avaliação, que orientam o diagnóstico de problemas que geram as discrepâncias	X	X	X			X
- participa juntamente com o pessoal do sistema da identificação de estratégias alternativas para solucionar os problemas	X	X	X			X
- participa juntamente com o pessoal do sistema da operacionalização dos objetivos da mudança a ser introduzida para solucionar os problemas tornando-os preciosos e mensuráveis.	X	X	X			X
- participa juntamente com o pessoal do sistema da especificação de critérios para avaliar o alcance dos objetivos	X	X	X			X
- participa juntamente com o pessoal do sistema da identificação dos recursos humanos e materiais disponíveis na Agência para orientar a escolha da melhor estratégia		X				

ATIVIDADES	CRITÉRIOS					ATIVIDADES QUE SERÃO TREINADAS
	D	U	C	F	N	
- participa juntamente com o pessoal do sistema da identificação de "designs" de procedimentos (p/ a implementação da estratégia escolhida)	X	X				
- mantém um registro constante dos procedimentos tal como ocorrem	X		X		X	X
- detecta ou prediz defeitos do "design" de procedimentos		X	X			X
- fornece realimentação periódica sobre o desempenho do sistema	X	X	X			X
- mede alcance de objetivos	X	X	X			X
- compara efetividade com padrões pré-estabelecidos	X	X	X			X
- elabora Plano de Avaliação para orientar a execução de suas atividades						
- planeja coleta de dados incluindo fontes de informação	X	X	X			X
- planeja coleta de dados incluindo procedimentos de amostragem	X	X	X			X
- planeja coleta de dados incluindo instrumentais	X	X	X			X
- planeja coleta de dados incluindo o cronograma respectivo.		X				

ATIVIDADES	CRITÉRIOS					ATIVIDADES QUE SERÃO TREINADAS
	D	U	C	F	N	
- formula hipóteses testáveis	X	X	X			X
- especifica variáveis a serem medidas	X	X	X			X
- especifica dados para um teste de hipóteses rigoroso ou resposta inequívoca	X	X	X			X
- utiliza "designs" experimentais, pré-experimentais e quase-experimentais, de acordo com Campbell e Stanley.	X		X			X
- utiliza meios informais de investigação como estudo e visita a outros sistemas.						
- utiliza meios informais de investigação como participação em seminários e conferências.		X				
- utiliza procedimentos de investigação formal como técnica Delphi de convergência de opinião						

ATIVIDADES	CRITÉRIOS					ATIVIDADES QUE SERÃO TREINADAS
	D	U	C	F	N	
- utiliza procedimentos de investigação formal como técnica de pesquisa bibliográfica		X				
- utiliza técnicas de controle como: diagrama de marcos						
- utiliza técnicas de controle como: cronogramas mestres e parciais						
- utiliza técnicas de controle como: técnicas de O & M						
- aplica instrumentais de medida como: testes de aprendizagem	X		X		X	X
- aplica instrumentais de medida como: questionários		X				
- aplica instrumentais de medida como: escala de atitude						
- aplica instrumentais de medida como: fichas de observação						
- aplica instrumentais de medida como entrevistas		X				
- atribui pontos aos resultados obtidos na aplicação dos instrumentais de medida	X	X	X			X

ATIVIDADES	CRITÉRIOS					ATIVIDADES QUE SERÃO TREINADAS
	D	U	C	F	N	
- registra os dados obtidos através da aplicação dos instrumentais		X				
- para utilizar instrumentais você os seleciona						
- para utilizar instrumentais você os adapta						
- para utilizar instrumentais você os desenvolve	X	X	X			X
- com relação aos instrumentais de medida você faz: validação de conteúdo	X		X			X
- com relação aos instrumentais de medida você faz: validação de conteúdo						
- com relação aos instrumentais de medida você faz: validação preditiva						
- faz avaliação formativa de materiais	X		X			X
- faz avaliação somativa de materiais	X		X			X

ATIVIDADES	CRITÉRIOS					ATIVIDADES QUE SERÃO TREINADAS
	D	U	C	F	N	
- escolhe e emprega técnicas de Análise Estatística descritiva		X	X			
- escolhe e emprega técnicas de Análise Estatística inferencial	X		X		X	X
- utiliza técnicas analíticas como: gráficos, tabelas, perfis			X			
- utiliza, para a computação de dados, processamento manual		X				
- utiliza, para a computação de dados, o computador						
- interpreta dados em termos de um critério especificado	X	X	X			X
- elabora relatórios escritos e orais, para audiências especificadas, como por exemplo decisores, seus subordinados, clientela agências		X	X			

2.2 - PLANEJAMENTO DAS UNIDADES INSTRUÇÃOAIS

Após a identificação do universo de atividades a serem treinadas (bloco 2 da Figura VII.1), partimos para o planejamento das unidades instrucionais do curso, i.e.: elaboração dos objetivos do curso, estruturação e sequência do curso, especificação de conteúdos, identificação dos pré-requisitos, seleção de estratégias de treinamento e meios. Apresentamos, ainda, a carga horária por objetivo e número de alunos por instrutor.

Por razões práticas, não apresentamos a listagem dos objetivos na ordem em que foram gerados a partir do universo de atividades. Optamos pela apresentação da mesma, logo após a estrutura e sequência do curso, colocando os objetivos já na ordem das unidades e sub unidades instrucionais, evitando a colocação, no trabalho, de duas listas repetidas, diferindo, apenas, na sua ordenação.

A especificação do conteúdo para a consecução dos objetivos instrucionais se encontra no Apêndice F (Volume II).

2.2.1 - ESTRUTURA E SEQUÊNCIA DO CURSO

A estrutura e sequência do curso é apresentada na Figura II.1.

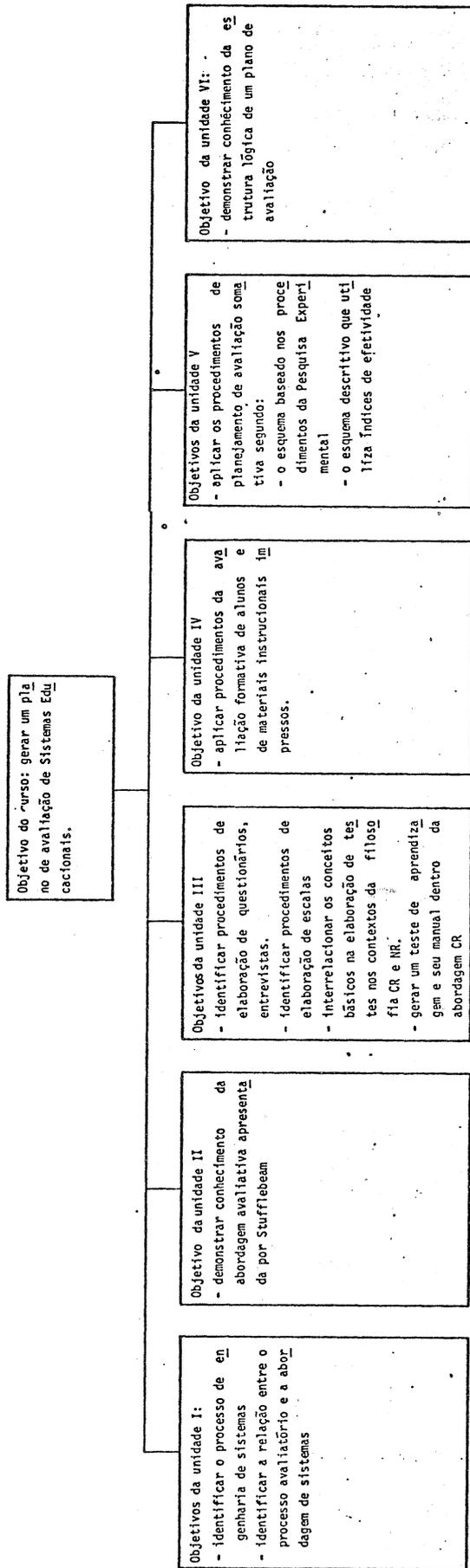


Fig. II.1 - Estrutura e Sequência do Curso

2.2.2 - OBJETIVOS ESPECÍFICOS DAS UNIDADES DO CURSO

A Tabela II.6, apresenta os objetivos específicos do curso. Os objetivos com asterístico deverão servir de base para a elaboração de testes. Os tipos de aprendizagem, de acordo com Gagné (1973), especificados nos objetivos, encontram-se identificados entre parênteses pelas seguintes legendas:

- I - informação verbal
- C - conceito
- P - princípio
- S - solução de problemas

TABELA II. 6

OBJETIVOS ESPECÍFICOS DAS UNIDADES DO CURSO

UNIDADE	SUB-UNIDADE	OBJETIVOS ESPECÍFICOS
I	1	O aluno deverá: <ul style="list-style-type: none">. apresentar as fases da Educação, explicando-as. (I).*. apresentar o planejamento educacional como um processo decisório, explicando-o. (I).. apresentar as vantagens e desvantagens do conhecimento especializado, explicando-as. (I).. apresentar as características básicas da teoria geral de sistemas, explicando-as. (I).

UNIDADE	SUB-UNIDADE	OBJETIVOS ESPECÍFICOS
I		<ul style="list-style-type: none">. apresentar as principais contribuições da teoria geral de sistemas, explicando-as. (I).*. identificar um sistema e seus elementos básicos, descrevendo-os. (C).*. apresentar as características básicas da abordagem sistêmica, explicando-as. (I).. apresentar os passos a serem seguidos na estruturação de um sistema, explicando-os. (I).
	2	<ul style="list-style-type: none">*. identificar as idéias básicas da filosofia de planejamento na qual se fundamenta o processo de Engenharia de Sistemas descrevendo-as. (C)*. identificar o porquê da sequência em que são utilizadas as questões: "O quê?" e o "como?" ao se planejar e executar um sistema, descrevendo-o. (C)*. apresentar as etapas do processo de planejamento de Engenharia de Sistemas e suas respectivas fases, explicando-as. (I).
	3	<ul style="list-style-type: none">*. apresentar as principais etapas do processo decisório, explicando-as. (I).*. apresentar a relação entre Avaliação e processo decisório, explicando-a. (I).

UNIDADE	SUB-UNIDADE	OBJETIVOS ESPECÍFICOS
I		<ul style="list-style-type: none">*. apresentar a relação entre processo decisório e planejamento de mudanças educacionais, explicando-a. (I).*. apresentar a relação entre Avaliação e o processo de planejamento de mudanças educacionais, explicando-a. (I).*. apresentar a relação entre abordagem sistêmica e o processo avaliatório, explicando-a. (I)*. identificar o uso da filosofia de planejamento do processo de Engenharia de Sistemas nos estudos de Avaliação, descrevendo-o. (C).
	4	<ul style="list-style-type: none">*. identificar os três maiores propósitos da Avaliação, descrevendo-os. (C).*. apresentar as abordagens avaliativas de Stake, Scriven, Provus e Stufflebeam explicando-as. (I).*. apresentar a Avaliação como medida, Avaliação congruência e Avaliação como julgamento, explicando-as. (I)
	5	<ul style="list-style-type: none">*. apresentar as diferenças básicas entre Avaliação e Pesquisa, explicando-as. (I)

UNIDADE	SUB-UNIDADE	OBJETIVOS ESPECÍFICOS
II	1	<ul style="list-style-type: none">. citar os elementos que caracterizam a Tomada de Decisão Educacional, <u>ex</u>pl<u>ic</u>ando-os. (I)..*apresentado um problema educacional, identificar o tipo de situação de <u>de</u>cisã<u>o</u> envolvido, no mesmo, <u>justifi</u>cando a resposta. (C)..*apresentado um problema educacional, identificar cada tipo de decisão <u>en</u>volvido no mesmo, justificando a <u>res</u>posta. (C).
	2	<ul style="list-style-type: none">.*descrever os tipos de Avaliação, <u>apre</u>sentados por Stufflebeam et al, seus propósitos e metodologias.(I)..*apresentado um problema educacional, identificar os tipos de Avaliação <u>en</u>volvidos no mesmo, justificando a resposta.(C).
	3	<ul style="list-style-type: none">.*citar as premissas que embasam a <u>abor</u>dagem de Avaliação, apresentada por Stufflebeam et al (I)..*descrever o relacionamento entre os tipos de decisão, as situações de <u>de</u>cisã<u>o</u> e os tipos de Avaliação dentro da abordagem de Avaliação, <u>apresenta</u>da por Stufflebeam et al.(I).

UNIDADE	SUB-UNIDADE	OBJETIVOS ESPECÍFICOS
III	1	<ul style="list-style-type: none">. * conceituar o questionário e a entrevista quanto ao conteúdo, tipos, formatos de questões, modos de respostas, vantagens e desvantagens. (I).. apresentar os passos para a elaboração de questionários e entrevistas, explicando-os. (I).. apresentar os procedimentos de validade e fidedignidade de questionários e entrevistas, explicando-os. (I).
	2	<ul style="list-style-type: none">. identificar os tipos de escalas de classificação pelas suas características. (C).. identificar as vantagens e desvantagens das escalas de classificação. (C).. * Dada uma descrição de uma escala de atitude, identificar o tipo de escala caracterizado pela descrição. (C).
	3	<ul style="list-style-type: none">. dada uma descrição de um tipo de teste, identificar o tipo de teste caracterizado pela descrição. (C).. apresentar as vantagens e desvantagens dos formatos de itens. (I).. apresentar os cuidados a serem tomados na elaboração dos formatos de itens. (I).

UNIDADE	SUB-UNIDADE	OBJETIVOS ESPECÍFICOS
		<ul style="list-style-type: none"><li data-bbox="758 360 1366 450">.* construir itens de teste paralelos. (P).<li data-bbox="758 477 1366 566">.* identificar as características da abordagem de teste NR. (C).<li data-bbox="758 593 1366 728">. apresentar as deficiências da abordagem NR para a Avaliação da instrução. (I).<li data-bbox="758 754 1366 844">.* explicar os procedimentos de elaboração de teste NR. (I).<li data-bbox="758 871 1366 1005">.* dados os resultados empíricos para um teste NR, selecionar os melhores itens. (P).<li data-bbox="758 1032 1366 1122">.* identificar as características da abordagem de teste CR. (C).<li data-bbox="758 1149 1366 1238">. apresentar as vantagens do teste CR para a Avaliação da instrução. (I).<li data-bbox="758 1265 1366 1355">.* explicar os procedimentos de elaboração de teste CR. (I).<li data-bbox="758 1382 1366 1516">. aperfeiçoar itens para um teste CR considerando as técnicas aplicáveis. (P).<li data-bbox="758 1543 1366 1677">.* Dadas situações de validação de teste, identificar o tipo de validade indicado. (C).<li data-bbox="758 1704 1366 1794">.* relacionar os tipos de validade às abordagens NR e CR. (I).<li data-bbox="758 1821 1366 1910">.* relacionar o escore observado com o escore-universo. (C).

UNIDADE	SUB-UNIDADE	OBJETIVOS ESPECÍFICOS
		<ul style="list-style-type: none">. * descrever os procedimentos para <u>es</u>timar a confiabilidade de teste. (I).. determinar a consistência interna de um teste. (P).. relacionar os métodos de estimar a confiabilidade às abordagens NR e CR. (I).. transformar um escore bruto em <u>per</u>centil. (P).. transformar um escore bruto em <u>esco</u>re padrão. (P).. normalizar escores. (P).. citar o propósito das normas. (I).. apresentados os dados empíricos de um teste, determinar se os objetivos foram atingidos. (P).. * atribuir pontos tanto para a <u>Avalia</u>ção para decisão comparativa como absoluta. (P).. descrever o processo de <u>administra</u>ção de testes incluindo as condições de testagem, procedimentos de <u>contro</u>le e direção dos sujeitos. (I).
IV	1	<ul style="list-style-type: none">. * identificar o propósito da <u>Avalia</u>ção Formativa e o momento em que ela ocorre, descrevendo-o. (C).. * apresentar fontes de informação <u>pa</u>ra a Avaliação Formativa, <u>descreven</u>do-as. (C).

UNIDADE	SUB-UNIDADE	OBJETIVOS ESPECÍFICOS
	2	<ul style="list-style-type: none">*: identificar o propósito da Avaliação Formativa de materiais instrucionais impressos, descrevendo-o. (C).*: apresentar evidências úteis para tomada de decisão relativa a melhoria de materiais instrucionais impressos, explicando-as. (I).*. apresentar variáveis de sala de aula que interferem no desempenho dos alunos, explicando-as. (I).*. identificar procedimentos para a Avaliação Formativa de materiais instrucionais impressos, descrevendo-os. (C).
	3	<ul style="list-style-type: none">*. identificar o propósito da Avaliação Formativa de alunos, descrevendo-o. (C).*. identificar procedimentos para a Avaliação Formativa de alunos, descrevendo-os. (C).*. apresentar algumas concepções errôneas sobre a Avaliação Formativa, descrevendo-as. (I).
V	1	<ul style="list-style-type: none">*. apresentar o tipo de informação fornecido ao decisor pelo uso do Esquema de Avaliação Somativa baseado nos procedimentos da Pesquisa Experimental. (I).

UNIDADE	SUB-UNIDADE	OBJETIVOS ESPECÍFICOS
		<ul style="list-style-type: none">. dada uma série de questões, identificar as que constituem problemas de Avaliação Somativa assinalando-as. (C).*. aplicar as regras para definição de problemas de Avaliação Somativa, formulando um. (P).. dada uma série de afirmativas, identificar as que constituem hipóteses para um problema de Avaliação Somativa, assinalando-as. (C).*. dado um problema, aplicar as regras para formulação de hipóteses, escrevendo uma. (P).*. dada uma hipótese, identificar as variáveis envolvidas na mesma, classificando-as. (C).. dada uma hipótese, identificar a variável independente como sendo de tratamento ou de classificação, nomeando-a. (C).. dada uma variável e seus níveis, identificá-los como: fixos, aleatórios, qualitativos ou quantitativos, nomeando-os. (C).*. dada uma hipótese, aplicar as regras para definição operacional de suas variáveis, definindo cada uma delas. (P).

UNIDADE	SUB-UNIDADE	OBJETIVOS ESPECÍFICOS
		<ul style="list-style-type: none">. apresentar as características básicas da amostragem que tornam seu uso adequado para estudos de Avaliação, explicando-as.(I).. apresentar os esquemas de amostragem probabilística e não probabilística, mais usuais, explicando-os. (I).*. dada uma situação que envolva o uso de uma amostra para representar a população, identificar o esquema de amostragem mais adequado no caso, justificando.(C).*. dada uma hipótese, identificar a população a que ela se refere, descrevendo-a.(C).. dada uma população e indicado o tamanho da amostra, extrair a amostra mais adequada.(P).. apresentar as fontes de invalidez externa e interna que podem influir nos resultados de um estudo de Avaliação, explicando-as. (I).. apresentar as principais técnicas para o controle de variáveis não experimentais que possam contaminar o experimento, explicando-as. (I).. apresentar os procedimentos dos "designs" de coleta de dados:<ul style="list-style-type: none">- um grupo pós-teste- um grupo pré-teste e pós-teste

UNIDADE	SUB-UNIDADE	OBJETIVOS ESPECÍFICOS
		<ul style="list-style-type: none">- grupo de controle prē-teste pōs-teste- grupo de controle somente pōs-teste- "designs" fatoriais- grupo de controle nōo equivalente- um grupo sērie no tempo:(I). <ul style="list-style-type: none">. apresentar a utilidade de cada "design" de coleta de dados:(I).*. dado um relatōrio de Avaliaçōo Somativa que utiliza "designs" experimentais ou quase-experimentais para coleta de dados, identificar as vantagens e desvantagens do "design" no controle das fontes de invalidade externa e interna, descrevendo-as.(C).*. dada uma situaçōo que envolva o uso de um "design" de coleta de dados, aplicar conhecimentos relativos a vantagens e desvantagens de cada "design" no controle das fontes de validade, e sua utilidade, selecionando o "design" mais adequado a esta situaçōo. (P).. apresentar o conceito de teste de hipōtese, explicando o seu procedimento geral. (I).. apresentar os passos do processo de testagem de hipōtese, explicando-os. (I).

UNIDADE	SUB-UNIDADE	OBJETIVOS ESPECÍFICOS
		<ul style="list-style-type: none">. dadas hipóteses alternativas, reescrevê-las sob a forma de sentenças matemáticas.(P).. dada uma hipótese alternativa, aplicar regras para transformá-la numa hipótese nula, escrevendo esta última sob a forma de uma sentença matemática.(P).. aplicar conhecimentos relativos a seleção do teste de hipótese. (P).. apresentar o conceito de erro do tipo I e erro do tipo II.(C).. dadas hipóteses alternativas e nulas sob a forma de sentenças matemáticas, aplicar as regras para determinação das regiões críticas. (P).. aplicar os procedimentos para elaboração de um plano de testagem de hipótese. (S).*. dadas hipóteses alternativas e nulas em teste, e os resultados teóricos e observados da estatística do teste empregado, aplicar as regras de decisão acerca de tais hipóteses. (P).
	2	*. apresentar o tipo de informação fornecido ao decisor através do uso do esquema descritivo de Avaliação Somativa. (I).

UNIDADE	SUB-UNIDADE	OBJETIVOS ESPECÍFICOS
		<ul style="list-style-type: none">. apresentar o conceito de <u>efetividade</u>.(I).*. apresentar os passos do esquema <u>des</u>critivo de <u>Avaliação Somativa</u> (I).*. fornecidos dados hipotéticos, aplicar regras para determinação dos <u>Índices</u> de efetividade. (P).
VI		<ul style="list-style-type: none">*. explicar a estrutura <u>lógica</u> de um <u>Plano de Avaliação</u>. (I).. descrever o relacionamento entre os passos do <u>Plano de Avaliação</u> com a <u>definição da Avaliação</u> apresentada por <u>Stufflebeam et al</u> (I).

2.2.3 - PRÉ-REQUISITOS DO CURSO

Estão listados abaixo, sob forma de objetivos comportamentais, os pré-requisitos do curso.

O aluno deverá ser capaz de:

- . identificar as escalas - nominal, ordinal, intervalo e razão
- . construir uma distribuição de frequência:
 - agrupar dados
 - amplitude
 - intervalo de classe
 - frequências - absoluta, acumulada, relativa e relativa acumulada
 - limites
 - ponto médio
- . calcular as medidas de tendência central: média e mediana
- . calcular as medidas de variabilidade - desvio padrão, variância
- . representar graficamente a distribuição de frequência
- . interpretar dados
- . calcular índices de correlação
- . identificar eventos aleatórios
- . aplicar os princípios básicos de probabilidade
- . aplicar os teoremas de probabilidade:
 - teorema da adição
 - teorema da multiplicação
- . identificar a distribuição binomial
- . aplicar o modelo binomial
- . identificar a distribuição normal

- . transformar a curva normal em normal reduzida
- . utilizar a tabela "áreas sob a curva normal"
- . identificar distribuições amostrais e erro padrão de medida
- . identificar outros modelos matemáticos baseados nos princípios de probabilidade:
 - distribuição t de Student
 - distribuição F. de Snedecor
- . consultar as tabelas:
 - valores t de Student
 - valores F. de Snedecor
- . construir objetivos instrucionais comportamentais

2.2.4 - SELEÇÃO DE MEIOS

Utilizando os procedimentos de seleção de meios propostos por Bretz (1971), foi constatado que a grande maioria dos objetivos instrucionais do curso conduziu à classe V, Visual Estático, como a mais apropriada. Esta classe inclui os seguintes meios:

- páginas impressas
- "filmstrip"
- conjunto de imagens
- microformas
- arquivo de vídeo

Qualquer um destes meios, pode ser utilizado para o alcance dos objetivos identificados. Porém, considerando que o problema de seleção de meios, no caso, prendeu-se à identificação do meio ou dos meios, que deverão ser utilizados num curso de treinamento; direcionado para suprir uma necessidade de avaliadores, a nível nacional, a seleção final refletiu a necessidade de meios de fácil acessibilidade, nos diferentes lugares. Fácil acessibilidade tem aqui o significado de materiais que podem ser produzidos, entregues e apresentados com menor

dispêndio de recursos humanos e materiais. Assim sendo, o material impresso é, aqui, sugerido por sua praticabilidade para a maioria dos objetivos. Este meio será utilizado basicamente para:

- comunicar, aos alunos, os objetivos da aprendizagem;
- apresentar informações básicas, ou seja, toda informação essencial para o alcance dos objetivos propostos.

O material impresso apresentando o conteúdo básico deixará o professor livre para: corrigir as tarefas dirigidas dos alunos, diagnosticar seus problemas e orientá-los individualmente.

Para alguns objetivos, porém, nenhum meio de comunicação foi julgado mais adequado em termos de sua propriedade e praticabilidade do que o professor, elemento do sistema instrucional, não identificado como um meio de comunicação. Neste caso, o professor utilizará materiais impressos e transparências como meios auxiliares.

As Figuras II.2, II.3 e II.4 apresentam um exemplo de seleção de meios, segundo o modelo de Bretz, para a consecução de um objetivo instrucional: "O aluno deverá apresentar as abordagens avaliativas de Stake, Scriven, Provus e Stufflebeam et al, explicando-as". A linha mais grossa indica a direção seguida para a seleção do meio. Note-se que o resultado da análise, neste exemplo, levou à classe V - Visual Estático.

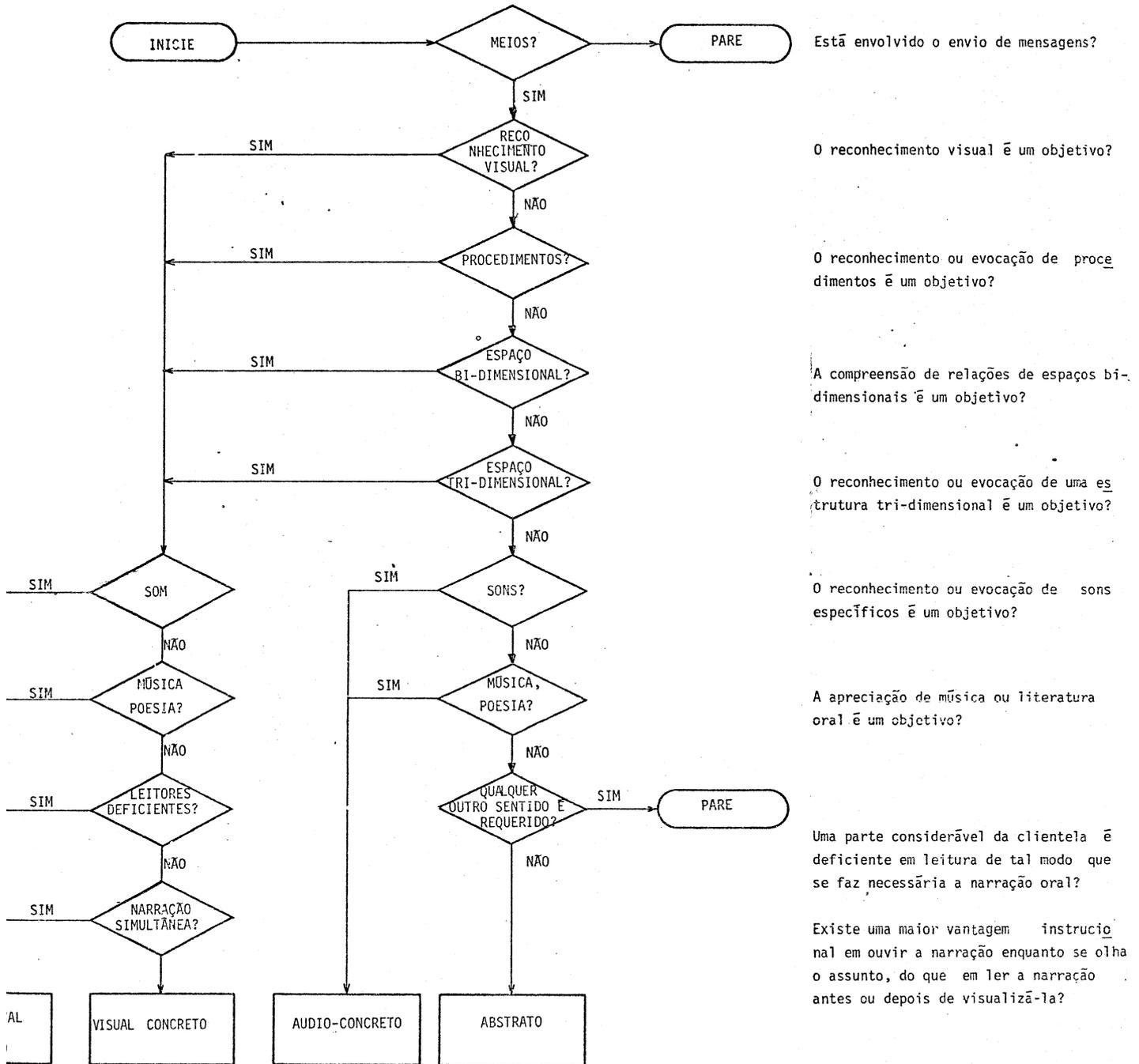


Fig. II.2 - Critério para Distinção entre Conteúdo Concreto e Abstrato (Bretz, 1971)

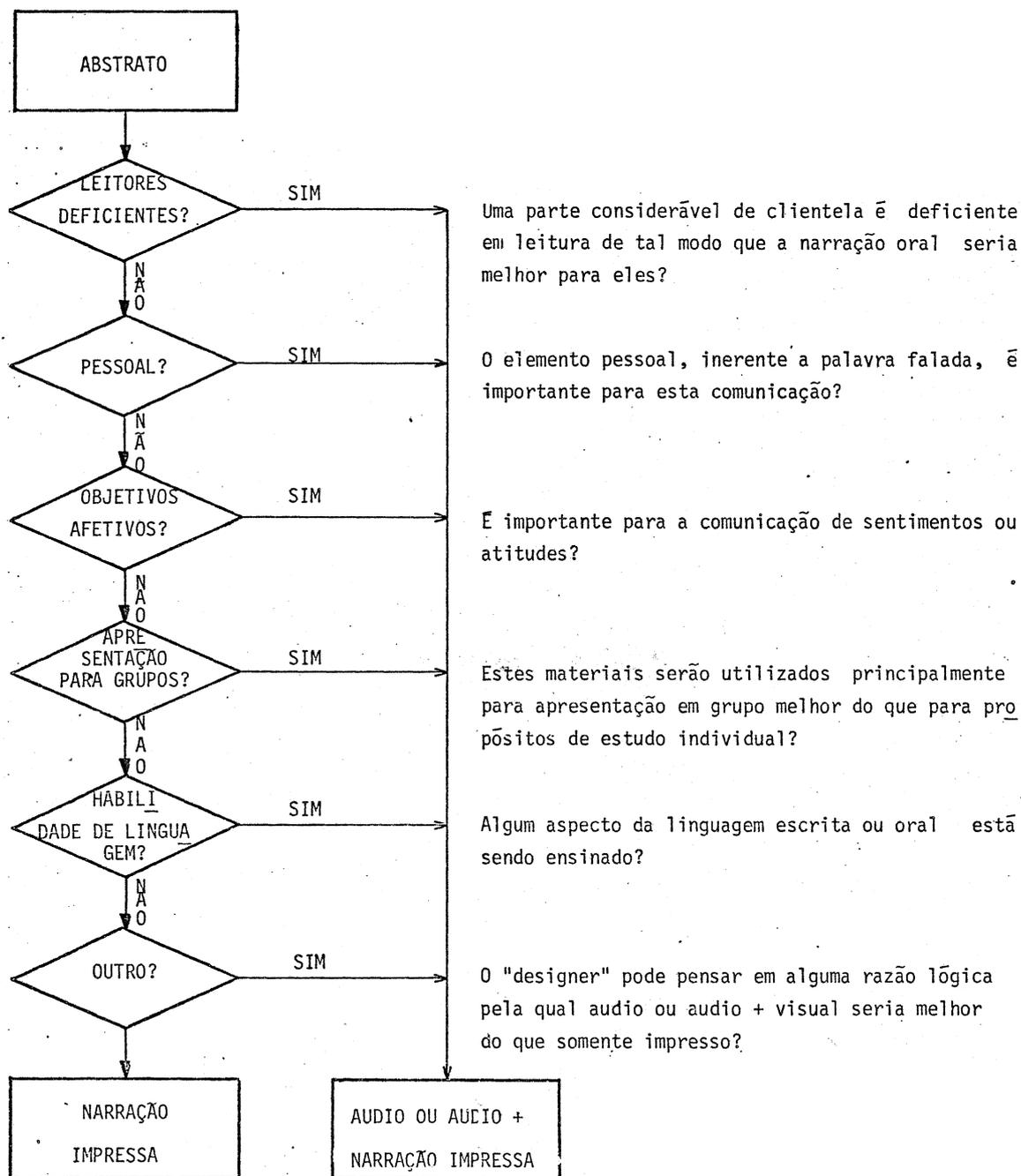


Fig. II.3 - Escolha dos Meios de Narração (Bretz, 1971)

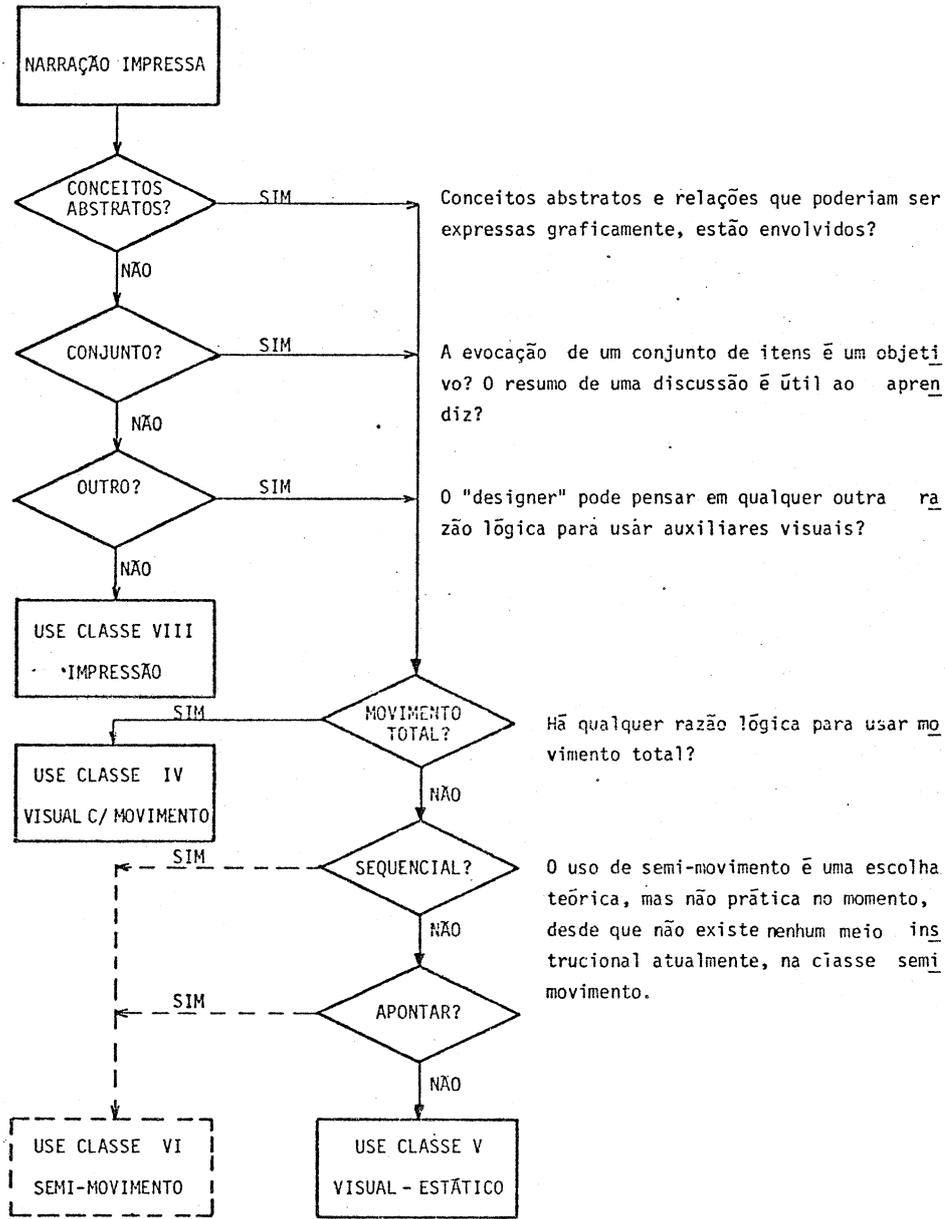


Fig. II.4 - Visualização da Narração Impressa (Bretz, 1972).

Os resultados da utilização do modelo de Bretz (1971) para seleção dos meios instrucionais são apresentados na Tabela II.7, mostrada no tópicó seguinte. Para cada objetivo instrucional, foi repetido o procedimento exemplificado.

2.2.5 - SELEÇÃO DAS ESTRATÉGIAS INSTRUCIONAIS, MEIOS, E DISTRIBUIÇÃO DE HORAS DE AULA PARA A CONSECUÇÃO DE CADA OBJETIVO INSTRUCIONAL

A Tabela II.7 apresenta o planejamento instrucional das unidades do curso de treinamento de Avaliadores Educacionais. Para cada objetivo (2a. coluna) são indicados a estratégia (3a. coluna), o meio instrucional selecionado (4a. coluna) e a carga horária (horas de aula) correspondente (5a. coluna).

As estratégias instrucionais encontram-se descritas na página 46.

A legenda abaixo facilita a leitura da Tabela II.7:

EDI = Estudo Dirigido Individual

EDG = Estudo Dirigido em Grupo

AE = Aula Expositiva

TDG = Tarefa Dirigida em Grupo

TDI = Tarefa Dirigida Individual

MI = Material Impresso

OBJETIVOS, ESTRATÉGIAS INSTRUCIONAIS, MEIOS E HORAS/AULA DO CURSO DE TREINAMENTO DE AVALIAÇÃO

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCIONAIS	MEIOS	HORAS DE AULA
I	<p>Aplicação do Pré-Teste</p> <p>O aluno deverá:</p> <ul style="list-style-type: none"> • apresentar as fases da Educação, explicando-as. • apresentar o conceito de planejamento educacional como um processo decisório, explicando-o. • apresentar as vantagens e desvantagens do conhecimento especializado, explicando-as. • apresentar as características básicas da teoria geral de sistemas, explicando-as. • apresentar as principais contribuições da teoria geral de sistemas, explicando-as. 		MI	2:00
		EDI	MI	00:40
		EDI	MI	00:40
		EDI	MI	00:40
		EDI	MI	00:30
		EDI	MI	00:30

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
I	<ul style="list-style-type: none"> - identificar um sistema e seus elementos básicos, descrevendo-os. - apresentar as características básicas da abordagem sistêmica, explicando-as. - apresentar os passos a serem seguidos na estruturação de um sistema, explicando-os - identificar as idéias básicas da filosofia de planejamento na qual se fundamenta o processo de Engenharia de Sistemas, descrevendo-as. 	<p>EDG + TDG</p> <p>EDI</p> <p>EDI</p> <p>EDG + TDG</p>	<p>MI</p> <p>MI</p> <p>MI</p> <p>MI</p>	<p>3:00</p> <p>2:00</p> <p>2:00</p> <p>2:00</p>

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
I	<ul style="list-style-type: none">- identificar o porquê da seqüência em que são utilizadas as questões: "O quê?" e o "Como?" ao se planejar e executar um sistema, descrevendo-o.- apresentar as etapas do processo de planejamento de Engenharia de Sistemas e suas respectivas fases, explicando-as.- apresentar as principais etapas do processo decisório, explicando-as.- apresentar a relação entre avaliação e processo decisório, explicando-a.	<p>EDG</p> <p>EDG</p> <p>EDG</p> <p>EDG</p>	<p>MI</p> <p>MI</p> <p>MI</p> <p>MI</p>	<p>3:00</p> <p>6:00</p> <p>1:00</p> <p>00:40</p>

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
I	<ul style="list-style-type: none">- apresentar a relação entre processo decisório e planejamento de mudanças educacionais, explicando-a.- apresentar a relação entre Avaliação e o processo de planejamento educacionais, explicando-a.- apresentar a relação entre abordagem sistêmica e o processo avaliatório, explicando-a.- identificar o uso da filosofia de planejamento do processo de Engenharia de Sistemas nos estudos de Avaliação, descrevendo-o	<p>EDG</p> <p>EDG</p> <p>EDG</p> <p>EDG + TDG</p>	<p>MI</p> <p>MI</p> <p>MI</p> <p>MI</p>	<p>00:40</p> <p>00:40</p> <p>1:00</p> <p>1:30</p>

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUIONAIS	MEIOS	HORAS DE AULA
I	<ul style="list-style-type: none"> - identificar os três maiores propósitos da Avaliação, descrevendo-os. - apresentar os conceitos de Avaliação de Stake, Scriven, Provus e Stufflebeam, explicando-os. 	EDG + TDI	MI	00:40
	<ul style="list-style-type: none"> - apresentar a Avaliação como medida, Avaliação como congruência e Avaliação como julgamento, explicando-as. 	EDG	MI	3:00
	<ul style="list-style-type: none"> - apresentar as diferenças básicas entre Avaliação e Pesquisa, explicando-as. 	EDG	MI	3:00
	Aplicação do Pós-Teste (forma paralela do Pré-teste)			3:00

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
II	<p>Aplicação do Pré Teste</p> <ul style="list-style-type: none"> . citar os elementos que caracterizam a tomada de Decisão Educacional, explicando-os. . apresentado um problema educacional, identificar o tipo de situação de decisão envolvido, no mesmo, justificando. . apresentado um problema educacional, identificar cada tipo de decisão envolvido, no mesmo, justificando. . descrever os tipos de Avaliação, apresentado por Stufflebeam et al (1971), seus propósitos e metodologias. 	<p>EDG</p> <p>TDI</p> <p>TDI</p> <p>EDG</p>	<p>MI</p> <p>MI</p> <p>MI</p> <p>MI</p>	<p>02:00</p> <p>05:00</p> <p>00:30</p> <p>00:30</p> <p>10:00</p>

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS/AULA
	Aplicação do Pós-Teste	(Forma Paralela do Pré-Teste)		03:00
III	Aplicação do pré-teste			2:00
	conceituar o <u>questionário</u> e a <u>entrevista</u> quanto ao conteúdo, tipos, formatos de questões, modos de respostas, vantagens e <u>desvantagens</u> .	EDI	MI	2:00

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
III	<ul style="list-style-type: none">• apresentar os passos para a elaboração de questionários e entre vistas, explicando-os.• apresentar os procedimentos de validade e fidedignidade de questionários e entrevistas, explicando-os.• identificar os tipos de escalas de classificação pelas suas características.• identificar as vantagens e desvantagens das escalas de classificação.	EDI EDI EDI EDI	MI MI MI MI	0:20 0:20 0:20 0:10

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
III	<ul style="list-style-type: none"> • dada uma descrição de uma escala de atitude identificar o tipo de escala caracterizado pela descrição. • dada uma descrição de um tipo de teste, identificar o tipo de teste caracterizado pela descrição. • identificar as vantagens e desvantagens dos formatos de itens. • identificar os cuidados a serem tomados na elaboração dos formatos de itens. • construir itens de testes paralelos. 	<p>EDG</p> <p>EDG</p> <p>EDI</p> <p>EDI</p> <p>EDI + TDI</p>	<p>MI</p> <p>MI</p> <p>MI</p> <p>MI</p> <p>MI</p>	<p>2:30</p> <p>2:00</p> <p>1:30</p> <p>1:30</p> <p>4:00</p>

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
III	<ul style="list-style-type: none"> • identificar as características da abordagem de teste NR. • apresentar as deficiências da abordagem NR para a Avaliação da instrução. • explicar os procedimentos de elaboração de teste NR. • Dados os resultados empíricos para um teste NR, selecionar os melhores itens. • identificar as características da abordagem de teste CR. 	<p>EDI</p> <p>EDI</p> <p>EDI</p> <p>TDI</p> <p>EDI</p>	<p>MI</p> <p>MI</p> <p>MI</p> <p>MI</p> <p>MI</p>	<p>1:00</p> <p>0:15</p> <p>2:00</p> <p>2:00</p> <p>1:00</p>

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
III	<ul style="list-style-type: none"> • apresentar as vantagens do teste CR para a Avaliação da instrução. • explicar os procedimentos de elaboração de teste CR. • aperfeiçoar itens para um teste CR considerando as técnicas aplicáveis. 	<p>EDI</p> <p>EDI</p> <p>TDI</p>	<p>MI</p> <p>MI</p> <p>MI</p>	<p>0:15</p> <p>2:00</p> <p>1:00</p>
	<ul style="list-style-type: none"> • Dadas situações de avaliação de teste, identificar o tipo de validade indicado. 	<p>EDG</p>	<p>MI</p>	<p>3:00</p>
	<ul style="list-style-type: none"> • relacionar os tipos de validade às abordagens NR e CR. 	<p>EDG</p>	<p>MI</p>	<p>2:00</p>

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
III	<ul style="list-style-type: none"> • apresentar o <u>relacionamento</u> entre o <u>escore observado</u> e o <u>escore-universo</u>. • <u>descrever</u> os <u>procedimentos</u> para <u>estimar</u> a <u>confiabilidade</u> de teste. • <u>determinar</u> a <u>consistência interna</u> de um teste • <u>relacionar</u> os <u>métodos</u> de <u>estimar</u> a <u>confiabilidade</u> às <u>abordagens</u> NR e CR. • <u>transformar</u> um <u>escore bruto</u> em <u>percentil</u>. • <u>transformar</u> um <u>escore bruto</u> em <u>escore padrão</u>. 	<p>AE</p> <p>EDI</p> <p>TDI</p> <p>EDG</p> <p>EDI + TDI</p> <p>EDI + TDI</p>	<p>MI (meio auxiliar)</p> <p>MI</p> <p>MI</p> <p>MI</p> <p>MI</p> <p>MI</p> <p>MI</p>	<p>0:30</p> <p>3:00</p> <p>1:00</p> <p>2:00</p> <p>0:20</p> <p>0:30</p>

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUIONAIS	MEIOS	HORAS DE AULA
III	<ul style="list-style-type: none">normalizar escorescitar o propósito das normas,apresentados os dados empíricos de um teste, determinar se os objetivos foram atingidos.atribuir pontos tanto para a Avaliação para decisão comparativa como absoluta.descrever o processo de administração de testes incluindo as condições de testagem, procedimentos de controle e direção dos sujeitos.	EDI + TDI EDI AE + TDI TDI EDI	MI MI MI + Transparência (meios auxiliares) MI MI	0:30 0:20 2:00 1:30 1:30

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
III	construir um teste de aprendizagem e seu manual (objetivo geral da subunidade 3).	TDG		10:00 OBS.: O teste deve ser elaborado durante o desenvolvimento do curso
	Aplicação do pós-teste	(forma paralela do pré-teste)		3:00
IV	Aplicação do pré-teste			2:00
	identificar o propósito da Avaliação Formativa e o momento em que ela ocorre, descrevendo-o.	EDI + TDI	MI	1:00
	apresentar fontes de informação para a Avaliação Formativa, descrevendo-as.	EDI	MI	2:00

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
IV	<p>- identificar o propósito da Avaliação Formativa de materiais instrucionais impressos, explicando-os.</p> <p>- apresentar algumas evidências úteis para tomada de decisão relativa a melhoria de materiais instrucionais impressos, explicando-as.</p> <p>- apresentar variáveis de sala de aula que interferem no desempenho dos alunos, explicando-as</p>	<p>EDI</p> <p>EDI</p> <p>EDI</p>	<p>MI</p> <p>MI</p> <p>MI</p>	<p>00:40</p> <p>1:00</p> <p>1:30</p>

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
IV	<ul style="list-style-type: none"> - identificar procedimentos para a Avaliação Formativa de materiais instrucionais impressos, descrevendo-os. - identificar o propósito da Avaliação Formativa de alunos, descrevendo-o. - identificar procedimentos para a Avaliação Formativa de alunos, descrevendo-os. - apresentar algumas concepções errôneas de Avaliação Formativa, descrevendo-as. 	<p>EDG</p> <p>EDI + TDI</p> <p>EDG</p> <p>EDI</p>	<p>MI</p> <p>MI</p> <p>MI</p> <p>MI</p>	<p>10:00</p> <p>1:00</p> <p>8:00</p> <p>00:20</p>

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
IV	<p>- aplicar procedimentos de Avaliação Formativa de materiais instrucionais impressos e de alunos (objetivo geral da unidade IV)</p>	TDG	MI	12:00
	Aplicação do pós-teste			3:00
V	<p>Aplicação do pré-teste</p> <p>- apresentar o tipo de informação fornecido ao decisor através do uso do esquema de Avaliação Somativa baseado nos procedimentos da Pesquisa Experimental.</p> <p>. Dada uma série de questões, identificar as que constituem problemas de Avaliação Somativa, assinalando-as.</p>	EDI	MI	02:00 00:30 00:20

UNIDADE	OBJETIVOS	ESTRATEGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
V	<ul style="list-style-type: none">• aplicar as regras para definição de Problemas de Avaliação Somativa, formulando um,• dada uma série de afirmativas, identificar as que constituem hipóteses para um problema de Avaliação Somativa, assinalando-as.• dado um Problema, aplicar as regras para formulação de hipóteses, escrevendo uma.• dada uma hipótese, identificar as variáveis envolvidas na mesma, classificando-as.	EDI + TDI EDI + TDI EDI + TDI EDI + TDI	MI MI MI MI	00:40 00:10 00:20 01:30

V	<ul style="list-style-type: none">. apresentar as <u>características</u> básicas da amostragem que tornam seu uso <u>adequado</u> para estudos de Avaliação, explicando-as.. apresentar os <u>esquemas</u> de amostragem <u>probabilística</u> e <u>não probabilística</u>, mas usuais, explicando-os.. dada uma situação que <u>envolva</u> o uso de amostragem <u>representativa</u>, identificar o <u>esquema</u> de amostragem mais <u>adequado</u> no caso, <u>explicando-os</u>.. dada uma hipótese, <u>identificar</u> a população a que ela se refere, <u>descrevendo-a</u>.. dada uma população e <u>indicado</u> o tamanho da amostra, <u>extrair</u> a amostra mais adequada.	EDII EDII TDI EDII + TDI TDG	MI MI MI MI	00:30 02:00 00:30 00:30 00:30
---	--	--	----------------------------------	---

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
V	<ul style="list-style-type: none">- grupo de controle pré-teste pós-teste- grupo de controle somente pós-teste- "designs" fatoriais- grupo de controle não equivalente- um grupo série no tempo. <p>apresentar a utilidade de cada "design" de coleta de dados;</p> <p>dado um relatório de Avaliação Somativa que utiliza "designs" experimentais ou quase experimentais, identificar as vantagens e desvantagens do "design" no controle das fontes de validade externa e interna, descrevendo-as.</p>	EDI EDI + TDG	MI MI	01:30 05:00

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
V	<ul style="list-style-type: none"> • dada uma hipótese alternativa, aplicar regras para transformá-la numa hipótese nula, escrevendo esta última. • aplicar conhecimentos relativos a seleção do teste de hipótese. • apresentar o conceito de erro do tipo I e erro do tipo II • dadas hipóteses alternativas e nulas sob a forma de sentenças matemáticas, aplicar as regras para determinação das regiões críticas. 	<p>AE + EDI + TDI</p> <p>AE + EDI + TDI</p> <p>AE + EDI</p> <p>AE + EDI + TDI</p>	<p>MI (meio auxiliar)</p> <p>MI (meio auxiliar)</p> <p>MI (meio auxiliar)</p> <p>MI (meio auxiliar)</p>	<p>00:30</p> <p>04:00</p> <p>02:00</p> <p>02:00</p>

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
V	<ul style="list-style-type: none">• aplicar os procedimentos de elaboração de um plano de testagem de hipótese.• dadas hipóteses alternativas e nulas em teste, e os resultados teóricos e observados da estatística do teste empregado, aplicar as regras de decisão acerca de tais hipóteses.• aplicar procedimentos para planejar uma AV. Somativa, utilizando o esquema baseado nos procedimentos da Pesquisa Experimental.	TDG AE + EDI + TDI TDG	MI (meio auxiliar)	02:00 03:00 06:00

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUIONAIS	MEIOS	HORAS DE AULA
V	<ul style="list-style-type: none"> aplicar procedimentos para planejar uma Av. Somativa utilizando o esquema descritivo, que utiliza índices de efetividade. 	TDG		03:00
	Aplicação do pós-teste			03:30
VI	Aplicação do pré-teste			02:00
	<ul style="list-style-type: none"> explicar a estrutura lógica de um Plano de Avaliação. 	EDG	MI	05:00
	<ul style="list-style-type: none"> descrever o relacionamento entre os passos do Plano de Avaliação, com a definição de Avaliação, apresentada por Stufflebeam, et al (1971). 	EDI	MI	00:30

UNIDADE	OBJETIVOS	ESTRATÉGIAS INSTRUCCIONAIS	MEIOS	HORAS DE AULA
VI	<ul style="list-style-type: none"> dada uma situação envolvendo o planejamento da Avaliação, discutir a aplicação da estrutura lógica do Plano de Avaliação na situação. (Objetivo Geral da Unidade VI). 	TDG		05:00
	<ul style="list-style-type: none"> Aplicação do pós-teste (Forma Paralela do Pré-Teste) gerar um plano de avaliação (objetivo terminal do curso). 	TDG		03:00 40:00

2.2.6 - DURAÇÃO DO CURSO

A duração total prevista para o curso é de 279 horas. Considerando seu funcionamento com 6 horas diárias, serão necessários 47 dias úteis.

2.2.7 - NÚMERO DE ALUNOS POR PROFESSOR

Sugerimos que deva haver um professor para cada grupo de vinte e cinco alunos. Este número é sugerido com base na experiência profissional das autoras.

Neste capítulo foi apresentado o universo de atividades que serão treinadas, e foi desenvolvido o planejamento instrucional das unidades do curso. Este último engloba: a estrutura e sequência do curso, objetivos específicos das unidades, pré-requisitos do curso, seleção de estratégias e meios instrucionais, carga horária relativa a cada objetivo específico, duração total do curso, e proporção número de alunos por professor com que o mesmo deverá funcionar.

O próximo capítulo trata do planejamento do teste piloto e da instalação do curso de treinamento.

CAPÍTULO III

PLANEJAMENTO DO TESTE PILOTO E INSTALAÇÃO DO CURSO DE TREINAMENTO

Com base no esquema apresentado na Figura I.7, foi feito o planejamento da obtenção e operação do Ensaio e da Instalação do curso, utilizando técnicas de Engenharia de Sistemas.

O Ensaio consiste na implementação do Curso com o objetivo de validá-lo, ou seja, avaliar procedimentos planejados e consecução dos objetivos propostos - validação intrínseca - e se estes são congruentes com as atividades do trabalho - validação extrínseca.

Mais especificamente, a validação intrínseca e extrínseca têm os seguintes objetivos:

- Detetar pontos falhos na instrução e/ou nos testes de aprendizagem;
- Verificar a adequação:
 - . da estrutura e sequência das unidades instrucionais do curso;
 - . da relação entre distribuição de horas de aula diárias e duração total do curso;
 - . da frequência do uso das estratégias instrucionais;
 - . das estratégias de Avaliação utilizadas durante o ensaio do curso;
 - . da carga horária planejada para cada objetivo instrucional;
 - . da estratégia instrucional planejada para cada objetivo.
- Verificar a congruência entre informação transmitida pelo instrutor e informação contida nos materiais instrucionais;
- Verificar a efetividade do curso em termos da aprendizagem dos alunos;

- Verificar a utilidade do curso, em termos das atividades desempenhadas pelos alunos, no trabalho subsequente.

A Instalação do Curso consiste na utilização do mesmo, depois de comprovada sua validade intrínseca e extrínseca.

Os procedimentos de validade extrínseca são mantidos continuamente como normais do curso, garantindo a atualização do mesmo.

Sempre que reformulações substanciais no curso forem necessárias, o decisor deverá julgar se, após estas reformulações, será conveniente um novo ensaio ou, então, a instalação imediata. Se a decisão for por um novo ensaio, deverão ser utilizados os mesmos procedimentos identificados para esta situação.

Este planejamento é dirigido para Órgãos ou Instituições Educacionais que se ocupam com treinamento de recursos humanos. Assim sendo, no planejamento do Ensaio e da Instalação do Curso, foi considerado que a Instituição que o executará dispõe dos seguintes requisitos básicos:

- . prédio e instalações;
- . pessoal técnico, administrativo e de apoio;
- . materiais permanentes e de consumo.

Tendo em vista estes requisitos, serão tratadas, aqui, apenas as atividades e recursos diretamente ligados ao processo instrucional, para garantir a adequação dos mesmos aos propósitos deste trabalho.

Considerando que a Instituição, para quem está dirigido este planejamento, dispõe de infra-estrutura própria para o desenvolvimento de cursos de treinamento, não será abordado, também, o custo deste curso específico, uma vez que encargos diversos encontram-se diluídos na Instituição.

3.1 - ENSAIO DO CURSO

3.1.1 - OBTENÇÃO DO ENSAIO

A obtenção do ensaio consiste na execução de todas as atividades necessárias para conseguir os recursos para a operação do sistema.

3.1.1.1 - ESTRUTURA DE DIVISÃO DE ATIVIDADES (EDA)

A Figura III.1 apresenta a Estrutura de Divisão de Atividades para a obtenção do ensaio do curso.

3.1.1.2 - ESTRUTURA DE DIVISÃO DE RECURSOS (EDR)

A Figura III.2 apresenta a Estrutura de Divisão de Recursos para a obtenção do ensaio do curso.

3.1.1.3 - ESPECIFICAÇÃO DOS RECURSOS PARA A OBTENÇÃO DO ENSAIO DO CURSO

Da Figura III.2 serão especificados apenas os recursos referentes ao pessoal técnico e a amostra da clientela. O pessoal de apoio e os recursos relativos aos materiais permanentes e de consumo não serão, aqui, especificados, visto que poderão ser utilizados aqueles disponíveis na Instituição. Quanto ao pessoal técnico e a amostra da clientela, deverão ser consideradas as especificações que se seguem.

1.1 - PESSOAL TÉCNICO

Refere-se ao pessoal necessário para as atividades de Planejamento, Avaliação e Administração do curso.

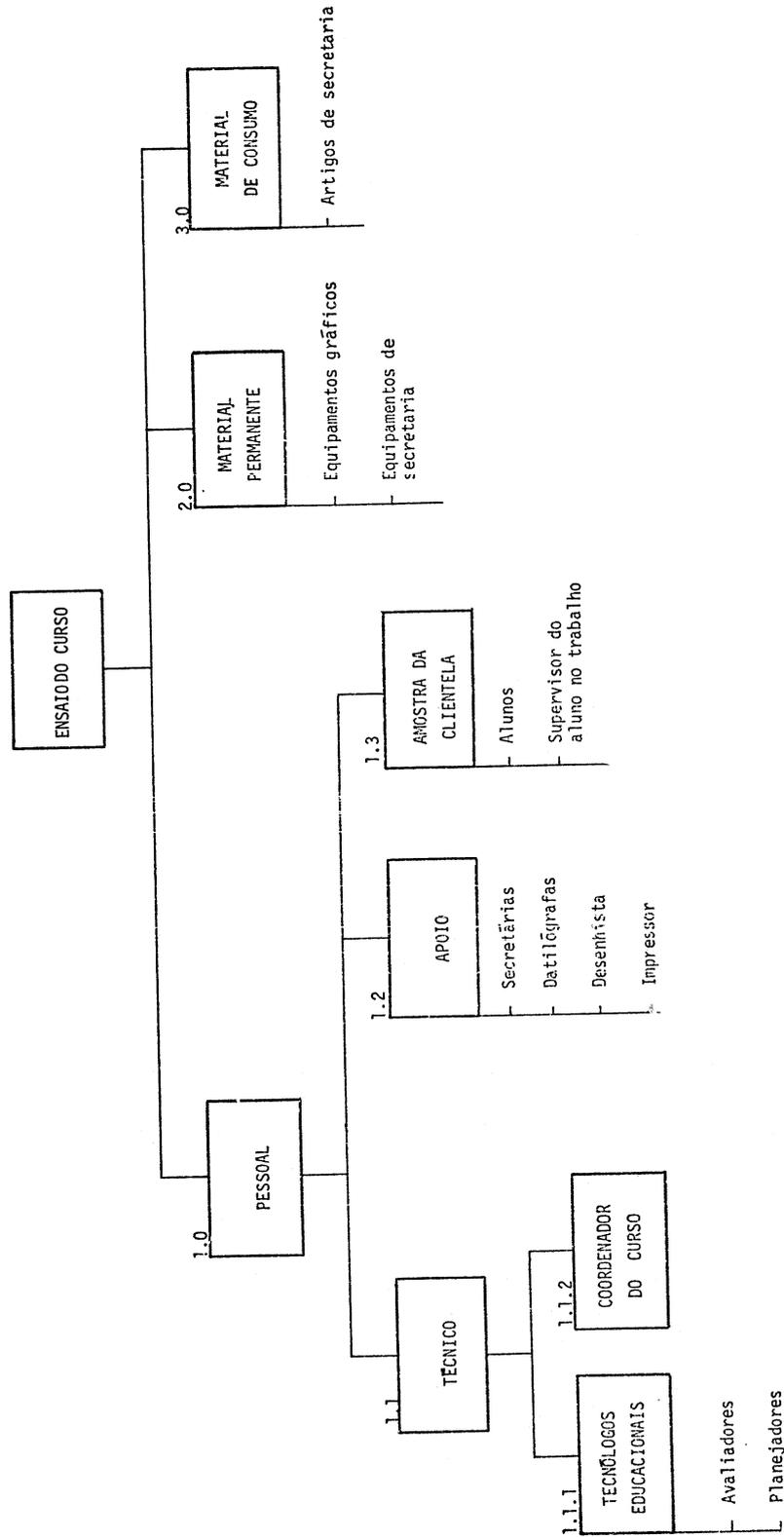


Fig. III.2 - Estrutura de Divisão de Recursos para a Obtenção do Ensaio do Curso

1.1.1 - TECNÓLOGOS EDUCACIONAIS

a) Avaliadores

- Nº: dois avaliadores

- Características:

- . especialização a nível de mestrado em Tecnologia Educacional na área de Avaliação, com conhecimentos relevantes sobre Avaliação de Projetos Educacionais;
- . dois anos de experiências em atividades avaliativas em Projetos Educacionais.

- Atividades

- . Desenvolvimento dos testes de aprendizagem e dos outros instrumentos de medida.

b) Planejadores

- Nº: um planejador

- Características:

- . especialização a nível de mestrado em Tecnologia Educacional na área de Planejamento Instrucional;
- . dois anos de experiência na área.

- Atividades

- . elaboração das tarefas de aprendizagem a partir dos objetivos instrucionais especificados no Capítulo II (página 87);
- . organização dos materiais instrucionais;
- . reformulação dos materiais instrucionais.

1.1.2 - COORDENADOR

- Nº: um coordenador

- Características:

- . curso superior completo, em qualquer área de conhecimento;

- . dois anos de experiência em coordenação de curso de treinamento ou equivalente.

- Atividades:

- . Responsável pelas decisões administrativas relativas à execução deste planejamento, na fase de obtenção do ensaio, no que diz respeito à: aquisição de recursos materiais e humanos, cumprimento das atividades em tempo previsto pelo cronograma.

1.3 - AMOSTRA DA CLIENTELA

- Nº de alunos: seis
- Nº de supervisores: três
- Nº de observadores: dois

Características

- alunos

- . nível superior

- supervisores

- . indivíduos que exerçam a função de supervisão de atividades de Avaliação em Instituições Educacionais.

3.1.1.4 - DIAGRAMA DE FLUXO DE TRABALHO (DFT)

A Figura III.3 apresenta o Diagrama de Fluxo de Trabalho para a obtenção do ensaio do curso.

3.1.1.5 - DESCRIÇÃO DAS ATIVIDADES DO DFT (FIGURA III.3) OBTENÇÃO DO ENSAIO DO CURSO.

1.0 - Recrutar pessoal para o ensaio do curso

Consiste no recrutamento (realocação ou contratação) de pessoal conforme indicado na Estrutura de Divisão de Recursos da Obtenção do ensaio do curso (Figura III.2).

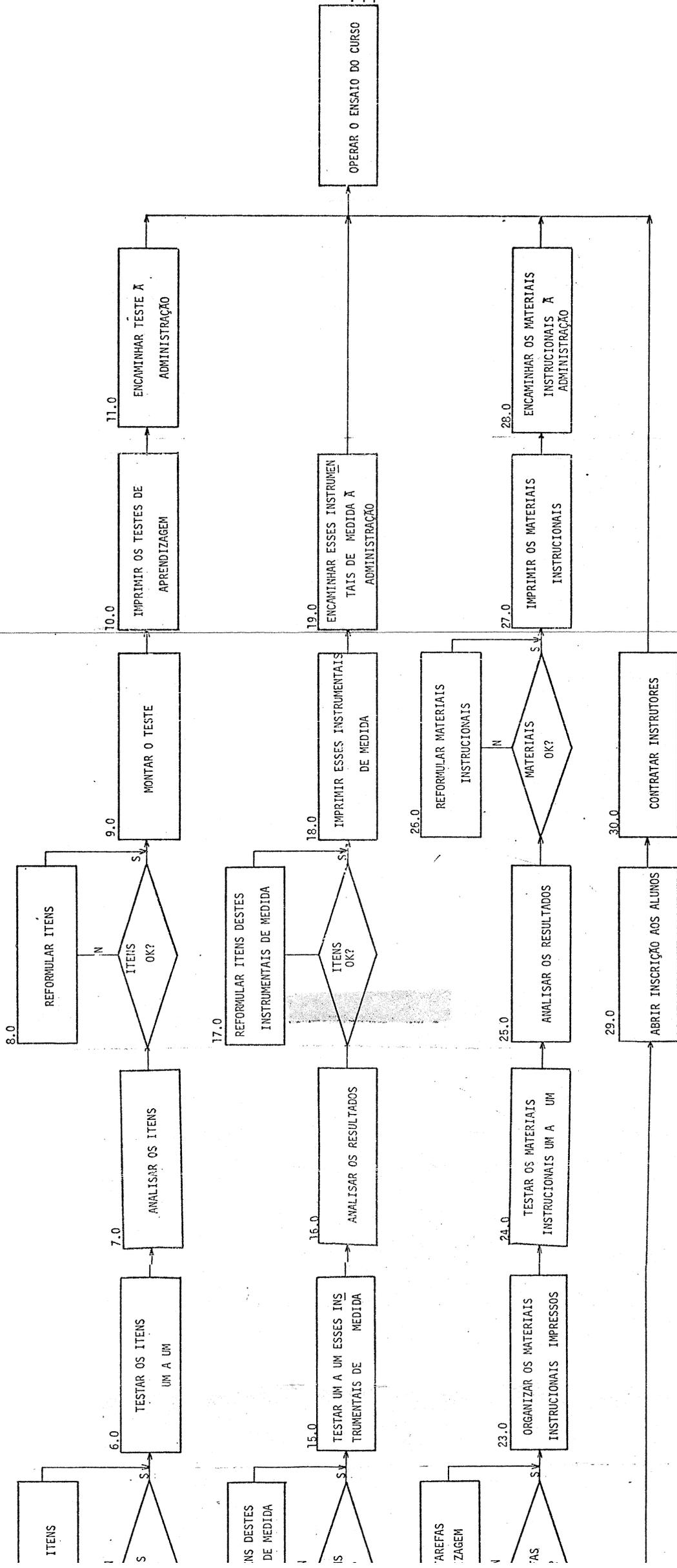
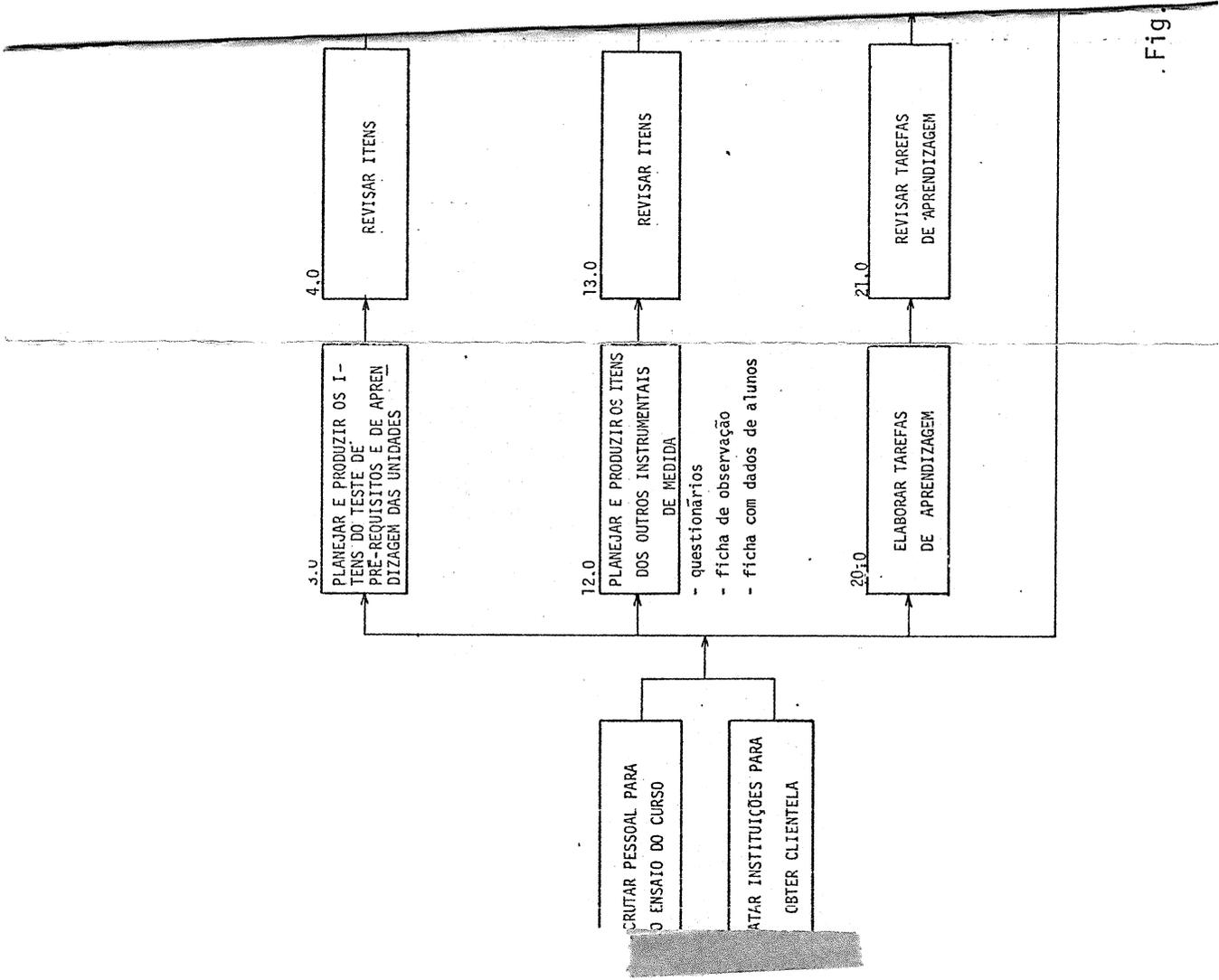


Diagrama de Fluxo de Trabalho para a Obtenção do Ensaio do Curso



. Fig.

2.0 - Contatar instituições para obter clientela

Consiste em uma interface com:

- a) - instituições ligadas ao desenvolvimento de projetos educacionais que estejam envolvidas com atividades avaliativas tais como: Avaliação de projetos educacionais, Desenvolvimento de instrumentais de medida, Avaliação formativa de materiais e alunos, Avaliação somativa dos produtos do sistema em termos de aprendizagem, com objetivo de recrutar alunos (clientela) para receberem o curso em sua fase de ensaio. A exigência de contar com alunos que exercem as diferentes atividades avaliativas citadas tem o propósito de poder comparar os objetivos instrucionais do curso com o trabalho destes indivíduos quando da realização do "follow-up".
- b) - instituições escolares, com objetivo de fornecer alunos que constituirão a clientela para o desenvolvimento de um trabalho específico dos alunos do curso propriamente dito.

3.0 - Planejar e produzir os itens dos testes de aprendizagem das unidades

Esta atividade consiste inicialmente na análise dos objetivos instrucionais das Unidades que foram indicadas no Capítulo II, como a fonte para geração dos itens do teste de aprendizagem e dos objetivos identificados como pré-requisitos. Esta análise tem o propósito de determinar o formato e o número de itens mais adequados para verificar o alcance de cada um destes objetivos. Em seguida são produzidos os itens, construindo-se um primeiro esboço do teste.

4.0 - Revisar itens:

Consiste no controle de qualidade dos itens pelo especialista em Avaliação, buscando a congruência entre item e objetivo instrucional.

cional, especificidade, clareza e exatidão do mesmo. Objetiva-se com isto a qualidade técnica e de conteúdo do curso.

5.0 - Reformular itens

Nesta atividade, de acordo com os resultados de 4.0, os itens com falhas serão reformulados.

6.0 - Testar os itens um a um

Consiste na aplicação do esboço de cada teste em três alunos representativos da clientela a que se destina o curso de treinamento. Esta testagem deve ser realizada face a face com os alunos, com o objetivo de obter realimentação imediata dos pontos falhos dos itens.

7.0 - Analisar os itens

Consiste na identificação dos pontos falhos, nos itens. A atividade 6.0 fornece base empírica para tal identificação.

8.0 - Reformular itens

Consiste em reescrever os itens que apresentaram falhas de acordo com as informações obtidas na atividade 7.0.

9.0 - Montar o teste

Esta atividade refere-se à integração dos itens numa forma inicial do teste, ou seja, a que será utilizada na operação do ensaio do curso.

10.0 - Imprimir os testes de aprendizagem

Consiste na reprodução dos testes, de forma impressa, em quantidade suficiente para a utilização na operação do ensaio do curso.

11.0 - Encaminhar o teste à administração

Consiste no encaminhamento das cópias do teste à administração para que sejam estocados para a utilização na operação do ensaio do curso.

12.0 - Planejar e produzir os itens dos outros instrumentais de medida

Trata do planejamento e produção dos seguintes instrumentais:

- Questionário Final junto aos Alunos

A fonte para a geração dos itens deste instrumental serão os seguintes objetivos:

Verificar a adequação

- . da estrutura e sequência das unidades instrucionais;
- . da relação entre distribuição de horas de aula diárias e duração total do curso;
- . da frequência do uso das estratégias instrucionais;
- . das estratégias de avaliação, utilizadas durante o ensaio do curso.

Sugere-se, para medida do alcance do primeiro objetivo, que seja dada ao aluno uma listagem das unidades do curso para que ele indique a sequência que considera mais adequada. Para os demais objetivos, sugere-se questões abertas que possibilitem, ao aluno, emitir opiniões a respeito do planejamento nos pontos referentes as variáveis especificadas nestes objetivos.

- Questionário de "Follow-up" junto aos Treinados e seus supervisores

A fonte para a geração dos itens deste instrumental será o seguinte objetivo:

- . Verificar a utilidade do curso em termos das atividades desempenhadas, pelo aluno, no trabalho (Validade Externa do Curso);

Sugere-se para medida deste objetivo:

- . uma listagem das atividades, treinadas no curso, para que o aluno juntamente com seu supervisor, identifique aquelas que ele executa no trabalho e dentre estas, aquelas para as quais considera que o curso foi útil.
- . uma questão aberta que possibilite ao aluno especificar atividades que ele executa e que não foram treinadas no curso.

- Ficha de observação

A fonte para a geração dos itens deste instrumental serão os seguintes objetivos:

Verificar:

- . a adequação da carga horária planejada para cada objetivo instrucional;
- . a adequação da estratégia instrucional planejada para cada objetivo instrucional;
- . a congruência entre informações transmitidas pelo instrutor e informação contida nos materiais instrucionais.

Sugere-se, para medida destes objetivos, a utilização de uma ficha de observação sob o formato de uma tabela. Na coluna indicadora da mesma serão listados os objetivos instrucionais das unidades e na cabeça os indicadores, carga horária, estratégia instrucional e congruência entre informação transmitida pelo instrutor e informação contida nos materiais instrucionais, explícitos nos objetivos identificados para esta ficha.

O observador, após assistir às aulas, preencherá a ficha de observação, juntamente com o professor, especificando para cada objetivo a adequação do planejado em relação a cada um dos indicadores.

- Ficha com dados pessoais e profissionais do aluno

A fonte para a geração dos itens deste instrumental terá o seguinte objetivo:

- . Identificar as atividades desempenhadas pelos alunos do início do curso. As informações obtidas através desta ficha serão úteis para a posterior realização do "follow-up".

13.0 - Revisar Itens

Consiste no controle de qualidade dos itens planejados e produzidos na atividade 12.0 pelo especialista em Avaliação, buscando a congruência entre item e objetivo, clareza e exatidão do mesmo. Objetiva-se com isto a qualidade técnica dos itens.

14.0 - Reformular Itens

De acordo com os resultados da atividade 13.0, os itens com falhas serão descritos.

15.0 - Testar um a um

Consiste:

- . na aplicação da ficha com dados de alunos e dos questionários em três indivíduos representativos da clientela a que se destina o curso de treinamento;
- . na verificação do entendimento da ficha de observação pelo indivíduo que irá realizar esta atividade em sala de aula.

16.0 - Analisar os resultados

Consiste na identificação dos pontos falhos nos itens dos instrumentais de medida. A atividade 15.0 fornece base empírica para tal identificação.

17.0 - Reformular os itens dos instrumentais de medida produzidos em 12.0

Consiste na reformulação dos itens destes instrumentais de medida que apresentaram falhas de acordo com as informações obtidas na atividade 16.0.

18.0 - Imprimir os instrumentais de medida

Consiste na reprodução dos instrumentais de medida em quantidade suficiente para a utilização na operação do ensaio do curso.

19.0 - Encaminhar os instrumentais de medida à administração

Consiste no encaminhamento das cópias dos instrumentais à administração para que sejam estocados para sua utilização no ensaio do curso.

20.0 - Elaborar tarefas de aprendizagem

Consiste no desenvolvimento das tarefas de aprendizagem a partir dos objetivos instrucionais indicados no Capítulo II, a serem executadas pelos alunos, durante as aulas.

21.0 - Revisar tarefas de aprendizagem

Esta atividade trata do controle de qualidade das tarefas de aprendizagem pelo especialista na área de conhecimento, buscando a congruência entre as tarefas de aprendizagem e os objetivos instrucionais, clareza e exatidão das mesmas. Pretende-se com isto a qualidade técnica e de conteúdo do curso.

22.0 - Reformular tarefas de aprendizagem

De acordo com os resultados de 21.0 as tarefas com pontos falhos serão reformuladas.

23.0 - Organizar os materiais instrucionais impressos

Consiste na organização de cada material instrucional impresso em termos de objetivos instrucionais, especificação do conteúdo para o alcance desses objetivos e tarefas de aprendizagem.

24.0 - Testar os materiais instrucionais um a um

Consiste na aplicação de cada um dos materiais instrucionais em três alunos representativos da clientela a que se destina o curso de treinamento. Esta atividade deve ser realizada face a face com os alunos com o objetivo de obter realimentação imediata dos pontos falhos dos materiais instrucionais.

25.0 - Analisar os resultados

Trata da identificação dos pontos falhos nos materiais instrucionais. A atividade 24.0 fornece base empírica para tal identificação.

26.0 - Reformular materiais instrucionais

Consiste na reformulação dos materiais instrucionais que apresentaram falhas de acordo com as informações obtidas na atividade 25.0.

27.0 - Imprimir os materiais instrucionais

Consiste na reprodução dos materiais instrucionais de forma impressa, em quantidade suficiente para a utilização na operação do ensaio do curso.

28.0 - Encaminhar os materiais à administração

Consiste no encaminhamento das cópias dos materiais instrucionais à administração para que sejam estocados para a utilização na operação do ensaio do curso.

29.0 - Abrir inscrição aos alunos

Consiste na abertura de inscrição para alunos interessados em passar pelo curso durante a operação de seu ensaio.

30.0 - Contratar instrutores

Consiste na contratação de instrutores que atuarão durante a operação do ensaio.

3.1.1.6 - CRONOGRAMA DA OBTENÇÃO DO ENSAIO DO CURSO

A Figura III.4 que segue, mostra o cronograma para a obtenção do ensaio do curso.

3.1.2 - OPERAÇÃO DO ENSAIO

3.1.2.1 - ESTRUTURA DA DIVISÃO DE ATIVIDADES (EDA)

A Figura III.5 que segue, apresenta a Estrutura de Divisão de Atividades para a operação do ensaio do curso.

3.1.2.2 - ESTRUTURA DE DIVISÃO DE RECURSOS (EDR)

A Figura III.6 mostra a Estrutura de Divisão de Recursos para a operação do ensaio do curso.

3.1.2.3 - ESPECIFICAÇÃO DOS RECURSOS PARA A OPERAÇÃO DO ENSAIO DO CURSO

Da Figura III.6 serão especificados apenas os recursos referentes ao pessoal técnico. O pessoal de apoio e os recursos relativos aos materiais permanentes e de consumo não serão aqui especificados, visto que poderão ser utilizados aqueles disponíveis na Instituição. Quanto ao pessoal técnico, deverão ser consideradas as especificações que seguem.

ATIVIDADES	RESPONSÁVEL	TEMPO SEMANA					
		1a.	2a.	3a.	4a.	5a.	6a.
- Recrutar pessoal para o ensaio do curso.	Coordenador						
- Contatar Instituições para obter clientela	Coordenador						
- Planejar e Produzir os itens de teste C.R. das unidades	Avaliador						
- Revisar Itens	Avaliador						
- Reformular Itens	Avaliador						
- Testar os itens um a um	Avaliador						
- Analisar os itens	Avaliador						
- Reformular itens	Avaliador						
- Montar o teste	Avaliador						
- Imprimir os testes C.R.	Gráfica						
- Encaminhar teste à administração	Secretaria						
- Planejar e Produzir os itens dos outros instrumentais de medida	Avaliador						
- Revisar itens	Avaliador						
- Reformular itens destes instrumentais de medida	Avaliador						
- Testar um a um esses instrumentais de medida	Avaliador						
- Analisar os resultados	Avaliador						
- Reformular itens destes instrumentais de medida	Avaliador						
- Imprimir estes instrumentais de medida	Gráfica						
- Encaminhar estes instrumentais de medida à administração	Secretaria						
- Elaborar tarefas de aprendizagem	Planejador						
- Revisar tarefas de aprendizagem	Avaliador						
- Reformular tarefas de aprendizagem	Planejador						
- Organizar os materiais impressos	Planejador						
- Testar os materiais instrucionais, um a um	Avaliador						
- Analisar os resultados	Avaliador						
- Reformular os materiais instrucionais	Planejador						
- Imprimir os materiais instrucionais	Gráfica						
- Encaminhar os materiais instrucionais à administração	Secretaria						
29.0 - Abrir inscrição aos alunos	Secretaria						
30.0 - Contratar Instrutores	Coordenador						

Fig. III.4 - Cronograma para a Obtenção

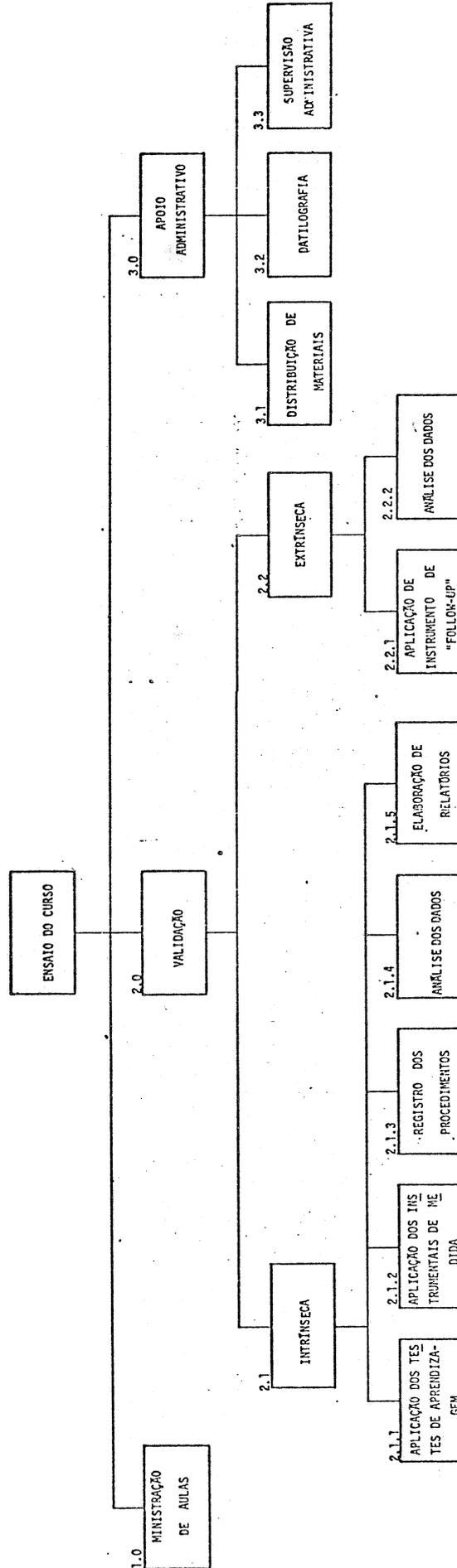


Fig. III.5 - Estrutura de Divisão de Atividades para a Operação do Ensaio do Curso

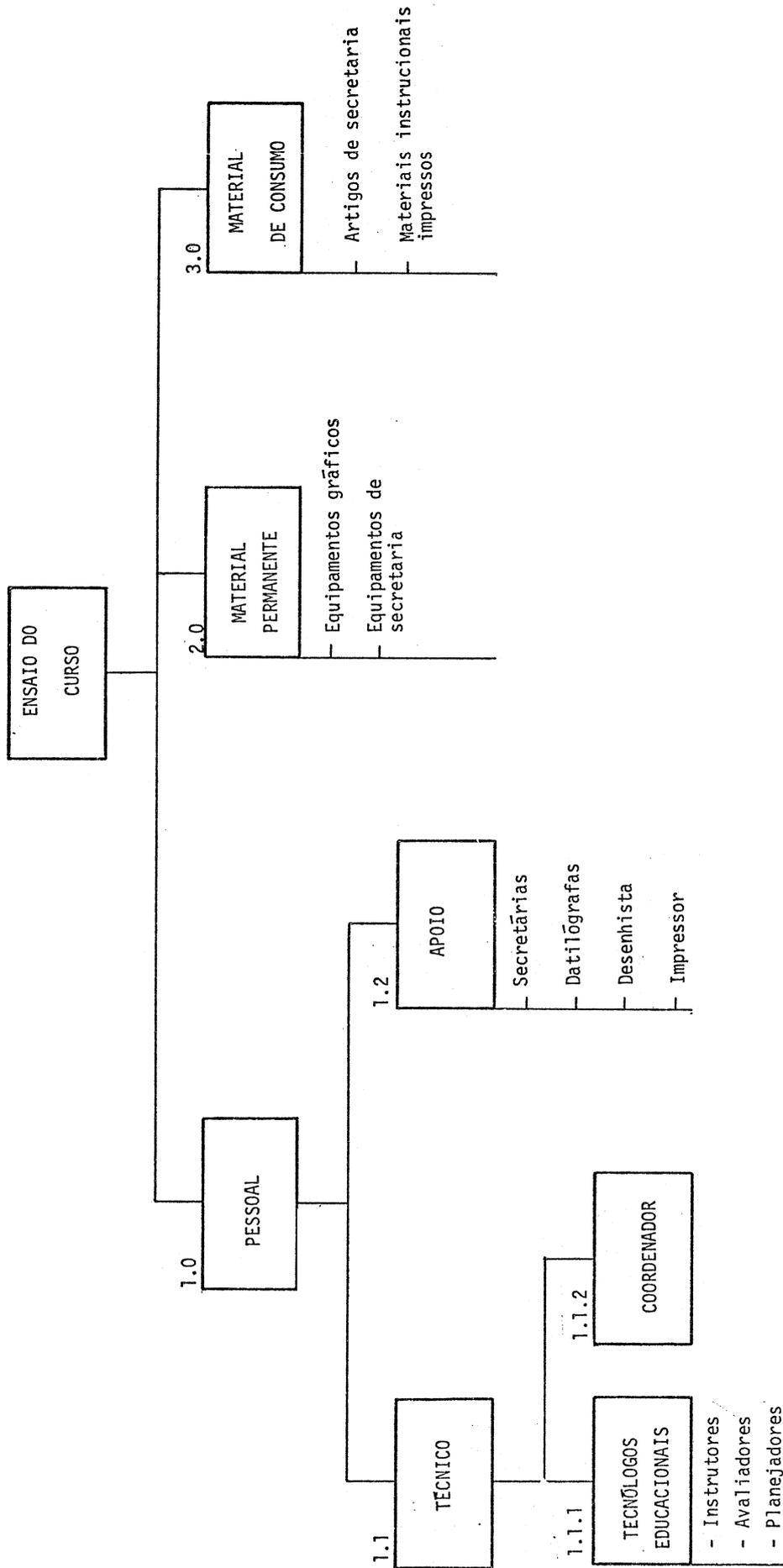


Fig. III.6 - Estrutura de Divisão de Recursos para a Operação do Ensaio do Curso

1.1 - PESSOAL TÉCNICO

Refere-se ao pessoal necessário para as atividades de docência, planejamento e avaliação.

1.1.1 - TECNÓLOGOS EDUCACIONAIS

a) Instrutores

Características:

- . especialização a nível de mestrado em Tecnologia Educacional na área de Avaliação e Pesquisa com conhecimentos relevantes sobre: Planejamento Sistemico, Avaliação de Projetos Educacionais (modelo de Avaliação de Stufflebeam et al,1971), Instrumentais de Medida, Avaliação Formativa e Somativa de Mudanças Educacionais;
- . dois anos de experiência no exercício da função de avaliador ou de docência na área.

Atividade:

- . ministrar aulas
- . orientar alunos
- . corrigir tarefas
- . aplicar pré e pós-testes de unidades e testes de pré-requisito.

b) Avaliadores

nº: dois avaliadores

Características:

- . especialização a nível de mestrado em Tecnologia Educacional na área de Avaliação, com conhecimentos relevantes sobre Avaliação de Projetos Educacionais;
- . dois anos de experiência em atividades avaliativas.

Atividades:

- . corrigir testes de pré-requisitos;
- . fazer observação sistemática em sala de aula;
- . fazer registro dos procedimentos tal como ocorrem;
- . fornecer retroalimentação periódica aos decisores;
- . aplicar instrumentais de medida;
- . analisar dados para Avaliação Formativa das unidades e do curso;
- . analisar dados para Avaliação Somativa do curso e do aluno;
- . aplicar os instrumentais do "follow-up";
- . analisar dados do
- . elaborar relatórios.

c) Planejadores:

nº: um planejador

Características:

- . especialização a nível de mestrado em Tecnologia Educacional, na área de Planejamento Instrucional;
- . dois anos de experiência na área;

Atividades

- . reformular os materiais instrucionais;
- . retomar o esquema do planejamento do curso, caso seja necessário reformulá-lo.

1.1.2 - COORDENADOR

nº: um coordenador

Características

- . curso superior completo, em qualquer área de conhecimento;
- . dois anos de experiência em coordenação de curso.

Atividades:

- Responsável pelas decisões administrativas relativas à execução deste planejamento, na fase de operação do ensaio, no que diz respeito à: aquisição de recursos materiais e humanos e cumprimento das atividades em tempo previsto pelo cronograma.

3.1.2.4 - DIAGRAMA DE FLUXO DE TRABALHO (DFT)

A Figura III.7 apresenta o Diagrama de Fluxo de Trabalho para a operação do ensaio do curso.

Note-se que não consta como uma atividade do DFT a Avaliação Formativa do aluno, visto que a ênfase na fase de ensaio do curso é a validação do curso. A Avaliação Somativa do aluno é incluída para atrair a clientela pelo fornecimento de um certificado de aproveitamento do curso.

3.1.2.5 - DESCRIÇÃO DAS ATIVIDADES DO DFT - OPERAÇÃO DO ENSAIO DO CURSO

1.0 - Aplicar o teste de pré-requisito

Consiste na administração do teste de pré-requisito, nos alunos inscritos para o ensaio do curso, com o objetivo de selecionar aqueles que passarão pela instrução planejada.

2.0 - Eliminar alunos do curso

Em função dos resultados das atividades 1.0, o aluno que não dominar os pré-requisitos identificados será eliminado do curso. Isto será feito pelos elementos responsáveis pela caixa de decisão A, que fixarão as regras de decisão.

3.0 - Integrar os alunos no curso

Em função dos resultados das atividades de 1.0, os alunos que dominarem os pré-requisitos serão integrados no curso, ou seja, serão matriculados e tomarão conhecimento da estrutura e funcionamento do curso através de um primeiro contato com os instrutores.

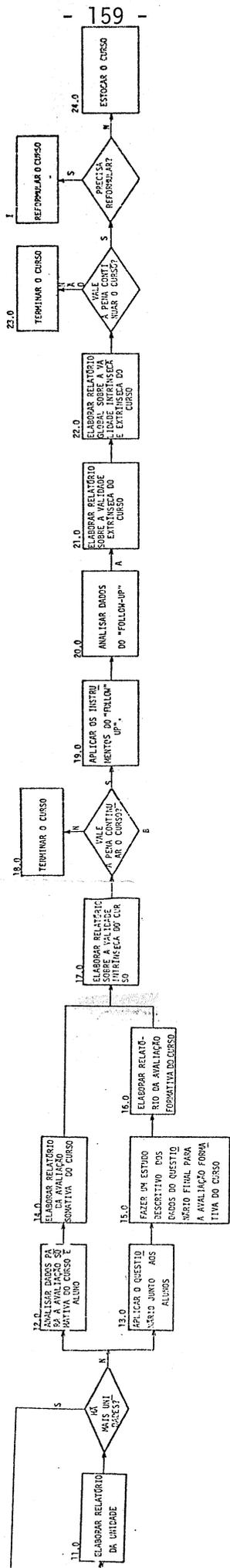
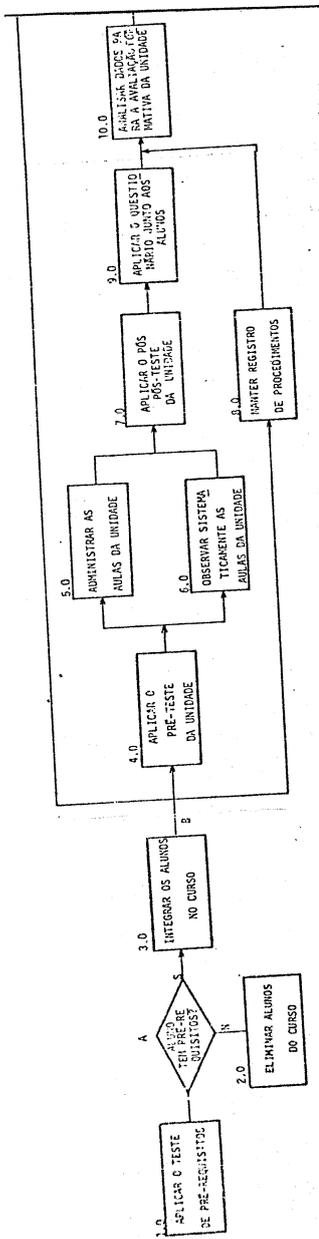


Fig. III.7 - Diagrama de Fluxo de Trabalho da Operação do Ensaio do Curso



4.0 - Aplicar o pré-teste da unidade

Esta atividade consiste na administração do pré-teste da unidade, junto aos alunos, com o objetivo de medir o seu nível de realização com relação aos objetivos da instrução, antes que esta ocorra.

5.0 - Administrar as aulas da unidade

Consiste na execução do planejamento instrucional da unidade conforme especificado na Capítulo II.

6.0 - Observação sistemática das aulas da unidade

Trata do preenchimento da ficha de observação, pelo avaliador, com dados referentes a cada objetivo instrucional. A observação deverá verificar a adequação da carga horária e da estratégia instrucional relativa a cada um dos objetivos, além da congruência entre informação transmitida pelo instrutor e informação contida nos materiais instrucionais. Isto, com o propósito de superar dificuldades presentes e/ou antecipar dificuldades futuras.

7.0 - Aplicar o pós-teste da unidade

Consiste na administração do pós-teste da unidade, junto aos alunos, com o objetivo de medir o seu nível de realização com relação aos objetivos propostos, após terem passado pela instrução.

8.0 - Manter registro dos procedimentos

Esta atividade consiste no registro de procedimentos relativos a: distribuição de materiais, presença de alunos no pré e pós-teste, na aplicação dos questionários e dificuldades na administração destes instrumentais. Esta atividade tem o objetivo de superar e antecipar dificuldades relativas aos procedimentos planejados.

9.0 - Aplicação do questionário junto aos alunos

Trata da administração do questionário, junto aos alunos, com o objetivo de coletar informações para a avaliação formativa da unidade.

10.0 - Analisar dados para a avaliação formativa da unidade

Consiste na análise dos dados obtidos através da aplicação do pré-teste, questionários, fichas de observação e registro dos procedimentos. Através dos resultados da aplicação do pré e pós-teste verificam-se eventuais pontos falhos na instrução e/ou testes. Para este propósito, deverão ser utilizados, para indicação dos pontos falhos, os seguintes índices:

$$\text{a) Índice de ganho por objetivo} = \frac{\left(\begin{array}{l} \text{nº de alunos que} \\ \text{atingiram o objetivo} \\ \text{no pós-teste.} \end{array} \right) - \left(\begin{array}{l} \text{nº de alunos que} \\ \text{atingiram o objetivo} \\ \text{no pré-teste} \end{array} \right)}{\left(\begin{array}{l} \text{nº máximo de alunos} \\ \text{que poderiam} \\ \text{ter atingido o objetivo} \\ \text{no pós-teste} \end{array} \right) - \left(\begin{array}{l} \text{nº de alunos que} \\ \text{atingiram o objetivo} \\ \text{no pré-teste} \end{array} \right)} \times 100$$

$$\text{b) percentagem de alcance por objetivo} = \frac{\left(\begin{array}{l} \text{nº de alunos que atingiram} \\ \text{o objetivo} \end{array} \right)}{\left(\begin{array}{l} \text{nº total de alunos} \end{array} \right)} \times 100$$

Sugerimos, para ambos os índices, o padrão de comparação 60% (Bloom, 1971). Os objetivos com baixo índice de ganho e/ou baixa percentagem de consecução (menor que 60%) indicam os pontos falhos da instrução. O passo seguinte será rever os itens de teste referentes a estes objetivos. Se nenhum problema for detectado nestes itens, a falha provavelmente estará na instrução. Considerando que a adequação da instrução está relacionada com: materiais instrucionais, estrutura e sequência dos objetivos instrucionais, domínio de conteúdo pelo professor, qualidade de suas orientações, estratégia instrucional, carga horária, relação carga

horária-teórica e atividades práticas, a causa específica da falha na instrução poderá estar em qualquer um destes elementos. Isto poderá ser detetado através dos resultados da aplicação dos questionários e ainda das fichas de observação e do registro de procedimentos.

11.0 - Elaborar relatório na unidade

Esta atividade consiste no fornecimento de informações relativas aos resultados da atividade 10.0. Estas informações deverão ser fornecidas aos decisores de uma maneira e forma que facilite o uso das mesmas, no momento oportuno.

12.0 - Analisar dados para Avaliação Somativa do curso e do aluno

- Avaliação Somativa do curso

Consiste na determinação da efetividade do curso, a partir dos resultados dos pré e pós-testes das seis unidades do curso. O grau de consecução dos objetivos instrucionais será determinado através da utilização de três índices:

a) Usando dados dos pós-testes:

$\frac{x}{y}$ em que x representa a percentagem de alunos que atingiu, nos pós-testes, determinada percentagem dos objetivos do curso como um todo e y a percentagem especificada de objetivos atingidos.

b) Usando dados dos pré e pós-testes:

$$\text{Índice de ganho (Individual)} = \frac{\left(\text{escore do aluno no pós-teste}^* \right) - \left(\text{escore do aluno no pré-teste} \right)}{\left(\text{escore máximo possível do aluno do pós-teste}^{**} \right) - \left(\text{escore do aluno no pré-teste} \right)} \times 100$$

* Será considerado como escore do aluno no pós-teste, o somatório de todos seus escores nos pós-testes das unidades.

** Será considerado como escore máximo possível do aluno no pós-teste, o somatório dos escores máximos possíveis nos pós-testes das unidades.

Esta fórmula deverá ser aplicada para cada aluno obtendo-se o ganho individual. A média dos ganhos individuais, isto é, o índice de ganho global considerando-se todos os alunos, refletirá a efetividade do curso.

c) Usando os dados dos pré e pós-testes:

$$\text{Índice de ganho por média grupal} = \frac{\left(\text{escore médio do grupo no pós-teste} \right) - \left(\text{escore médio do grupo no pré-teste} \right)}{\left(\text{escore médio máximo possível do grupo no pré-teste} \right) - \left(\text{escore médio do grupo no pré-teste} \right)} \times 100$$

Observe-se que a fórmula consiste numa variação da fórmula anterior. Ao invés de considerar os escores individuais, considera-se os escores médios do grupo.

Para verificar a validade interna do curso sugere-se, respectivamente para os índices a, b, e c, os seguintes padrões de comparação:

a) O valor entre $\frac{80\%}{60\%}$ a $\frac{100\%}{60\%}$ (80% a 100% dos alunos atingindo 60% dos objetivos).

b) Índice mínimo de ganho = 60%

c) Idem b

- Avaliação Somativa do aluno

Consiste na determinação do nível de realização do aluno em relação aos objetivos do curso, considerando-se os dados obtidos através da aplicação dos pós-testes e o trabalho final de desenvolvimento de um Plano de Avaliação. Sugere-se, para aprovação de cada aluno, o padrão de 60% de consecução dos objetivos instrucionais identificados nos pós-testes e um aproveitamento de 60% no trabalho relativo ao objetivo geral do curso. O responsável pela decisão de aprovação ou não do aluno é o professor.

13.0 - Aplicar questionário final junto aos alunos

Trata da administração do questionário, junto aos alunos, com o objetivo de coletar informações para a Avaliação Formativa do curso.

14.0 - Elaborar relatório de Avaliação Somativa do curso

Esta atividade consiste no fornecimento de informações relativas aos resultados da atividade 12.0. Estas informações deverão ser fornecidas aos decisores de uma maneira e forma que facilite o uso das mesmas, no momento oportuno.

15.0 - Fazer um estudo descritivo dos dados do questionário final para a Avaliação Formativa do curso

Consiste em um estudo descritivo dos dados obtidos através da aplicação do questionário final, com o objetivo de avaliar formativamente o curso como um todo, identificando pontos falhos no seu planejamento geral. Tendo em vista que existem aspectos do curso (estrutura e sequência das unidades, relação entre horas de aula e duração total do curso, frequência do uso das estratégias instrucionais, estratégias de Avaliação) que dependem da consideração das seis unidades como um todo, pode-se supor que falhas na instrução são devidas a qualquer um destes aspectos.

16.0 - Elaborar relatório da Avaliação Formativa do curso

Esta atividade consiste no fornecimento de informações relativas aos resultados da atividade 15.0. Estas informações deverão ser fornecidas aos decisores de uma maneira e forma que facilite o uso da mesma, no momento oportuno.

17.0 - Elaborar relatório global sobre a validade intrínseca

Consiste em integrar as informações contidas nos relatórios, elaborados nas atividades 11.0, 14.0 e 16.0, de modo a possibilitar ao decisor (Diretor do Centro de Treinamento) decidir se vale a pena continuar. Esta decisão será baseada no índice de efetividade obtido nesta fase de ensaio do curso e na quantidade de reformulações necessárias para melhoria do mesmo. (Obs.: embora na atividade 12.0 tenha sido sugerido um critério mínimo de efetividade, o decisor poderá adotar um outro critério que julgue mais adequado).

18.0 - Terminar o curso

Esta atividade consiste no encerramento dos esforços para a instalação do curso. O término do curso é resultante da decisão tomada em função das informações contidas no relatório elaborado em 17.0.

19.0 - Aplicar os instrumentos do "follow-up"

Consiste na aplicação do questionário junto aos treinados que estejam exercendo atividades de Avaliação, para ser respondido juntamente com seus supervisores, três meses após o término do curso. Esta atividade tem o propósito de obter informações que possibilitem comparar o desempenho dos treinados, no trabalho, com os seus desempenhos no curso e se existe correspondência entre ambos.

20.0 - Analisar dados do "follow-up"

Consiste na análise dos dados obtidos através da aplicação do questionário, junto aos alunos e supervisores, com o propósito de:

- Ampliar, restringir ou manter as atividades constantes do Universo;
- Reestruturar os objetivos instrucionais de modo que conduzam ao desempenho das atividades do Universo.

As atividades do Universo serão ampliadas se forem identificadas atividades que estão sendo executadas no trabalho, pelos alunos, e que não estão sendo treinadas no curso, e se estas atividades atenderem os critérios de universalidade, crucialidade, dificuldade, necessidade e frequência (apresentados no Capítulo II, página 76).

Serão mantidas no curso de treinamento aquelas atividades que o treinando executa e julga que o curso foi útil para o desempenho das mesmas. Se pelo menos dois treinandos entre três que as executem, julgarem que o curso foi útil para a realização das mesmas, elas serão mantidas. No caso em que menos de dois julgar que o curso não foi útil para o desempenho destas atividades, duas alternativas podem ter ocorrido:

- . a instrução não conduziu ao alcance dos objetivos (e isto será detetado através dos resultados da Avaliação Formativa intrínseca do curso);
- . a instrução conduz ao alcance dos objetivos, porém estes não conduzem ao desempenho das atividades. Neste caso, os resultados do "follow-up" indicariam a necessidade de reestruturação dos objetivos instrucionais.

As atividades que os treinandos não executam, no trabalho, serão eliminadas do curso, desde que o mesmo já tenha ocorrido um certo número de vezes, que o decisor julgue suficiente, para considerar que elas realmente não serão executadas, no trabalho, e, não atendam aos critérios de universalidade, frequência, dificuldade, necessidade, e crucialidade (apresentados no Capítulo II, página 76).

21.0 - Elaborar relatório sobre a validade extrínseca do curso

Consiste no fornecimento das informações obtidas através da atividade 20.0. Estas informações deverão ser fornecidas ao decisor de uma maneira e forma que facilite o uso das mesmas, no momento oportuno.

22.0 - Elaborar relatório global sobre a validade intrínseca e extrínseca do curso

Consiste em integrar as informações contidas nos relatórios elaborados nas atividades 17.0 e 21.0 de modo que possibilite ao diretor (Diretor do Centro de Treinamento) decidir se vale a pena continuar o curso.

23.0 - Terminar o curso

Esta atividade consiste no encerramento dos esforços para a instalação do curso. O término do curso é resultante da decisão tomada em função das informações contidas no relatório elaborado em 22.0

24.0 - Reformular o curso

Consiste em retomar o planejamento do curso segundo o esquema apresentado no Capítulo I, página 20, e reformulá-lo nos pontos indicados pelos resultados da avaliação.

25.0 - Estocar o curso

Esta atividade consiste na estocagem, em sua forma definitiva, dos testes de aprendizagem, outros instrumentais de medida, materiais instrucionais, planejamento das unidades instrucionais do curso, e das informações avaliativas obtidas no ensaio do curso.

3.1.2.6 - CRONOGRAMA DA OPERAÇÃO DO ENSAIO DO CURSO

A Figura III.8 mostra o cronograma para a operação do ensaio do curso.

3.2 - INSTALAÇÃO DO CURSO

3.2.1 - OBTENÇÃO DA INSTALAÇÃO

3.2.1.1 - ESTRUTURA DE DIVISÃO DE ATIVIDADES (EDA)

A Figura III.9 apresenta a Estrutura de Divisão de Atividades para a obtenção da instalação do curso.

3.2.1.2 - ESTRUTURA DE DIVISÃO DE RECURSOS (EDR)

A Figura III.10 mostra a Estrutura de Divisão de Recursos para a obtenção da instalação do curso. Não serão especificados, aqui, os recursos humanos, por serem os mesmos descritos nas páginas 139 e 142.

3.2.1.3 - DIAGRAMA DE FLUXO DE TRABALHO (DFT)

A Figura III.11 apresenta o Diagrama de Fluxo de Trabalho para a obtenção da instalação do curso.

3.2.1.4 - DESCRIÇÃO DAS ATIVIDADES DO DFT - OBTENÇÃO DA INSTALAÇÃO DO CURSO

1.0 - Contatar com Instituições

Esta atividade consiste em enviar às Instituições que desenvolvam atividades de Avaliação, informações relativas ao curso tais como: seus pré-requisitos, objetivos instrucionais e duração.

2.0 - Abrir inscrição aos alunos

Consiste em matricular os candidatos das Instituições interessadas.

3.0 - Recrutar pessoal para a instalação do curso

Consiste na realocação ou contratação de pessoal, conforme indicado na Estrutura de Divisão de Recursos da Obtenção da Instalação do curso.

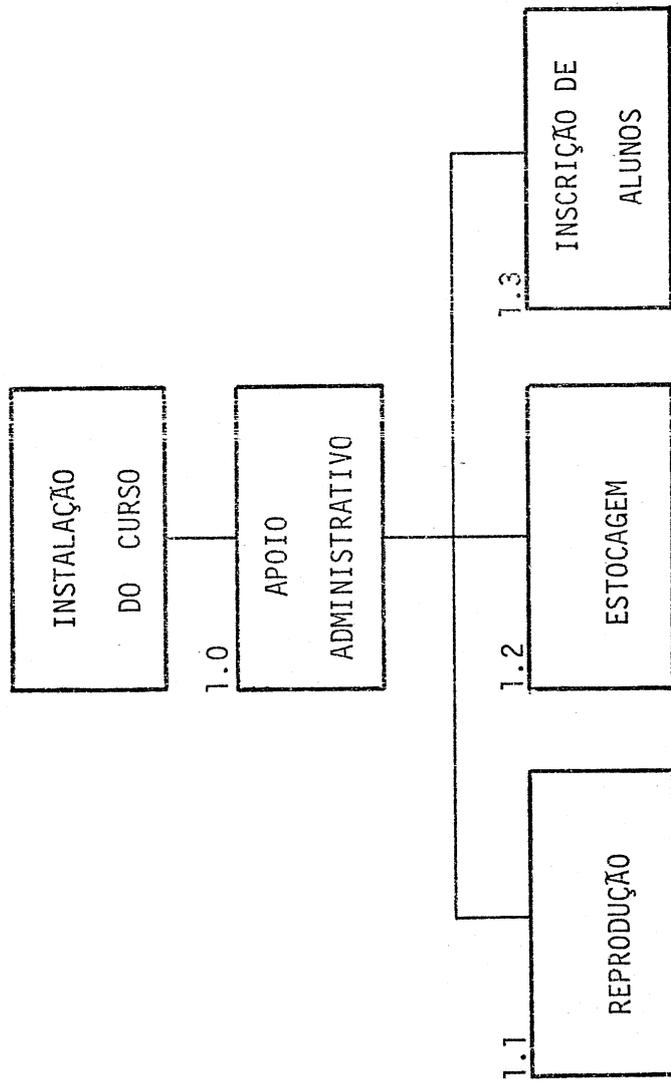


Fig. III.9 - Estrutura de Divisão de Atividades para a Obtenção da Instalação do Curso.

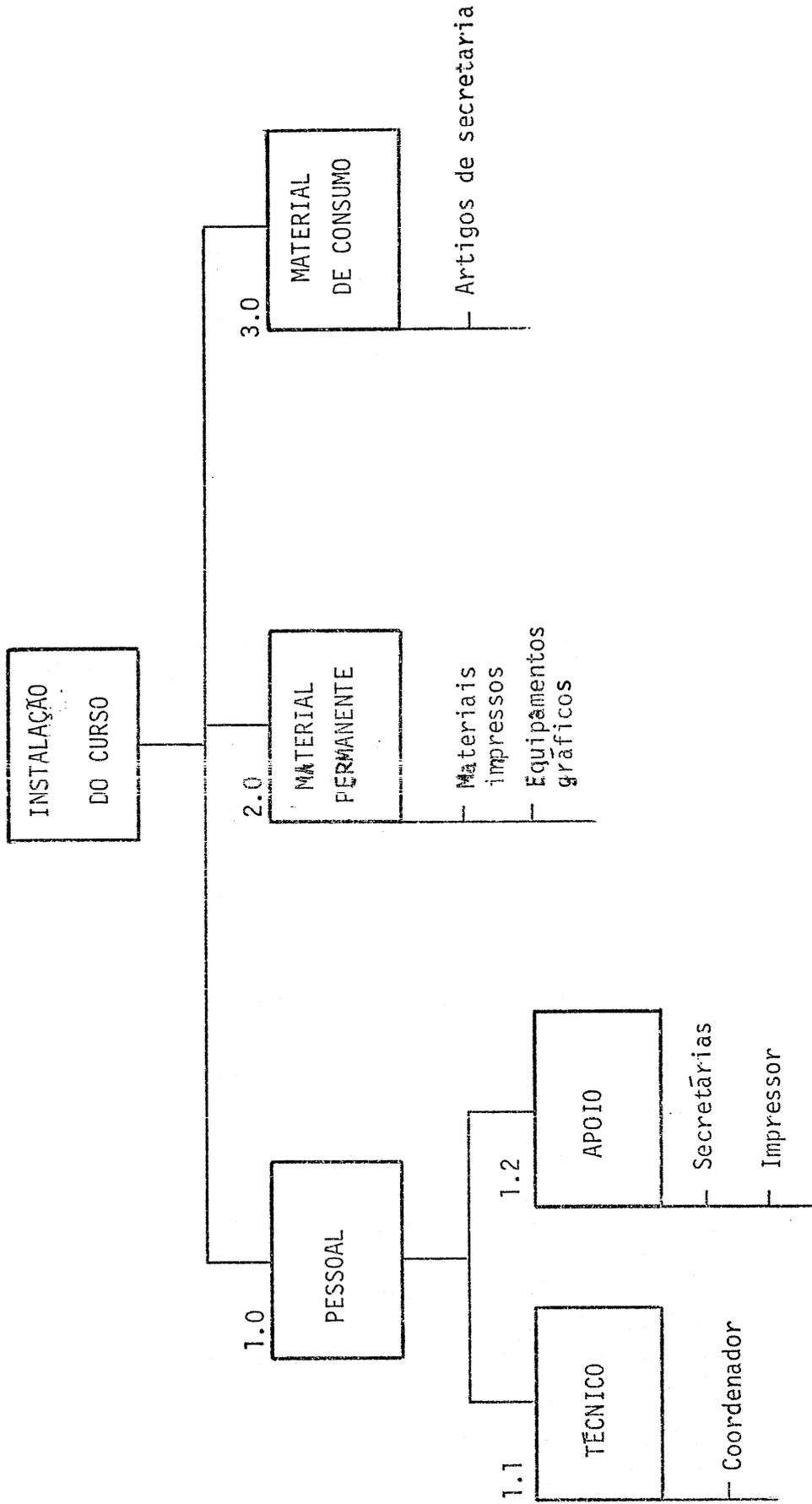


Fig. III.10 - Estrutura de Divisão de Recursos para a Obtenção da Instalação do Curso.

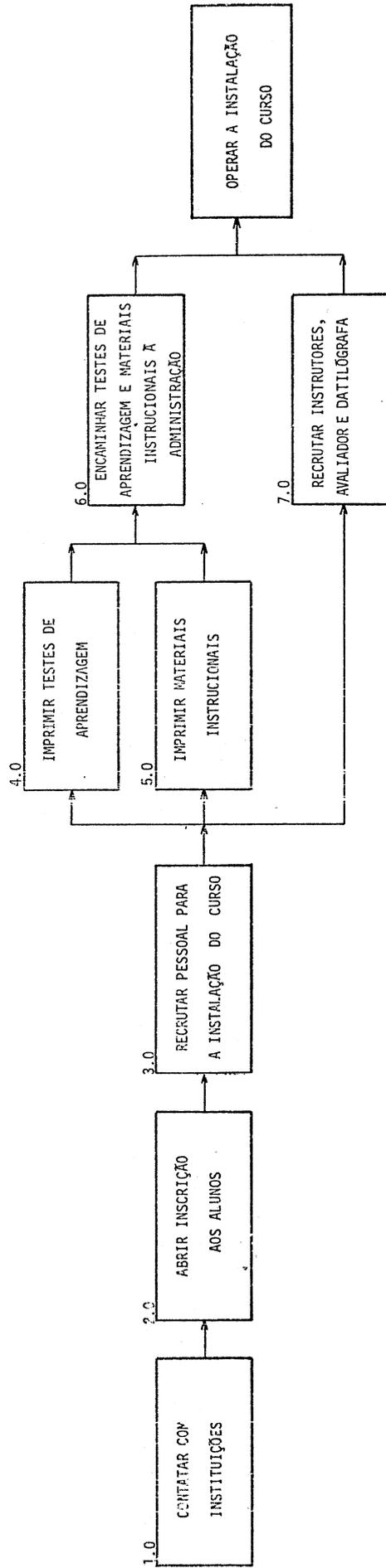


Fig. III.11 - Diagrama de Fluxo de Trabalho para a Obtenção da Instalação do Curso

4.0 - Imprimir testes de aprendizagem

Esta atividade consiste na reprodução gráfica dos testes C.R. em número correspondente as vagas disponíveis.

5.0 - Imprimir materiais instrucionais

Trata da reprodução gráfica dos materiais instrucionais em número correspondente às vagas disponíveis.

6.0 - Encaminhar testes de aprendizagem e materiais instrucionais, a administração

Consiste no encaminhamento das cópias dos testes e dos materiais instrucionais a administração para que sejam estocados para a utilização na operação da instalação do curso.

7.0 - Recrutar instrutores, avaliador e datilógrafa

Consiste na realocação ou contratação de instrutores e avaliador (conforme especificação na Estrutura de Divisão de Recursos para Obtenção do ensaio do curso, página 155) e datilógrafa que deverão atuar durante a operação da instalação do curso.

3.2.1.5 - CRONOGRAMA DA OPERAÇÃO DA INSTALAÇÃO DO CURSO

A Figura III.12 apresenta o cronograma para a obtenção da instalação do curso.

3.2.2 - OPERAÇÃO DA INSTALAÇÃO

3.2.2.1 - ESTRUTURA DE DIVISÃO DE ATIVIDADES (EDA)

A Figura III.13 mostra a Estrutura de Divisão de Atividades para a operação da instalação do curso.

ATIVIDADES	RESPONSABILIDADE	TEMPO (SEMANAS)							TOTAL EM DIAS
		1a.	2a.	3a.	4a.	5a.	6a.	7a.	
1.0 - CONTATAR COM INSTITUIÇÕES	COORDENADOR	■							12
2.0 - ABRIR INSCRIÇÃO AOS ALUNOS	SECRETARIA		■						6
3.0 - RECRUTAR PESSOAL PARA A INSTALAÇÃO DO CURSO	COORDENADOR			■					3
4.0 - IMPRIMIR TESTE DE APRENDIZAGEM	GRÁFICA				■				3
5.0 - IMPRIMIR MATERIAIS INSTRUCIONAIS	GRÁFICA				■	■			6
6.0 - ENCAMINHAR TESTES C.R. E MATERIAIS INSTRUCIONAIS À ADMINISTRAÇÃO	SECRETARIA					■			1
7.0 - RECRUTAR INSTRUTORES, AVALIADOR E DATILÓGRAFA	COORDENADOR						■		12

Fig. III.12 - Cronograma para a Obtenção da Instalação do Curso

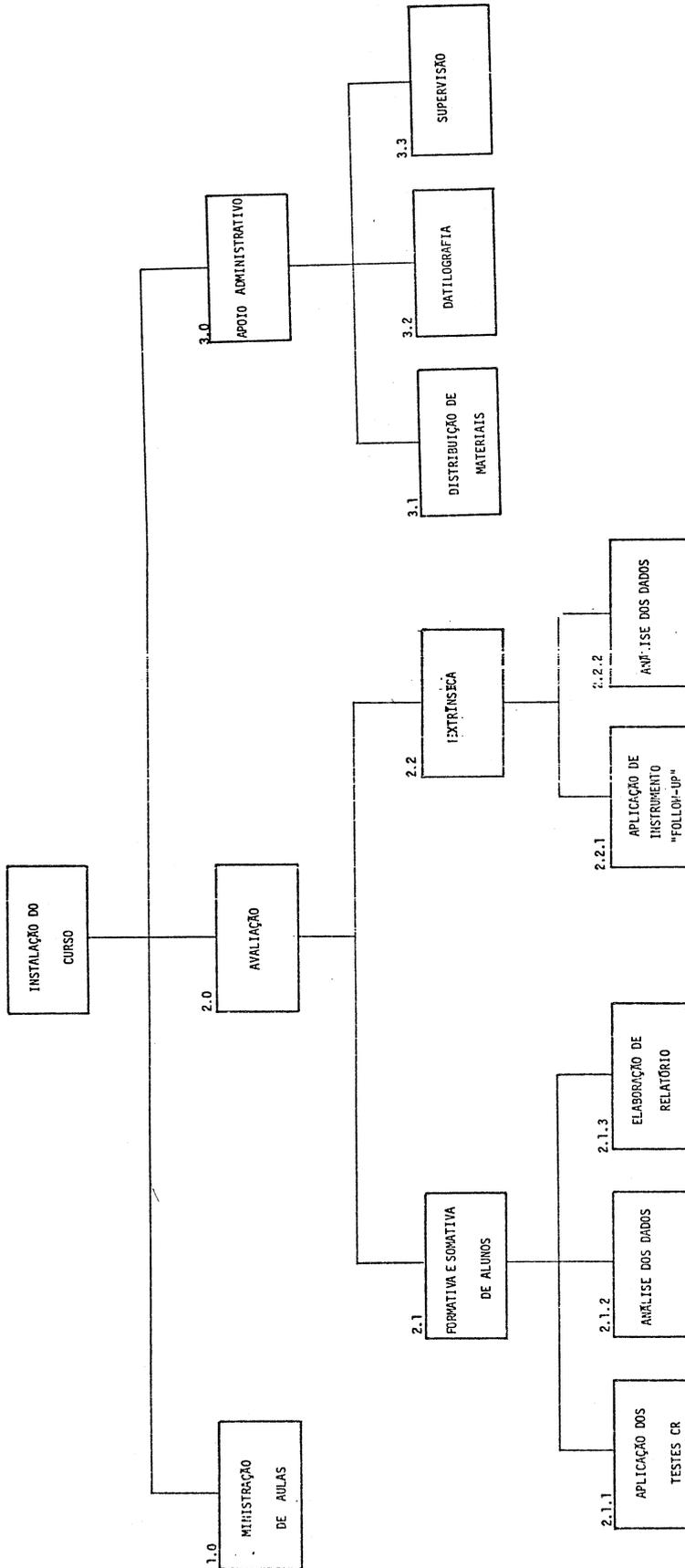


Fig. III.13 - Estrutura de Divisão de Atividades para a Operação da Instalação do Curso

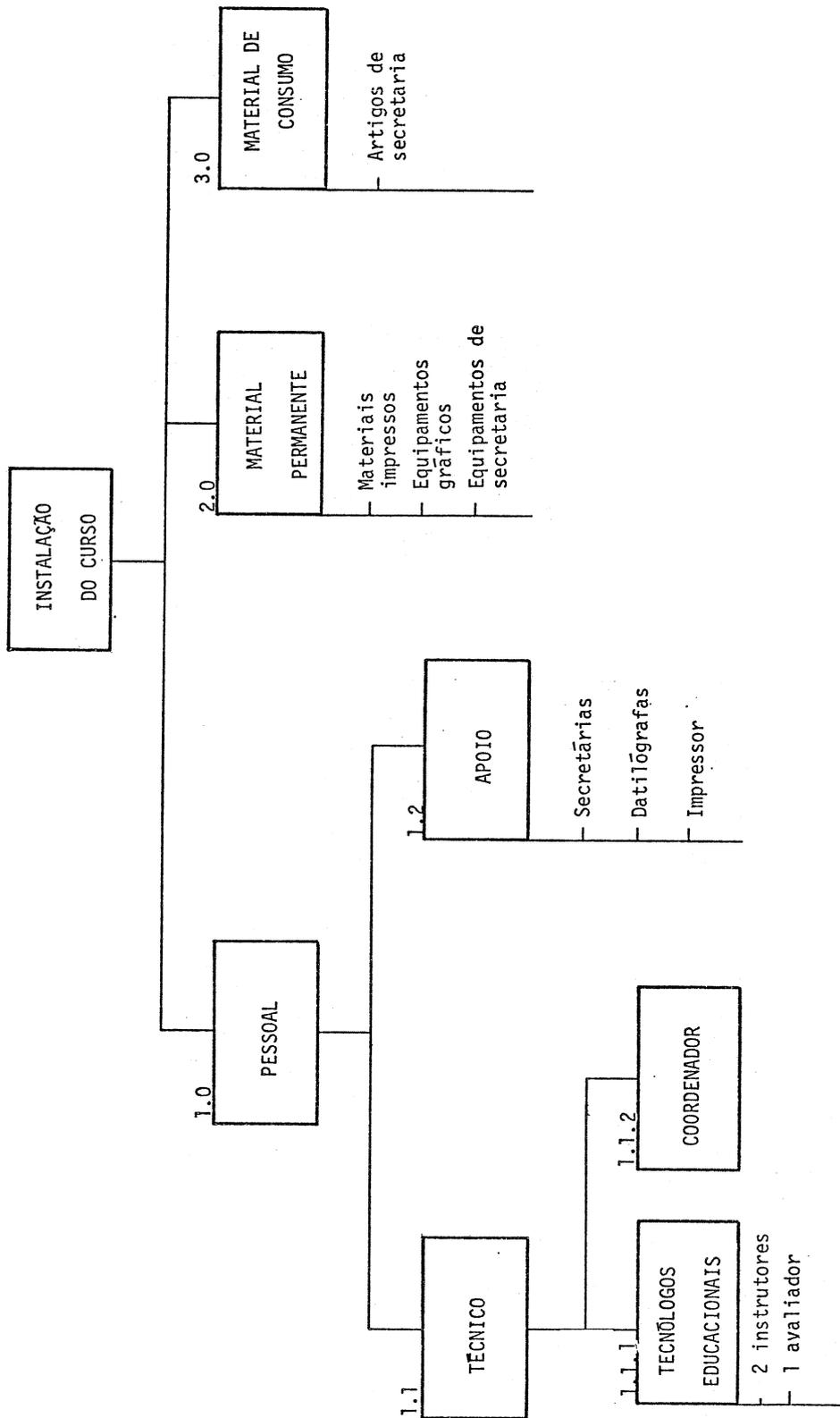
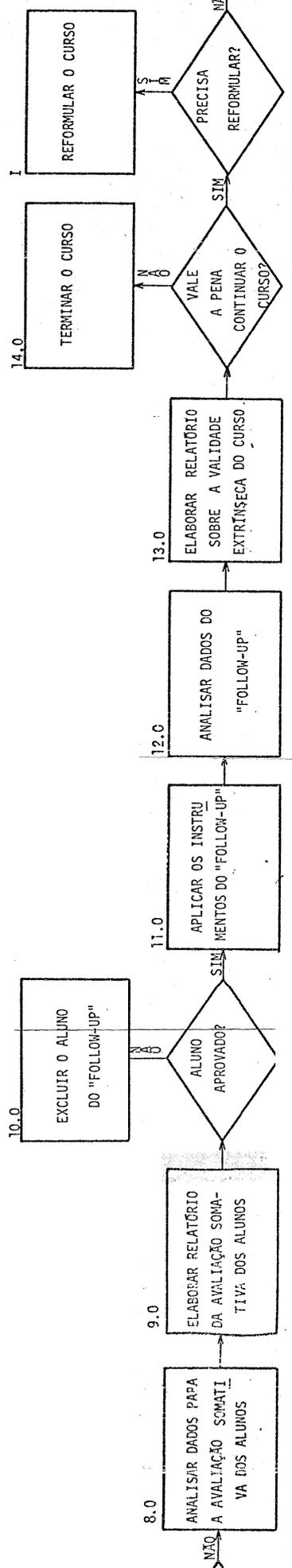
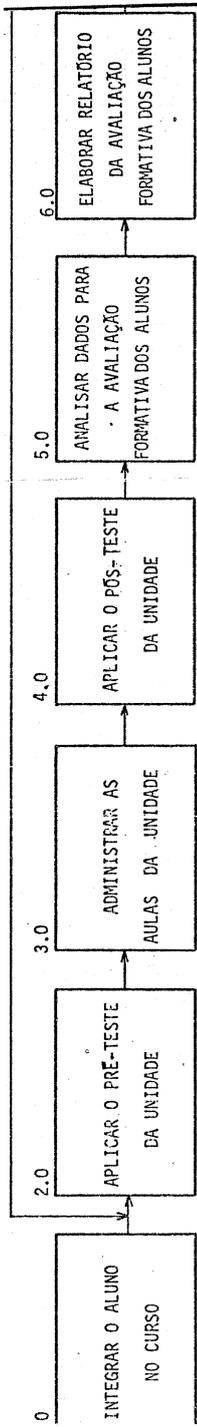


Fig. III.14 - Estrutura de Divisão de Recursos para a Operação da Instalação do Curso



Fluxo de Trabalho da Operação da Instalação do Curso



3.2.2.2 - ESTRUTURA DE DIVISÃO DE RECURSOS (EDR)

A Figura III.14 apresenta a Estrutura de Divisão de Recursos para a Instalação do Curso. Os recursos humanos não serão aqui especificados por serem os mesmos descritos na página 139.

3.2.2.3 - DIAGRAMA DE FLUXO DE TRABALHO (DFT)

A Figura III.15 mostra o Diagrama de Fluxo de Trabalho para a operação da instalação do curso.

3.2.2.4 - DESCRIÇÃO DAS ATIVIDADES DO DFT - OPERAÇÃO DA INSTALAÇÃO DO CURSO

1.0 - Integrar o aluno no curso

Nesta atividade, os alunos inscritos no curso serão integrados no mesmo, ou seja, tomarão conhecimento da estrutura e funcionamento do curso através do primeiro contato com os instrutores.

2.0 - Aplicar o pré-teste da unidade

Consiste na administração do pré-teste da unidade, junto aos alunos, com o objetivo de medir o seu nível de realização com relação aos objetivos da instrução, antes que esta ocorra.

3.0 - Administrar as aulas da unidade

Trata da execução do planejamento instrucional da unidade conforme especificado no Capítulo II.

4.0 - Aplicar o pós-teste da unidade

Esta atividade consiste na administração do pós-teste da unidade, junto aos alunos, com o objetivo de medir o seu nível de reali

zação, com relação aos objetivos propostos, após terem passado pela instrução.

5.0 - Analisar dados para a Avaliação Formativa do aluno

Consiste na análise dos dados obtidos através da aplicação do pós-teste com o propósito de identificar, para cada aluno, os objetivos não dominados depois de terem passado pelo processo instrucional.

6.0 - Elaborar relatório da Avaliação Formativa dos alunos

Trata do fornecimento de informações relativas aos resultados da atividade 5.0. Estas informações deverão ser fornecidas aos instrutores para que estes possam indicar, aos alunos, como superar suas falhas de aprendizagem.

7.0 - Fornecer retroalimentação aos alunos

Esta atividade consiste em informar, a cada aluno, quais os objetivos não dominados no pós-teste, de acordo com as informações contidas no relatório elaborado na atividade 6.0. Devem ser indicados também, os materiais instrucionais ou parte destes, que devem ser revistos pelos alunos.

8.0 - Analisar dados para a Avaliação Somativa dos alunos

Consiste na determinação do nível de realização do aluno, em relação aos objetivos do curso, considerando-se os dados obtidos através da aplicação dos pós-testes e o trabalho final de desenvolvimento de um Plano de Avaliação. Sugere-se, para a aprovação de cada aluno, o padrão de 60% de consecução dos objetivos instrucionais, identificados nos pós-testes, e um aproveitamento de 60% no trabalho relativo ao objetivo geral do curso.

9.0 - Elaborar relatório da Avaliação Somativa dos alunos

Trata do fornecimento de informações relativas aos resultados da atividade 8.0. Estas informações deverão ser fornecidas aos instrutores para que estes possam identificar alunos aprovados ou reprovados no curso.

10.0 - Excluir alunos do "Follow-up"

Consiste em eliminar do processo do follow-up os alunos reprovados no curso.

11.0 - Aplicar os instrumentais do "Follow-up"

Esta atividade consiste na aplicação do questionário junto aos treinados que estejam exercendo atividades de avaliação, para ser respondido juntamente com seus supervisores, três meses após o encerramento do curso. Esta atividade tem o propósito de obter informações que possibilitem comparar o desempenho dos treinados no trabalho, com os objetivos do curso e de verificar a qualidade deste desempenho.

12.0 - Analisar os dados do "Follow-up"

Consiste na análise dos dados obtidos através da aplicação do questionário, junto aos alunos e supervisores, com o propósito de:

- Ampliar, restringir ou manter as atividades constantes do Universo;
- Reestruturar os objetivos instrucionais de modo que conduzam ao desempenho das atividades constantes do Universo.

As atividades do Universo serão ampliadas se forem identificadas atividades que estão sendo executadas no trabalho, pelos alunos, e que não estão sendo treinadas no curso, e estas atividades atenderem aos critérios de: universalidade, crucialidade, dificuldade, necessidade e frequência (apresentados no Capítulo II, página 76).

Serão mantidas, no curso de treinamento, aquelas atividades que o treinado executa e julga que o curso foi útil para o desempenho das mesmas. Se pelo menos dois treinados entre três julgarem que o curso foi útil para a realização daquelas atividades, elas serão mantidas. No caso em que menos de dois julgar que o curso não foi útil para o desempenho das atividades que executa, duas alternativas podem ter ocorrido:

- . a instrução não conduziu ao alcance dos objetivos (e isto será detetado através dos resultados da Avaliação Formativa intrínseca do curso);
- . a instrução conduz ao alcance dos objetivos, porém estes não conduzem ao desempenho das atividades. Neste caso, os resultados do "follow-up" indicariam a necessidade de reestruturação dos objetivos instrucionais.

As atividades que os treinados não executam no trabalho, serão eliminadas do curso, desde que o mesmo já tenha ocorrido um certo número de vezes, que o decisor julgue suficiente para considerar que elas realmente não são executadas no trabalho, e não atendam aos critérios de universalidade, frequência, dificuldade, necessidade e crucialidade (apresentados no Capítulo II, página 76).

13.0 - Elaborar relatório sobre a validade extrínseca do curso

Consiste no fornecimento de informações obtidas através da atividade 12.0. Estas informações deverão ser fornecidas ao decisor para que este possa concluir sobre o término ou continuação do curso.

14.0 - Terminar o curso

Esta atividade é resultante da decisão tomada em função das informações contidas no relatório elaborado na atividade 13.0 e consiste em não mais utilizar o curso.

I - Reformular o curso

Consiste em retomar o planejamento do curso a partir da elaboração de objetivos, seguindo o esquema apresentado no Capítulo I, página 20, Figura I.7.

3.2.2.5 - CRONOGRAMA DE OPERAÇÃO DE INSTALAÇÃO DO CURSO

A Figura III.16 apresenta o cronograma para a operação da instalação do curso.

Este capítulo tratou do planejamento da obtenção e operação do ensaio do curso, para sua validação e do planejamento da obtenção e operação da sua instalação.

O Capítulo IV apresenta algumas considerações finais a respeito deste trabalho.

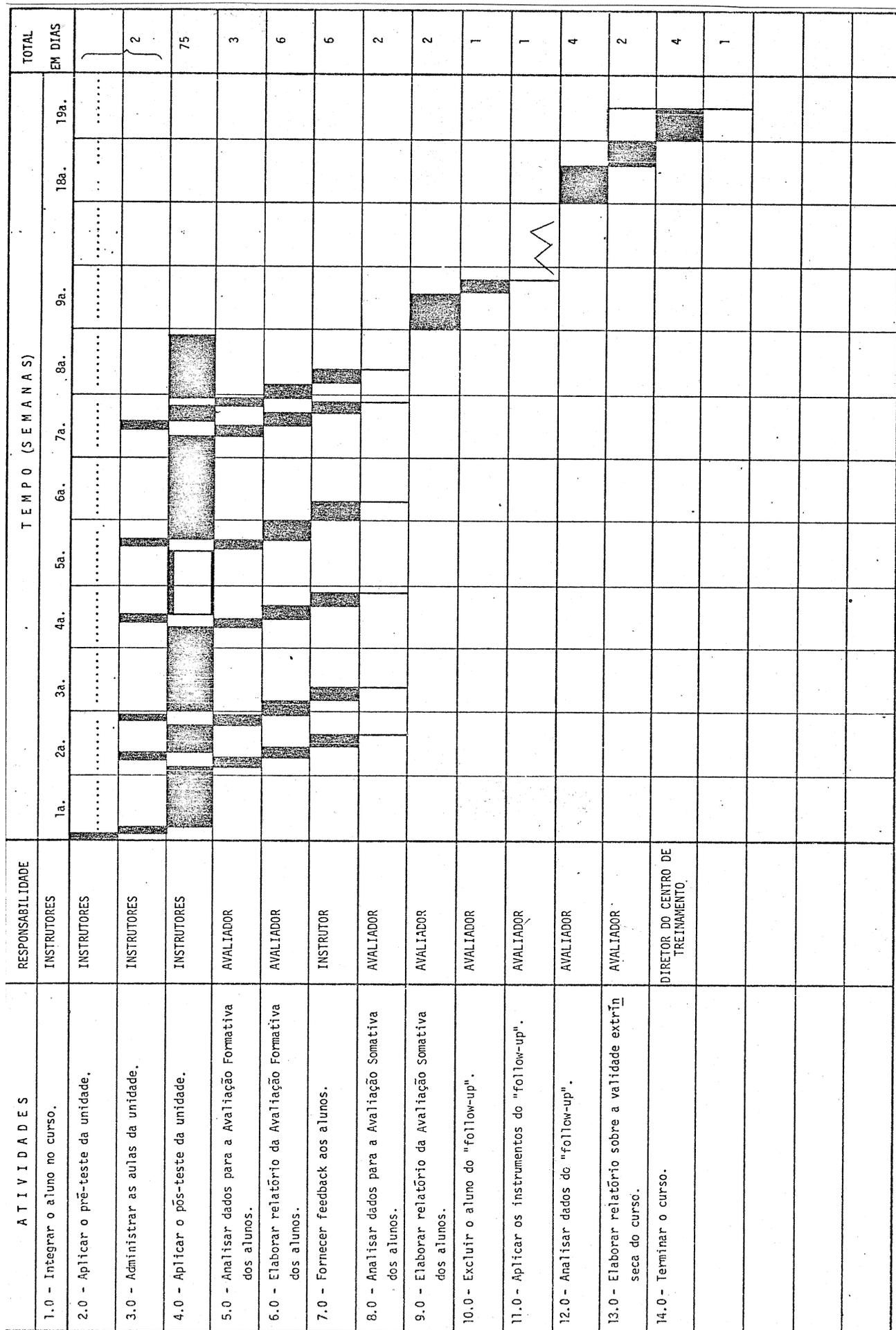


Fig. III.16 - Cronograma para a Operação da Instalação do Curso

CAPÍTULO IV

CONSIDERAÇÕES FINAIS

Este trabalho, que consiste no planejamento de um curso de treinamento de Avaliadores Educacionais, partiu da identificação de um universo de atividades que caracteriza a função, devido à inexistência de um consenso a respeito das atividades que a caracterizariam.

Após uma revisão de literatura de modelos instrucionais e de treinamento, é produzido um esquema, Figura I.7, página 20, a ser adotado para o planejamento do curso. Deste esquema, foram executadas, pelas autoras, neste trabalho, as atividades:

- 1.0 - Universo de atividades do trabalho;
- 2.0 - Universo de atividades a serem treinadas;
- 3.0 - Elaborar objetivos do curso;
- 6.0 - Estruturar e sequenciar o curso;
- 7.0 - Especificar conteúdo.

Foi ênfase deste trabalho a organização e estruturação lógica do conteúdo (Apêndice F, Volume II) de Avaliação Educacional, que se encontra disperso na literatura, pois isto implicará em economia de recursos para a Entidade que se propuser a colocar o curso de treinamento, aqui planejado, pronto para o funcionamento.

O Apêndice F pode ser utilizado como um manual para o Avaliador quando no exercício de sua função ou, ainda, como um livro texto em cursos da área.

- 8.0 - Identificar pré-requisitos do curso;
- 10.0 - Selecionar estratégias de treinamento;

11.0 - Selecionar meios.

As demais atividades, constantes no esquema citado, deverão ser executadas pelo Órgão ou Instituição Educacional que pretenda oferecer treinamento na área de Avaliação Educacional. Tal Instituição deverá possuir infra-estrutura própria para o desenvolvimento de cursos de treinamento. Para continuidade do trabalho, tal Entidade conta com o planejamento apresentado no Capítulo III. Sua tarefa, portanto, será executar tal planejamento que, especifica recursos, atividades, sequência e duração destas e atribuição de responsabilidades ao pessoal envolvido para:

- 4.0 - Construir testes de aprendizagem;
- 5.0 - Construir outros instrumentais de medida;
- 9.0 - Construir testes de pré-requisitos;
- 12.0 - Desenvolver materiais;
- 13.0 - Conduzir o ensaio do curso;
- 14.0 - Analisar resultados da administração de testes de aprendizagem;
- 15.0 - Analisar resultados da administração de outros instrumentais de medida;
- 16.0 - Comparar desempenho no curso com desempenho no trabalho.

A execução de tais atividades consistirá na obtenção e operação do ensaio do curso que ficará, portanto, pronto para a instalação.

O planejamento proposto pode parecer, em princípio, dispendioso, uma vez que a etapa de ensaio não é um procedimento comum. O que normalmente ocorre é a instalação imediata de cursos, buscando economia de tempo e recursos. Porém, somente através do ensaio é possível

aperfeiçoar um curso, garantindo uma maior efetividade. O ensaio se justifica, ainda mais, pela clientela potencial do curso de treinamento, o que implicará na repetição do mesmo inúmeras vezes. Isto, provocará um retorno à Instituição que oferecerá o curso, o que justificará os gastos envolvidos com a sua testagem.

AGRADECIMENTOS

Ao Instituto de Pesquisas Espaciais, entidade mantenedora do Curso de Mestrado em Tecnologia Educacional.

À Dra. Vathsala Iyenger Stone, pela valiosa orientação.

Ao MSc. Sérgio Suren Kurkdjian, pela orientação e incentivo constante.

À MSc. Margarida Ferreira da Costa Southard, pela orientação deste trabalho em sua fase preliminar, e pelas idéias fundamentais que possibilitaram a execução do mesmo.

À Edna Franceschetti Piza, pela sugestão do tema deste trabalho.

À Lígia de Oliveira Auricchio, pelas sugestões e estímulos recebidos.

Ao Dr. Ralf Gielow, pelo empenho no bom andamento deste trabalho.

Ao Carlos Alberto Bertti, Manoel Marteleto, Antônio Costa, Acir M. Soares e Odete Barros Cardozo, pelo apoio.

Às Instituições Educacionais, pelas informações recebidas.

À Secretaria Geral do INPE, nas pessoas de Marília Prado de Carvalho, Maria Dorotéa Cabral Gonzalez Blanco e Teresa Cristina A. do E. Santo, pela dedicação no trabalho de datilografia.

Enfim, a todos que direta ou indiretamente contribuíram para a execução deste trabalho.

BIBLIOGRAFIA

- ABEDOR, A.J. Second draft technology development and field test of a model for formative evaluation of self-instructional multi-media learning systems. In: *Bulletin of the school of education*. Indiana University, 48 (4): 9 - 44, Jul, 1972.
- ARANTES, M.S.; COSTA, S. Situações de decisões. In: ———. *Curso de programação, execução e avaliação de projetos educacionais*. São José dos Campos, INPE/CAPES (no prelo).
- . Tipos de decisão. In: ———. ———. São José dos Campos, INPE/CAPES (no prelo).
- . Modelo de avaliação de Stufflebeam. In: ———. ———. São José dos Campos. INPE/CAPES (no prelo).
- BANATHY, B.H. *Instructional Systems*. California, Fearon Publishers, 1967.
- BARRY, M. Evaluation technicians: who are they? Who needs them? *Educational technology*, 14 (2): 49-51, Fev. 1974.
- BART, W.M. Test validity and reliability from an ordering theoretic framework. *Educational technology*, 14(1):62-63, Jan. 1974.
- BLOOM, B.S.; ENGELHART, M.D.; FURST, E.J.; HILL, W.H.; KRATHWOHL, D.R. *Taxionomia de objetivos educacionais: domínio cognitivo*. Taxonomy of educational objectives: cognitive domain. Trad. Flávia M. Santana. Porto Alegre, Globo, 1973.

- BLOOM, B.S.; HASTINGS, J.T.; MADAUS, G.F. Formative evaluation.
In: ———. *Handbook on formative and summative evaluation of student learning*. New York, McGraw-Hill, 1971. p. 117-138. Cap. 6.
- BRANSON, R.K.; HANNUM, W.H.; RAYNER, G.T.; JOHNSON, B.F. *The instructional systems development model: a description of the processes, procedures and products*. Florida, Center for Educational Technology, 1973.
- BRASIL. Presidência da República. Revolução na educação. In *Metas e bases para ação do governo*. Brasília, 1970.
- . República Federativa. *II Plano nacional de desenvolvimento 1975/79*. Brasília, 1974. Cap. 14.
- . Leis e decretos. Lei 5692, 11 de agosto de 1971. *Diretrizes e bases para o ensino de 1º e 2º graus*. Brasília, Ago. 1971.
- . ———. Lei 5540, 28 de novembro de 1968. Fixa normas de organização e funcionamento do ensino superior e sua articulação com a escola média e dá outras providências. *Diário oficial*, Brasília, Nov., 1968.
- BRETZ, R.M. *The selection of appropriate communication media for instruction: a guide for designers of air force technical training programs*. California, Rand Corporation, 1971.
- BRIGGS, L.J.; GAGNÉ, R.M. Evaluating instruction. In ———. *Principles of instructional design*. New York, Holt Rinehart and Winston, 1974. P 231-255. Cap. 12.
- BRIGGS, L.J. *Handbook of procedures for the design of instruction*. Pittsburgh, American Institute for research, 1970.

CAMPBELL, D.T.; STANLEY, J.C. *Experimental and quasi-experimental designs for research.* Chicago, Rand Mc Nally, 1970.

CAPES. *Catálogo dos cursos de Pós-Graduação no Brasil.* Brasília, MEC, 1975.

———. *Novas perspectivas para o sistema de ensino superior: operacionalização e implantação do plano nacional de Pós-Graduação.* Brasília, MEC, 1975.

CARDOZO, G.B.; OLIVEIRA, M.L.N.; ARANTES, M.S.; COSTA, S. Tipos de avaliação, suas características e metodologias. In ———. *Curso de Programação, Execução e Avaliação de Projetos Educacionais.* São José dos Campos, INPE/CAPES (no prelo).

CARDOZO, G.B. Os quatro tipos de Avaliação: Contexto, input, processo, produto. In: ———. São José dos Campos, INPE/CAPES (no prelo).

CARDOZO, G.B.; OLIVEIRA M.L.N. Formato de itens. In: ———. *Curso de testes e medidas educacionais.* Natal, MEC/DSU/MINTER/SUDENE, 1975.

———. Abordagem CR e NR de medidas e sugestões de passos para elaboração de um teste CR. In: ———. Natal, MEC/DSU/MINTER/SUDENE, 1975.

———. Geração de itens. In: ———. Natal, MEC/DSU/MINTER/SUDENE, 1975.

CHEIN, I. Uma introdução à amostragem. In: SELTZ, C.; JAHODA M.; DEUSTSCH M.; COOK S. W. *Métodos de pesquisa nas relações sociais.* São Paulo, Pedagógica e Universitária, 1974. p. 571-611 Apêndice B.

- COLLINS, H.W.; JOHANSEN, J.H.; JOHNSON, J.A. *Educational measurement an evaluation: a worktext*. Illinois, Scott Foreman, 1969.
- COOMBS, P.H. Que é planejamento educacional. *Cadernos de pesquisa*, 4: 53-91, Out. 1972.
- CHURCHMAN, C.W. O que é um sistema. In:———. *Introdução a teoria dos sistemas*. 2 ed., Petrópolis, Vozes, 1972. p. 15-89. Cap. 1.
- CRONBACH, L.J. Who uses tests? In:———. *Essentials of psychological testing*. 3. ed. New York, Harper and Row, 1970. p. 3-21. cap. 1.
- . Purposes and types of tests. In:———.———. 3. ed. New York, Harper and Row, 1970, 22-43. cap. 2.
- . Administering tests. In:———.———. 3. ed. New York, Harper and Row, 1970. p. 44-47. cap. 3.
- . Scoring. In:———.———. 3. ed. New York, Harper and Row, 1970. p. 78-114. cap. 4.
- . Test validation. In:———.———. 3. ed. New York, Harper and Row, 1970. p. 115-150. cap. 5.
- . Other characteristic desired in tests. In:———.———. 3. ed. New York, Harper and Row, 1970. p. 151-193. Cap. 6.
- CUNNINGHAM, D. J. Comments on the case studies of formative evaluation the sources of information. In: *Bulletin of the school of education*, Indiana University, 48 (4): 111-118, Jul. 1972.
- CURLE, A. A profissão de planejador educacional. In: *Cadernos de Pesquisa*, 13: 73-95, Jun. 1975.

- DAVIES, I.K.; HUDSON, E.H.; DODD, B.; HARTHLEY, J.A. *Organização do treinamento*. São Paulo, McGraw Hill, 1976.
- DAVIES, I.K.; SCHWEN, T.M. Some process concerns about formative evaluation. In: *Bulletin of the school of education*, Indiana University, 48(4): 1-8, Jul. 1972.
- DAVIS, R.A. Gains of learning. In: ———. *Learning in the schools*. Belmont, Wadsworth, 1966. p. 11-35. cap. 1.
- DE CECCO, J.P. Evaluation and school use. In: ———. *Educational Technology*. New York, Holt Rinehart and Winston, 1964. p. 393-445. cap. 9.
- DEDERICK, W.F.; STURGE, H.H. An operational instructional systems model: US Naval Training. *Educational Technology*, 15 (6): 28-32, June, 1975.
- DELEON, A. Conceito atual de educação permanente e seu planejamento. *Revista brasileira de estudos pedagógicos*. 51 (113):19-31, Jan./Mar. 1969.
- DIAS, J.A. Sistema escolar brasileiro. In: Brejon, M. *Estrutura e funcionamento do ensino de 1º e 2º graus*. 4 ed., São Paulo, Pioneira, 1974. p. 71-91. cap. 5.
- FERGUSON, G.A. The design of experiments. In: ———. *Statistical analysis in psychology & education*. 3. ed., New York, McGraw-Hill, 1971. p. 197-288. parte 2.
- FERREIRA, P.P. *Treinamento de Pessoal: a técnico-pedagogia do treinamento*. São Paulo, Atlas, 1975.

- FINN, J.D. Institutionalization of evaluation. In: *The Educational technology review series: evaluation of education*. V. 11. Englewood Cliffs, Educational Technology, 1973. p. 10-19.
- FLIPPO, E.B. *Principles of personnel management*. New York, MacGraw-Hill, 1971.
- GAGNÉ, R.M. *Como se realiza a aprendizagem*. (The conditions of learning). Trad. Therezinha Maria Ramos Tovar. Rio de Janeiro, Ao Livro Técnico, 1973.
- . Curriculum research and the promotion of learning. In: TYLER, R. W. et al. *Perspectives of curriculum evaluation*. Chicago, Rand McNally, 1967. p. 19-38.
- GATTI, B.A.; FERES, N.L. Amostragem. In: ———. *Estatística básica para ciências humanas*. São Paulo, Alfa Omega, 1975. p. 46-52. cap. 3.
- GRIESSMAN, B.E. An approach to evaluating comprehensive social projects. In: *The Educational Technology Reviews Series: Evaluation of Education*. V. 11. Englewood Cliffs, Educational Technology, 1973. p. 90-93.
- GLONLUND, N.E. *A elaboração de testes de aproveitamento escolar*. Trad. de Erb Luiz Leite Cruz. São Paulo, Pedagógica e Universitária, 1974.
- GUBA, E.G. The failure of educational evaluation. In: *The Educational Technology Reviews Series: Evaluation of Education*. V. 11. Englewood Cliffs, Educational Technology, 1973. p. 5-9.

- GUILFORD, J.P. Statistical estimation and inferences. In: ———. *Fundamental Statistics in psychology and education*. 4. ed. New York, McGraw-Hill, 1965. p. 136-170. cap. 8.
- . Significance of differences. In: ———. ———. 4. ed. New York, McGraw-Hill, 1965. p. 136-170. cap. 8.
- . Hypothesis testing. In: ———. ———. 4. ed. New York, McGraw-Hill, 1965, p. 200-214. Cap. 10.
- . Introduction to analysis of variance. In: ———. ———. McGraw-Hill, 1965. p. 268-303. cap. 13.
- HICKS, C.R. *Fundamental concepts in the design of experiments*. New York, Holt Rinehart and Winston, 1963.
- HILL, J.E.; KERBER A. Problem and hypothesis. In: ———. *Models, methods, and analytical procedures in education research*. Detroit, Wayne State University Press, 1967. p. 23-36. cap. 3.
- INSTITUTO DE PESQUISAS ESPACIAIS (INPE). *Catálogo de Pós-Graduação: Programa da área de tecnologia da educação*. São José dos Campos, 1975/76. V. 8.
- JOHONSON, R.A.; KAST, F.E.; ROSENZWEIG, J.E. System theory and management. In: ———. *The theory and management systems*. New York, McGraw-Hill, 1967. p. 15-89. cap. 1.
- KING, R.B. Scientific sampling. In: *Bell telephone magazine*. New York, American Telephone and Telegraph Company, 1964, V. 43, N. 3, p. 16-26.
- KING, W.G.; CLELAND, D.I. *Systems analysis and project management*. New York, McGraw-Hill, 1968.

- LAFER, B.M. O conceito de planejamento. In: ———. *Planejamento no Brasil*. São Paulo, Perspectiva, 1970. p. 9-28.
- LIGHT, J.A.; REYNOLDS, L.J. Debugging product and testing errors procedures for the formative evaluation of an individualized mathematics curriculum. In: *Bulletin of the school of education*, Indiana University, 48(4): 45-78, Jul. 1972.
- LYMAN, H.B. *Test scores and what they mean*. Englewood Cliffs, Printice Hall, 1963.
- LYSAUGHT, J.P.; WILLIAMS, C.M. Evaluation. In: ———. *A guide for programmed instruction*. New York, John Wileyand Sons, 1968. p. 134-147. cap. 8.
- MAGER, R.F.; BEACH, K.M. *Developing vocational instruction*. Palo Alto, Fearon, 1967.
- MANAGEMENT INFORMATION SERVICE *Introduction to system analysis*. s.1, 1968.
- MEC. *Plano setorial de Educação e Cultura 72/74*. Brasília, 1971
- . *Plano setorial de Educação e Cultura*. 75/79. Brasília, 1974.
- NERICI, E.G. Métodos e técnicas de ensino. In: ———. *Introdução à didática geral*. 8. ed. Rio de Janeiro, Fundo de Cultura, 1969, p. 226-308. cap. 9.
- NETTO, A.D. Apresentação. In: ———. *Planejamento para o desenvolvimento econômico*. São Paulo, Editora da Universidade de São Paulo, 1966. p. 9-16.

- NUNNALLY, J.C. Introduction. In: ———. *Psychometric theory*.
New York, MacGraw Hill, 1976. p. 2-30. cap. 1.
- . Validity. In: ———.———. New York, MacGraw Hill, 1967.
p. 75-102. cap. 3.
- . Test construction. In: ———.———. New York, MacGraw Hill,
1967. p. 236-286. cap. 8.
- . Measurement of sentiments. In: ———.———. New York,
MacGraw Hill, 1967. p. 514-558, cap. 14.
- OLIVEIRA, J.B. *Tecnologia educacional: teorias da instrução*. Petrô
polis, Vozes, 1973.
- OLIVEIRA, J.B.; OLIVEIRA, M.R. A função da avaliação na tomada de
decisões educacionais. In: MEC *V Encontro de secretários e represen
tantes dos conselhos de educação*. Brasília, MEC, 1973.
- OTIS, J.L.; LEUKART, R.H. *Job evaluation*. Englewood Cliffs, Prentice
Hall, 1954.
- POPHAM, W.J. Expanding measurement alternatives. In: ———.
Educational Evaluation. Englewood Cliffs, Prentice Hall, 1975.
p. 90-101. cap. 5.
- . Classical measurement considerations. In: ———.———.
Englewood Cliffs, Prentice Hall, 1975. p. 104-125. cap. 6.
- . Criterion referenced measurement. In: ———.———. Englewood
Cliffs, Prentice Hall, 1975. p. 126-162. cap. 7.
- . Evaluation designs. In: ———.———. Englewood Cliffs,
Prentice Hall, 1975. p. 195-217. cap. 9.

- POPHAM, W.J. Sampling strategies. In: ———.————. Englewood Cliffs, Prentice Hall, 1975. p. 218-233. cap. 10.
- . Analysing evaluative data. In: ———.————. Englewood Cliffs, Prentice Hall, 1975. p. 234-257. cap. 11.
- . *Criterion referenced measurement: an introduction*. Englewood Cliffs, Educational Technology, 1973.
- PROVUS, M. Evaluation of ongoing programs in the public school system. In: WORTHEN, B.P.; SANDERS, J.R. *Educational evaluation: theory and practice*. Washington, D.C.; Charles A. Jones, 1973.
- RANDALL, R.S. An operational application of the CIPP model for evaluation. In: *The Educational Technology Reviews Series: Evaluation of Education*. Englewood Cliffs, Educational Technology, 1973. p. 43-47.
- REIS, J.R.; MONTEIRO, R.P.; CIMA, S.C.F. *Manual de engenharia de sistemas e projetos: uma abordagem prática*. São José dos Campos, (INPE-848-TPT/025, 1976).
- SABBAG, M.A. *A televisão educativa no Brasil: um estudo analítico*. Tese de Mestrado. Rio de Janeiro, PUC, 1974.
- SANDERS, J.R. Comments on the case studies of formative evaluation search for structure. In: *Bulletin of the school of education*, Indiana University, 48 (4): 119-132, Jul. 1972.
- SANDERS, J.R.; MURRAY, S.L. Alternatives for achievements testing. *Educational technology*, 16(3):17-23, Mar., 1976.

- SASLOW, M.G. Establishing the purpose for evaluation. In: PAULSON, C.F. Jr. *A strategy for evaluation design*. Oregon, 1970. p 1-55.
- SCRIVEN, M. The methodology of evaluation. In: TYLER, R.W. et al. *Perspectives of curriculum evaluation*. Chicago, Rand McNally, 1967. p. 39-83.
- SELLTIZ, C.; JAHODA, M.; DEUSTSCH, M.; COOK, S.W. Escolha e formulação de um problema de pesquisa. In:———. *Métodos de Pesquisa nas relações sociais*. São Paulo, Pedagógica e Universitária, 1974. p. 31-56. cap. 2.
- . Planejamento de Pesquisa. In:———.———. São Paulo, Pedagógica e Universitária, 1974. p. 57-90. Cap. 3.
- . Planejamento de pesquisa. In: ———.———. São Paulo, Pedagógica e Universitária, 1974. p. 91-162, cap. 4.
- . Coleta de dados: questionários e entrevistas. In: ———.———. São Paulo, Pedagógica e Universitária, 1974, p. 265-311. cap. 7.
- . Colocação de indivíduos em escalas. In: ———.———. São Paulo, Pedagógica e Universitária, 1974. p. 389-432, cap. 10.
- SIEGEL, S. The use of statistical tests in research. In: ———. *Nonparametric statistics for the behavioral sciences*. New York, McGraw-Hill, 1956. p. 6-17. cap. 2.
- STAKE, R.E. Evaluation design, instrumentation, data collection, and analysis of data. In: WORTHEN, B.R.; SANDERS, J.R. *Educational Evaluation: theory and practice*. Washington, D.C.; Charles A. Jones, 1973.

- STONE, V.I. *Partial results of evaluation of the project Pro-Cess materials: career education: an introduction.* Tallahassee, Florida State University, 1974. (Relatório não publicado).
- . *Developing criterion referenced tests: a review of the relevant literature.* (Relatório não publicado).
- . *Development of criterion referenced tests: an empirical investigation.* s.l., Florida State University, 1974. p. 5-30. cap. 2.
- STUFFLEBEAM, D.L.; FOLEY, W.J.; GEPHART, W.J.; GUBA, E.G.; HAMMOND, R.L.; MERRIMAN, H.O.; PROVUS, M.M. *Educational evaluation and decision making.* Itaska, Peacock, 1971.
- STUFFLEBEAM, D.L. An Introduction to the PDK Book: Educational Evaluation and Decision-Making. In: WORTHEN, B.R.; SANDERS, J.R. *Educational evaluation: theory and practice.* Washington, D.C., Charles A. Jones, 1973.
- . Evaluation as enlightenment for decision-making. In: WORTHEN, B.R.; SANDERS, J.R. *Educational evaluation theory and practice.* Washington, D.C. Charles A. Jones, 1973.
- . Toward a science of educational evaluation. In: *The educational technology reviews series: evaluation of education.* V. 11. Englewood Cliffs, Educational Technology, 1973. p. 20-27.
- THE EDUCATIONAL TECHNOLOGY PUBLICATION. *Evaluation of education.* s.l., Library of Congress Cataloging, 1973.
- TRACEY, W.R. *Designing training and development systems.* s.l., American Management Association, 1971.

TUCKMAN, B.W. *Conducting educational research*. New York, Harcourt Brace Jovanovich, 1972.

TURRA, C.M.G.; ENRICIONE, D.; SANT'ANNA, F.M.; ANDRE, L.C. *Planejamento de ensino e avaliação*. 4. ed. Porto Alegre, PUC, 1975.

U.S. CIVIL SERVICE COMMISSION - BUREAU OF TRAINING. *Instructional systems and technology: an introduction to the field and its use in federal training*. Washington, U.S. Government Printing Office, 1969. Training Systems and Technology, 1.

WORTHEN, B.R. Competencies for Educational Research and Evaluation. *Educational researcher*, 4(1):13-16, Jan., 1975.

———. Toward a taxonomy of evaluation designs. In: *The educational technology reviews series: evaluation of education*. V.11. Englewood Cliffs, Educational Technology, 1973. p. 28-35.

APÊNDICES

APÊNDICE A

LISTAGEM DE ENTIDADES EDUCACIONAIS

FORNECIDA PELA ABT

- . Estado Acre
 - Fundação Universidade Federal do Acre
 - SEEC - Departamento de Ensino Supletivo.
- . Estado de Alagoas - MEC - SRE - Projeto Minerva
- . Estado do Amazonas - Fundação TVE
- . Estado da Bahia
 - Instituto de Radiofusão Educativo da Bahia - IRDEB
 - Universidade Federal Vale do Canela
- . Estado do Ceará
 - TV Educativa
 - Secretaria de Educação
- . Distrito Federal
 - Universidade de Brasília
 - Ministério de Educação e Cultura
 - Fundação Educacional do Distrito Federal
 - Pimmo
- . Estado do Espírito Santo
 - SEEC - Divisão de Desenvolvimento de Recursos Humanos
- . Estado de Goiás
 - Faculdade de Educação da Universidade de Goiás
 - Universidade Católica de Goiás
- . Estado do Rio de Janeiro
 - PRONTEL
 - Universidade do Estado da Guanabara
 - DSU - Setor de TeleEducação
 - Fundação Centro-Brasileira de Televisão Educativa.
 - Universidade Federal Fluminense
 - SENAC
 - SENAI

- Projeto Minerva
- ABT
- C.N.R.H. - Centro Nacional de Recursos Humanos - do IPEA.
- Faculdade de Educação, Ciências e Letras - Mário H. Simonsen
- Pontifícia Universidade Católica do Rio de Janeiro
- Secretaria da Educação da Guanabara
- Instituto de Educação
- DSU - Rio de Janeiro
- Fundação Konrad Adenauer
- . Estado do Maranhão
 - TV Educativa do Maranhão
 - Fundação Maranhense de TV Educativa
 - Faculdade de Educação da Universidade do Maranhão
- . Estado de Minas Gerais
 - Fundação Pândia Calógeras
 - Projeto Minerva
 - SENAC
 - ULTRAMIG - Fundação de Educação para o Trabalho de Minas Gerais
 - Secretaria da Educação do Estado de Minas Gerais
 - Centro de Recursos Humanos João Pinheiro - CRHJP
 - Faculdade de Educação da U.F.M.G
 - Instituto Cultural Newton Paiva Ferreira
- . Estado de Paraíba
 - Secretaria de Educação e Cultura
- . Estado do Paraná
 - Universidade Federal do Paraná
- . Estado do Pernambuco
 - Departamento de Recursos Tecnológicos para a Educação.

- TV Universitária Canal 11
- Secretaria de Educação - DERTE
- Secretaria de Pernambuco - DDN
- . Estado do Piauí - Secretaria de Educação do Piauí
- . Estado do Rio Grande do Norte - Secretaria de Educação e Cultura do Estado do Rio Grande do Norte
- . Estado do Rio Grande do Sul - FEPLAM - Fundação Educacional Padre Landell de Moura
- Fundação para o Desenvolvimento de Recursos Humanos
- Faculdade de Educação da Universidade Federal do Rio Grande do Sul
- . Estado de Santa Catarina - Secretaria da Educação - Setor de Televisão Educativa.
- Instituto Estadual de Educação.
- . Estado de São Paulo - Fundação Padre Anchieta
- INPE
- Fundação CENAFOR
- Fundação Carlos Chagas
- . Estado de Sergipe - Secretaria da Educação e Cultura
- Universidade Federal de Sergipe

APÊNDICE B

FORMULÁRIO

1.0 - Instituição _____

2.0 - Existe atividade de avaliação na Instituição?

Sim

Não

3.0 - Em caso positivo, quantos elementos estão envolvidos nessas atividades?

4.0 - Desses elementos, quantos têm:

Curso de doutoramento?

Curso de mestrado?

Curso universitário?

5.0 - Quantos desses elementos tiveram cursos específicos de avaliação?

6.0 - Descreva brevemente as atividades de avaliação executadas nesta
Instituição.

Informante: _____

Data: _____

Cargo: _____

APÊNDICE C

INSTRUMENTAL DE COLETA DE DADOS PARA ANÁLISE OCUPACIONAL
DO AVALIADOR EDUCACIONAL

Nome da Instituição: _____

Departamento: _____

Nome do Respondente: _____

Função que exerce: _____

Trabalho Supervisionado por: _____

Entrevistador: _____

Data da Entrevista: _____

FORMAÇÃO ACADÊMICA		
NOME DA INSTITUIÇÃO (por extenso)		Sigla do Estado
CURSO (por extenso)	Nível	Ano de Término
NOME DA INSTITUIÇÃO (por extenso)		Sigla do Estado
CURSO (por extenso)	Nível	Ano de Término

EXPERIÊNCIA PROFISSIONAL		INÍCIO		TÉRMINO	
		MÊS	ANO	MÊS	ANO
EMPRESA OU INSTITUIÇÃO	FUNÇÃO				

- Para o exercício de suas atividades de Avaliação você passou por um treinamento?

Sim

Não

- Se sim, quanto tempo de treinamento?

- Indique a natureza do seu treinamento

treinamento em serviço (dias, semanas, meses de trabalho supervisionado).

treinamento através de cursos

ambos acima

- Dentre as afirmativas listadas a seguir, assinale sim ou não, conforme façam ou não parte de seu trabalho como avaliador.

1. Participa, juntamente com o pessoal do sistema (líderes de grupo, professores, chefes de departamento, gerentes, diretores, planejadores, etc...)

- da definição do sistema (podendo se constituir num Currículo, um Programa, um Curso, um Projeto, etc.): seus limites, elementos e características destes elementos.

Sim Não

- da identificação dos objetivos do sistema

Sim Não

- da determinação de padrões e sua aplicação sobre os objetivos do sistema para estimar o valor dos mesmos

Sim Não

- da determinação da viabilidade dos objetivos do sistema em relação aos recursos disponíveis.

Sim Não

- do estudo e descrição detalhada do sistema corrente em termos das seguintes variáveis:

- populacionais (aluno, professor, especialistas, família e comunidade)

Sim Não

- instrucionais (métodos, conteúdo, organização, facilidades e custo)

Sim

Não

- comportamentais (cognitivo, afetivo e psicomotor)

Sim

Não

2. Compara a situação atual do sistema com a desejada, identificando discrepâncias.

Sim

Não

3. Identifica, fora do sistema, oportunidades não usadas por este

Sim

Não

4. Levanta e responde questões, através de um estudo de Avaliação, que orientam o diagnóstico de problemas que gera as discrepâncias (entre o real e o desejado).

Sim

Não

5. Participa juntamente com o pessoal do sistema:

- da identificação de estratégias alternativas para solucionar os problemas

Sim

Não

- da operacionalização dos objetivos, da mudança a ser introduzida para solucionar os problemas, tornando-os precisos e mensuráveis.

Sim

Não

- da especificação de critérios para avaliar o alcance dos objetivos

Sim

Não

- da identificação dos recursos humanos e materiais disponíveis na agência, para orientar a escolha da melhor estratégia

Sim

Não

- da identificação de "designs" de procedimentos (para implementação da estratégia escolhida)

Sim

Não

6. Mantém um registro constante dos procedimentos tal como ocorrem

Sim

Não

7. Deteta ou prediz defeitos do "design" de procedimentos

Sim

Não

8. Fornece realimentação periódica sobre o desempenho do sistema

Sim

Não

9. Mede alcance de objetivos

Sim

Não

10. Compara dados obtidos com critérios pré-estabelecidos

Sim

Não

11. Elabora Plano de Avaliação para orientar a execução de suas atividades

Sim

Não

12. Planeja coleta de dados incluindo:

- fontes de informação sim não

- procedimentos de amostragem sim não

- instrumentais sim não

- cronograma para coleta de dados sim não

13. Formula hipóteses testáveis

Sim

Não

14. Especifica variáveis a serem medidas

Sim

Não

15. Especifica dados para um teste de hipóteses rigoroso ou resposta inequívoca

Sim

Não

16. Utiliza "designs" experimentais, pré-experimentais ou quase-experimentais, de acordo com Campbell e Stanley.

Sim

Não

17. Utiliza outros tipos de "designs" de Avaliação como por exemplo _____

18. Utiliza meios informais de investigação como

. estuda e visita outros sistemas Sim Não

. participa de seminários e conferências Sim Não

19. Utiliza procedimentos de investigação formal como:

. técnica Delphi de convergência de opinião Sim Não

. técnica de pesquisa bibliográfica Sim Não

20. Utiliza técnicas de Controle como:

- . Diagrama de Marcos Sim Não
- . Cronogramas Mestres e Parciais Sim Não
- . Técnicas de O & M Sim Não

21. Aplica instrumentais de medida tais como:

- . Testes de Aprendizagem Sim Não
- . Questionários Sim Não
- . Escalas de atitude Sim Não
- . Fichas de observação Sim Não
- . Entrevistas Sim Não

22. Atribui pontos aos resultados obtidos na aplicação dos instrumentais de medida

- Sim Não

23. Registra os dados obtidos através da aplicação dos instrumentais

- Sim Não

24. Para utilizar instrumentos você os:

- Seleciona Sim Não
- Adapta Sim Não
- Desenvolve Sim Não

28. Que tipo de material é elaborado na Instituição e que procedimentos utiliza para Avaliação destes materiais?

29. Escolhe e emprega técnicas de Análise Estatística?

- descritiva Sim Não
- inferencial Sim Não

30. Quais técnicas são mais frequentes?

31. Utiliza técnicas analíticas como: gráficos, tabelas, perfis.

- Sim Não

32. Utiliza, para a computação dos dados:

- processamento manual Sim Não
- computador Sim Não

33. Interpreta dados em termos de um critério especificado.

Sim

Não

34. Elabora relatórios escritos e orais, para audiências especificadas como por exemplo: decisores, seus subordinados, clientela, agências.

Sim

Não

35. Indique as atividades que você desempenha:

- diariamente:

- semanalmente:

- mensalmente:

- trimestralmente:

- semestralmente:

- anualmente:

- ocasionalmente: (atividades desempenhadas em intervalos irregulares).

36. Quais as atividades que, em sua opinião, constituem a parte mais difícil ou complexa de seu trabalho como avaliador. Indique se são atividades diárias, semanais, mensais, etc.

ATIVIDADES MAIS DIFÍCEIS OU COMPLEXAS	FREQUÊNCIA

APÊNDICE D

INSTRUMENTAL DE COLETA DE DADOS, JUNTO AOS DECISORES,
PARA ANÁLISE OCUPACIONAL DO AVALIADOR

Nome da Instituição _____
Nome do Entrevistado _____
Função que Exerce _____
Entrevistador _____
Data da Entrevista _____

1.0 - Quais os objetivos gerais e específicos da Instituição?

(se o espaço não for suficiente, utilize o verso da folha).

2.0 - Qual a importância da equipe de avaliação para a consecução destes objetivos?

3.0 - Quais as atividades que a equipe de avaliação realiza?

4.0 - Que outras atividades você acha que ela deveria realizar?

5.0 - Se houvesse um curso de treinamento para seus avaliadores, quais as atividades que, na sua opinião, deveriam ser enfatizadas?

APÊNDICE E

FREQUÊNCIA DE EXECUÇÃO DAS ATIVIDADES

A T I V I D A D E S	FREQUÊNCIA COM QUE É EXECUTADA * %
- Definir o sistema . .	100
- Identificar os objetivos do sistema .	100
- Determinar o valor e viabilidade dos objetivos .	91
- Descreve o sistema corrente em termos de variáveis:	
Populacionais	91
Instrucionais	75
Comportamentais	75
- Comparar a situação atual do sistema com a desejada identificando discrepâncias.	100
- Identificar, fora do sistema, oportunidades não usadas por este.	8
- Levantar questões que orientam o diagnóstico dos pro blemas que geram as discrepâncias (entre o real e o de sejado).	91
- Identificar estratégias alternativas, para relacionar o problema.	83
- Operacionalizar os objetivos da mudança a ser introdu zida, para solucionar o problema, tornando-os precisos e mensuráveis.	83
- especificar critérios para avaliar o alcance dos obje tivos.	100

* Esta percentagem é baseada nas respostas de 26 avaliadores.

A T I V I D A D E S	FREQUÊNCIA COM QUE É EXECUTADA %
- Identificar recursos materiais e humanos disponíveis, para orientar a escolha da melhor estratégia.	83
- Identificar "designs" de procedimentos (para implementação da estratégia escolhida).	83
- Manter um registro constante dos procedimentos, tal como ocorrem.	16
- Detetar ou predizer defeitos no "design" de procedimentos.	100
- Fornecer realimentação periódica sobre o desempenho do sistema.	100
- Medir o alcance dos objetivos.	100
- Comparar dados obtidos com critérios pré-estabelecidos.	100
- Elaborar plano de avaliação para orientar a execução das atividades.	100
- Planejar coleta de dados incluindo:	
. Fontes de informação	83
. Procedimentos de amostragem	83
. Instrumentais	83
. Cronograma de coleta de dados	83
- Formular hipóteses testáveis.	75
- Especificar variáveis a serem medidas.	83
- Especificar dados para um teste de hipótese rigoroso.	75
- Utilizar "designs" pré-experimentais, quase-experimentais e experimentais, de acordo com Campbell Stanley.	00

A T I V I D A D E S	FREQUÊNCIA COM QUE É EXECUTADA %
- Utilizar meios informais de investigação como:	
. Estudar e visitar outros sistemas	58
. Participar de seminários e conferências.	91
- Utilizar procedimentos de investigação formal, como:	
. Técnica Delphi de convergência de opinião	00
. Técnica de pesquisa bibliográfica.	83
- Utilizar técnicas de controle como:	
. Diagrama de Marcas	00
. Cronogramas mestres e parciais	67
. Técnicas de O & M .	50
- Aplicar instrumentais de medida tais como:	
. Testes de aprendizagem	42
. Questionários	83
. Escalas de atitude	08
. Fichas de observação	58
. Entrevistas.	91
- Atribuir pontos aos resultados obtidos com a aplicação de instrumentais de medida.	83
- Registrar dados	75
- Estabelecer os propósitos do uso do instrumental de medida
- Selecionar instrumental	67
- Adaptar instrumental	67
- Desenvolver instrumental.	75

A T I V I D A D E S	FREQUÊNCIA COM QUE É EXECUTADA %
- Realizar validação de instrumental	
. Validação de conteúdo	58
. Validação de construto	17
. Validação preditiva.	50
- Utilizar para computação dos dados	
. Processamento manual	75
. Computador.	41
- Escolher e empregar técnicas de análise estatística:	
. Descritiva	83
. Inferencial.	17
- Utilizar técnicas analíticas tais como: gráficos, ta belas, perfis.	00
- Interpretar dados em termos de critérios especificados.	100
- Identificar audiência
- Elaborar relatórios escritos e orais	100
- Faz avaliação de materiais.	
. Avaliação Formativa	00
. Avaliação Somativa.	25

APÊNDICE F

ESPECIFICAÇÃO DO CONTEÚDO PARA O CURSO DE TREINAMENTO DE AVALIADORES EDUCACIONAIS

A partir da análise ocupacional, foram detetadas as atividades que caracterizam a função do avaliador. Em seguida, com base nessas atividades, foram gerados os objetivos instrucionais do curso de treinamento.

Este apêndice trata da especificação do conteúdo para a consecução dos objetivos instrucionais propostos, e está assim apresentado:

- Unidade 1 - Contexto do Processo de Avaliação - Planejamento Sistemático e Processo Decisório.
- Unidade 2 - Uma Abordagem Avaliativa - O Modelo de Stufflebeam et al.
- Unidade 3 - Obtenção de Informações - Instrumentais de Medida.
- Unidade 4 - Obtenção de Informações - Avaliação Formativa de Materiais Instrucionais Impressos e Alunos.
- Unidade 5 - Obtenção de Informações - Avaliação Somativa dos Produtos da Mudança em Termos de Aprendizagem.
- Unidade 6 - A Estrutura de um Plano de Avaliação.

1.0 - UNIDADE 1: O CONTEXTO DO PROCESSO DE AVALIAÇÃO: PLANEJAMENTO SISTÊMICO E PROCESSO DECISÓRIO

1.1 - SUB-UNIDADE 1: PLANEJAMENTO, TEORIA GERAL DE SISTEMAS, ABORDAGEM DE SISTEMAS: VISÃO GERAL

1.1.1 - PLANEJAMENTO

Até por volta de 1929, a política econômica mundial fundamentava-se na filosofia do "laissez-faire". Acreditava-se que o funcionamento do mercado era automático e que qualquer intervenção governamental, na economia, seria nociva por romper a competição. O papel do Estado era exclusivamente de mediador, objetivando assegurar liberdade para as atividades econômicas. Assim sendo, qualquer tipo de planejamento seria desnecessário, desde que o próprio mercado realizava as funções de um órgão central de planificação. O seu mecanismo era uma consequência da relação entre oferta e demanda. Se havia grande procura de um certo produto, havia maior fluência de indivíduos para aquela área de produção, até que oferta e demanda se equilibrassem. Quando isto ocorria, os lucros diminuam e os indivíduos procuravam um novo mercado que lhes pudessem oferecer maiores vantagens financeiras. Neste mercado permaneciam, até que novamente houvesse um equilíbrio entre oferta e demanda, diminuindo seus lucros. Isto ocorreu por volta de 1929 quando houve uma superprodução em todos os mercados. Em decorrência disto agravou-se um sério problema: o desemprego. Esta instabilidade do sistema econômico, com crises cíclicas de desemprego e a nova ênfase dada ao desenvolvimento econômico, mostravam que as simples forças do mercado, com pequena intervenção do Estado, eram incapazes de conduzir aos resultados pretendidos pela sociedade. Compreendeu-se que a filosofia do "laissez-faire" não possibilitava o alcance dos objetivos sociais. Recursos do Governo deveriam ser alocados no mercado para que este readquirisse o equilíbrio perdido. Porém, desde que o Governo resolveu aplicar seus recursos, no mercado, sentiu necessidade de planejar para isto. Perdeu-se, assim a crença no auto

matismo do mercado; a filosofia do "laissez-faire" foi abandonada, e adotou-se a técnica de planejamento.

A técnica de planejamento, em suas linhas gerais, consiste em assegurar o equilíbrio, entre os níveis de produção e a demanda dos bens, dada a oferta de fatores de produção de forma a atingir certos objetivos básicos (Lafer, 1970).

O planejamento, em si mesmo, é neutro, ele pode ser utilizado para fortalecer a economia do mercado ou para substituí-la; pode ser restrito às áreas tradicionais da atividade governamental ou pode ampliá-la; pode ser usado com objetivos sociais dignos ou para beneficiar uma classe em detrimento da outra (Netto, 1966).

O planejamento em qualquer setor possibilita, com maior segurança, um equilíbrio entre oferta e demanda. Com a aplicação dos recursos do Governo, no mercado, este foi reativado com a volta dos consumidores, e surgiram novas oportunidades de emprego. Um novo equilíbrio, entre oferta e demanda, foi então obtido através de um planejamento racional para alocação ótima dos recursos existentes.

A intervenção governamental fez-se necessária, não só na área econômica, como também na educação. Surgiu, assim, a necessidade do planejamento educacional sobre o qual se descreverá a seguir.

PLANEJAMENTO EDUCACIONAL

a) Algumas Considerações

O planejamento educacional sempre foi uma necessidade, existindo através dos tempos com diferentes características, para atender a objetivos particulares. Assim, vários países utilizaram-se de diferentes abordagens de planejamento em relação ao seu alcance, obje-

tivos e complexidade.

Em épocas anteriores à 2ª Grande Guerra, os procedimentos de planejamento eram, geralmente, bastante simples e centralizados apenas no próprio processo educacional. Havia pouca ou nenhuma preocupação com as necessidades de sua clientela ou do seu meio-ambiente. Porém, os procedimentos, então utilizados, mostravam-se apropriados no sentido de que conseguiam manter um equilíbrio entre os componentes do sistema educacional e o seu meio ambiente. Este se desenvolvia e se modificava em ritmo lento; e poucas exigências fazia à educação.

Com o passar do tempo, profundas e rápidas modificações ocorreram em todas as partes do mundo. Nas últimas décadas do nosso século ocorreram transformações científicas, políticas e econômicas que afetaram profundamente a educação, pressionando-a a assumir novas e mais complexas funções. Na tentativa de assumir seu novo papel, a educação passou a enfrentar sérios problemas de planejamento. Constatou-se a ineficiência dos antigos procedimentos, e a necessidade de criar um novo tipo de planejamento.

Tal necessidade tornou-se premente logo após a 2ª Grande Guerra.

b) Fases da Educação e a Necessidade de Planejamento

Segundo Combs (1972), as nações industrializadas, no período de pós-guerra, passaram por três fases educacionais de 1954 a 1970, e encontram-se agora numa quarta fase um tanto desnorteante. Cada uma delas trouxe consigo sérios problemas de planejamento. Elas se identificam como:

1. Fase da Reconstrução
2. Fase da Falta de Mão-de-obra

3. Fase da Expansão Desenfreada
4. Fase da Inovação.

A primeira fase de reconstrução foi caracterizada por uma séria desorganização dos sistemas educacionais. Porém, a educação deveria retornar às suas funções normais; as escolas deveriam ser reconstruídas. Como conseguir isto? Viu-se, aqui, a necessidade de planejar esta reconstrução, porém com maior preparo e elaboração. Vários países, na tentativa de atingir tal objetivos de forma mais rápida e efetiva, lançaram mão de novos métodos de planejamento, contudo nem sempre bem sucedidos.

Enquanto se procurava solucionar o problema de reconstrução de escolas, a economia dos vários países recuperava-se rapidamente e absorvia o total de recursos humanos especializados que existia na época. Este fato gerou um novo problema: a falta de mão-de-obra. E, à educação, coube a tarefa de solucioná-lo; ou seja, de formar mão-de-obra especializada. Para que esta nova exigência fosse cumprida, um planejamento mais cuidadoso e minucioso deveria ser feito. Paralelamente, surgia um outro problema de maior preocupação para o governo: a incapacidade do sistema educacional para atender à crescente demanda da população escolar. Mais uma vez, evidenciou-se a necessidade de planejar a educação, tendo em vista a explosão desta demanda.

Esta situação permaneceu até que surgiu uma quarta fase de pós-guerra, a atual, da inovação, objetivando a efetiva integração da educação com o seu meio-ambiente. Esta integração, constituindo-se de inovações, exige, necessariamente, refinamento no próprio planejamento educacional para que tal objetivo seja atingido.

Os países em desenvolvimento, também, enfrentaram, em grande parte, os problemas acima relacionados. Propunham-se a atingir

metas educacionais bastante otimistas, o que exigia um planejamento cuidadoso, tendo em vista a escassez de recursos humanos e materiais disponíveis. Porém, não tinham conhecimentos básicos e nem técnicos especializados em condição de planejar a educação em direção às suas novas metas.

Sérios desequilíbrios começaram a surgir no sistema. Por exemplo: quando havia número suficiente de edifícios escolares, havia carência de professores. Os recursos, que o sistema educacional possuía, não eram suficientes para atender o número de alunos que crescia cada vez mais, atraídos pelos objetivos propostos pela educação. A viabilidade econômica de tais objetivos passou a ser vista como um sério problema. Outras dificuldades de caráter administrativo, recrutamento e formação de professores, constituíam grandes limitações para o alcance das metas educacionais. À medida em que estes problemas foram sendo solucionados, outros foram surgindo. Assim, ocorreu o problema do desemprego, como consequência da excessiva produção educacional (em algumas regiões mais adiantadas dos países em desenvolvimento), porém, não em todas as categorias especializadas.

Da existência de tantos problemas, podia-se, logicamente, inferir a necessidade de formas mais adequadas de planejamento, que atentasse para todos os aspectos relevantes da educação. Começou-se a questionar sobre o tipo de educação ministrado pelas escolas. Era realmente o correto? Adaptava-se às necessidades das nações em desenvolvimento? Sentiu-se que uma solução poderia ser encontrada em um planejamento cuidadoso de um sistema educacional, que atendesse às reais necessidades do meio-ambiente e otimizasse o aproveitamento dos recursos materiais e humanos existentes. Confirmou-se, também, que as abordagens, até então utilizadas, mostravam-se ineficientes diante do novo papel assumido pela educação. Iniciou-se, assim, uma busca de novos procedimentos e técnicas para enfrentar os problemas bastante difíceis do desenvolvimento educacional.

O planejamento passou a ser visto dentro de novas e mais amplas dimensões.

c) O Novo Conceito de Planejamento Educacional

Em nossos dias, o planejamento educacional pode ser visto como um conjunto de ações, organizadas de forma consciente e racional, com o objetivo de integrar a educação no desenvolvimento nacional. Tendo em vista tal propósito, o planejamento deve ocorrer dentro de uma dimensão bastante ampla na qual todas as variáveis inter-relacionadas possam ser focalizadas e vistas como partes de um todo orgânico e dinâmico, ou seja, como um sistema. Deve-se atentar para as condições existentes na sociedade, suas necessidades e metas, para as exigências feitas à educação em relação a estas metas, para a natureza e características de sua clientela, para o conhecimento disponível, para as funções da educação e para outros aspectos importantes no contexto educacional.

Coombs (1972), afirma que as pessoas envolvidas em planejamento educacional enfrentam as seguintes questões básicas:

- quais devem ser os objetivos prioritários e funções do sistema educacional e de cada um dos seus subsistemas?
- qual a melhor abordagem para o sistema educacional?
- quanto do orçamento nacional deve ser aplicado à educação?
- até que ponto a sociedade deve financiar a educação?
- qual a alocação mais eficiente do orçamento educacional?

Estas questões mostram que a nova visão de planejamento implica em ampla abordagem do processo educacional, que é visto como um processo dinâmico:

- integrado nos planos de um desenvolvimento econômico, político e social mais amplo,
- que prevê resultados a longo, médio e curto prazo,
- abrangente,
- que utiliza meios alternativos para decidir sobre o melhor caminho a ser seguido, tendo em vista os objetivos propostos,
- que se preocupa não só com aspectos quantitativos, mas também com os qualitativos do desenvolvimento da educação.

Considerando que o planejamento deve atentar para todos os aspectos relevantes da educação e que novas técnicas, procedimentos, recursos e conhecimentos estão disponíveis, podendo ser utilizados para atingir os objetivos educacionais, o máximo de informações a respeito dos mesmos, deve ser obtido. Tais informações devem possibilitar melhores decisões a respeito dos meios alternativos identificados para o alcance dos objetivos propostos. As decisões tomadas devem conduzir a maior eficácia e eficiência do planejamento. Dentro desta perspectiva, o planejamento pode ser visto como um processo de tomada de decisão, entre vários cursos de ação direcionados para o alcance dos resultados pretendidos. Segundo Curle (1975), o planejamento é, em essência, um esforço organizado, consciente e contínuo no sentido de selecionar as melhores alternativas para se atingir objetivos particulares.

Concluindo: o planejamento como processo decisório, caracteriza-se por uma dinâmica que é ativada na medida em que os que são por ele responsáveis tomam decisões que definem exatamente:

- . o que se deseja;
- . como obter o que é desejado;
- . o que e como analisar a situação para verificar se se conseguiram os resultados desejados.

Planeja-se para que decisões mais conscientes e objetivas possam ser tomadas. E tais decisões pretendem, sempre, formas mais eficientes e econômicas de obter os resultados desejados. O planejamento do sistema educacional ocorre em vários níveis e, portanto, decisões são tomadas não só pelas altas autoridades, ministros e secretários de educação, como também pelos inspetores e diretores escolares e até mesmo pelos professores.

Existem várias abordagens que podem ser utilizadas no planejamento da educação. Uma delas é a abordagem sistêmica que se aplica aos problemas educacionais se estes se apresentam como problemas complexos, envolvendo um conjunto de variáveis inter-relacionadas, ou seja, se a educação é vista dentro da idéia do sistema.

Uma visão sumária desta abordagem será apresentada a seguir. Para melhor compreensão da mesma, serão descritos, de início, rapidamente, seus fundamentos filosóficos; e, em seguida, suas características e princípios básicos e, finalmente, sua aplicação prática.

1.1.2 - TEORIA GERAL DE SISTEMAS

a) Origem e Conceito de Teoria Geral de Sistemas

Partindo-se da época da filosofia pré-socrática e até a época atual, ver-se-á que o conhecimento humano passou por um processo de enriquecimento, que teve como consequência a autonomia de várias ciências. Estas constituíam, antes, simples ramificações da filosofia, ciência única que se preocupava com problemas globais. No momento em que este processo atingiu um alto grau de complexidade, fez-se necessário o conhecimento especializado devido a impossibilidade de uma só pessoa dominar todo o conhecimento acumulado até então.

Surgiu, assim a especialização com algumas vantagens e desvantagens tais como:

- . possibilitou o aprofundamento de conhecimentos em uma área específica da ciência;
- . possibilitou um estudo mais detalhado de todos os aspectos de um problema com base no conhecimento detalhado de cada um destes aspectos;
- . introduziu um grande risco de, ao se estudar um problema, perder-se a visão global do mesmo.
- . introduziu um problema de comunicação entre os especialistas que, adquirindo uma terminologia muito específica de sua área de conhecimento, não se faziam entender por seus colegas de outras áreas.

Considerando suas vantagens, pode-se dizer que a especialização trouxe uma grande contribuição para a tarefa de solucionar problemas. Devido à amplitude e complexidade dos problemas atuais, faz-se necessário um estudo isolado de cada uma de suas partes. Porém, o estudo independente de cada um dos seus aspectos só é significativo, na medida que resulta em informações que contribuam para a compreensão e para a solução total dos mesmos. A solução total de um problema exige que as soluções parciais, encontradas no estudo isolado das partes, sejam integradas visando solucionar o problema no seu todo. Isto evidencia que as especializações são pouco úteis, se consideradas isoladamente, e contribuem eficazmente para a solução de problemas amplos, se consideradas em um inter-relacionamento.

A partir deste ponto de vista, começou a surgir um grande interesse por estudos interdisciplinares, que mostrassem uma relação entre os vários campos de conhecimentos. Passou a existir um grande incentivo para o desenvolvimento de corpos de conhecimentos que pudessem ser organizados em um todo complexo e dentro do qual sub partes

estivessem inter-relacionadas (Johnson et al, 1967). Assim, surgiram a Psicologia Social, a Sociologia Econômica e algumas outras áreas de estudo procurando enfatizar a relação entre disciplinas, isoladas em consequência da especialização. Todo este interesse traduziu-se numa tendência que se concretizou na Teoria Geral de Sistemas. Esta teoria, uma das ciências de sistemas, procurou desenvolver um esquema teórico para descrever relações do mundo empírico (Johnson et al, 1967). Assim seu objetivo último é identificar uma estrutura que evidencie uma relação significativa entre disciplinas.

A teoria geral de sistemas foi inicialmente sistematizada por Bertalanffy em 1945. Seu esquema teórico e sua metodologia estão ainda em fase de organização. Bertalanffy propôs um método para construção de uma Teoria Geral que consiste em analisar os vários tipos de sistemas e, a partir deste estudo, deduzir princípios gerais ou comuns a todos eles; ou seja, deduzir princípios aplicáveis a muitos fenômenos complexos, seja qual for a sua natureza, social, biológica, etc..

A utilidade desta teoria se evidencia, principalmente, através de algumas das consequências da especialização, tais como:

- . problemas de comunicação que começaram a surgir entre os vários especialistas de áreas.
- . perda da visão global no estudo de problemas, para os quais uma solução deveria ser encontrada, tendo-se em vista que tal solução exigia a consideração do problema como um todo e não apenas de suas partes isoladamente.

Assim, a Teoria Geral de Sistemas pode ser vista como um conjunto de leis e princípios que procura dar uma visão geral das ciências, através da explicação de fatos comuns a todas elas. Fundamenta-se na crença de que existem semelhanças na construção teórica das várias disciplinas e procura, então, caracterizar a parte que é co

mum a elas. A Teoria Geral de Sistemas, tendo em vista a unificação das ciências, contribui principalmente para eliminar algumas consequências não desejáveis da especialização:

- procurando estabelecer uma linguagem comum entre as várias disciplinas que constituem o conhecimento humano;
- criando condições para que seja possível um estudo detalhado do planejamento e das decisões relativas a uma situação problemática. Permite que esta seja analisada de diferentes maneiras no seu todo e em cada uma de suas partes; e ainda
- contribui para evitar problemas de duplicação de esforços de correntes da crescente especialização, desenvolvendo grandes quadros de referências para sintetizar resultados de pesquisas relativos às várias ciências.

b) Alguns Conceitos de Sistema

A teoria geral do sistemas pode, então, ser vista como um produto do desenvolvimento do conhecimento humano ou, mais precisamente, da especialização. O desenvolvimento do conhecimento humano, por sua vez, está relacionado com os conceitos de sistemas que foram sendo adotados no decorrer dos anos. Assim falar-se-á rapidamente sobre sistemas.

Um sistema é uma reunião ou combinação de coisas ou partes, formando um todo complexo e unitário (Johnson et al, 1967).

Um sistema é um conjunto de partes coordenadas para realizar um conjunto de finalidades (Churchman, 1972).

Ou pode-se ainda, definir sistema como um todo formado de partes interdependentes que, interagindo continuamente entre si e com o todo, executam suas funções visando um objetivo último comum a

todas elas. As partes interagem de acordo com um plano ou princípio previamente identificado.

Em se tratando de sistema educacional, encontram-se também várias definições.

Gagné (1973) define sistema educacional como um conjunto de recursos humanos e materiais organizados e interagindo, com o propósito de provocar mudanças duradouras no comportamento do indivíduo, mudanças estas que são atribuídas ao processo de aprendizagem.

Brejon (1974) diferencia entre Sistema de Educação, Sistema de Ensino e Sistema Escolar. Identifica Sistema de Educação com sociedade. Assim Sistema Educacional incluiria todas as instituições sociais que educam de maneira formal ou informal tais como: família, associações recreativas, empresas, escolas. O sistema de ensino seria um sistema menos amplo, incluindo escolas e outras pessoas, ou mesmo instituições que transmitam educação sistemática. E o sistema escolar que se constituiria de uma rede de escolas e sua estrutura de sustentação (estrutura administrativa do sistema escolar, abrangendo normas, metodologia e conteúdo de ensino; entidades mantenedoras e os organismos que tem, por finalidade a administração do sistema).

Segundo Banathy (1968) sistemas são reuniões de partes (ou elementos) que são projetadas e construídas pelo homem dentro de papéis organizados, para o alcance de objetivos específicos. O propósito de um sistema é realizado através do processo no qual os seus componentes interagem para produzir um determinado resultado.

c) Caracterização de um Sistema

Basicamente, um sistema se caracteriza pela interdependência de suas partes e pela natureza das relações que se verificam entre elas. Um sistema é também caracterizado em função de quatro

elementos básicos: entradas, conteúdo, saídas e realimentação.

- Entradas: todos os recursos necessários para o desenvolvimento do sistema.
- Conteúdo: refere-se a todos os procedimentos e atividades que, na operacionalização do sistema, possibilitam que as entradas sejam transformadas nos produtos desejados. Abrange todos os componentes materiais e humanos identificados e selecionados como componentes necessários para efetuar o processo. Este inclui todas as ações específicas que permitem a dinamização do sistema. Implica em uma série de passos sucessivos, claramente definidos e formalizados por lei.
- Saídas: São todos os resultados esperados e identificados para o sistema. É exatamente o que se pretende.
- Realimentação: As comparações entre os produtos desejáveis e os resultados podem conduzir a modificações no sistema. Estas comparações e modificações implicam no processo de realimentação.

Apenas como exemplo, caracterizando um sistema de ensino/aprendizagem através destes quatro elementos (que constituem apenas uma das várias maneiras de se ver um sistema) tem-se:

- Entrada: recursos financeiros, alunos, conteúdo cultural, filosofia educacional expressa na legislação escolar.
- Conteúdo: professores, pessoal administrativo, técnicos em educação, método de ensino, materiais didáticos como livros, mapas, recursos tecnológicos, conhecimentos. O processo seria o funcionamento do

sistema através da execução dos procedimentos de ensino e da utilização de todos os recursos selecionados para o alcance dos objetivos. Ou seja, seriam os professores, conteúdo cultural, recursos tecnológicos, métodos de ensino, interagindo de maneira a produzir modificações no comportamento do estudante que está sendo processado. Tais estudantes sairiam do sistema com diferentes tipos de educação e treinamento.

- Saídas: pessoal qualificado em vários níveis, novas descobertas científicas e, conseqüentemente, o enriquecimento do conhecimento, o aumento do nível cultural da sociedade.
- Realimentação: o processo do sistema é constantemente avaliado em seus recursos, procedimentos, relações e interações entre eles, assim como os resultados destas interações. As informações obtidas são utilizadas para realimentar o sistema, modificar, se necessário as suas entradas ou seu processo.

Estes quatro elementos constituem um modelo básico a través do qual um sistema pode ser representado. O esquema deste modelo seria o seguinte:

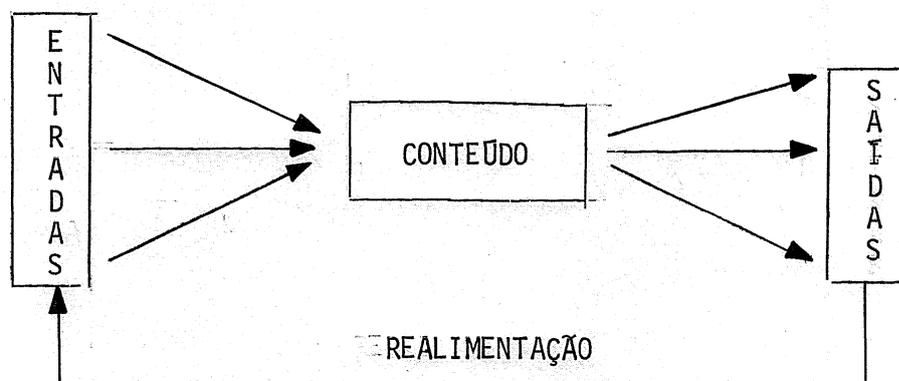


Fig. F.1 - Caracterização de um Sistema

É este um modelo sistêmico. Dentro deste enfoque, são consideradas todas as partes do sistema; são estudadas todas as relações e interações das partes entre si; e entre elas e o todo. As atividades de cada componente são relacionadas com o rendimento total do sistema.

Existem vários modelos através dos quais os sistemas podem ser representados. Utilizando modelos, poder-se-á fazer experimentos e estudos de situações que, em circunstâncias reais, não seriam possíveis.

Muitas outras maneiras de se caracterizar um sistema existem.

Churchman (1972) diz, por exemplo, que uma das maneiras de se pensar a respeito de um sistema é ter em mente:

1. objetivos e medidas de rendimento do sistema total;
2. ambiente do sistema;
3. recursos do sistema;
4. componentes do sistema, suas atividades e medidas de rendimento;
5. administração do sistema.

Um exemplo, incluindo estes elementos, é apresentado a seguir:

Objetivo: Propiciar a formação necessária, para a habilitação profissional de 60% dos estudantes na faixa etária entre 16 e 18 anos de idade.

Ambiente do Sistema: Tempo previsto para o alcance do objetivo proposto, restrições financeiras (limite de verba), políticas (Filosofia educacional) e sociais.

Recursos do Sistema: Verba disponível para processar o sistema, edifícios escolares, bibliotecas, nº de professores qualificados, administradores, técnicos educacionais, recursos tecnológicos.

Componentes: Conjunto de atividades pedagógicas e administrativas (executadas pelo grupo de professores, alunos, diretor, pessoal de apoio).

Medida de Rendimento: Será feita através de testes de aprendizagem para medir o grau de atingimento dos objetivos. O indicador da medida de rendimento seria a percentagem de 60% expressa no objetivo.

Administração do Sistema: Tomar decisões quanto ao nº de professores a serem contratados, tipos e quantidade de equipamentos, carga horária, nº de alunos que serão atendidos pelo sistema, tipo e quantidade de materiais didáticos, recursos tecnológicos tendo em vista as limitações do sistema.

Assim:

- os objetivos de um sistema se traduziriam naquilo que se de seja conseguir;
- o ambiente do sistema inclui tudo aquilo que está situado fora de seus limites e está, também, fora de seu controle. Como diz Churchman (1972), são todas as coisas e pessoas "fixadas" ou "dadas" do ponto de vista do sistema.

- os recursos de um sistema referem-se a tudo aquilo que se encontra dentro de seus limites e que possibilita a realização de suas funções.
- os componentes de um sistema referem-se às suas partes, ou subsistemas. São identificados pelas suas atividades.
- a administração implica na elaboração de planos para o sistema, considerando seus objetivos, ambiente, recursos e componentes. Controla a execução do planejamento e o rendimento do sistema.

1.1.3 - A ABORDAGEM DE SISTEMAS

a) Conceito

A abordagem sistêmica é uma metodologia que conduz, logicamente, o pensamento através de uma análise formal e global do problema em estudo utilizando-se das premissas básicas da Teoria Geral de Sistemas. É aplicada quando a situação problemática se caracteriza por uma certa complexidade; ou seja, se o problema é crítico e complexo.

b) Características Básicas

A abordagem sistêmica fornece uma direção segura e eficiente para o planejamento e desenvolvimento de sistemas. Através de sua aplicação, chegar-se-á à identificação de uma solução que conduz aos objetivos propostos dentro das limitações previstas. Tal abordagem apresenta as seguintes características, que refletem suas idéias básicas:

- preocupa-se com a otimização do todo e não somente das partes consideradas isoladamente, pois, nem sempre, a otimização das partes, somente, conduz à otimização do todo;

- identifica a necessidade de se utilizar equipes interdisciplinares para resolver um determinado problema, criar ou otimizar um sistema considerando-os como um todo;
- busca uma linguagem comum às diversas disciplinas, que constituem o conhecimento humano, para que possa existir a necessária comunicação entre os especialistas;
- faz amplo uso de modelos;
- utiliza-se do método dedutivo;
- possui um processo de avaliação contínuo, ou seja, na definição do sistema, meio-ambiente e nas suas inter-relações. São avaliados, a cada passo, os resultados obtidos. Esta avaliação objetiva o aumento da eficácia e eficiência no planejamento e desenvolvimento do sistema.

Concluindo: O enfoque sistêmico, visando a otimização do sistema como um todo, procura, sempre, considerar todas as relações e interações existentes entre os seus vários elementos e entre os objetivos de cada um deles, Enfatiza, sobretudo, a integração de todas as partes. Para tanto, coloca a necessidade de se ter um conhecimento de cada uma das áreas que se inter-relacionam dentro do sistema total. Este conhecimento é obtido através de uma equipe interdisciplinar que procura estudar cada um dos vários aspectos do problema sem, entretanto, perder a visão global do mesmo.

c) Etapas Principais

A utilização da abordagem sistêmica implica em:

- . obter uma definição exata do problema;
- . determinar uma solução para o mesmo.

A estruturação de sistemas (para solucionar um problema) deve conter uma resposta a três questões básicas:

- . o que conseguir?
- . como conseguir?
- . quando conseguir?

Assim, na estruturação de um sistema devem ser claramente descritos:

- 1) O que se quer conseguir, ou seja, os objetivos do sistema;
- 2) Todas as características que o sistema deve apresentar para que seja possível a realização dos seus objetivos, ou seja, todos os seus requisitos;
- 3) As especificações dos requisitos do sistema e outras informações relevantes para sua definição;
- 4) Todas as fases e grupos de trabalho envolvidos no sistema;
- 5) Uma sequência de passos a serem executados para a consecução dos objetivos pretendidos. Neste fluxo, devem estar, também, claramente identificadas as inter-relações entre os diferentes grupos de trabalho;
- 6) A duração de cada tarefa e as datas em que os resultados devem ser apresentados.

A definição do sistema deve ser bastante trabalhada, chegando-se a um nível de detalhamento que possibilite uma visão bastante compreensível e abrangente do problema considerado.

Uma definição com tais características conduzirá, mais facilmente, à determinação de uma solução diretamente relacionada com os objetivos especificados.

Para a identificação desta solução, várias alternativas são analisadas e é selecionada a melhor dentre todas, de acordo com critérios pré-estabelecidos. Para a escolha da melhor alternativa, será utilizado um conjunto de técnicas que constitui o processo de Análise de Sistemas. A metodologia da abordagem de sistemas conta, inicialmente, com um mínimo de informações (os objetivos), que aumenta à medida em que o problema é claramente definido e descreve a solução encontrada.

d) Áreas de Aplicação da Abordagem de Sistemas

A abordagem de sistemas tem sido aplicada em muitos campos do conhecimento: econômico, administrativo, social, educacional.

No campo educacional ela tem sido utilizada com o objetivo de ampliar e renovar as perspectivas relacionadas ao processo de ensino/aprendizagem. Tem sido também muito aplicada ao planejamento de sistemas, sobretudo administrativos.

Este enfoque, aplicado ao planejamento pode ser considerado como um método logicamente consistente, objetivando reduzir uma grande parte da complexidade do problema.

Ele pode ser utilizado pelo decisor em combinação com outras considerações, para se chegar a melhores decisões (King, 1968).

Se utilizado no planejamento de sistemas, o enfoque sistêmico deve ser visto como um processo que visa a tomada de decisões.

Decisões são sempre tomadas se soluções devem ser encontradas para problemas críticos e complexos. Tais decisões implicam na busca ou alcance de objetivos pretendidos.

A consecução dos objetivos requer um processo de transformação do sistema: de um estado real existente a um estado ideal previsto. Para a modificação da realidade, realidade com a qual não se está satisfeito necessita-se concebê-la como um sistema e organizá-la como tal (Turra et al, 1975).

O enfoque sistêmico se opõe, assim, à maneira de descrever problemas com base em observações e informações de outras pessoas. Tal enfoque dinamiza os processos mentais, colocando-os em ação a partir do primeiro momento em que se começa a descrever o que se está planejando.

1.2 - SUB-UNIDADE 2: O PROCESSO DE ENGENHARIA DE SISTEMAS

a) Conceito

A abordagem de sistemas é também aplicada ao planejamento e execução de projetos. Neste caso ela é operacionalizada através de um conjunto de passos sequenciais, bem definidos, que constitui o processo de Engenharia de Sistemas. Este trata do detalhamento e integração de todas as partes de um sistema. Cuida da sua criação, modificação, implementação, análise e avaliação.

b) Idéias Básicas

O processo de Engenharia de Sistemas fundamenta-se em uma filosofia de planejamento cuja idéia básica refere-se à sequência com que duas questões são colocadas ao se planejar e executar um sistema.

Estas questões são:

- O que?
- Como?

O "o que?" refere-se a especificação detalhada e exata daquilo que se quer conseguir. Assim, se pretende um curso de treinamento, é preciso pensar em tudo aquilo que o caracteriza em termos de sua operação.

"O que?" implica em pensar no curso em processo, realizando atividades e utilizando recursos, ou seja, é necessário pensar no curso em funcionamento, em operação.

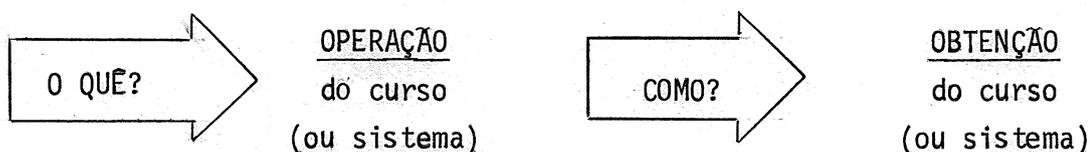
O "Como?" refere-se a especificação de atividades e recursos necessários para obter ou construir aquilo que se pretende.

"Como?" implica em pensar em tudo que é necessário para colocar, o curso, em operação.

Assim, ao planejar um curso, deve-se, primeiro, caracterizá-lo detalhadamente, em termos de suas entradas, saídas e atividades. Pensar exatamente, no curso operando; ou seja, no "o que?"

Em seguida deve-se pensar em como obter tudo que seria necessário para sua construção; ou seja no "Como?".

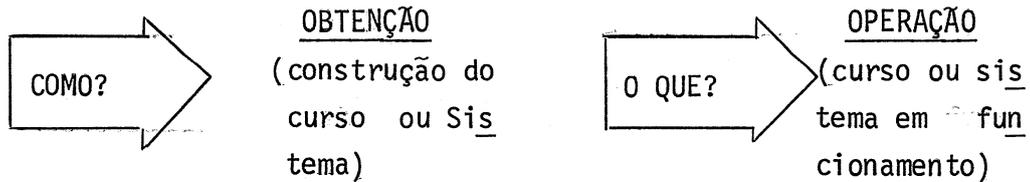
No Planejamento do Curso tem-se então a seguinte sequência de passos:



Depois de planejada a Operação e a Obtenção do sistema, passa-se para a etapa seguinte: a Execução do Curso. Na execução, a sequência de passos ("o que?" = operação e "como?" = obtenção) se inverte. Primeiro obtém-se o curso e depois ele é colocado

em operação; ou seja, primeiro pensa-se no "Como?" e depois no "O Que?"

Assim na Execução do curso tem-se a sequência:



Na etapa de execução, o que foi planejado é então concretizado. São realizadas as diversas atividades e consumidos os recursos para construir ou obter o curso. Construído ou obtido, ele entrará em operação.

A utilização deste processo leva à identificação e separação de atividades e recursos para a obtenção e operação do sistema nas etapas de Planejamento e Execução.

Em resumo tem-se:

P L A N E J A M E N T O	1) <u>OPERAÇÃO</u> (O QUE?) (o que é o curso ou sistema operando)	E X E C U Ç Ã O	1) <u>OBTENÇÃO</u> (COMO?) (construção do curso ou sistema)
	2) <u>OBTENÇÃO</u> (COMO?) (como será construído o curso ou sistema)		2) <u>OPERAÇÃO</u> (O QUE?) (curso ou sistema em funcionamento ou operando)

Fig. F.2 - Sequência de Passos no Planejamento e Execução de um Sistema

O processo de Engenharia de Sistemas abrange um conjunto de atividades de caráter técnico e administrativo. Assim sendo, basicamente, dois grupos devem existir dentro deste processo: um grupo denominado Grupo de Controle Administrativo (GCA) que será o responsável pelas atividades de caráter administrativo e o outro denominado Grupo de Engenharia de Sistemas (GES) que será o responsável pelas atividades de planejamento do ponto de vista técnico.

Mais especificamente o GCA, exerce uma função de assessoria administrativa ao gerente do sistema. É responsável pelo planejamento e controle do sistema em relação a tempo, custo e documentação. Suas tarefas referem-se à definição do Quando (tempo) e do Quanto (custo). É ele também que, com informações do GES, elabora a estrutura organizacional do sistema, que define níveis de autoridade e responsabilidade.

O grupo de Engenharia de Sistemas é o responsável pelo "O QUE?" e pelo "COMO" do sistema. Executa as tarefas necessárias para caracterizar o sistema (identifica suas entradas, saídas e atividades), cuida da integração de todas as suas partes e avaliação do mesmo. Procura ter uma visão geral do problema.

c) Etapas Principais

O processo de planejamento de Engenharia de Sistemas (Reis et al, 1976) compreende as seguintes etapas:

- 1º) Identificação da situação;
- 2º) Planejamento Preliminar;
- 3º) Planejamento Final;
- 4º) Execução do Projeto;
- 5º) Teste e Operação do Sistema;
- 6º) Controle (acompanhamento, avaliação e realimentação).

Será dada, em seguida, uma visão bastante geral sobre cada uma destas etapas mostrando como a filosofia de planejamento é utilizada dentro do processo de Engenharia de Sistemas.

19) Identificação da Situação

Inicialmente é formada uma equipe interdisciplinar, incluindo especialistas em áreas de conhecimentos diretamente relacionadas com o problema em estudo.

Este grupo é o chamado Grupo de Engenharia de Sistemas, que dará início ao trabalho, coletando e fornecendo informações relevantes, que possibilitam uma definição clara e precisa do problema, uma análise do meio-ambiente do sistema e a limitação de sua área de atuação. Estas atividades são básicas neste processo pois nelas se fundamentam todas as etapas seguintes. Devem, portanto, ser executadas por verdadeiros especialistas no assunto.

Ainda nesta etapa inicial, são delineados as primeiras alternativas para a solução do problema.

Por exemplo, tem-se um problema que é a carência de mão-de-obra especializada para a indústria. Para resolvê-lo, serão necessárias, inicialmente, informações mais detalhadas sobre ele. O problema deve então ser estudado, por uma equipe de especialistas, dentro das necessidades reais da situação que se pretende definir.

Poder-se-ia ter uma equipe formada por administradores, pedagogos, economistas que, coletariam informações sobre oferta e demanda da mão de obra especializada para a indústria, recursos e procedimentos didáticos utilizados na formação destes profissionais, nível de qualificação dos trabalhadores neste campo, número de empresas industriais que contribuem financeiramente para formação destes indivíduos, legislação escolar que dispõe sobre o ensino profissionalizante.

Depois que o problema foi precisamente caracterizado, algumas alternativas de solução poderiam ser identificadas, como por exemplo:

- Ministrar cursos rápidos de treinamento de mão-de-obra especializada para a indústria, utilizando recursos tecnológicos.

Após a definição clara e precisa do problema é a identificação de possíveis alternativas de solução, inicia-se a segunda etapa do processo: planejamento preliminar.

2º) Planejamento Preliminar

Nesta etapa cada uma das alternativas de solução, a apresentadas, é estudada detalhadamente. No estudo de cada uma delas é feito um planejamento onde são identificados seus resultados (ou saídas), seu meio-ambiente, sua operação e obtenção. Isto é feito através de três fases sequenciais:

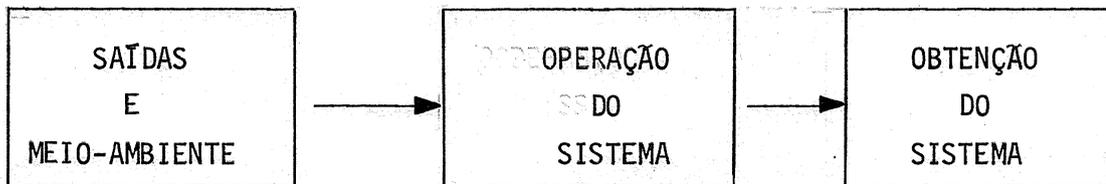


Fig. F.3 - Fases Sequenciais do Planejamento Preliminar

1ª FASE: Saídas e Meio-Ambiente

Nesta fase as SAÍDAS ou RESULTADOS são identificadas através da definição de objetivos para o sistema. São também apresentadas as medidas de rendimento que permitem avaliar os resultados obtidos.

Os objetivos devem ser definidos de forma clara e precisa para possibilitar a avaliação das saídas ou resultados do sistema. Isto é possível pela definição das medidas de rendimento, em termos quantitativos, para expressar seu caráter qualitativo. Para que o objetivo geral do sistema fique bastante claro e possa assim ser avaliado, ele deve ser detalhado em objetivos específicos.

Uma das maneiras de se identificar objetivos específicos é relacionar os componentes (ou partes) do sistema e definir objetivos para cada um deles. Estaria sendo utilizada, aqui, a abordagem sistêmica onde os resultados obtidos pelos componentes são inter-relacionados e contribuem para o alcance do objetivo geral do sistema. Assim poder-se-ia ter como objetivo geral de um curso de treinamento:

- Habilitar, para o trabalho industrial, no período máximo de dois anos, 50% dos alunos em nível de 2º grau.

E, um objetivo específico para o subsistema de ensino/aprendizagem seria:

- Desenvolver, em 90% dos alunos, as habilidades necessárias para sua qualificação profissional, através do processo de ensino/aprendizagem.

É importante observar que os objetivos específicos podem ser quebrados em outros que apresentam maior nível de detalhes, facilitando, assim, uma melhor identificação e avaliação dos resultados pretendidos.

Todos os objetivos e medidas de rendimento do sistema e de cada um dos seus componentes devem ser então identificados clara e precisamente. (No objetivo acima exemplificado as medidas de rendimento seria o dado percentual 90%).

O AMBIENTE refere-se a todas as influências do meio, sobre o sistema, que não são possíveis de serem controladas. Estas influências, também denominadas RESTRIÇÕES, interferem, no funcionamento, na obtenção e nos resultados do sistema. No alcance dos objetivos, tais restrições devem ser respeitadas; deve-se dar a elas grande atenção. Devem ser identificadas tanto para a Operação como para a Obtenção do sistema. Assim, por exemplo, a Legislação Escolar referente à formação e habilitação profissional de indivíduos poderia ser uma restrição para o funcionamento ou operação de um curso (ou sistema) profissionalizante. Na sua operação, o curso deve obedecer as normas prescritas na legislação. A verba de que dispõe o sistema seria uma restrição para sua obtenção. As dimensões do edifício escolar, o número de pessoas a serem contratados, o número de alunos a serem matriculados no curso, teria uma relação direta com a quantidade de verba disponível.

A determinação precisa das restrições depende da quantidade de informações que se tem do ambiente. O máximo de informações sobre o meio permitirá ao planejador determinar, precisamente, quais são estas influências (ou restrições) e até que ponto elas são fixas ou possíveis de serem alteradas.

É importante também que sejam identificadas, separadamente, aquelas restrições referentes à operação e à obtenção do sistema, nas etapas de planejamento e execução.

As restrições de um sistema dependem da sua natureza. Assim algumas restrições mais comuns aos sistemas sociais seriam: políticas, econômicas, sociais, financeiras, geográficas.

A identificação de objetivos gerais e específicos e de restrições são dois pontos bastante importantes na caracterização de um sistema. Portanto é uma tarefa que deve ser cuidado-

samente realizada.

2ª FASE: Operação do Sistema

Para entrar em operação, o sistema necessita de um conjunto de atividades e recursos que, processados, são transformados no produto desejado (Saídas do Sistema).

Estas atividades são realizadas pelas partes que compõem o sistema total.

Assim, nesta fase de operação, devem ser identificados todos os componentes (ou partes) do sistema; as atividades e recursos necessários para sua realização (ou operação).

Estas atividades e recursos são também descritos detalhadamente; ou seja, são bem especificados possibilitando uma maior compreensão dos mesmos.

Considerando um curso (ou sistema) de treinamento de mão-de-obra especializada, poder-se-ia, por exemplo, identificar para sua Operação atividades como:

- transmitir instrução;
- avaliar aprendizagem;
- coordenar reuniões pedagógicas;
- controlar o cumprimento das normas fixadas pela Legislação escolar.

E, como recursos para sua operação:

- livros técnicos;
- professores;
- administrador;
- maquinários.

Pode surgir, aqui, a necessidade de se construir um fluxo de atividades que visualize a sequência da operação do sistema. Pode ainda ser estimado o custo e o tempo de operação, o que possibilita uma posterior análise das alternativas de solução e a escolha da melhor dentre elas.

Em resumo, tem-se:

OPERAÇÃO DO SISTEMA
- identificação e especificação de: <ul style="list-style-type: none">. recursos. atividades
- elaboração de um fluxo de atividades
- estimativa de tempo e custo

Fig. F.4 - Atividades da Fase de Operação do Sistema

Relembrando as idéias básicas ("O QUE?" e o "COMO?") da filosofia de planejamento, utilizada pelo processo de Engenharia de Sistemas, é importante destacar, aqui, sua aplicação.

Na 1ª e 2ª Fases do Planejamento Preliminar (Saídas, Meio-Ambiente e Operação do Sistema) a primeira tarefa é definir o "O QUE?"; ou seja caracterizar exatamente o sistema em termos de suas atividades, recursos, funcionamento e saídas.

Para isto é feito um planejamento da Operação do sistema identificando-se os resultados (ou saídas) esperados, os recursos e atividades necessários para seu funcionamento. É então definido o "o que?" ou planejada a "Operação" do sistema.

Em seguida é planejado o "Como?", ou seja a Obtenção do Sistema.

3ª FASE: Obtenção do Sistema

Nesta fase é planejada a construção do sistema através da identificação das atividades e recursos necessários a sua obtenção.

Os recursos e atividades são descritos detalhadamente e um fluxo de atividades é elaborado para visualizar a sequência que deve ser considerada para a obtenção do sistema.

Por exemplo, poder-se-ia identificar como atividades para a obtenção de um curso (ou sistema) de treinamento:

- estudar a situação existente;
- estabelecer convênios financeiros com indústrias;
- contratar pessoal qualificado para o magistério.

E, como recursos:

- maquinários;
- livros;
- pessoal de apoio: secretário, escriturário.

Para se visualizar todo o trabalho a ser realizado nesta etapa, (Planejamento Preliminar) deve ser elaborado um único fluxo mostrando a sequência da operação e obtenção do sistema, suas interrelações e realimentações. São também estimados

nesta 3ª Fase (de Obtenção) o tempo e o custo para a construção ou obtenção do sistema.

Em resumo, tem-se:

OBTENÇÃO DO SISTEMA
- identificação e especificações de: <ul style="list-style-type: none">. recursos. atividades
- elaboração de fluxo de atividades
- estimativa de tempo e custo

Fig. F.5 - Atividades da Fase de Obtenção do Sistema

Identificados os recursos e atividades para a construção do curso tem-se definido o "Como?" ou planejado a "Obtenção" do sistema.

Para a identificação e especificações de atividades e recursos, elaboração de fluxo de atividades e estimativa de tempo para a operação e obtenção do sistema, são utilizadas várias técnicas de planejamento, tais como: Diagrama de fluxo de trabalho (DFT), cronogramas mestres e parciais, PERT, Estrutura de Divisão de Atividades (EDA), Estrutura de Divisão de Recursos (EDR).

Depois que todas as alternativas foram planejadas detalhadamente em relação às suas entradas, saídas e atividades, elas serão estudadas utilizando-se a metodologia de Análise de Sistemas.

Através deste processo, chega-se à identificação de uma alternativa de solução, considerada a melhor de acordo com critérios estabelecidos. É, ele, um processo lógico que possibilita a tomada de decisão relativa ao problema considerado.

A metodologia de Análise de Sistemas constitui-se, portanto, num processo que ocorre depois que todas as possíveis alternativas de solução foram identificadas e planejadas. A aplicação desta metodologia é vista como uma etapa independente do processo de Engenharia de Sistemas, mas é utilizada dentro dele em um determinado momento; ou seja no momento da tomada de decisão relativa a situação problemática. Seus principais passos são apresentados na Figura F.6.

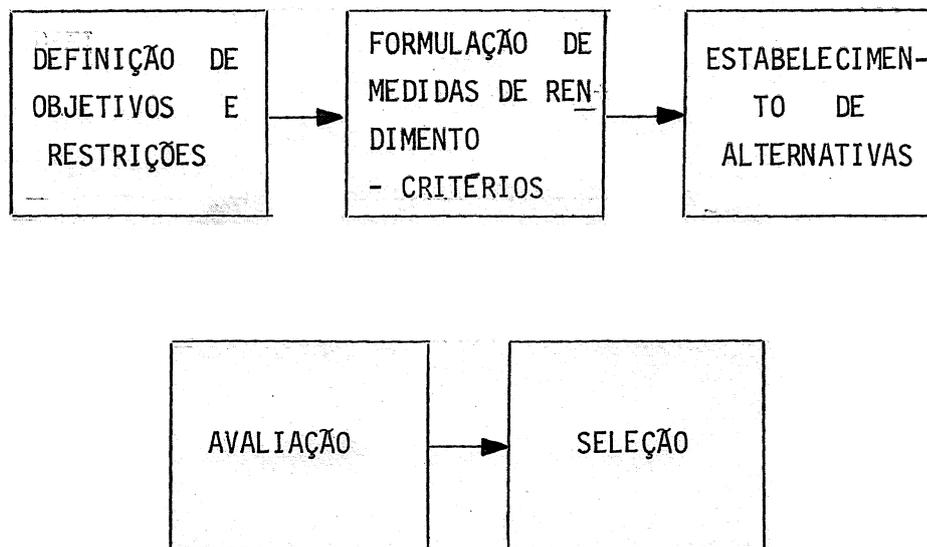


Fig. F.6 - Passos Gerais do Processo de Análise de Sistemas

Para a seleção da melhor alternativa, a Análise de Sistemas utiliza uma variedade de técnicas como por exemplo: programação matemática, teoria das filas, teoria dos jogos, análise de decisões.

Os critérios de decisão, que orientarão a escolha da alternativa ótima, podem ser, entre outros, a efetividade, custo-tempo, custo/efetividade.

Assim, se o critério especificado foi custo/efetividade, a decisão será a favor daquela alternativa que tiver a menor relação custo/efetividade.

Selecionada a alternativa ótima, através da metodologia de Análise de Sistemas, o processo de Engenharia de Sistemas têm, então, continuidade iniciando-se a sua próxima etapa, o Planejamento Final.

39) Planejamento Final

Nesta etapa do processo, a alternativa escolhida é estudada detalhadamente visando o seu planejamento final.

Assim, são novamente estudados os objetivos e restrições identificados anteriormente, são revistas e apresentadas, de maneira mais detalhada, todas as especificações de recursos e atividades assim como os fluxos de obtenção e operação do sistema. É elaborado um fluxo visualizando todas as atividades e a sequência em que elas devem ser realizadas, estimado, separadamente, o tempo das atividades e o custo dos recursos das fases de obtenção e operação. É ainda estimada uma relação entre tempo e custo.

Depois de bem detalhadas ou bem especificadas, as atividades e recursos para o alcance dos objetivos pretendidos, são identificadas as responsabilidades, determinando-se quem é responsável por o quê, e o relacionamento entre os componentes do sistema to-

tal. As pessoas e recursos são assim colocados em uma estrutura de nominada Estrutura Organizacional do Sistema.

Nesta etapa, são ainda planejadas atividades de preparação de equipamentos e instalações para a execução do sistema.

O planejamento final implica em uma repetição de atividades já realizadas no planejamento preliminar. Tal repetição visa fornecer informações em um nível de detalhes bem maior, possibilitando assim a execução do plano, ou ainda, em casos extremos, concluir pela não viabilidade da alternativa escolhida.

A necessidade de se descer ao maior nível de detalhamento possível no planejamento da alternativa selecionada, se justifica, também, pelo fato de que nem sempre quem planeja é quem executa.

Terminou, aqui, o conjunto de etapas do Planejamento (Identificação da Situação, Planejamento Preliminar e Planejamento Final) o qual referiu-se tanto ao planejamento da operação como da obtenção do sistema.

Resumindo, nesta etapa tem-se:

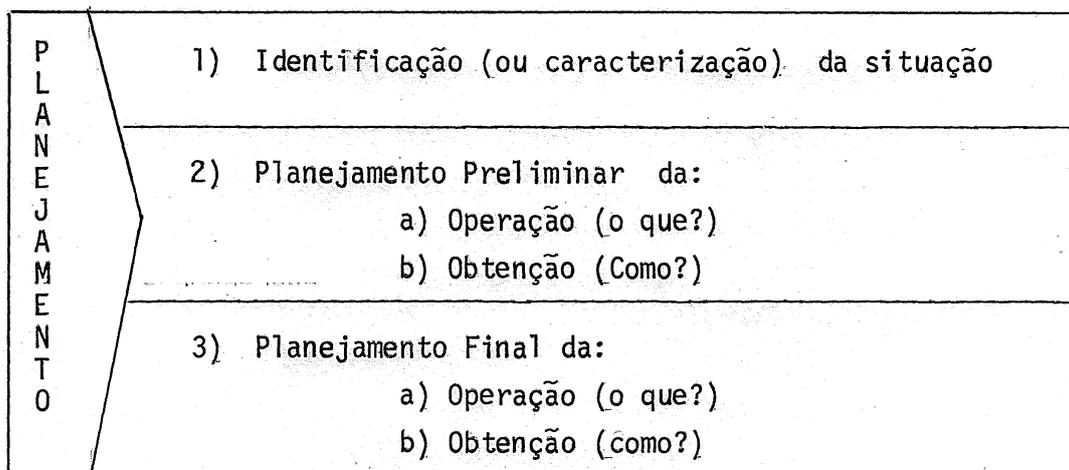


Fig. F.7 Primeiro Conjunto de Etapas do Processo de Engenharia de Sistemas: Planejamento

A Engenharia de Sistemas abrange dois conjuntos de etapas: 1) Planejamento e 2) Execução.

Findo o planejamento do sistema, passa-se para a Execução (segundo conjunto de etapas) que inclui: execução; teste e operação; e controle do sistema.

Neste segundo conjunto de etapas (Execução) destaca-se novamente a aplicação da filosofia de planejamento. Cuida-se primeiramente do "Como?", ou seja constrói-se o sistema e depois do "o que?", coloca-se o sistema em operação.

49) Execução do Projeto

São aqui realizadas todas as atividades para a obtenção ou construção do sistema, adquiridos, instalados e testados os recursos necessários ao seu posterior funcionamento ou operação.

Nesta etapa é também introduzido, no processo, o grupo denominado Grupo de Controle Administrativo (GCA). A inclusão deste Grupo se justifica principalmente pela quantidade de tarefas administrativas que já se apresentam nesta etapa do processo. Este grupo trabalha junto com o grupo de Engenharia de Sistemas e, como já foi visto, tem responsabilidade de assessoria ao gerente do sistema; de definição de custos e tempo e de elaboração da estrutura organizacional do sistema.

O CGA, em seu trabalho, utiliza várias técnicas, tais como PERT/CPM, cronogramas mestre e parcial, cronograma de desembolso, matriz de custos, fluxogramas.

Nesta fase são então executadas as atividades e obtidos os recursos planejados para a construção do sistema. Cuidou-se aqui do "como?" ou da "Obtenção" do sistema.

As etapas seguintes serão referentes a "operação" do sistema, ou seja ao "o que?".

59) Teste e Operação do Sistema

É a etapa de ensaio do sistema.

Nem sempre o sistema total é colocado em operação. As vezes é necessário que sejam testados apenas os pontos mais relevantes do mesmo.

Se o sistema não é muito complexo, a sua operação de= definitiva já pode ser iniciada.

São aqui executadas as principais (ou se possível e necessário, todas) atividades de operação do sistema para detetar pos síveis falhas no seu funcionamento antes que o sistema total entre em operação definitiva.

Para testar o funcionamento de sistemas, a Engenharia de Sistemas utiliza modelos (uma representação ou abstração simplifi cada da realidade, considerando suas variáveis mais importantes) que permitem experimentar e estudar situações que não podem ser testa das na realidade.

Existem vários tipos de modelos que são adequados ou não as várias situações. Na escolha de um modelo é necessário que se considere a natureza da situação em estudo. Dependendo do tipo de es tudo haverá um modelo mais adequado que deverá ser o selecionado.

60) Controle

O controle apesar de ter sido apresentado como a 6ª e- etapa deste processo, é uma atividade contínua, ocorrendo em todos os momentos.

O controle abrange atividades de acompanhamento, avaliação e realimentação e é feito em relação a tempo, custo e pessoal. Tais atividades objetivam a melhoria do trabalho e conseqüentemente a maior eficácia e eficiência de planejamento e execução.

Esquematisando as informações apresentadas na 4^a, 5^a, e 6^a etapas tem-se:

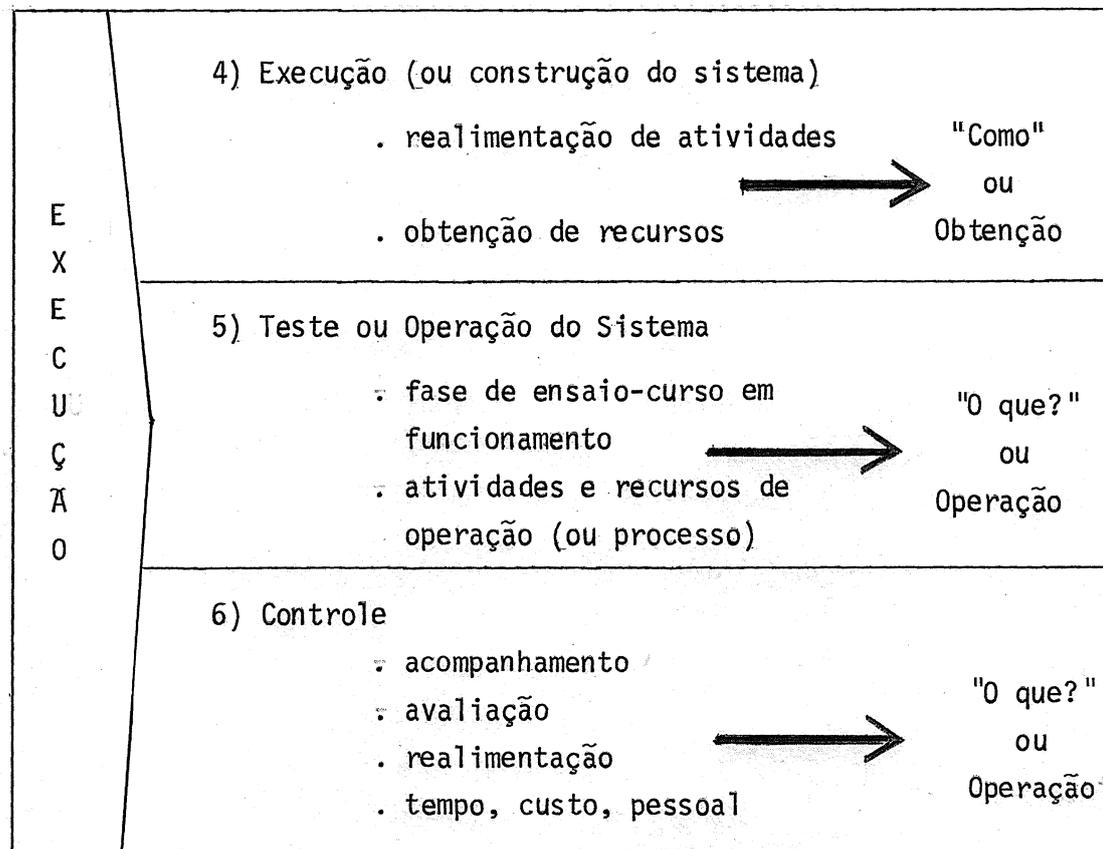


Fig. F.8 - Segundo Conjunto de Etapas do Processo de Engenharia de Sistemas

A figura F.9 foi apresentada por Reis et al (1976) e visualiza o que se acabou de informar sobre o processo de Engenharia de Sistemas.

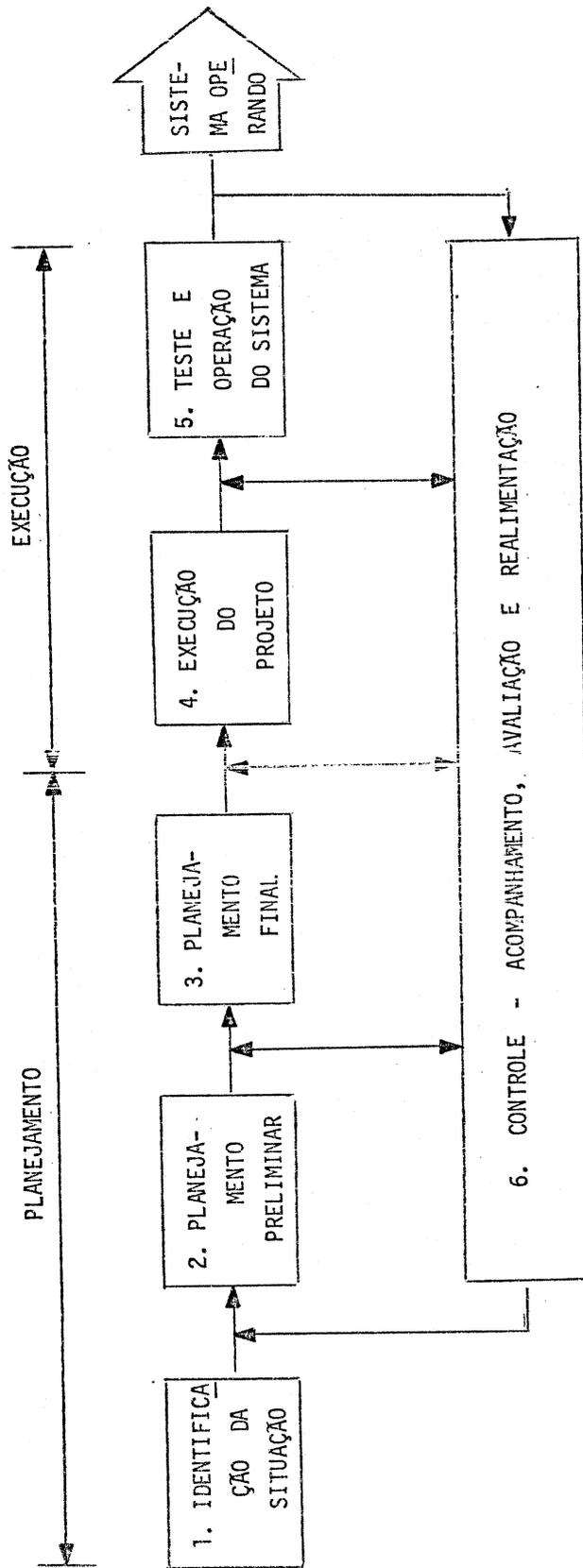


Fig. F.9 - Sequência de Etapas no Processo de Engenharia de Sistemas

1.3. - SUB-UNIDADE 3: A AVALIAÇÃO EDUCACIONAL DENTRO DE UMA ABORDAGEM SISTÊMICA

Foi visto que o Planejamento Educacional, dentro de sua nova abordagem é visto como sendo, essencialmente, um processo de tomada de decisão. Tal processo deve gerar decisões de boa qualidade que permitam alcançar, satisfatoriamente, os objetivos pretendidos. Para tanto são necessárias informações úteis que conduzam ao aprimoramento destas decisões. O fornecimento destas informações é função básica da Avaliação que pretende, assim, facilitar o processo decisório de tal forma que melhores decisões possam ser tomadas.

Antes de se abordar diretamente a Avaliação Educacional, serão descritos alguns aspectos do processo de tomada de decisão, tendo em vista sua relação com os estudos de Avaliação.

1.3.1 - O PROCESSO DE TOMADA DE DECISÃO EDUCACIONAL

Não se tem, aqui, a pretensão de descrever este processo com toda a complexidade que o caracteriza. Será apresentada apenas uma rápida visão de seus pontos básicos:

a) O Processo Decisório em Quatro Etapas

O processo decisório se apoia em quatro etapas maiores que são dependentes e ocorrem na mesma sequência em que serão apresentadas a seguir. São elas:

1. Conscientização do problema ou necessidade.
2. "Design" ou especificação de planos de ação.
3. Escolha de um plano de ação.
4. Ação.

1ª ETAPA: Conscientização do Problema ou Necessidade

Este 1º Passo, do processo decisório implica em concluir sobre a existência de um problema que precisa ser solucionado.

Informações sobre a situação atual da educação podem mostrar que:

- existem desequilíbrios no funcionamento do sistema educacional (problemas não solucionados, necessidades não encontradas);
- existem fatores do seu meio-ambiente (legislação ou política educacional, exigências sociais, restrições econômicas) pressionando-o;
- existem novos recursos (métodos, técnicas inovativas, instrumentais, conhecimentos teóricos no campo educacional ou em áreas relacionadas) que podem ser utilizados para melhorar o seu funcionamento.

Através destas informações, fornecidas pela Avaliação, o decisor pode conscientizar-se de que alterações ou inovações devem ser feitas na situação existente. Ocorre assim a 1ª Etapa do processo decisório: conscientização de um problema ou necessidade.

2ª ETAPA: "Design" ou Especificação de planos de ação

Desde que a existência de um problema tenha sido confirmada, surge a necessidade de identificar meios alternativos para solucioná-lo e de escolher um dentre estes. Assim, várias alternativas podem ser consideradas tais como: introduzir um novo método de ensino, treinar professores, aumentar a carga horária de aulas, diminuir o número de alunos, aumentar o número de professores, etc.. Todas as alternativas identificadas são especificadas em termos operacionais de tal forma que possam ser avaliadas e selecionadas uma dentre elas. Tem-se, aqui, uma situação de decisão que deve ser descrita de maneira bastante clara e precisa. Devem ser:

- identificados os decisores com diferentes níveis de autoridade (Secretário da Educação, Inspetores, Diretores, Professores, etc.)
- formulados explicitamente os vários meios alternativos (introduzir um novo método de ensino, treinar professores, etc.) para solucionar o problema.
- especificados os critérios que serão utilizados para pesar as alternativas e as regras de decisão para a seleção da alternativa que será instalada (custo, efetividade, custo/efetividade, etc.)
- estimado o tempo em que cada decisão será tomada.

Assim, também nesta etapa, o decisor precisa de muitas informações e é o avaliador que deve provê-lo de tal necessidade, de modo a ajudá-lo no julgamento e seleção da melhor alternativa para solucionar o problema.

3ª ETAPA: Escolha de um Plano de Ação

Identificamos os meios alternativos de solução, deve
rão ser realizadas as seguintes atividades:

1. Julgar as alternativas identificadas, com base em informa
ções (fornecidas pela avaliação) sobre cada uma delas.
2. Aplicar as regras de decisão para identificar a alternati
va a ser implantada.
3. Questionar a decisão tomada; ou seja refletir sobre a al
ternativa escolhida, (Esta decisão foi realmente a mais a-
certada? Os critérios e regras de decisão são realmente
válidos?).
4. Confirmar a decisão tomada, rejeitar a alternativa escolhi
da ou reciclar.

Através de maiores informações, o decisor poderá, por exemplo, alterar seus critérios, regras de decisão, conscienti
zar-se da existência de novas alternativas de ação ou ainda de
cidir que os procedimentos em uso devem ser mantidos.

4ª ETAPA: Ação

Escolhida a alternativa de solução outras tarefas de
vem ser executadas. São elas:

- Atribuir responsabilidade para instalar a alternativa sele
cionada. Isto implica em conscientizar os decisores, dos n
íveis inferiores, da necessidade de decidir sobre como im-
plantar a ação pretendida.

- Apresentar em termos operacionais a alternativa escolhida, identificando procedimentos, facilidades, orçamentos, pessoal, de modo que as ações especificadas possam ser executadas.
- Questionar sobre a eficácia da alternativa, apresentada em termos operacionais. (É válida? É possível de ser realizada? É eficaz?). Aqui, a reflexão é dos decisores dos diversos níveis.
- Executar o plano de ação, tal como foi operacionalizado, rejeitá-lo ou reciclá-lo.

Como pode ser observado, o processo decisório, em cada uma de suas etapas, é alimentado por informações. Tais informações são providas pela Avaliação e servem aos decisores dos diferentes níveis. Elas objetivam, sempre, a redução de incertezas na tomada de decisão educacional, quando melhorias ou mudanças são pretendidas para o sistema, ou seja, quando problemas ou necessidades existem. A Avaliação está assim diretamente relacionada com o processo decisório, na medida em que visa o aperfeiçoamento do mesmo.

Para cumprir sua função, a Avaliação precisa conhecer detalhadamente a situação de decisão para a qual irá fornecer informações. Ou seja, precisa saber exatamente:

- quem necessita de informações (decisores);
- em que nível de detalhamento deverão ser fornecidas as informações;
- quando a informação será necessária;
- que tipos de decisão serão tomados (decisões sobre: melhoria dos procedimentos em uso, compra de materiais, aprovação de alunos, adoção de um novo currículo);

- o que será avaliado (aluno? programas? materiais instrucionais?);
- quais são os usos potenciais das informações coletadas;
- que informações estão disponíveis.

Desde que a situação de decisão tenha sido exatamente definida, a Avaliação terá condições de acompanhar todo o processo decisório, auxiliando decisores a fazer escolhas com o maior grau de confiança.

Falou-se, aqui, da relação existente entre Avaliação Educacional e o processo de tomada de decisão. Em seguida, será identificada a relação entre processo decisório e Planejamento Educacional para melhor compreensão da abordagem avaliativa.

b) O Processo Decisório no Planejamento de Mudanças Educativas

O processo de tomada de decisão está implícito no planejamento e execução de novos procedimentos educacionais. Ao se planejar e executar mudanças, com o objetivo de melhoria, várias decisões se fazem necessárias. São elas relativas a:

- . objetivos alternativos que possibilitam a obtenção de melhores resultados para o sistema,
- . ações alternativas para escolha de uma delas, tendo em vista os resultados pretendidos,
- . execução dos procedimentos implantados,
- . modificações, continuidade ou término das atividades planejadas.

Estas decisões foram caracterizadas por Stufflebeam et al (1971) respectivamente como decisões de Planejamento, Estruturação, Implementação e Reciclagem (e encontram-se melhor conceituadas na unidade 2.0).

Em cada um destes quatro tipos de decisão (planejar, estruturar, implementar, reciclar) ocorrem as quatro etapas do processo de decisório (conscientização, design, decisão, ação). Por exemplo, para se tomar decisões de planejamento, são necessárias:- informações sobre uma necessidade do sistema (conscientização); informações sobre que objetivos alternativos podem conduzir à satisfação desta necessidade ("design") -decidir qual destes objetivos é o melhor para satisfazê-la (escolha), e, finalmente, ordenar a execução do objetivo escolhido (ação). De um modo geral, é necessário decidir sobre:

- quais serão os objetivos da mudança;
- que cursos de ação alternativos identificar;
- qual deles selecionar;
- como implantar a alternativa escolhida;
- como está, ela, sendo processada;
- são necessárias reformulações;
- quais são seus resultados.

Como se pode observar, várias decisões são tomadas e a Avaliação serve a todas elas, desloca-se através de todas as etapas do processo de mudança, fornecendo informações sobre cada um dos pontos considerados. Assim, pode-se concluir que auxiliando o processo decisório, a Avaliação ajuda às decisões de planejar, estruturar, implementar e reciclar procedimentos de mudanças educacionais.

Concluindo tudo que foi visto sobre tomada de decisão, tem-se que o processo decisório é ativado na medida em que pro

blemas ou necessidades são confirmadas e mudanças ou melhorias desejadas. Neste caso ações alternativas devem ser identificadas e uma delas escolhida para atingir os objetivos da mudança. Esta escolha implica em decisão e para tanto, são necessárias informações. Quem informa é a avaliação e quem escolhe ou decide é a pessoa chamada decisor. As decisões são tomadas com base:

- em certos critérios implícitos ou explícitos que refletem os valores pessoais do decisor ou os da instituição,
- e em informações avaliativas referentes a cada alternativa.

O processo avaliatório pretende, assim, facilitar a identificação, pelo decisor, da melhor maneira de provocar as mudanças pretendidas. Tendo em vista tal propósito a Avaliação se preocupa em informar sobre o sistema total e/ou cada um dos seus componentes. Para isto utiliza-se de uma abordagem sistêmica que lhe permite executar, apropriadamente, suas funções.

Será falado, em seguida, sobre a Avaliação Educacional dentro de sua nova abordagem.

1.3.2 - USO DA ABORDAGEM SISTÊMICA NOS ESTUDOS DE AVALIAÇÃO EDUCACIONAL

Considerando que estão sendo adotadas novas maneiras de tratar o planejamento de mudanças educacionais, uma consequência natural deste fato é a modificação dos antigos procedimentos de Avaliação. Desde que o conjunto de atividade avaliatórias é visto como um subsistema do sistema total de planejamento deve aquele estar em consonância com este. Se o sistema se modifica, seus elementos (ou subsistemas) devem ajustar-se às modificações para que o equilíbrio do sistema total possa ser mantido. Assim, não mais existe justificativa para uma Avaliação restrita apenas a um dos aspectos da educação, ou a cada um deles considerados individualmente. Não mais

se justifica uma Avaliação que não utiliza procedimentos científicos e que forneça informações desprovidas de objetividade.

A nova abordagem utilizada pela Avaliação é abrangente, são considerados todos os aspectos relevantes de um dado problema; a situação é vista como um todo e assim avaliada. Dentro desta abordagem, os estudos avaliatórios pretendem fornecer o máximo de informações que auxiliem as decisões, a serem tomadas numa situação de mudanças, quer sejam relativas ao ambiente do sistema, para identificar necessidades e escaloná-las; à estruturação de cursos de ação alternativos, para encontrar as necessidades prioritárias; à instalação de novos procedimentos ou à modificações nos mesmos. Assim temos avaliação do ambiente do sistema, de suas entradas, do seu processo e dos seus resultados.

. Avaliação do Ambiente do Sistema

Esta Avaliação informa sobre recursos disponíveis; limites ou restrições do sistema; seus problemas ou necessidades; atividades que estão sendo, por ele, desenvolvidas; suas capacidades para provocar modificações nos procedimentos em uso.

. Avaliação das Entradas do Sistema

Tendo em vista as modificações desejadas, várias maneiras de provocá-las, devem ser estudadas. A Avaliação irá ajudar na identificação, estruturação e análise de vários cursos de ação ligados aos propósitos do sistema e na escolha daquele que constituiu o melhor caminho para obter os resultados pretendidos. Aqui, a Avaliação se preocupa com todos os elementos considerados necessários para que as mudanças sejam processadas; ou seja, preocupa-se com as entradas do sistema.

. Avaliação do Processo do Sistema

Durante a implantação da mudança, a Avaliação irá informar sobre a execução dos procedimentos planejados. Preocupar-se-á com todos os aspectos do processo em andamento, com a maneira como os recursos estão sendo consumidos, as atividades desenvolvidas, com o desempenho do pessoal. Irá procurar informações que possibilitem decisões mais acertadas a respeito dos procedimentos adotados.

. Avaliação dos Resultados do Sistema

E, ainda, é preocupação da avaliação informar os decisores sobre os resultados do sistema, quer sejam intermediários (obtidos após cada etapa do processo) ou finais. Assim, esta Avaliação acompanha todo o processo fornecendo informações sobre o atingimento de tais objetivos e possibilitando decisões a respeito de modificações em qualquer ponto do sistema, tendo em vista os resultados obtidos.

Em resumo, a Avaliação, utilizando-se da abordagem sistêmica, assume efetivamente suas funções. Presta auxílio contínuo às decisões de:

- planejar,
- julgar e selecionar alternativas de ação,
- executar a alternativa selecionada,
- continuar com os novos procedimentos adotados, modificá-los ou rejeitá-los,

preocupando-se com:

- . o AMBIENTE (Avaliação do ambiente do sistema ou diagnóstica)
- . as ENTRADAS (Avaliação das entradas do sistema)
- . o PROCESSO (Avaliação do processo do sistema)

os RESULTADOS (Avaliação dos resultados do sistema),
do sistema.

Pensando ainda em abordagem sistêmica para estudos avaliatórios julgou-se importante relacionar, aqui, a utilização da filosofia de Engenharia de Sistemas no planejamento e execução da Avaliação Educacional.

Anteriormente, foi visto que o processo de planejamento da Engenharia de Sistemas se fundamenta em uma filosofia que, essencialmente, refere-se à sequência com que duas questões são colocadas ao se planejar e executar um sistema: "o que?" e "como?".

Ao se planejar e executar o processo de Avaliação, esta mesma filosofia pode ser adotada.

Na 1ª Etapa, PLANEJAMENTO, temos a seguinte sequência de perguntas: "O QUE?" e depois o "COMO?". Na 2ª Etapa, EXECUÇÃO, a sequência se inverte, cuida-se primeiro do "COMO?" e depois de "O QUE?".

PLANEJAMENTO: 1ª ETAPA

No planejamento da Avaliação, o "o que?" refere-se a especificação em detalhes, do processo avaliatório; ou seja, do propósito da Avaliação, da natureza das informações que serão fornecidas, relativas a quais variáveis, que dados serão coletados, que atividades serão executadas, que recursos serão utilizados. É necessário aqui, a determinação exata de tudo aquilo que caracteriza o processo de Avaliação; é preciso pensar neste processo em funcionamento. Por exemplo, a Avaliação objetiva fornecer informações que auxiliem no diagnóstico de uma situação existente, tendo em vista possíveis modificações no sistema. Para isto, um conjunto de atividades devem ser

executadas. Quais seriam estas atividades? Como seria a Avaliação operando de modo a fornecer as informações necessárias? Através de contatos com outras Instituições, entrevistas, utilização de instrumentos, estudos das áreas de conhecimento relacionadas com o sistema (política, econômica, social) e outros procedimentos de investigação dentro de limites especificados, a Avaliação iria:

- Verificar discrepâncias entre resultados pretendidos e obtidos pelo sistema;
- Detetar problemas que estariam prejudicando o seu funcionamento e impedindo-o de alcançar seus objetivos;
- descobrir outros recursos que poderiam ser utilizados para melhorar sua atuação;
- Auxiliar na identificação de objetivos para a mudança pretendida no sistema.

O "Como?" seria a especificação dos procedimentos, atividades e recursos necessários para conseguir a Avaliação. Por exemplo, seria preciso:

- contratar mais avaliadores;
- contratar outros especialistas;
- comprar materiais e contratar pessoas para imprimir os instrumentais ou outros materiais gráficos.
- uma bibliografia sobre literatura pertinente;
- instituições à disposição para receberem visitas.

Os passos a serem seguidos para o planejamento de estudos de Avaliação são apresentados na unidade 6, página F.363.

EXECUÇÃO: 2ª ETAPA

Na execução, o "Como?" seria a realização das atividades necessárias para que a Avaliação seja processada. Por exemplo:

- seriam contratados mais avaliadores e organizada a equipe;
- adquiridos os materiais e contratadas pessoas para imprimir os instrumentais;
- identificadas as instituições que seriam visitadas;
- localizada em alguma biblioteca ou adquirida a bibliografia para o estudo exploratório da literatura relacionada;
- identificadas as pessoas dentro do sistema que seriam contata das pela Avaliação.

O "O que?" seria a Avaliação sendo processada; ou seja, avaliadores:

- contatando com outras instituições;
- elaborando e aplicando instrumentais;
- contatando com elementos do sistema para coletar dados ou para prover as informações necessárias, estudando e analisando literatura pertinente;
- enfim, fornecendo informações úteis que permitam um diagnóstico da situação e, conseqüentemente, a identificação de objetivos de mudanças pretendidas para o sistema.

Resumindo, no PLANEJAMENTO dos estudos avaliatórios, é necessário saber exatamente o que será avaliado, com que propósito, que procedimentos serão utilizados. Caracteriza-se, todo o processo avaliatório e depois define-se a maneira como este processo será obtido, ou seja, identifica-se tudo aquilo que será necessário para processar a avaliação (instrumentais, avaliadores, materiais impressos, pessoal de apoio, etc.).

Na EXECUÇÃO será obtido tudo o que foi planejado (instrumentais serão elaborados, pessoal contratado, materiais gráficos elaborados e impressos, etc.). E, obtido, o sistema de Avaliação, ele será processado; ou seja entrará em funcionamento.

Assim no planejamento e execução da Avaliação, seguindo a filosofia de planejamento da Engenharia de Sistema tem-se a sequência:

- no Planejamento: "O QUE?" "COMO?"
- na Execução: "COMO?" "O QUE?"

Concluindo, pode-se observar que o processo avaliatório não é independente do processo de planejamento, mas o primeiro existe em função do segundo. A Avaliação serve ao planejamento e seu propósito está diretamente relacionado com os tipos de decisões a serem tomadas. Planejamento e Avaliação coexistem em todos os momentos em que tentativas de solucionar problemas, estão sendo feitas. Caminham juntos, são processos paralelos. Isto pôde significar que ambos os processos, Planejamento e Avaliação, devem fundamentar-se numa mesma filosofia, seja ela sistêmica ou outra qualquer.

1.4 - SUB-UNIDADE 4: ALGUMAS ABORDAGENS DE AVALIAÇÃO EDUCACIONAL

a) O que Avaliar?

Em educação avaliam-se mudanças:

- no processo de ensino/aprendizagem;
- nas escolas;
- nos currículos escolares;

- nos objetivos de um curso;
- no sistema total.

Avaliam-se também:

- administradores;
- professores;
- alunos;
- procedimentos em uso.

A Avaliação de cada um destes elementos educacionais envolve uma ou mais variáveis que devem ser cuidadosamente identificadas e analisadas para que os dados coletados, a respeito das mesmas, possam ser transformados em informações úteis ao processo decisório.

b) Com que Propósito?

Foram identificados, aqui, tres propósitos maiores da Avaliação; ou seja, DIAGNOSTICATIVO, FORMATIVO e SOMATIVO.

A Avaliação com propósito diagnóstico fornece informações que irão auxiliar na identificação de causas de problemas que estão prejudicando o processo das atividades educacionais.

A Avaliação com propósito formativo fornece informações que irão auxiliar na identificação de falhas no sistema total ou em alguns dos seus componentes. Tais informações são utilizadas para melhorar o sistema no seu todo ou em algumas de suas partes.

A Avaliação com propósito somativo fornece informações que irão auxiliar no julgamento final de um determinado sistema ou de um componente educacional. Vai dizer o quanto ele é satisfatório.

c) Novas Abordagens Avaliativas

Dentre as várias abordagens avaliativas que surgiram atualmente, são apresentadas aqui, aquelas de Stake (Sanders, 1973), Scriven (Sanders, 1973), Provus (Sanders, 1973) e Stufflebeam et al (1971).

1) A Abordagem Avaliativa de Stake

Stake (Sanders, 1973) define a Avaliação como um processo de descrever e julgar programas educacionais através de investigação formal. Nesta abordagem, descrição e julgamento de dados são consideradas atividades básicas do avaliador. Descrição e julgamento são termos chaves nesta definição e seus significados são descritos a seguir:

DESCRIÇÃO:

O avaliador descreve ambientes, atitudes e realizações. Isto significa descrever:

- o desempenho do estudante em relação aos objetivos inter-disciplinares e extra curriculares;
- variáveis que se encontram especificadas nos objetivos e outras que possam estar relacionadas com possíveis causas e efeitos da qualidade do ensino;
- conteúdo;
- os efeitos desejados e outros que possam ocorrer;
- a lógica do programa.

Na descrição de dados, o avaliador deve identificar três conjuntos de informações: antecedentes, transações e resultados.

Tais elementos são assim definidos:

Antecedentes são todas as condições que existem antes do processo de ensino, tais como estudantes, equipe, materiais, facilidades, que podem relacionar-se com os resultados.

Transações são todas as atividades realizadas para o alcance dos objetivos. Inclui relações e interações entre estudantes, professor e estudante, equipe do programa e de todo o pessoal envolvido no processo educacional. É o próprio processo.

Resultados referem-se aos produtos da aprendizagem que podem ser identificados como habilidades, realizações, atitudes e aspirações de estudantes. Inclui não só aqueles que são evidentes, como fins da aprendizagem, mas também aplicação, transferência e efeitos de reaprendizagem. Os resultados podem ser consequências imediatas ou a longo prazo, da educação.

Identificados estes três conjuntos de informações (antecedentes, transações e resultados), eles são descritos em termos de "Intenções" e "Observações".

O termo intenção é aqui considerado como sinônimo daquilo que é pretendido, seja em relação a antecedentes, transações ou resultados. A descrição das intenções é feita através de afirmativas explícitas das mesmas.

O termo observação refere-se ao que realmente ocorre em termos de antecedentes, transações e resultados. A descrição das observações, pelo avaliador, pode ser feita direta e pessoalmente, ou através do uso de instrumentais adequados a cada situação.

Os dados descritos são processados de duas maneiras:

- identificando contingências entre antecedentes, transações e resultados, pretendidos e observados;
- procurando congruência entre "intenções" e "observações".

Contingência - refere-se a lógica do programa. Há contingência quando existe uma conexão lógica entre antecedentes, transações e resultados.

Congruência - existe congruência quando o que estava sendo pretendido, ocorre realmente, ou seja, quando não há discrepância entre antecedentes, transações e resultados pretendidos e observados.

JULGAMENTO:

Julgamento é aqui definido como o ato de decidir sobre que padrões considerar, para a Avaliação de programas educacionais.

Para julgar as características de um programa, o avaliador pode considerar duas bases: 1) padrões relativos que se referem às características de programas alternativos; 2) padrões absolutos que se referem a julgamentos pessoais.

O julgamento final, feito com base em dados relativos ou absolutos, deve prover uma resposta imediata para a tomada de decisões educacionais.

Nesta abordagem, a Avaliação é orientada mais sobre programas educacionais do que sobre produtos educacionais.

2) A Abordagem Avaliativa de Scriven

Para Scriven (Sanders, 1973), a Avaliação consiste:

- na coleta sistemática de dados de desempenho, na combinação destes dados com um conjunto de metas valoradas, e na análise desta informação com o propósito de julgar o valor ou mérito de um programa educacional;
- em justificar:
 - . instrumentos para a coleta de dados;
 - . atribuição de pesos aos objetivos;
 - . a seleção de critérios e padrões.

A Avaliação está diretamente relacionada com os objetivos do programa. O avaliador identifica estes objetivos, julga o seu valor e verifica o seu alcance.

Segundo Scriven, dois tipos de padrões podem ser considerados na Avaliação de programas educacionais: padrões absolutos e relativos. Porém, segundo Scriven, mesmo na Avaliação absoluta está implícita uma comparação.

Scriven (Sanders, 1973) faz uma distinção entre objetivo e papel da Avaliação.

O objetivo da Avaliação é julgar o mérito ou valor de alguma coisa. Responde a questões de seleção, adoção e valor de materiais e atividades educacionais.

Em relação ao seu papel, a Avaliação pode ser formativa ou somativa, dependendo se é feita para conduzir à melhoria ou à determinação do mérito de um programa educacional.

A Avaliação Formativa consiste no fornecimento de informações, a serem utilizadas no melhoramento de componentes do programa ou deste como um todo. É uma avaliação orientada para padrões e critérios do elaborador do programa.

A Avaliação Somativa consiste no fornecimento de informações sobre o valor de um programa instrucional. Vai atender aos consumidores deste material dizendo o quanto ele é bom.

Scriven relaciona, ainda a estes papéis formativo e somativo, dois tipos de Avaliação: intrínseca, relativa ao conteúdo do material ou currículo, e extrínseca relativa aos seus efeitos.

A tabela que segue, apresentada por Sanders (1973), visualiza estas relações.

TABELA F.1

DISTINÇÃO ENTRE AVALIAÇÃO INTRÍNSECA E
EXTRÍNSECA E AVALIAÇÃO FORMATIVA E SOMA
TIVA

	INTRÍNSECA	EXTRÍNSECA
FORMATIVA	Julgamento da integridade (Ex.: estrutura e sequência de conteúdo) do conteúdo.	Julgamento dos efeitos intermediários para fornecer realimentação aos elaboradores de materiais
SOMATIVA	Julgamento final dos materiais	Julgamento final dos efeitos dos materiais, em alunos, professores, pais, administradores, etc..

3) A Abordagem Avaliativa de Provus

Para Provus (Sanders, 1973) "Avaliação é o processo de:

- . conduzir a um acordo em relação aos padrões do programa;
- . determinar se existe discrepância entre algum aspecto do programa, e os padrões que governam aquele aspecto;
- . usar a informação de discrepância para identificar os pontos fracos do programa",

com o propósito de monitorar e eventualmente julgar a efetividade de programas, determinando se devem ser melhorados, conservados como estão, ou eliminados.

De acordo com esta abordagem, a Avaliação é conduzida por uma equipe, que coleta informações essenciais para serem utilizadas, na melhoria do programa e no aconselhamento à administração. Ela é contínua, parte dos objetivos e da formulação de instrumentos, para avaliar estes objetivos. Exige também uma definição explícita de padrões.

Através da comparação entre padrões e desempenhos do programa, obtem-se a informação de discrepância. Esta comparação constitui um instrumento para avaliar o desempenho do programa.

Provus (1973) distingue três tipos de programas:

- de instalação instantânea;
- "enlatados";
- cuidadosamente projetados pelo seu próprio sistema escolar.

O modelo de Avaliação, proposto por Provus, aplica-se aos programas do primeiro tipo, com pessoal já alocado e em desenvolvimento. Tais programas contêm quatro estágios maiores de desenvolvimento, e o processo de Avaliação se desloca através de cada um deles: 1º) definição, 2º) instalação, 3º) processo e 4º) produto.

No primeiro estágio a tarefa do avaliador é obter uma definição do programa, com base numa taxionomia de conteúdo (um modelo geral que auxilia a definição do programa em termos de seus inputs, processos e outputs). A definição obtida, torna-se a informação do desempenho do programa a ser comparado com a taxionomia que constitui o padrão de desempenho deste estágio.

No segundo estágio, a informação do desempenho do programa é obtida, através de observações de campo, relativas à instalação dos componentes do programa. Neste estágio, o padrão para comparação é a definição do programa, obtida no primeiro estágio.

No terceiro estágio, em que o programa é lançado à população alvo, a informação, sobre o desempenho do programa, é obtida em termos de seus processos e efeitos iniciais de tratamento. Neste estágio, o padrão de comparação é aquela parte da definição, que descreve relações entre processos do programa e objetivos de capacitação.

No quarto estágio, as informações de desempenho do programa são relativas às medidas de alcance de objetivos terminais do projeto, sendo seu padrão de comparação, aquela parte da definição do programa, referente aos objetivos terminais.

A informação de discrepância é utilizada em todos os estágios de desenvolvimento e leva sempre a uma decisão que pode ser de:

- . ir ao estágio seguinte;
- . reciclar o estágio depois de mudanças nos padrões ou operações do programa;
- . reciclar para o primeiro estágio;
- . interromper o programa.

Assim a Avaliação é um processo contínuo de fornecimento de informações de discrepância, pertinentes e confiáveis, que servem como base de decisões administrativas.

4) A Abordagem Avaliativa de Stufflebeam

Stufflebeam et al (1971), definem a Avaliação como "um processo de delinear, obter e prover informações úteis para o julgamento de alternativas de decisão.

A Avaliação é, então, vista como um processo contínuo, que fornece informações relevantes em todas as fases e momentos do desenvolvimento de um programa. Quando, porém, acontece de a Avaliação não ter sido conduzida desde o início do programa, ela pode atuar em qualquer um dos estágios.

A Avaliação tem, essencialmente, uma função de assistência aos decisores e administradores pelo fornecimento de informações úteis a serem utilizadas na tomada de decisão educacional. Nesta decisão está implícito um julgamento de valor. O decisor possui um sistema de valores, e com base neste, especifica critérios de decisão. Cada alternativa é então pesada, em

função destes critérios. E, para a determinação dos pesos de cada uma das alternativas, em função dos critérios especificados, o avaliador delinea, obtém e provê informações úteis, ao decisor, para que ele possa julgar cada uma das ações em relação aos seus valores.

Nesta definição, oito termos são chaves:

- . Processo: é uma atividade contínua, sequencial, iterativa, envolvendo muitos métodos, técnicas, passos e operações, respeitando uma lei.
- . Delinear: é identificar a informação avaliativa necessária para a tomada de decisão. No delineamento, o avaliador está em constante interface com o decisor, para a identificação das alternativas e critérios a serem considerados.
- . Obter: é coletar, organizar e analisar os dados através de modos tradicionais como medida, processamento de dados e análise estatística. A obtenção é o aspecto mais técnico da avaliação.
- . Prover: é juntar as informações em sistemas ou subsistemas e relatá-las ao decisor.
- . Informação: são dados descritivos ou interpretativos, referentes a entidades e suas relações, em termos de algum propósito. As informações servem para diferenciar as alternativas envolvidas na situação de decisão. Assim as informações tornam-se meios que reduzem as incertezas, que rodeiam a decisão. Então, quanto mais informação se tem sobre as alternativas, menos arriscada será a decisão.
- . Útil: a informação tem utilidade quando satisfaz a 3 classes de critérios: Científicos, Práticos e de Eficiência.

Critérios Científicos incluem os seguintes:

- Validade interna - refere-se a correspondência entre a informação e o fenômeno que ela representa.
- Validade externa - refere-se a generalização da informaçãõ da amostra para a população.
- Confiabilidade - refere-se a consistência da informação isto é, se as informações forem colhidas em tempos diferentes chegar-se-ã aos mesmos resultados, ou se forem utilizados instrumentos equivalentes, também chegar-se-ã aos mesmos resultados.
- Objetividade - refere-se a "publicidade" da informação. A informação é objetiva se todos os especialistas que forem chamados, a julgar, concordarem com o significado dos da dos (convergência de interpretação).

Critérios Práticos incluem os seguintes:

- relevância - refere-se a relação entre informação e o objetivo da avaliação.
- importância - refere-se a focalização, dentre as informações relevantes, daquelas que são mais significativas para o objetivo da avaliação.
- escopo - refere-se a obtenção de um esquema completo de informações relacionado a uma avaliação. Possibilita a visão geral do fenômeno.
- credibilidade - refere-se a confiança que o usuário da informação tem no avaliador.
- ocasião - refere-se ao fornecimento da informação em tempo hábil. A melhor informação é inútil, se chega muito tarde.
- penetração - refere-se à difusão da informação avaliati

va. A utilidade da informação se prende também ao fato de que se todas as pessoas, que necessitam da informação, realmente as recebem.

Critério de Eficiência referem-se a obtenção da informação, levando-se em consideração o tempo, custo e pessoal necessários.

- juízo - é o ato de escolher entre as várias alternativas. É o ato de decidir. Juízo é o termo central da definição de Avaliação.
- alternativas de decisão - são duas ou mais ações diferentes, que poderiam ser tomadas, em resposta a alguma situação requerendo ação alternativa.

Nesta definição pode ser observado que um grande destaque é dado ao processo de juízo, e ao objetivo com que é feito este juízo, isto é, a tomada de decisão.

Stufflebeam et al (1971) identificam 4 tipos de decisão dentro de um programa educacional:

- . decisões de planejamento para determinar os objetivos;
- . decisões de estruturação para planejar procedimentos, que conduzam ao alcance dos objetivos;
- . decisões de implementação para utilizar, controlar e refinar procedimentos.
- . decisões de reciclagem para julgar e reatuar as obtenções.

Relacionado a cada um destes tipos de decisão, ele propõe respectivamente quatro tipos de Avaliação: Contexto, Input, Processo e Produto.

A Avaliação de Contexto provê uma lógica para determinação dos objetivos de um programa.

A Avaliação de Input provê informações para determinar como utilizar recursos, para encontrar os objetivos do programa. Ela identifica e avalia as estratégias relevantes, as capacidades e procedimentos da implementação.

A Avaliação de Processo provê realimentação periódicas às pessoas responsáveis pelos procedimentos e planos de implementação, a fim de superar dificuldades relativas a procedimentos.

A Avaliação de Produto verifica a extensão na qual os fins estão sendo alcançados, através da medida e interpretação dos efeitos da mudança.

Os dados obtidos através destes quatro tipos de Avaliação, desde que obedeam aos critérios científicos, práticos e de eficiência, resultam em informações úteis que podem ser consideradas essenciais em toda situação que envolva mudanças educacionais.

d) Outras Abordagens de Avaliação Educacional

Considerando que, atualmente em nossas escolas, a Avaliação tem-se constituído num processo um tanto limitado, prendendo-se apenas a alguns aspectos do processo educacional, julgou-se útil apresentar, aqui, uma análise feita por Stufflebeam et al (1971) de alguns conceitos de Avaliação que surgiram em épocas passadas e que estão sendo utilizados ainda hoje. Estes conceitos não são totalmente desprovidos de utilidade, porém não reflêtem a real função dos estudos avaliatórios. São eles:

- 1ª) Avaliação como Medida;
- 2ª) Avaliação de Congruência;
- 3ª) Avaliação como Julgamento.

1ª) Avaliação como medida

Esta definição surgiu na década de 1920-1930, identificando Avaliação como Medida. A Avaliação era vista apenas como um processo de desenvolvimento e interpretação de instrumentos. Assim, no processo instrucional, somente aquelas variáveis passíveis de mensuração, podiam ser avaliadas. Outras variáveis, que não podiam ser medidas, eram consideradas irrelevantes.

A Avaliação dentro desta definição tem, portanto, um enfoque limitado e uma abordagem mecânica. Uma de suas vantagens é que fornece informações confiáveis e objetivas.

2ª) Avaliação de Congruência

Esta definição foi proposta por Tyler por volta de 1950. A Avaliação é considerada como um processo de determinar o grau de consecução dos objetivos educacionais, que um curso ou programa procura alcançar. Em outras palavras, a Avaliação procura determinar a congruência entre desempenho do aluno e objetivos comportamentais estabelecidos.

Uma das vantagens desta definição é que a Avaliação proverá informações para serem utilizadas na tomada de decisão relativa ao programa educacional, focalizando o estudante, o currículo e todos os procedimentos instrucionais. Uma outra vantagem desta abordagem é que a Avaliação é feita sobre refe

rências concretas ou seja objetivos e critérios descritos explicitamente.

Porém, uma vez que o que se deve avaliar é a congruência entre desempenho e objetivo, a Avaliação, aqui tem sido entendida comumente como uma técnica terminal, embora não tenha sido esta a intenção do seu autor.

3ª) Avaliação como Julgamento

Esta definição enfatiza a Avaliação como um processo de julgamento de profissionais. Um grupo de especialistas, em diferentes áreas, emite julgamentos de valor a respeito de um currículo, um estudante, uma escola, etc..

Esta Avaliação tem a vantagem de ser rapidamente conduzida e facilmente implementada. Entretanto, tem uma grande desvantagem, sua confiabilidade e objetividade são questionáveis. E, na maioria das vezes não existem critérios bem definidos para o julgamento, além dos objetivos, quase sempre, ser descritos de maneira ambígua.

Assim, pode-se observar que dentre essas três definições, somente a Avaliação como julgamento leva a determinação do mérito ou valor daquilo que está sendo avaliado. Porém, nesta abordagem somente algumas vezes o julgamento do grupo de especialistas, em forma de recomendações, é fornecido ao grupo responsável pelas decisões finais. Também, esta Avaliação não possui nenhum procedimento sistemático. Este fato além de causar dificuldades na tarefa do avaliador não conduz aos reais objetivos do processo avaliatório.

Descritas algumas abordagens de Avaliação Educacional será apresentado a seguir um paralelo entre Avaliação e Pesquisa, objetivando diferenciá-las através de alguns pontos fundamentais que

as caracterizam. Estas duas atividades embora sejam ambas atividades de investigação disciplinada, diferenciam-se basicamente e algumas destas diferenças serão aqui apresentadas desde que é necessário que se faça uma clara distinção entre ambas.

1.5 SUB-UNIDADE 5: PESQUISA X AVALIAÇÃO

Avaliação e pesquisa são atividades de investigação disciplinar. Esta se caracteriza, principalmente, como uma investigação sistemática e objetiva.

Inclui a investigação empírica, filosófica e histórica.

Sanders (1973) propõe a Figura F.10 que se segue, para mostrar as inter-relações entre os três tipos de investigação disciplinada.

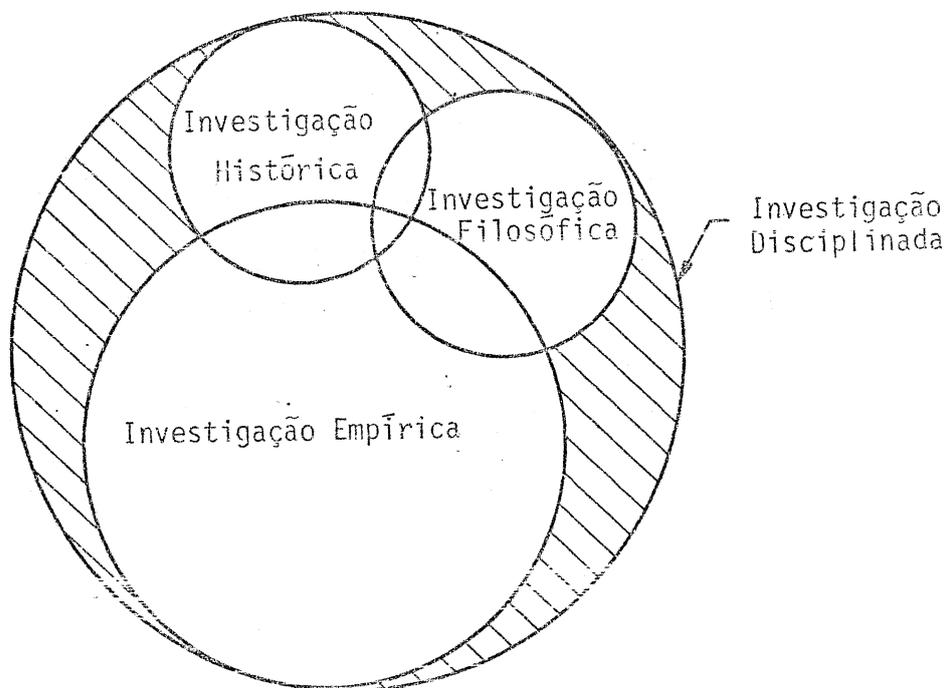


Fig. F.10 - Domínio da Investigação

Estes três tipos de investigação encontram-se assim, definidos:

- Investigação Empírica é aquela que se realiza através de observações e experimentações, com a finalidade de descrever condições existentes, verificar ou refutar asserções, afirmações ou hipóteses sobre relações entre variáveis.
- Investigação Histórica refere-se ao estudo do desenvolvimento de organizações de pessoas, de vidas particulares, movimentos ou sequências de eventos.
- Investigação Filosófica está relacionada com análise racional baseada em lógica formal e semântica.

Existem algumas atividades que exigem mais de um tipo de investigação. Por exemplo, técnicas empíricas são utilizadas para estudar os svários significados atribuídos a certas palavras por pessoas com diferentes orientações filosóficas.

Na Figura F.10 estas atividades encontram-se representadas pelas áreas de superposição.

A investigação disciplinada inclui atividades que são comuns em educação; ou seja, atividades de Pesquisa, Avaliação, Desenvolvimento e Difusão.

- Pesquisa consiste em um conjunto de procedimentos sistemáticos ou empíricos, objetivando a produção de conhecimentos universalmente válidos.

Se os estudos de Pesquisa são realizados com o objetivo de aplicar conhecimentos ou de fazer algo, de forma mais eficiente temos a Pesquisa Aplicada.

Se o objetivo da pesquisa é desenvolver modelos ou teorias que identificam as variáveis relevantes em um ambiente e suas relações, temos a Pesquisa Básica.

- Avaliação refere-se a julgamentos de valor com base em informações úteis obtidas através de um estudo do problema. Em educação atribui-se um valor a um programa, procedimentos, produtos, objetivos.
- Desenvolvimento na área educacional refere-se a produção e testagem de materiais e meios utilizados na instrução.
- Difusão refere-se ao planejamento, "design" e condução de atividades que resultam na aplicação de conhecimentos e produtos de pesquisa e de desenvolvimento.

A abordagem, aqui, será limitada às atividades de Pesquisa e Avaliação.

Apesar de terem propósitos diferentes, Pesquisa e Avaliação, como atividades de investigação disciplinadas, apresentam algumas semelhanças sobrepondo-se em alguns aspectos.

A Figura F.11 que se segue, apresentada por Sanders (1973), mostra a superposição entre estas duas atividades.

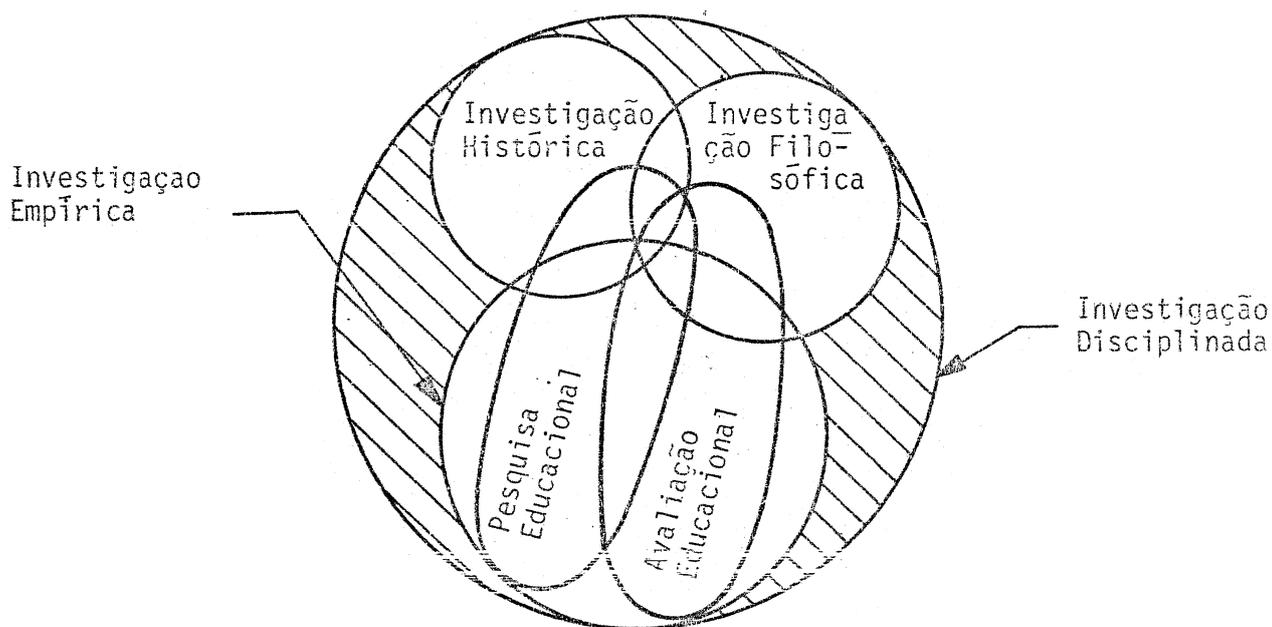


Fig. F.11 - Domínio da Investigação: Avaliação e Pesquisa Educacional

Tanto Pesquisa como Avaliação dependem de técnicas e métodos empíricos e utilizam-se da investigação histórica e filosófica, embora com propósitos diferentes.

Como a Avaliação está preocupada com questões de valores, utiliza-se mais a investigação filosófica e pouco da histórica. O inverso ocorre com Pesquisa cuja preocupação é produzir um novo

conhecimento que possa ser generalizado para várias situações e a través do tempo.

Apesar de, algumas vezes, a Avaliação ter sido conside rada simplesmente como uma forma de Pesquisa Aplicada, elas possuem alguns aspectos que as diferenciam de maneira bastante clara.

Identificaremos aqui algumas diferenças gerais entre am bas.

- O propósito da Pesquisa é a obtenção de novos conhecimen tos, universalmente válidos. Procura criar leis universais. Está preocupada com a verdade científica.
- O propósito da Avaliação é auxiliar o processo de tomada de decisão através do fornecimento de informações úteis que permitam julgar o valor ou utilidade social de um fenômeno.
- Na Pesquisa, julgamentos de valor estão relacionados apenas com a seleção do problema a ser estudado. Valor não é o instrumento básico do pesquisador, mas somente uma preocu pação indireta.

Na Avaliação, julgamentos de valor são feitos na sele ção do problema, durante o seu desenvolvimento e implementação.

Sem julgamentos de valor não existe Avaliação. Aqui há uma preocupação direta com valor.

- A Pesquisa procura descobrir a relação entre duas ou mais variáveis ou fenômenos.
- A Avaliação procura simplesmente descrever um fenômeno particular com relação a uma ou mais escalas de valores.

- A Pesquisa procura explicar o "como" e o "porquê" dos fenômenos.

A explicação do "como" e do "porquê" não é objetivo primeiro da Avaliação, mas sim fornecer informações úteis que permitam julgar o valor de um fenômeno.

- Tendo em vista seus propósitos, a Pesquisa busca informações utilizando-se de "laboratórios" que permitem manipular ou controlar variáveis que interferem no fato ou fenômeno que está sendo estudado. Nesta situação torna-se possível eliminar interferências de variáveis não desejáveis, estudando-se apenas os efeitos daquelas que são de interesse da Pesquisa ou randomizar o processo de entrada das variáveis.

O uso de laboratório permitindo um alto controle da situação em estudo, constitui um meio bastante apropriado para os estudos de Pesquisa que pretendem produzir conhecimentos desvinculados do contexto e que possam ser generalizados.

Os estudos de Avaliação são realizados nas condições naturais em que o problema ocorre. A Avaliação, por exemplo, não está interessada em coletar informações a respeito do desempenho de um grupo de alunos numa situação artificial de laboratório, mas sim numa situação real onde ocorrem interferências dos vários fatores que possam estar relacionados com o desempenho dos estudantes.

Assim, as situações de Avaliação são pouco controladas, o que limita o poder de generalização dos seus resultados. Isto significa que os conhecimentos produzidos pela Avaliação, são relativos a um contexto particular e portanto a ele, vinculados. Não pretende generalizar resultados.

- Como foi visto, o pesquisador intervém na situação a fim de arranjar as condições necessárias para analisar as interações entre as variáveis (dependentes e independentes) que estão

sendo objeto da pesquisa. Este fato tem como consequência uma perda de informações sobre o fenômeno desde que são analisadas apenas as influências de variáveis específicas. Outras possíveis fontes de interferências não são, aqui, consideradas.

As informações obtidas através de um estudo de Pesquisa não atendem aos propósitos da Avaliação. Esta se interessa por dados coletados em situações reais onde procura obter o máximo de informações úteis que possibilitem julgar o fenômeno em estudo.

O avaliador, ao contrário do pesquisador, procura atrair interferências de todos os fatores que possam vir a influenciar a situação analisada.

Assim enquanto a Avaliação controla as variáveis estudando-as, a pesquisa as controla, eliminando-as.

- Os procedimentos utilizados pela Pesquisa, principalmente pe la Pesquisa Experimental, são rígidos e altamente controla dos a fim de que conhecimentos universalmente válidos possam ser obtidos.

Embora não seja propósito da Avaliação produzir verdade científica, os mesmos procedimentos de Pesquisa podem e são utilizados em estudos avaliativos (embora nem todos os estudiosos do assunto concordem com tal afirmação).

A Avaliação utiliza-se de técnicas de Pesquisa, principalmen te se tem como objetivo fornecer informações para um julga mento final de valor dos resultados ou produtos de um estu do, ou seja, quando seu propósito é somativo.

Isto porque os procedimentos de Pesquisa são direcionados apenas para os efeitos terminais ou resultados do estudo, pro duzindo dados somente no final do experimento.

Porém o uso exclusivo de técnicas de Pesquisa não satisfazem plenamente as necessidades da Avaliação.

O avaliador diante de seus propósitos (diagnóstico, somativo ou formativo) deve selecionar os procedimentos mais adequados para atingi-los.

- Pesquisa e Avaliação diferenciam-se também em relação aos critérios que são utilizados para julgar a adequação de tais estudos.

A adequação dos estudos de Avaliação é julgada em relação a dois critérios, aqui, considerados mais importantes: isomorfismo (a extensão na qual a informação obtida corresponde a realidade) e credibilidade (a extensão na qual a informação é confiável).

A adequação dos estudos de Pesquisa é julgada em relação a dois critérios básicos: validade externa (a extensão na qual os resultados podem ser generalizados para outras situações) e validade interna (os resultados devem ser função dos procedimentos utilizados e não de outras causas que não estejam sistematicamente relacionadas com o estudo). Se os procedimentos utilizados não possibilitam validade interna (o que causou os efeitos observados?) a validade externa ficará prejudicada e não será possível generalizar a partir dos resultados obtidos.

Em resumo, as informações fornecidas por estudos de Avaliação devem atender as necessidades de um decisor e não aos interesses particulares de quem faz a indagação (como é o caso da Pesquisa).

A Avaliação, ao contrário da Pesquisa, não tem o propósito de contribuir com a literatura científica produzindo conhecimentos universais, mas sim de prover informações úteis a um decisor espe

cífico, numa situação e tempo, também, específicos. Cada contexto de Avaliação é único e não está sob o controle total do avaliador.

A pesquisa procura desvincular-se de contextos e obter o máximo controle sobre a situação em estudo.

Assim enquanto a Avaliação é flexível em sua abordagem, a pesquisa necessita de grande controle para executar facilmente sua metodologia e produzir leis universais.

A Avaliação envolve um conjunto de fatores que não se encontram bem definidos no seu contexto. A Pesquisa abrange um conjunto de fatores bem definidos e controlados.

A Avaliação procura testar e fazer uso específico de todos os fatores que interferem no estudo com o propósito de desenvolver, analisar e fazer inferências para a melhoria do projeto.

A Pesquisa procura testar e fazer uso geral dos fatores que são objetos de seu estudo com o propósito de descrever, analisar e fazer inferências sobre o sistema de conhecimentos (Sašlow, 1970).

Sanders (1973) diz que uma forma de diferenciar Pesquisa e Avaliação é perguntar se a investigação que se está conduzindo seria vista como um fracasso se não produzisse informação sobre se o fenômeno estudado é útil ou não. Um pesquisador provavelmente responderia que não.

Foi visto que novas abordagens de Planejamento e consequentemente de Avaliação, tornaram-se necessárias, diante das transformações científicas, políticas, econômicas ocorridas em todas as partes do mundo e que afetaram profundamente a educação. Falou-se, aqui, de uma abordagem atual, a abordagem sistêmica, utilizada para

planejar e avaliar soluções propostas para eliminar problemas complexos. Um modelo de Avaliação Educacional, utilizando esta abordagem, é descrito a seguir na unidade 2.

2.0 - UNIDADE 2 - UMA ABORDAGEM AVALIATIVA - O MODELO DE STUFFLEBEAM ET AL

2.1 - SUB-UNIDADE 1 - A TOMADA DE DECISÃO EDUCACIONAL

Considerando a definição de Avaliação, apresentada por Stufflebeam et al (1971), como o processo de delinear, obter e fornecer informação útil para a tomada de decisão, a lógica geral da abordagem proposta para a Avaliação é mostrada na Figura F.12.

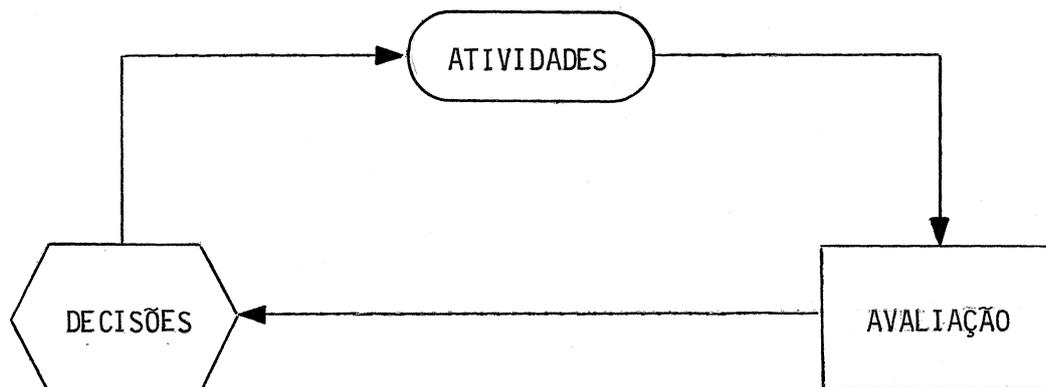


Fig. F.12 - Relação entre a Avaliação e a Tomada de Decisão (Stufflebeam et al, 1971; pág. 216)

Nesta abordagem, as operações (atividades) do Sistema são avaliadas para influenciar decisões que, por sua vez, influenciam ou tra vez as atividades que são de novo avaliadas, e assim por diante.

Assim, avaliar Sistemas e Projetos para determinar o valor dos esforços dispendidos nos mesmos, é uma necessidade. E, considerando esta necessidade surgem as indagações:

- Como avaliar tais Sistemas e Projetos? Existem esquemas, guias gerais que podem ser seguidos em qualquer situação, ou o esquema varia basicamente com esta?
- Que caminho seguir?

Planejar é essencial em estudos de Avaliação Sistemática. Sem bons esquemas, estudos de Avaliação podem facilmente ser desenvolvidos sem direção, e, conseqüentemente sem ponto final.

Retomando a definição de Avaliação com ênfase na tomada de decisão é preciso, inicialmente, se fazer uma descrição desta, para que se possa embasar uma metodologia de Avaliação relevante.

Uma descrição da tomada de decisão educacional é feita caracterizando-se os seguintes elementos:

- As situações dentro das quais ocorrem as decisões.
- Os tipos de decisão que são feitos.

SITUAÇÕES DE DECISÃO

Uma situação de decisão é o conjunto total de circunstâncias que governam a análise e a escolha no processo de tomada de decisão. Estas circunstâncias incluem o grau de mudanças que poderia resultar de uma escolha, a urgência requerida, os meios alternativos e/ou critérios que são gerados, a predição de conseqüências para as diferentes opções de decisão, os custos e riscos associados com estas opções, atitudes da comunidade sobre possíveis mudanças e modos de realizá-las, etc..

As situações de decisão são diferenciadas através de dois contínuos: pequena X grande mudança e alto X baixo nível de informação. A combinação destes 2 contínuos possibilita a conceitualização de 4 classes de situações de decisão educacional, da forma apresentada na Tabela F.2.

TABELA F.2

SITUAÇÕES DE DECISÃO

ALTA	QUANTIDADE	<u>HOMEOSTÁTICA</u> Atividade: restaurativa. Propósito: manter o equilíbrio. Base: padrões técnicos e controle de qualidade.	<u>METAMÓRFICA</u> Atividade: utópica. Propósito: mudança total. Base: "Teoria abrangente do Saber".
		<u>INCREMENTAL</u> Atividade: desenvolvimento. Propósito: melhoramento contínuo do processo. Base: julgamento de especialistas e investigação estruturada.	<u>NEOMOBILÍSTICA</u> Atividade: inovativa. Propósito: investigação, testagem e difusão de soluções para problemas significantes. Base: conceitualização, investigação heurística e estruturada.
BAIXA	INFORMAÇÃO	NÍVEL DE MUDANÇA	
		PEQUENO	GRANDE

(Fonte: Stufflebeam, et al 1971, página 62).

Para uma maior compreensão da classificação apresentada na Tabela F.2 é necessário maior detalhamento dos contínuos usados na mesma, isto é, graus de mudança e nível de informação, tal como segue.

a) Mudança Pequena X Grande

A mudança em educação é axiomática. Mesmo em muitos sistemas escolares conservadores, os processos educacionais estão em contínua mudança, sendo, porém, menos perceptíveis. Em outros sistemas, mudanças no processo educacional são abruptas, deliberadas e altamente visíveis. Elas são tão revolucionárias como caso das máquinas de ensinar, que causaram uma reestruturação bastante significativa no ensino nos Estados Unidos.

Braybrooke e Lindblom (Stufflebeam, et al (1971) distinguiram pequena e grande mudança, como segue:

o tamanho da mudança (grande ou pequena) é consequência direta do valor que é dado a ela. E este valor pode variar de pessoa a pessoa. Porém, a noção de "pequena" ou "grande" não é tão subjetiva e pessoal, isto porque em qualquer sociedade, desenvolve-se uma forte tendência em relação a convergência de opinião. Esta convergência dá ao julgamento de valor da mudança uma conotação objetiva.

Deste modo, mudanças potenciais devem ser classificadas em termos de duas condições: a sociedade vê as variáveis a serem alteradas como importantes? E a sociedade vê a própria mudança proposta, como trivial ou como importante? Se não há variáveis importantes envolvidas, a mudança é pequena. Se estão envolvidas variáveis importantes, mas são manipuladas de maneira não importante, a mudança também é pequena. Contudo, se variáveis importantes são manipuladas de maneira importante, a mudança é grande.

Portanto, o sistema de Braybrooke e Lindblom nos permite classificar uma mudança como pequena ou grande, a partir da percepção de um público especificado.

b) Alta X Baixa Quantidade de Informação

O segundo maior contínuo, proposto por Lindblom e Braybrooke, que caracteriza as situações de decisão, é a quantidade de informação disponível ou seja, o grau no qual os decisores supõem entender todos os aspectos do problema que eles enfrentam.

O entendimento da situação inclui dois elementos que são importantes na tomada de decisão:

. informação relevante e capacidade intelectual do decisor, para utilizar as informações adequadas para resolver problemas práticos.

Tendo presente esta distinção entre os 2 contínuos - quantidade de mudança e nível de informação - pode-se compreender melhor as classes gerais das situações de decisão, apresentadas na Tabela F.2, ou seja:

- Situações de Decisão Metamórficas,
- Situações de Decisão Homeostáticas,
- Situações de Decisão Incremental,
- Situações de Decisão Neomobilísticas.

1) Situações de Decisão Metamórfica

Estas situações implicam em grandes mudanças com um alto nível de informações relevantes. São caracterizadas por atividades utópicas, para produzir mudanças completas no Sistema Educacional. Aqui, é necessário que o decisor:

. Tenha um conhecimento teórico necessário e suficiente de todos os detalhes da mudança proposta e um entendimento completo de todas as suas manifestações.

. Seja capaz de considerar todas as variáveis relevantes, coletar, analisar e sintetizar dados de desempenho sobre elas.

Obviamente, tais situações de decisão educacionais utópicas são teóricas, não sendo consideradas nos estudos de Avaliação.

2) Situações de Decisão Homeostática

Estas situações implicam em pequenas mudanças com um alto nível de informações.

A tomada de decisão homeostática é caracterizada por atividade restaurativa visando o equilíbrio normal de um Sistema Educacional. Estas situações de mudança são guiadas por padrões técnicos, por uma rotina e por um sistema cíclico de coleta de dados.

Assim, as mudanças efetuadas por estas decisões são pequenas e remediais e exigem um sistema de coleta de dados de controle de qualidade tais como: escores de testes de aprendizagem, informações pessoais do aluno, registros de frequência, graus de curso, informações sobre a comunidade

. Exemplo de uma Situação de Decisão Homeostática

A escola X está com um problema de ausência maciça dos alunos, das 3as. e 4as.séries do 1º Grau, às aulas em determinados dias da semana. Esta ausência, ao que tudo indica, está sendo motivada por um Parque de Diversões que está funcionando no horário das aulas e está localizado perto da escola e, além disso, a maioria dos pais trabalha e não verifica se seus filhos estão frequentando as aulas.

Para a tomada de decisão relativa a este problema, ou seja, a ausência maciça dos alunos, o diretor (o decisor no caso) conta com um alto nível de informação:

- . sabe exatamente em que dia e quais alunos faltam às aulas;
- . sabe da existência, localização e horário de funcionamento do Parque de Diversões.
- . sabe que os pais geralmente não verificam se seus filhos estão indo às aulas, mas fazem questão que eles as frequentem.

E, assim, ele possui os elementos necessários para identificar as estratégias relevantes para alcançar a mudança desejada e conhecer os níveis de riscos dos resultados das mesmas. Neste exemplo, as estratégias podem ser identificadas como segue:

- . conversar com os pais.
- . convocar os alunos
- . conversar com o dono do Parque de Diversões solicitando sua cooperação.

Conhecendo as pessoas envolvidas e como possivelmente elas reagirão a estas soluções, o decisor, pode, facilmente, selecionar uma mistura satisfatória de estratégias, sem precisar de maiores informações.

Quanto a mudança a ser realizada, ela é, sem dúvida, pequena e remedial, ou seja:

- . se os pais fazem questão que seus filhos frequentem as aulas em seu horário normal, qualquer medida que garanta esta frequência será considerada normal e rotineira.

3) Situações de Decisão Incremental

Estas situações implicam em pequenas mudanças, com um baixo nível de informações. A tomada de decisão incremental é caracterizada por atividade desenvolvimentista, com o propósito de melhoria contínua de um Sistema. Tal atividade é comumente sustentada por julgamento de especialistas e investigação sistemática da eficácia do sistema presente e das mudanças recomendadas.

As Decisões Incrementais mudam o Sistema em direção a um novo equilíbrio, baseado numa série de pequenas melhorias. Estudos especiais, emprego de consultores especializados e a formação de equipes, caracterizam os muitos esforços para introduzir a mudança incremental. As decisões incrementais tentam melhorar o Sistema atuante sem riscos de uma falha maior. Os ajustamentos são poucos e podem ser feitos à medida

em que os problemas são detetados. Tais mudanças são baseadas em ensaaios e erro, são iterativas e seriadas. O tipo de mudança, aqui, envolvido tem sentido desenvolvimental, mais do que restaurativo e inovativo.

. Exemplo de uma Situação de Decisão Incremental

O Secretário de Educação do Estado X está às voltas com a necessidade de melhorar os níveis de aprendizagem de leitura dos alunos das Escolas Públicas de 1º Grau. Ele recebeu informações que revelaram que os níveis reais de aprendizagem de leitura estão muito aquém do que o Sistema Escolar Estadual espera deles. Foi informado, também, de vários problemas causadores desta discrepância, sendo, porém, ao que tudo indica, o mais crucial, a má qualidade dos materiais instrucionais que estão sendo utilizados para o ensino da leitura. Porém, não tem conhecimento a respeito de estratégias que poderiam ser utilizadas para atacar o problema.

Para a tomada de decisão relativa a este problema, ou seja, os baixos níveis de aprendizagem de leitura dos alunos do 1º Grau, o Secretário (o decisor no caso) do Estado X, conta com baixo nível de informação:

- . não tem conhecimento a respeito de estratégias que poderiam ser utilizadas para atacar o problema.

E, não sabendo como resolver a situação, ele precisa de informações para que possa decidir qual o procedimento a ser utilizado.

Quanto à mudança a ser realizada, ela provavelmente será percebida como pequena pela comunidade (pessoas direta ou indiretamente relacionadas à esta situação). Isto, porque a mudança se prende apenas aos materiais instrucionais, sem implicar em mudanças de estrutura ou organização do Sistema, e que, provavelmente, não exigirá reformulação dos valores já existentes ou aquisição de novos.

4) Situações de Decisão Neomobilística

Estas situações implicam em grandes mudanças com baixo nível de informação. A tomada de decisão neomobilística é caracterizada por atividade inovativa para inventar, testar e difundir novas soluções para problemas significantes. A mudança é grande, mas sustentada por pouca teoria ou conhecimento existente. Nos primeiros estágios de tentativa de mudança, as investigações não são rigorosas e os estudos são frequentemente exploratórios e heurísticos. Contudo, nos últimos estágios de esforços de mudança, o rigor torna-se o "sine qua non". Assim, os esforços de mudança neomobilística começam com uma prova exploratória que, depois, é transformada em uma engenharia rigorosa e numa investigação de pesquisa de mercado. Eles são, conseqüentemente, mais extensos, mais caros e mais propensos a riscos do que as estratégias homeostáticas ou incrementais.

. Exemplo de uma Situação de Decisão Neomobilística

O Secretário da Educação do Estado X tem a necessidade de melhorar o nível de aprendizagem dos alunos nas matérias básicas (como Português, Matemática e Ciências) das 4 séries iniciais do 1º Grau. Ele recebeu informações, que revelaram que os níveis reais de aprendizagem dos alunos, está muito aquém do que o Sistema Escolar Estadual espera deles. Foi informado, também, de que a causa mais provável deste problema é o nível do próprio professorado, pois num concurso para o magistério, ocorrido há pouco tempo, estes mesmos professores foram candidatos e houve 75% de reprovação, por não terem alcançado a nota mínima. E mais, em 38 dos 64 municípios do Estado, não houve qualquer aprovação em matérias básicas, como Português, Matemática e Ciências.

Com estas informações, o decisor (o Secretário da Educação) sabe que o problema mais urgente a ser atacado é o da formação de seus professores, o que implicará, inevitavelmente, numa reestruturação completa do Ensino Normal e numa reciclagem dos professores já formados. Porém, ele não tem conhecimento de qual a melhor abordagem para resolvê-lo.

Para a tomada de decisão relativa a este problema, o Secretário conta com um baixo nível de informação, ou seja, não tem conhecimento a respeito de estratégias que poderiam ser utilizadas para atacar o problema. Não sabendo como resolver a situação, ele precisa de informações para que possa decidir qual o procedimento a ser utilizado.

Quanto à mudança a ser realizada, ela provavelmente será percebida como grande pela comunidade. Isto porque, provavelmente, a mudança estará relacionada com aspectos da estrutura e organização do sistema e exigirá reformulação dos valores já existentes ou aquisição de novos.

A utilidade da classificação de situações de decisões Educativas do relacionamento entre os dois contínuos - tamanho da mudança e nível de informação - discutido anteriormente, é uma consequência direta da definição de Avaliação proposta por Stufflebeam et al. (1971) como o processo de delinear, obter e prover informação útil para a tomada de decisão. Primeiro, a importância da decisão a ser servida determina o rigor e a extensão de uma Avaliação; segundo, a medida em que delinea a informação que obterá e fornecerá, o avaliador deve ter em mente a informação, já disponível, e a habilidade de seu cliente para usá-la como está. Assim, os sistemas de Avaliação são necessários para sustentar ambos os tipos de mudança: grande e pequena. A informação é necessária, obviamente, para fornecer avisos prévios de condições de emergência; informação avaliativa é necessária em esforços científicos e tecnológicos, voltada para o desenvolvimento e compreensão de problemas educacionais significativos, para criar e planejar soluções.

Em suma, dependendo da situação de decisão a ser servida, medidas avaliativas diferentes são exigidas, em termos de maior ou menor estruturação das mesmas.

A seguir será caracterizado o 2º elemento da tomada de decisão educacional, ou seja, tipos de decisão.

TIPOS DE DECISÃO

O conhecimento das quatro situações de tomada de decisão é necessário, mas não é condição suficiente para formular um Modelo de Avaliação, capaz de servir a tomada de decisão. É preciso uma tipologia ou taxionomia de decisões, cujas categorias sejam exaustivas de todas as decisões educacionais possíveis, embora sejam, também, mutuamente exclusivas. Sob estas circunstâncias, planos de Avaliação generalizáveis que se adequam a todos os tipos de decisão, dentro de categorias similares, tornam-se viáveis.

A Tabela F.3 que segue, apresenta a base conceitual a partir da qual a tipologia de decisões, apresentada por Stufflebeam et al (1971) foi gerada. Todas as decisões educacionais devem, então, ser classificadas exaustiva e não ambigualmente pelos:

- . fins pretendidos (objetivos)
- . meios pretendidos (planos de procedimentos)
- . meios reais (procedimentos em uso)
- . fins reais (obtenções)

TABELA F.3

TIPOS DE DECISÃO

	PRETENDIDO	REAL
FINS	Decisões de Planejamento para determinar os objetivos	Decisões de Reciclagem para julgar e reatuar as obtenções.
MEIOS	Decisões de Estruturação para determinar os procedimentos.	Decisões de Implementação para utilizar, controlar e aperfeiçoar procedimentos.

(Fonte: Stufflebeam, et al , 1971, página 80)

a) Decisões de Planejamento

As Decisões de Planejamento especificam as mudanças que devem ser processadas em um Sistema. A necessidade de decisões de Planejamento vem da:

- . conscientização da discrepância entre o que se pretende que o Sistema seja e o que ele é realmente.
- . conscientização da discrepância entre o que o Sistema pode tornar-se e o que é possível que se torne.

Em ambos os casos, as decisões podem ou não mudar tanto as pretensões quanto a realidade, relacionadas com os meios e os fins. Qualquer decisão para introduzir mudança deve resultar no estabelecimento de objetivos da mesma. Estes objetivos são baseados nas necessidades e problemas detetados no Sistema e também nas oportunidades que podem ser utilizadas para melhorá-lo.

No entanto, pode ser que o decisor se sinta seguro, com as informações que possui, para persistir com o Sistema tal como está.

b) Decisões de Estruturação

As Decisões de Estruturação especificam os meios de executar os fins estabelecidos como resultado das decisões de planejamento. Consideram variáveis tais como: método, conteúdo, organização, pessoal, tempo, facilidades e orçamento.

As 3 fontes para estruturar um plano de ação para atingir os objetivos desejados são:

- . as decisões de Planejamento que especificam o que deve ser executado.
- . a existência de estratégias alternativas disponíveis para atingir os resultados especificados.
- . pontos fortes e fracos das alternativas disponíveis.

É importante notar que decisões de Estruturação podem resultar na modificação dos objetivos estabelecidos. Pois, embora os objetivos sejam, inicialmente, baseados nas necessidades, problemas e oportunidades, eles podem ser e frequentemente são modificados devido a limitações realísticas dos meios disponíveis para assegurar seu cumprimento.

Um plano de ação baseado na decisão de Estruturação é uma exposição abrangente dos resultados a serem alcançados, do trabalho a ser realizado, dos recursos e tempo a ser utilizados. Os resultados especificados são aqueles fornecidos pelas decisões de Planejamento, possivelmente modificados pelas decisões de Estruturação na seleção de meios.

Quando a mudança desejada é uma inovação ou ainda não foi adequadamente testada, não podendo, portanto, ser instalada no Sistema, decisões Estruturais conduzem a algum tipo de fase de Ensaio ou Projeto Piloto. Entretanto, quando não é este o caso, as decisões Estruturais podem dispensar o Projeto Piloto e conduzir à instalação de procedimentos de mudança no sistema.

No todo, tais decisões fornecem as linhas de operação necessárias para responder efetivamente ao Planejamento.

As decisões que seguem, ou seja, de Implementação e de Reciclagem, só existem quando é necessário o desenvolvimento de um Projeto Piloto.

c) Decisões de Implementação

As decisões de Implementação são aquelas envolvidas no desenrolar do plano de ação. São provenientes de duas fontes:

- . conhecimento das especificações de procedimentos do Projeto Piloto, decididas na fase anterior - Estruturação.

- . conhecimento contínuo da relação entre especificações de procedimentos e procedimentos reais do Projeto Piloto.

As decisões de Implementação envolvem muitas escolhas com relação às mudanças de procedimentos no processo. Assim, envolvem decisões a respeito de retreinamento da equipe, busca de novos recursos, redesignação de responsabilidades, modificação de esquemas e similares.

d) Decisões de Reciclagem

Estas decisões são usadas na determinação da relação da consecução dos objetivos e a continuação, término, desenvolvimento ou modificação completa da atividade. O essencial neste tipo de decisão é o conhecimento da natureza e da ocasião em que os resultados especificados serão fornecidos.

Comumente, pensa-se que tais decisões ocorrem depois de um ciclo de uma atividade, mas esta é uma visão limitada. Concebidas mais apropriadamente, as decisões de Reciclagem ocorrem através de uma atividade como planos de qualidade ou controle do produto. É importante enfatizar que as decisões de Reciclagem referem-se a consecução em qualquer ponto de um Projeto e não só aos resultados de um ciclo completo do projeto.

Portanto, decisões de Implementação dependem da extensão na qual os procedimentos estão operando como pretendido, e decisões de Reciclagem dependem da extensão na qual os fins desejados estão sendo ou foram alcançados.

2.2 - SUB-UNIDADE 2 - TIPOS DE AVALIAÇÃO

Considerando os conceitos de situações e tipos de decisão, que caracterizam a tomada de decisão educacional, na lógica geral da abordagem proposta para a Avaliação, mostrada na Figura F.12, pág. F.79 são apresentados por Stufflebeam et al (1971), correspondendo aos

quatro tipos de decisão (Planejamento, Estruturação, Implementação e Re ciclagem), quatro tipos de Avaliação - Contexto, Input, Processo e Pro duto.

A Avaliação de Contexto serve às decisões de Planejamento para determinar os objetivos; a Avaliação de Input serve às decisões de Estruturação para determinar planos de procedimentos; a Avaliação de Processo serve às decisões de Implementação para controlar as operações do Projeto Piloto; e a Avaliação de Produto serve às decisões de Recicla gem para julgar e reatuar as obtenções do Projeto Piloto.

Agora serão vistos, mais detalhadamente, cada um dos qua tro tipos de Avaliação.

a) Avaliação de Contexto

A Avaliação de Contexto delinea, obtém e provê infor mações úteis com o propósito de fornecer uma lógica para a determinação de objetivos para mudança no Sistema, ou para a manutenção do "status quo".

A Avaliação de Contexto é sistemática, monitora o Siste ma, mantendo uma base constante de informação sobre ele.

Para isso, este tipo de Avaliação conta com 2 métodos bá sicos: método de Contingência e de Congruência.

1) Método de Contingência

A Avaliação de Contexto procura por oportunidades e pre ssões fora do Sistema imediato, para promover melhorias dentro dele. Oportunidades podem ser entendidas como novos programas educacionais, meios ou materiais instrucionais, estrutura e funcionamento de equipes, siste mas de financiamento, enfim tudo que estiver sendo utilizado por outros Sistemas que podem otimizar o Sistema que está sendo avaliado. Pressões

podem ser entendidas como valores e necessidades sociais, atitudes, prioridades, avanços tecnológicos, tendências políticas, desenvolvimento econômico, estatísticas populacionais, em suma, tudo o que está emergente na sociedade e que deve ser considerado pelo Sistema Educacional em questão. Isto envolve investigação exploratória dentro de limites livremente definidos, estudo e visitas a outros Sistemas, estudo da Literatura, consultores de fora, avaliações de valores da comunidade, valores, atitudes, prioridades, etc.. Envolve, também, pesquisa em termos de futuro considerando projeções de necessidades e valores sociais, avanços tecnológicos, tendências políticas, desenvolvimento econômico, estatísticas populacionais, etc..

Deste modo, o método de contingência provê o Sistema com informações para responder a situações que envolvem questões do tipo se-então, tais como:

- . Se há necessidade crescente de potencial humano para preencher papéis tecnológicos, o Sistema Educacional atual pode então atender esta necessidade?
- . Se os professores de um Sistema Escolar entrarem em greve, o que pode então ser feito para fornecer serviços educacionais ininterruptos à população estudantil?
- . Se se torna possível mecanizar totalmente o programa instrucional, quais serão então os benefícios educacionais e econômicos e, para que propósito os professores são necessários?
- . Se houver algum tumulto dentro de uma situação escolar, como então o Sistema Escolar será afetado?
- . Se o orçamento de uma escola (componente do Sistema) for aumentado em 30%, como ele será usado e no que resultará?

Assim, o método de contingência procura a discrepância entre o estado provável e o estado de um sistema em algum ponto no futuro.

2) Método de Congruência

O Método de Congruência compara o desempenho real do Sistema com o desempenho pretendido. Permite, assim, a monitoração do Sistema determinando se os objetivos estão ou não sendo atingidos como pretendido, mantendo uma base dinâmica de informações sobre o seu desempenho.

Este método possibilita a identificação de necessidades, não atendidas, através da informação de discrepância entre o real e o pretendido.

Com as informações de contingência e congruência, a Avaliação de Contexto faz, ainda, o diagnóstico dos problemas que impedem o atendimento das necessidades e a utilização das oportunidades.

Operacionalmente, os Sistemas ou Instituições Educacionais devem manter mecanismos de Avaliação de Contexto que forneçam dados de congruência e contingência. Dados os objetivos de um Sistema é importante investigar se ou quão bem estes objetivos estão sendo satisfeitos (congruência). Numa sociedade em mudança rápida, as necessidades de pessoas e localidades estão sujeitas a constantes mudanças, consequentemente, os sistemas devem perguntar constantemente as questões se-então (contingência) para determinar se objetivos e políticas devem ser mudados.

Os métodos de contingência e congruência são diferentes em muitos aspectos. Congruência, procura conformidade; Contingência, procura melhoria. No entanto, estas duas maneiras são, obviamente, complementares.

Muitas técnicas são úteis na condução da Avaliação de Contexto. As técnicas de Planejamento de Engenharia de Sistemas tais como a estrutura de quebra do trabalho encontrada no PERT (Program Evaluation and Review Technique) podem ser de considerável assistência

na conceitualização e descrição da natureza do Sistema a ser avaliado. Técnicas de Opiniões e Levantamentos Amostrais podem ser usados para determinar as percepções de vários grupos em relação às prioridades, oportunidades não usadas e necessidades não atendidas. Além disso, dados sobre necessidades e oportunidades podem ser obtidos, através do uso de testes padronizados, escalas de atitude, levantamentos diagnósticos, perfil de escolas, estudos e visitas a outros Sistemas, levantamentos de literatura de pesquisa, estudos contínuos de oportunidades e financiamento.

Técnicas analíticas simples podem ser usadas para interpretar os dados de Contexto - estas incluem retratos gráficos tais como gráficos de barras e perfis, comparações dos dados de testes padronizados com normas, comparação de resultados observados com um padrão absoluto especificado.

Resumindo, a Avaliação de Contexto localiza o componente do Sistema em que a mudança deve ocorrer, descreve as condições desejadas e reais pertencentes ao Sistema, identifica necessidades não atendidas e oportunidades não usadas e diagnóstica os problemas que impedem o atendimento das necessidades e a utilização das oportunidades. O diagnóstico dos problemas provê uma base essencial para a fixação de objetivos, cujo alcance resultará na melhoria do Sistema.

Exemplificando, o componente do Sistema em que a mudança deve ocorrer pode ser definido como sendo o Programa de Português nas quatro séries básicas das Escolas de 1º Grau do Estado. Um estudo de tal componente pode revelar que os níveis reais de aprendizagem de leitura das crianças está muito aquém do que o Sistema Escolar Estadual espera deles. Assim, estaria identificada uma necessidade, isto é, a Avaliação de Contexto teria revelado que os níveis de aprendizagem de leitura das crianças deve ser elevado. A próxima etapa seria a identificação das razões de tal necessidade, ou seja, o diagnóstico dos problemas. Os problemas devem ser entendidos como dificuldades potenciais

que impedem o alcance dos objetivos desejados e então provocam a existência de necessidades. No caso do exemplo do Programa de Português, eles poderiam ser operacionalizados através das seguintes questões:

- . Os alunos estão recebendo instrução adequada?
- . Os materiais instrucionais são adequados?
- . Há alguma barreira de linguagem?
- . Há um alto índice de ausência às aulas?
- . As expectativas com relação aos alunos não estão muito altas?

Até o momento, as informações foram obtidas pelo método de congruência (discrepância entre o real e o pretendido). As informações de contingência podem revelar a utilização de métodos ou materiais instrucionais ou treinamento de professores em outros Sistemas que podem melhorar o Sistema que está sendo avaliado.

O fornecimento destas informações aos decisores do Sistema, permitirá, a estes, a determinação dos objetivos da mudança (Decisões de Planejamento) para atender as necessidades e oportunidades.

A Tabela F.4 visualiza detalhes de informação pertinentes a Avaliação de Contexto.

Depois dos decisores terem determinado os objetivos, a Avaliação de Contexto, ainda os analisa para determinar a quantidade de mudança a ser causada e a quantidade de informação disponível para sustentá-la, ou seja, qual o tipo de situação de decisão (Homeostática, Incremental ou Neomobilística) envolvido. Com isso ela fornece uma base inicial para o próximo tipo de decisão - Decisão de Estruturação - indicando a necessidade ou não de uma Avaliação de Input, mais ou menos estruturada, para servir a este tipo de decisão, ou seja:

- . Se a situação de decisão identificada for Homeostática (muita informação e pequena mudança) pode-se passar direto às decisões de Estruturação, sem necessidade de uma Avaliação de Input estruturada, coletando informações para servir à determinação de pro

TABELA F.4
AVALIAÇÃO DE CONTEXTO

QUESTÕES DE PLANEJAMENTO	PERGUNTAS ORIENTADORAS DO DELINEAMENTO DA INFORMAÇÃO A SER OBTIDA PELA AVALIAÇÃO DE CONTEXTO.	INFORMAÇÕES DELINEADAS QUE DEVEM SER OBTIDAS E FORNECIDAS AOS DECISORES PELA AVALIAÇÃO DE CONTEXTO.	TÉCNICAS QUE PODEM SER USADAS PARA OBTER AS INFORMAÇÕES.
<ul style="list-style-type: none"> . Devemos mudar ou sustentar nossa linha de ação presente? . Quais são as necessidades prioritárias que nosso Projeto deve atender? . Se devemos mudar, quais os elementos que podemos utilizar, nesta mudança? . Que objetivos, orientados para a melhoria, devem ser focalizados a fim de que as necessidades encontradas no Sistema sejam satisfeitas? . Que Objetivos orientados para a melhoria receberão o endosso e o apoio da comunidade? . Quais os objetivos mais viáveis? . Que tipo de situação de decisão os objetivos do Projeto, para a melhoria, envolvem? 	<ul style="list-style-type: none"> . O desempenho real do sistema está de acordo com o desempenho planejado? . Que necessidades não satisfeitas existem dentro do sistema? . Quais as oportunidades existentes, fora do Sistema, que podem contribuir para a sua melhoria? . Quais são as características dos problemas que devem ser resolvidos no encontro das necessidades prioritárias a serem satisfeitas pelo Projeto? . Quais as percepções da comunidade, em que atua o Sistema, em relação às possíveis mudanças a serem introduzidas nele? . Quais os padrões que devem ser aplicados, sobre os objetivos, para a determinação de sua viabilidade? . Qual o nível de mudança envolvido nos novos objetivos e quanta informação existe disponível? 	<ul style="list-style-type: none"> . Os objetivos do Sistema e o desempenho real do Sistema e discrepância entre os dois. . Diagnóstico dos problemas que contribuem para a discrepância entre os objetivos do Sistema e o seu desempenho. . Informações relativas à operação de Sistemas similares. . Diagnósticos dos problemas que contribuem para a discrepância entre os objetivos do Sistema e o seu desempenho. . Descrições e análises, buscando um consenso, dos valores da comunidade relativos a objetivos possíveis que podem ser pensados como recursos para a melhoria do sistema. . Consenso a respeito dos padrões que devem ser considerados na determinação da viabilidade dos objetivos. . Nível de mudança de acordo com as percepções da Sociedade servida pelo Sistema. . Quantidade de informação dos decisores, o grau no qual os decisores supõem entender todos os aspectos do problema que eles enfrentam. 	<ul style="list-style-type: none"> . Análise de Sistemas. . Técnica de Congruência. . Estudo de Caso . Perfis de Escolas . Escalas de Atitudes . Entrevistas . Questionários de Opinião . Estudos e visitas a outros Sistemas. . Técnica de Contingência . Estudo de Caso . Perfis de Escolas . Escalas de Atitude . Entrevistas . Questionário de Opinião . Escalas de atitude . Questionários de opinião . Entrevistas. . Reunião com especialistas (equipe de decisão do sistema). . Estudo dos valores da comunidade relativos aos objetivos determinados para a melhoria do Sistema. . Reunião com Decisores.

cedimentos. Isto porque as informações que o decisor precisa, para as decisões estruturais estão disponíveis na própria Avaliação de Contexto. E, depois disto, nesta situação, a mudança pode ser imediatamente instalada no Sistema, sendo reajustado o mecanismo de Avaliação de Contexto (para avaliar, também, os objetivos da mudança introduzida no Sistema). A Avaliação de Contexto, então, continua a fazer o seu trabalho até que, novamente, sejam identificadas necessidades não atendidas ou oportunidades não utilizadas, levando à possibilidade de mudança no Sistema.

- . Se a situação de decisão identificada for Incremental (pouca informação e pequena mudança) é necessário que haja uma Avaliação de Input, para sustentar as decisões de Estruturação. Isto porque, as informações necessárias ao decisor, para as decisões de Estruturação, não estão disponíveis.
- . Se a situação de decisão identificada for Neomobilística (pouca informação e grande mudança) é necessário também que haja uma Avaliação de Input, para sustentar as decisões de Estruturação, pelo mesmo motivo colocado na situação anterior.

O tipo de Avaliação de Input necessário a estas duas situações (incremental e neomobilística) será detalhado no próximo tópico - Avaliação de Input.

b) Avaliação de Input

A Avaliação de Input delinea, obtém e provê informações úteis para determinar como utilizar recursos, para atingir os objetivos da mudança. Isto é conseguido através da identificação e Avaliação:

- 1) das capacidades relevantes da agência responsável;
- 2) das estratégias para executar os objetivos da mudança, e
- 3) dos planos para implementar a estratégia selecionada.

Esta informação é essencial para estruturar planos específicos de procedimentos para a consecução dos objetivos da mudança.

Portanto, a Avaliação de Input reúne as informações para a máxima utilização dos recursos em relação aos objetivos e avalia as estratégias relevantes, as capacidades e os procedimentos de implementação.

O produto final da Avaliação de Input é a análise de um ou mais planos de procedimentos em termos de custos e efetividade. Especificamente, planos alternativos são avaliados em relação a: equipes, tempo e despesas requeridas; dificuldades potenciais de procedimentos; as consequências da não superação destas dificuldades, e as possibilidades e custos para superá-las; relevância dos planos propostos para os objetivos da mudança; e a capacidade potencial dos planos para realizar os objetivos.

Essencialmente, este tipo de Avaliação provê informações para decidir:

- . Se outros auxílios são requeridos para atingir os objetivos;
- . Como os objetivos devem ser afirmados operacionalmente;
- . Que estratégia geral deve ser empregada (por exemplo, a adoção de soluções disponíveis ou o desenvolvimento de novas);
- . Que planos de procedimentos devem ser empregados para implementar a estratégia selecionada.

As propostas de projetos exemplificam os resultados das decisões baseadas em Avaliação de Input.

Enquanto a Avaliação de Contexto é principalmente sistêmica e macroanalítica, a Avaliação de Input é essencialmente específica, particular e microanalítica. Sua ênfase é dada aos objetivos para a mudança, estabelecidos devido a necessidades, oportunidades e problemas no Contexto. Sua função é determinar como melhor atingir os objetivos estabelecidos. Analisa os objetivos, procurando identificar as estratégias alternativas relevantes para o encontro dos mesmos.

A Avaliação de Input, mais detalhadamente, trata de informações como as que seguem:

- . operacionalização dos objetivos.
- . designação da responsabilidade pela seleção ou desenvolvimento de uma estratégia para a consecução dos objetivos.
- . estabelecimento de critérios a serem utilizados na seleção de uma determinada estratégia.
- . existência de estratégias potencialmente relevantes para a efetivação dos objetivos.
- . custo e efetividade das estratégias em julgamento.
- . características e efeitos da operação das estratégias em julgamento sob condições piloto.
- . viabilidade da implementação da estratégia em termos dos recursos disponíveis.
- . quantidade de treinamento exigido pela equipe antes que possa implementar efetivamente a estratégia selecionada.
- . procedimentos específicos a serem utilizados na implementação da estratégia selecionada.
- . esquema específico de atividades e eventos para guiar a implementação da estratégia.
- . plano de Avaliação de Processo exigido para uma implementação efetiva e eficiente da estratégia.
- . plano de Avaliação de Produto exigido para uma implementação efetiva e eficiente da estratégia.

Estas e outras informações similares indicam a importância da Avaliação dos dados de Input: se os objetivos do programa devem ser atingidos é essencial selecionar e estruturar estratégias, com o compromisso de alcançar satisfatoriamente os objetivos em um nível razoável de gastos. A Avaliação de Input é planejada para ajudar a equipe do Sistema a responder questões que são importantes na seleção e estruturação de planos de Projetos.

A metodologia de Avaliação de Input varia grandemente, dependendo do envolvimento de grande ou pequena mudança e disponibilidade de alta ou baixa quantidade de informação para sustentar a mudança.

É importante notar que as informações com as quais a Avaliação de Input trata são as mesmas em qualquer situação de decisão (Homeostática, Incremental ou Neomobilística). A diferença de metodologia é no que se refere à maneira de coletá-las em relação à identificação de estratégias e procedimentos a serem utilizados para a consecução dos objetivos.

A Tabela F.5 visualiza detalhes de informação pertinentes a Avaliação de Input.

Num conjunto Homeostático (pág. F.84, onde pequena mudança é necessária e onde muitas informações estão disponíveis para sustentá-la, exige-se, usualmente, pouca Avaliação formal, como no exemplo já apresentado à pág. F.84, em que para resolver o problema o Diretor da Escola X não precisou de uma quantidade significativa de dados de Avaliação de Input a respeito das estratégias de solução, ou seja:

- . conversar com os pais;
- . convocar os alunos
- . conversar com o dono do Parque de Diversões solicitando sua cooperação.

Conhecendo as pessoas envolvidas e como possivelmente elas reagirão a estas soluções, ele pôde, facilmente, selecionar uma mistura satisfatória de estratégias.

Contudo, quando um conjunto de mudanças é incremental ou neomobilístico, a situação é inversa.

No conjunto Incremental (pág. F.85, o decisor quer efetuar uma pequena mudança, mas ele possui pouco conhecimento a respeito da melhor abordagem para realizá-la. O decisor quer levar a pequenas

TABELA F.5
AVALIAÇÃO DE INPUT

QUESTÕES DE ESTRUTURAÇÃO (ORIENTADORAS DO DELINEAMENTO DA IN FORMAÇÃO A SER OBTIDA PELA AVALIA ÇÃO DE INPUT)	INFORMAÇÕES DELINEADAS QUE DEVEM SER OBTIDAS E FORNECIDAS AOS DECISORES PELA AVALIAÇÃO DE INPUT.	TÉCNICAS QUE PODEM SER USADAS PARA OBTER AS INFORMAÇÕES
<ul style="list-style-type: none"> . Os objetivos do Projeto estão de finidos operacionalmente, ou se já, de forma precisa e mensurável? . Quem será responsável pela seleção ou desenvolvimento de uma estratégia para efetivar os objetivos? . Que critérios serão usados na seleção de uma determinada estratégia? . Que estratégias já existem, potencialmente relevantes para a efetivação dos objetivos? . Qual o custo e efetividade das estratégias em julgamento? . Quais são as características e efeitos da operação das estratégias em julgamento, sob condições piloto? . É lógico acreditar que a estratégia selecionada pode realizar os objetivos? . A implementação da estratégia é realística em termos dos recursos disponíveis? . Quanto treinamento exigirá a equipe antes que possa implementar efetivamente a estratégia selecionada? . Que procedimentos específicos serão utilizados na implementação da estratégia selecionada? . Que esquema específico de atividade e eventos pode guiar a implementação da estratégia? . Que Plano de Avaliação de Processo é exigido para uma implementação efetiva e eficiente da estratégia? . Que Plano de Avaliação de Produto é exigido para implementação efetiva e eficiente da estratégia? 	<ul style="list-style-type: none"> . Afirmativas de julgamento dos responsáveis pelas decisões dos elementos da Avaliação. . Determinação de quais são os decisores responsáveis pela Estruturação, ou seja, determinação de procedimentos. . Listagem dos critérios a serem usados na seleção de uma estratégia. . Análise das Estratégias que estão operando em Sistemas Similares. . Descrição do Custo e Efetividade das Estratégias. . Dados comparativos relativos aos Custos/Efetividade das estratégias. . Afirmativas de julgamento de especialistas. . Análise dos requisitos de procedimentos, equipamentos, recursos humanos da estratégia, dos recursos à mão no Sistema e dos que podem ser procurados, e disponibilidade dos mesmos. . Requisitos de implementação da estratégia selecionada e possibilidades dos membros da equipe e discrepância entre as duas informações. . EDA, EDR, DFT e Cronogramas Mestres e Parciais. . Identificação de atividades e eventos do Projeto, desenvolvimento de uma rede de trabalho para mostrar as interrelações entre eventos e atividades e estimar o tempo para as atividades. . O Plano de Avaliação de Processo. . O Plano de Avaliação de Produto. 	<ul style="list-style-type: none"> . Revisão da Proposta de Objetivos através de Reunião com especialistas. . Reunião com os responsáveis pelo Sistema. . Reunião com os responsáveis pelo Sistema. . Estudo e visita a outros sistemas . Revisão de Literatura. . Estudo das informações relativas a cada uma das estratégias. . Estudo comparativo das informações relativas a cada uma das estratégias. . Reuniões com especialistas para estudo da estratégia selecionada. . Reunião com especialistas. . Estudo comparativo dos requisitos da implementação da estratégia selecionada e possibilidades dos membros da equipe. . Reunião com especialistas. . PERT/CPM . Revisão de Literatura sobre esquemas de Planos de Avaliação. . Reunião com os decisores do Sistema. . Revisão de Literatura sobre esquemas de Planos de Avaliação. . Reuniões com os decisores do Sistema.

mudanças, diferentes do "status quo", apenas em grau. Ele está procurando mais desenvolver seu Sistema do que restaurá-lo ou inová-lo. Nesta situação, a metodologia de Input consiste antes na análise do problema e aproximações sucessivas de uma solução, do que na consideração de todas as alternativas possíveis, tentando encontrar a melhor solução.

Tomando o exemplo já apresentado à pág. F.86, o método de Avaliação de Input na situação Incremental se tornará mais claro.

No exemplo apresentado anteriormente, diante da informação de Contexto indicando a má qualidade dos materiais instrucionais, como uma das grandes causas dos baixos níveis de aprendizagem de leitura no 1º Grau, o Secretário (o decisor) optou por um programa para melhorar a qualidade destes materiais, visto que ele não contava com recursos para atacar, também, outros problemas. Com esta decisão, ele trabalhará numa situação Incremental, com apenas uma variável e fazendo mudanças por ensaio e erro sem o perigo de riscos que poderiam ser dispendiosos.

Assim, a metodologia de Avaliação de Input, na situação Incremental, seguirá o seguinte caminho para a identificação da estratégia de solução:

- . identificação de um ou alguns tipos de materiais instrucionais potencialmente relevantes para a efetivação dos objetivos.
- . características e efeitos da operação do tipo de material identificado (um ou alguns) e em julgamento sob condições piloto.
- . viabilidade da implementação do tipo de material instrucional selecionado em termos dos recursos disponíveis.

Com estas informações e as especificadas na pág. F.101, o decisor partirá para a estruturação do plano de procedimentos, ou seja, um plano operacional detalhado especificando objetivos, métodos, esquemas, facilidades, equipes, planos de Avaliação de Processo e de Produto e Orçamento.

Na situação Incremental, o decisor poderá ou instalar diretamente a mudança no Sistema, quando ele tiver uma margem muito grande de segurança de que ela surtirá o efeito desejado, ou, poderá partir para um Ensaio ou Projeto Piloto que nesta situação deve ser intermitente, isto é, testa-se a alternativa escolhida, melhorando-a, testa-a novamente, até que possa ser instalada.

Note-se que, apesar do Secretário precisar de um avaliador de Input, a Avaliação não foi exaustiva, em termos de identificação de alternativas, e nem foi dispendiosa ou consumiu muito tempo. Contudo, ela, provavelmente, foi adequada, dado o conjunto incremental para a mudança.

No exemplo dado, a situação foi designada como Incremental. Isto não foi somente por causa das necessidades a serem satisfeitas e dos problemas a serem resolvidos, que requeriam somente pequenas mudanças, mas porque o Secretário decidiu trabalhar apenas com uma parte das necessidades e problemas. Ele limitou bastante o problema a ser atacado - se focalizou nos Materiais Instrucionais.

No conjunto Neomobilístico (pág. F.87, o decisor quer efetuar uma grande mudança, mas ele possui pouco conhecimento a respeito da melhor abordagem para realizá-lo. Nesta situação, a metodologia de Avaliação de Input consiste na análise do problema e na consideração de todas as alternativas possíveis, tentando encontrar a melhor solução.

Tomando o exemplo já apresentado à pág. F.87, o método de Avaliação de Input, nesta situação, se tornará mais claro.

Neste exemplo, com as informações do Contexto, o Secretário (o decisor) do Estado X tomou conhecimento de que o problema a ser atacado é o da formação de seus professores, e decidiu por um programa para melhorar o nível do professorado. Este programa deve implicar numa reestruturação completa do Ensino Normal e numa reciclagem dos professores já formados. Porém, o decisor tem pouco conhecimento a respeito de

estratēgias de soluçāo, ou seja, nāo sabe de que modo reestruturar o Ensino Normal e reciclar os professores jā formados.

Neste caso, com esta decisāo, o Secretārio trabalharā numa situaçāo de Decisāo Neomobilīstica. Ele estā propondo uma grande mudança e necessita de muita informaçāo sobre como o programa poderā ser melhor determinado e conduzido. Deste modo, a metodologia de Avaliaçāo de Input, na Situaçāo Neomobilīstica, para identificaçāo da estratēgia de soluçāo, deve ser a que segue:

- . identificaçāo de todas as estratēgias possīveis, potencialmente relevantes para a efetivaçāo dos objetivos, ou seja, maneiras de reciclar os professores e reestruturar o Ensino Normal.
- . caracterīsticas e efeitos da operaçāo de todas as estratēgias em julgamento, sob condiçōes piloto.
- . viabilidade da implementaçāo da estratēgia selecionada, em termos dos recursos disponīveis.

Com estas informaçōes e as especificadas na pāg. F.101, o decisor partirā para a estruturaçāo do plano de procedimentos, ou seja, um plano operacional detalhado especificando objetivos, mētodos, esquemas, facilidades, equipes, planos de Avaliaçāo de Processo e de Produto, e orçamento.

Na Situaçāo de Decisāo Neomobilīstica, o decisor poderā, tambēm, ou instalar a mudançā direto no Sistema, quando ele se sente seguro dos efeitos da mesma, ou partir para um Projeto Piloto ou Ensaio. É mais prudente, quando é grande a mudançā envolvida, que se faça o Ensaio.

Note-se que, nesta situaçāo, o decisor precisou de um avaliador de Input, a Avaliaçāo foi exaustiva em termos de identificaçāo de alternativas, foi tambēm dispendiosa e consumiu tempo.

É bom observar que existe certa dificuldade em classificar as situaçōes em termos de generalizaçāo, isto é, uma mesma situaçāo

pode ser classificada como Incremental numa sociedade e como Homeostática ou Neomobilística em outra, é preciso ter sempre em mente as percepções das pessoas envolvidas. E, dependendo da situação envolvida, a Avaliação de Input, obviamente, pode ser um empreendimento simples ou complexo. Evidentemente, os conceitos do tipo de situação de decisão (Homeostática, Incremental ou Neomobilística) são importantes em considerar como será o planejamento e a execução da Avaliação de Input.

As técnicas de coleta de informações para Avaliação de Input são deficientes no campo da Educação. Práticas correntes incluem apelos para a literatura profissional e serviços de consultores (assessoria). Muitas técnicas fora do campo da Educação também merecem investigação - o Sistema de Orçamento e Planejamento do Programa, a Técnica de Convergência de Opinião e análise Custo/Efetividade. A técnica PERT, já validada por Desmond L. Cook (Stufflebeam, et al 1971), tem utilidade conhecida: sua abordagem pela análise de um trabalho em uma estrutura de tarefas menores, ajuda a organizar uma estratégia em áreas do projeto, e a sua apresentação em rede de atividades, provavelmente é a melhor que existe para mostrar as interrelações entre atividades e eventos necessários para a consecução dos objetivos especificados.

Assim, como se pode ver pelos exemplos dados, decisões baseadas na Avaliação de Input, usualmente resultam nas especificações de procedimentos, materiais, facilidades, equipamentos, programas, esquemas organizacionais, requisitos de equipe e orçamento. E, estando decidida a necessidade do Ensaio, passa-se, então, para o Planejamento da Avaliação de Processo e Produto.

c) Avaliação de Processo

Uma vez que um plano de procedimento foi aprovado (Decisão de Estruturação), e começada a Implementação (Ensaio ou Projeto Piloto), a avaliação de Processo é necessária para prover realimentação periódica para pessoas responsáveis pelos procedimentos e planos de implementação.

A Avaliação de Processo tem três objetivos principais:

- 1) Detetar ou predizer defeitos no plano de procedimentos ou sua implementação durante os estágios de implementação.
- 2) Prover informações para decisões programadas.
- 3) Manter um registro dos procedimentos tal como eles ocorrem.

Hã três estratégias a serem seguidas na Avaliação de Processo. A primeira é identificar e controlar continuamente as fontes potenciais de falhas no Projeto. Estas fontes incluem, mas não são limitadas a, o que segue: relações interpessoais entre equipes e estudantes, canais de comunicação, logística, compreensão e agrado do propósito do Projeto por pessoas envolvidas e afetadas por ele, adequação de recursos, facilidades físicas, equipe e tempo de programa. A segunda estratégia, envolve planejamento e ajuda a decisões pré-programadas para serem feitas pelos decisores do Projeto, durante a implementação do mesmo. Um exemplo destas decisões pré-programadas é encontrado em Projetos de desenvolvimento de testes, onde é necessário escolher uma amostra específica de escolas, numa experiência com novos itens de teste. Tal decisão pode ser projetada antes do tempo real, quando a escolha está para ser feita, até que seja desenvolvida e enviada pelo correio uma lista de escolas, estas sejam contatadas, e sejam feitas determinações relativas a estas escolas que estão desejando participar do experimento. Então, o autor do teste deve selecionar aquelas escolas cujas combinações atendam melhor aos requisitos do plano de amostragem (uma decisão estruturada previamente), e ele deve planejar as especificações relativas à administração do experimento das formas do teste. Tais situações de decisão são numerosas em muitos projetos e denotam requisitos de dados de Avaliação de Processo explícita (de outro modo eles poderiam ser feitos durante a estruturação dos planos do Projeto). Se a Avaliação de Input e Decisões de Estruturação tiverem sido implementadas efetivamente, o avaliador de processo terá pequena dificuldade em delinear decisões pré-programadas e seus requisitos de informação associados. Tal delineamento formará a base para o plano de Avaliação de Processo original.

A terceira estratégia da Avaliação de Processo é registrar os principais feitos do plano de procedimentos do Projeto. Esta informação será especialmente útil, mais tarde, ao se determinar porque os objetivos foram ou não alcançados - o fato de certos objetivos não terem sido alcançados poderá ou não ser falha do plano de procedimentos ou do desempenho do mesmo.

Na Avaliação de Processo o avaliador essencialmente trabalha com instrumentos, para descrever o processo, encontros regulares com o pessoal envolvido no Projeto a atualização frequente dos dados do Plano de Avaliação de Processo. O avaliador de processo deve-se encontrar periodicamente com os decisores do Projeto, discutir com o pessoal do programa para identificar relações e resultados sobre atividades futuras do Projeto. Enquanto o avaliador de processo deve permanecer como espectador, na maioria das vezes, ele também deve prover informações sobre operações passadas, para identificar resultados relativos a operações futuras. Depois do pessoal de decisão ter identificado os resultados e questões, o avaliador de processo deve novamente proporcionar qualquer informação importante que ele tenha. Embora outros dados possam interessar, o avaliador deve reservá-los até o tempo apropriado. Durante as reuniões de realimentação da Avaliação de Processo, os decisores devem definir as condições de relevância para a realimentação da informação, e o avaliador deve fornecer informação dentro desta estrutura. O avaliador deve, também, ajudar o pessoal do Projeto a identificar as questões que necessitam ser respondidas na próxima reunião de realimentação; ele pode fazer isto, resumizando o plano de coleta de dados para o período de intervenção imediata e respondendo aos decisores se o plano proverá a informação que eles vão necessitar. Ou, então, o avaliador pode trabalhar mais a fundo com o pessoal do programa, para identificar as questões que eles querem respondidas e os dados que eles necessitam.

Na coleta de informação, o avaliador de processo deve contar com ambos os procedimentos - formal e informal - análises de interação, entrevistas, escalas de classificação, escalas de atitude, registros de reuniões de equipes e decisores, redes PERT atualizadas. Enquan

to a avaliador de processo deve-se focalizar em variáveis teoricamente importantes, e usar abordagens sistemáticas para controlá-las, ele deve também procurar por eventos não previstos mas significantes. Assim, ele deve usar quaisquer meios que estejam disponíveis para investigar os problemas identificados.

Dois pontos são importantes, com respeito à relação da Avaliação de Processo com outros tipos de Avaliação. Primeiro, a Avaliação de Processo é uma função da extensão na qual a Avaliação de Input e Contexto foram realizadas adequadamente - quanto mais adequadas as Avaliações de Input e Contexto, mais certo o decisor do Projeto pode estar de quão bem operará o seu plano de procedimentos, e menos crítica é necessária na Avaliação de Processo. Quando a lógica para os objetivos especificados (informação de Contexto) é vaga, e o plano de procedimentos (informação de Input) também o é, o Projeto, provavelmente, estará caminhando em direção a dificuldades e, talvez, ao insucesso. Para identificar e remediar, logo, os problemas durante o processo, é essencial ter realimentação contínua sobre como o Projeto está funcionando. Assim, neste caso, a Avaliação de Processo pode desempenhar uma função vital.

Segundo, a Avaliação de Processo é mais importante do que a Avaliação de Produto nos primeiros estágios do desenvolvimento do Projeto, mas esta tendência muda a medida que os planos do desenvolvimento do Projeto tornam-se menos exploratórios e mais estruturados. Há uma interdependência entre a Avaliação de Processo e de Produto. Se por um lado, a Avaliação de Processo é necessária, para ajudar na interpretação dos resultados; por outro lado, a necessidade para mudanças no processo presente (informação da Avaliação de Processo) não pode ser determinada sem conhecimento de que efeito elas estão produzindo (informação da Avaliação de Produto). Assim, decisões para reestruturar planos ou procedimentos, que estão baseados na Avaliação de Processo, devem ser sustentadas também por informações da Avaliação de Produto.

Em suma, a Avaliação de Processo delinea, obtém e relata a informação tão frequentemente quanto o pessoal do projeto requeira,

diariamente, se necessário, especialmente durante os primeiros estágios de um Projeto. Isto provê os decisores do Projeto não somente com informação necessária para antecipar e superar dificuldades de procedimentos, mas também com um registro de informações de processo para interpretar o logramento do Projeto.

A Tabela F.6 visualiza detalhes de informação pertinentes à Avaliação de Processo.

d) Avaliação de Produto

A Avaliação de Produto é a fase final de Avaliação de Projeto. Seu propósito é medir e interpretar o logramento do Projeto, não somente no fim de um ciclo, mas tão frequentemente quanto necessário, durante o período do Projeto.

Os resultados do Projeto podem ser classificadas em termos de:

1) Mudanças Pretendidas:

- . Mudanças em indivíduos - alterações nos alunos com relação aos 4 domínios de aprendizagem: informação verbal, domínio psicomotor, habilidades intelectuais e estratégias cognitivas (Oliveira, 1973).
- . Mudanças em relações sociais e Instituições - em termos de atitudes de alunos, professores e pessoal do programa, "backgrounds" de todos elementos envolvidos no programa.

2) Mudanças não Pretendidas

- . Frequentemente são geradas atividades que não foram antecipadas, ou seja, constantes no plano do procedimento original do Projeto. Se isto ocorre, os resultados destas atividades podem ter significado direto para a consecução dos objetivos do Projeto.

O método geral da Avaliação de Produto inclui a definição operacional de objetivos, critérios de medida associados com objetivos da atividade, comparação destas medidas com padrões relativos ou absolutos pré-determinados e elaboração de interpretações lógicas de resultados, usando as informações registradas por Contexto, Input, Processo.

O tratamento que os dados sofrem, em sua análise, deve ser ditado pelo tipo de dados que devem ser obtidos, o tempo e recursos disponíveis e necessidades dos decisores.

AVALIAÇÃO DE PROCESSO

QUESTÕES DE IMPLEMENTAÇÃO	PERGUNTAS ORIENTADORAS DO DESEMPENHO DA INFORMAÇÃO A SER OBTIDA PELA AVALIAÇÃO DO PROCESSO.	INFORMAÇÕES DELINEADAS QUE DEVEM SER OBTIDAS E FORNECIDAS AOS DECISORES PELA AVALIAÇÃO DE PROCESSO.	TÉCNICAS QUE PODEM SER USADAS PARA OBTER AS INFORMAÇÕES.
<ul style="list-style-type: none"> • Que atividades devem ser mudadas em função do tempo planejado? • Deve-se redesignar responsabilidades a equipe? • O esquema deve ser modificado? • Devem ser instituídos novos procedimentos? 	<ul style="list-style-type: none"> • O desempenho do Projeto, no tempo, está de acordo com o planejado? • A equipe deve ser reorientada ou reorientada para que possa concluir o ciclo do Projeto em andamento? • As facilidades e os materiais estão sendo usados adequadamente e apropriadamente? • Quais as principais dificuldades de procedimentos que precisam ser superados, durante o ciclo do Projeto em desenvolvimento? 	<ul style="list-style-type: none"> • Comparação das datas planejadas com as reais dos eventos já realizados no Projeto em desenvolvimento. • Relatório relativo à extensão na qual a equipe compreende seu papel, está motivada para desempenhá-lo e o está desempenhando. • Relatório relativo a extensão na qual os materiais e facilidades estão sendo usados da maneira e na quantidade planejada. • Relatório representando as percepções da equipe e participantes do Projeto, relativas às dificuldades que eles pensam existir e que devem ser superadas. 	<ul style="list-style-type: none"> • PERT/CPM. • Observação. • Entrevistas. • Observação em sala de aula. • Entrevistas. • Listagem do uso de materiais e facilidades. • Entrevistas. • Reuniões sistematizadas para discutir o Projeto. • Questionário de Opinião.

A Avaliação de Produto em termos de aprendizagem dos alunos pode ser feita utilizando-se, para a obtenção de dados, o esquema baseado nos procedimentos da Pesquisa Experimental ou o esquema descritivo baseado em Índices de efetividade (conforme a Unidade 5).

A Tabela F.7 visualiza detalhes de informação pertencentes à Avaliação de Produto.

2.3 - SUB UNIDADE 3 - O MODELO DE AVALIAÇÃO SEGUNDO STUFFLEBEAM ET AL

De acordo com a 1a. e 2a. sub unidades, Stufflebeam et al (1971), combinam quatro tipos de Avaliação e conceitos relacionados que constituem os elementos de um modelo generalizável. Deste modo, tem-se:

- a) Delineamento, obtenção e fornecimento de informação como os três passos mais importantes, dentro de cada um dos quatro tipos de Avaliação (Contexto, Input, Processo e Produto);
- b) Situação de decisão Homeostática, Incremental e Neomobilística formando as 3 classes de Situações de mudança, que determinam o tipo ou tipos de Avaliação mais adequados para delinear, obter e fornecer informação útil para a tomada de decisão;
- c) Os quatro tipos de decisão: Planejamento, Estruturação, Implementação e Reciclagem, para os quais existe uma correspondência direta dos quatro tipos de Avaliação: Contexto, Input, Processo e Produto, respectivamente.

A partir destes conceitos e formando a base para o modelo de Avaliação que será apresentado, estão as premissas que seguem:

- 1) Uma vez que o propósito da Avaliação é fornecer informações para a tomada de decisão, é necessário conhecer as decisões a serem tratadas;
- 2) Para a Avaliação ser relevante para a tomada de decisão, o avaliador deve ser orientado para as decisões a serem tratadas e para a sua função dentro desta orientação;

TABELA F.7

AVALIAÇÃO DE PRODUTO

QUESTÕES DE RECICLAGEM.	PERGUNTAS ORIENTADORAS DO DELINEAMENTO DA INFORMAÇÃO A SER OBTIDA PELA AVALIAÇÃO DE PRODUTO.	INFORMAÇÕES DELINEADAS QUE DEVEM SER OBTIDAS E FORNECIDAS AOS DECISORES. PELA AVALIAÇÃO DE PRODUTO.	TÉCNICAS QUE PODEM SER USADAS PARA OBTER INFORMAÇÕES.
<ul style="list-style-type: none"> . Os problemas foram resolvidos conforme se pretendia? . Os resultados do Projeto A foram melhores que os do Projeto B? . Até que ponto as necessidades dos alunos foram atendidas? . Houve um ganho significativo no desempenho dos alunos? . A efetividade do Projeto foi atingida com o custo especificado? 	<ul style="list-style-type: none"> . Os objetivos foram atingidos? . Que afirmativas probabilísticas podem ser feitas sobre a relação entre especificações e procedimentos e consequências do desempenho do Projeto? . Até que ponto as necessidades individuais dos alunos foram atendidas? . A efetividade do Projeto foi atingida com o custo especificado? 	<ul style="list-style-type: none"> . Comparação das medidas de alcance com o desempenho de um grupo de controle. . Inferência sobre a relação causal entre meios e dados de resultados. . O número de sucessos de cada estudante no Projeto, em termos de suas necessidades individuais. . Custo e efetividade do Projeto. 	<ul style="list-style-type: none"> . Comparação de medidas de alcance com padrões absolutos e relativos. . "Design" Experimental de Coleta de Dados. . Estudo de Caso de uma amostra aleatória de casos. . Análise Custo/Efetividade de.

- 3) Um modelo de Avaliação válido deve ser fundamentado nas conceitualizações das diferentes situações de mudança (Homeostática, Incremental e Neomobilística) e nas diferentes maneiras de abordá-las;
- 4) Os diferentes tipos de decisão (Planejamento, Estruturação, Implementação e Reciclagem) exigem diferentes tipos de estratégias de Avaliação (Contexto, Input, Processo e Produto), e um modelo de Avaliação generalizável e eficiente deve ser conceitualizado;
- 5) Enquanto diferentes estratégias de Avaliação variam no conteúdo, um simples conjunto de passos generalizáveis (delineamento, obtenção e fornecimento) pode ser seguido;
- 6) Para responder questões feitas pelos decisores, os planos para estudos de Avaliação devem satisfazer critérios de adequação científica (validade interna e externa, confiabilidade e objetividade), de utilidade prática (relevância, importância, escopo, credibilidade, oportunidade e penetração) e prudência (eficiência):
- 7) A tomada de decisão exige potencialmente informação avaliativa, assim, a relação entre Avaliação e tomada de decisão é simbiótica;
- 8) Desde que os requisitos da tomada de decisão são sujeitos a mudanças, os planos de Avaliação devem ser flexíveis e capazes de encontrar os requisitos de mudança.

Portanto, para a interpretação do modelo de Avaliação, Figura F.7, é necessário que se considere essas premissas.

Este modelo contém as relações básicas entre os elementos: atividades, Avaliação e decisões, que são inerentes a qualquer estudo de Avaliação.

Em seguida, para melhor entendimento do Modelo, serão resumidas algumas idéias já apresentadas.

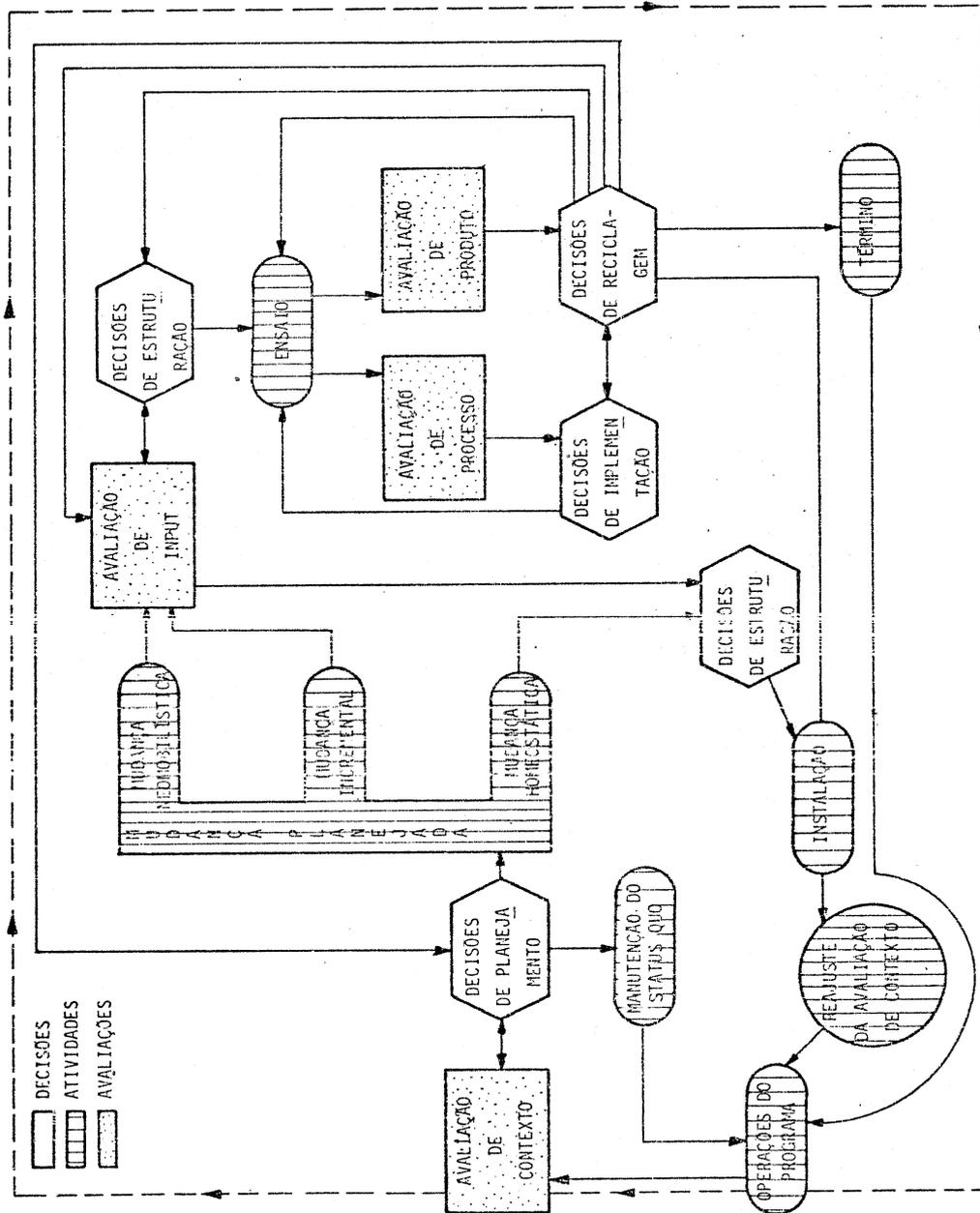


Fig. F. 13 - Um Modelo de Avaliação

O círculo mais de fora apresenta um mecanismo contínuo, sistemático que provê dados do Contexto, tanto de Contingência como de Congruência. Este mecanismo delinea, obtém e provê informações para o corpo de planejamento de um Sistema, para torná-lo capaz de tomar decisões tanto para mudar o Sistema, como para manutenção do "status quo", ou seja, continuar com os procedimentos presentes, porque eles estão servindo efetiva e eficientemente objetivos importantes. Se a Avaliação de Contexto não indica nenhuma discrepância entre o pretendido e o real ou entre o possível e o provável, o pessoal de planejamento pode se sentir confiante para continuar no mesmo nível. Entretanto, se a Avaliação indica a deficiência ou identifica algumas oportunidades não utilizadas para melhoria, um corpo de tomada de decisão racional decidirá pela execução de mudanças apropriadas. Tais mudanças podem ser de três tipos:

- 1) Mudança Homeostática: baseada em decisões para efetuar uma pequena mudança, sustentada por um alto nível de informações relevantes;
- 2) Mudança Incremental: baseada em decisões para efetuar uma mudança pequena, sustentada por um nível inicialmente baixo de informações relevantes;
- 3) Mudança Neomobilística: baseada em decisões para efetuar uma grande mudança sustentada por um nível inicialmente baixo de informações relevantes.

Dependendo do tipo de mudança resultante de planos de decisão, podem ser exigidas medidas avaliativas bem diferentes.

Em resposta à mudança homeostática, onde a informação adequada para sustentar a tomada de decisão já está disponível, em literatura de Pesquisa e/ou no mecanismo de Avaliação de Contexto, um estudo avaliativo para fornecer informação é redundante. Portanto, a figura que representa o modelo mostra que:

- . os decisores devem fazer decisões estruturais, relativas aos meios necessários para executar mudança homeostática, sem qualquer mecanismo de Avaliação de Input formal intervindo, a não ser aquele da Avaliação de Contexto;

- . as decisões estruturais vão conduzir diretamente à instalação da mudança no Sistema e subseqüente ajustamento do mecanismo de Avaliação de Contexto, para prover, através de controle rotineiro, o novo aspecto do Sistema, pela Avaliação de Contexto sistemática.

Se são exigidas mudanças neomobilísticas ou incrementais, mecanismo de Avaliação de Input especificados para sustentá-las são explicitamente necessários, desde que tanto o mecanismo de Avaliação de Contexto, quanto a literatura de pesquisa fornecem um suprimento de informação inadequado para embasar estes tipos de mudança.

Primeiro, um estudo de Avaliação de Input deve ser feito para que sejam identificadas, avaliadas estratégias e procedimentos, para que possam ser feitas as mudanças que se pretende. A informação avaliativa do Input deve auxiliar o decisor a tomar decisões ao estruturar procedimentos de mudança. Como se pode notar, Avaliações de Input neomobilisticamente orientadas são muito mais estruturadas do que Avaliações de Input incrementalmente orientadas. Para completar, decisões estruturais conduzem, usualmente, a algum tipo de fase de ensaio (piloto) porque a mudança desejada é ainda uma inovação e não foi adequadamente testada, portanto, não está pronta para ser instalada no Sistema total. Entretanto, quando não é este o caso, as decisões estruturais podem dispensar o Projeto Piloto e conduzir à instalação de procedimentos de mudança no Sistema.

As avaliações de processo e de Produto estão incluídas proximamente para ajudar em decisões relativas à fase de ensaio. A Avaliação de Processo provê informações para decisões envolvidas na implementação do ensaio, incluindo as modificações de decisões estruturais prévias, conforme a necessidade. A Avaliação de Produto ocorre, simultaneamente, através do ensaio em conjunção com a Avaliação do Processo e sustenta Decisões de Reciclagem, conduzindo a:

- . uma reformulação da mudança a ser executada, modificação tanto na estratégia como no procedimento;

- . ao término dos esforços de mudança, ou
- . à instalação da mudança no Sistema.

No caso de instalação, novamente o mecanismo de Avaliação de Contexto será ajustado, para permitir controle sistemático do novo elemento no Sistema e, com isso, a Avaliação de seu impacto.

É importante notar que a abordagem proposta por Stufflebeam et al (1971), é cíclica, isto é, a realimentação é continuamente fornecida ao decisor e, as novas informações, podem conduzir ao reexame de decisões anteriores. Assim, é possível, por exemplo, que a informação fornecida durante a Avaliação de Processo possa conduzir o decisor a reconsiderar uma decisão de estruturação que ele fez anteriormente e, então, iniciar uma nova Avaliação de Input ou revisar a que foi feita.

Porém, ainda que manejável na teoria, tal abordagem conduz o avaliador para uma tarefa bastante difícil na prática. Principalmente com relação à identificação de situações de decisão, que não existe, como se pode notar, uma forma operacionalizada para tal. E, também com relação ao próprio processo de tomada de decisão, pois, como se sabe, decisores são seres humanos e, não raras vezes, se vêem às voltas com a necessidade de tomar decisões político-administrativas, deixando de lado as informações fornecidas pelo avaliador. Estas são algumas restrições que sofre a teoria quando aplicada à situações particulares. Assim sendo, o avaliador, no seu trabalho, deve ousar criar, com base na abordagem adotada, procedimentos que superem, ou, pelo menos, minimizem as dificuldades advindas de restrições como as citadas.

Tendo em vista a definição de Avaliação que embasa o Modelo CIPP (Contexto, Input, Processo e Produto), ou seja, o processo de definir, obter e prover informação útil para a tomada de decisão, e considerando a suposição de que o papel técnico do avaliador se caracteriza pela obtenção de informações, as próximas unidades darão ênfase a este aspecto, isto é, as maneiras adequadas de obter informações. Assim, tem-se a seguir:

- 3.0 - UNIDADE 3 - : Obtenção de Informações - Instrumentais de Medida.
- 4.0 - UNIDADE 4 - : Obtenção de Informações - Avaliação Formativa de Materiais Impressos e de Alunos.
- 5.0 - UNIDADE 5 - : Obtenção de Informações - Avaliação Somativa dos Produtos da Mudança em termos de Aprendizagem.

3.0 - UNIDADE 3 - OBTENÇÃO DE INFORMAÇÕES: INSTRUMENTAIS DE MEDIDA

Quando se trabalha com pessoas, toma-se decisões continuamente. Um gerente de uma indústria decide a quem contratar, um médico decide qual o melhor tratamento para seu cliente, um professor decide se seus alunos estão aptos para estudar a próxima unidade. O decisor, que se baseia nas melhores informações antes de tomar decisões, obterá os melhores resultados.

A obtenção da informação corresponde ao papel técnico do avaliador, em que este usa os instrumentais de medida para obtê-la.

A medida consiste de regras para atribuir números aos objetos para representar quantidades de atributos (Nunnally, 1967). O termo regra indica que os procedimentos para designar números devem ser formulados explicitamente. O termo atributo indica que a medida refere-se a alguma característica de objetos. Por exemplo, não se mede um indivíduo, mas uma característica sua, como inteligência, criatividade e outras.

É possível que o atributo seja hipotético (por exemplo percepção extra-sensorial), ou ainda que uma mesma medida refira-se a uma mistura de atributos.

Os números são usados para representar quantidades. A quantificação refere-se à quantidade de um atributo representado em um objeto. A quantificação parte do pressuposto que se existe alguma coisa, existe em alguma quantidade e se existe em alguma quantidade, pode ser medida.

A vantagem da quantificação é que:

- . os índices numéricos tornam possível relatar os resultados em detalhes maiores do que seria em julgamento qualitativos;
- . permite o uso de métodos poderosos de análise matemática;

Como dito anteriormente, usa-se instrumentais para obter informações para a tomada de decisões. Os instrumentos auxiliam na tomada de muitos tipos de decisões (Cronbach, 1970), incluindo:

- . seleção e classificação de indivíduos;
- . Avaliação de procedimentos educacionais ou tratamentos;
- . rejeição ou não de hipóteses científicas.

Na seleção, decide-se aceitar alguns indivíduos e rejeitar outros. Admitir um calouro em uma Universidade, quando o número de vagas é limitado, é uma questão de seleção.

Para a decisão de classificação, determina-se qual designação ou tratamento, um indivíduo deve receber. Quando uma pessoa é designada a diferentes níveis de trabalho tem-se uma decisão de classificação.

Na Avaliação de tratamentos, toma-se decisões sobre sua efetividade. Quando se administra um teste de desempenho de Português, pode-se testar a instrução. Se os resultados forem baixos, deve-se, provavelmente, alterar o método.

Os testes são também usados para medir resultados de experimentos científicos a fim de verificar as hipóteses levantadas.

As informações podem ser coletadas de forma escrita (através de uso de testes escritos) ou oral (entrevistas), através da aplicação dos instrumentais em um grupo ou em um indivíduo de cada vez. Ainda, tais informações podem estar ligadas com o domínio afetivo do

indivíduo (atitudes, interesses) ou cognitivo (aprendizagem, inteligência) ou uma mistura dos dois domínios (opinião).

Serão discutidas, nesta unidade, algumas técnicas úteis de obtenção de informações. Elas serão limitadas ao uso e elaboração de:

- . questionários e entrevistas;
- . escalas de atitude;
- . testes de aprendizagem.

Com relação aos instrumentais de medida, deve-se decidir ou sobre a utilização dos existentes ou no desenvolvimento de novos, para uma dada situação.

A seleção de instrumentos envolve a verificação, dentre os existentes, daquele que é congruente com os objetivos da medição. Muitas vezes, após a seleção, é necessário fazer adaptações nele. No caso de se precisar de muitas reformulações, é preferível desenvolver um instrumento novo.

Deve-se lembrar que a importância do uso dos instrumentais está no fato de se obter informações úteis e confiáveis para a tomada de decisões.

3.1 - SUB UNIDADE 1 - QUESTIONÁRIO E ENTREVISTA

Esta sub-unidade é baseada em Selltiz et al (1974).

Os questionários e entrevistas são usados para transformar os "dados" fornecidos diretamente pelos indivíduos em "informações". Estas técnicas tornam possível medir:

- . o que o indivíduo sabe (conhecimento ou informação);
- . do que a pessoa gosta ou desgosta (valores e preferências);
- . o que a pessoa pensa (atitudes e crenças).

Podem ser usados, também, para descobrir a ocorrência de experiências (biografia) e o que está ocorrendo no momento.

Tanto o questionário quanto a entrevista conseguem descrições verbais do indivíduo, apenas do que ele pode ou está disposto a relatar. Muitas autodescrições verbais são aceitas como válidas, mas em muitas situações, pode-se duvidar delas, principalmente quando as pesoas tendem a mostrar uma imagem diferente do que realmente são. Alguns não conseguem descrever, abertamente, suas crenças, sentimentos, motivações, por não serem conscientes de muitos deles e não saberem fazer um julgamento a seu próprio respeito. O questionário e a entrevista confiam na validade das autodescrições.

O que se obtém em um questionário é limitado às respostas escritas e questões pré-estabelecidas.

A entrevista permite maior flexibilidade para a obtenção de informações, uma vez que estão reunidos o entrevistador e o entrevistado, no momento em que as perguntas são apresentadas e respondidas.

CONTEÚDO DO QUESTIONÁRIO E ENTREVISTA

Quando o conteúdo do questionário ou entrevista é dirigido para a verificação de fatos, a melhor maneira de obtê-los é contatar-se diretamente com pessoas que estão em condições de conhecê-los e que têm acesso a eles, para que dêem descrições destes.

A aquisição de informações sobre os fatos precisa ser avaliada através de sua credibilidade. Uma boa fonte de compreensão dos fatores que influem na credibilidade são as regras de prova, através de perguntas:

- como foi que o indivíduo tomou conhecimento do fato?
- que motivos o indivíduo pode ter para descrevê-lo?
- até que ponto a lembrança que a pessoa tem do fato é exata?

No caso de se estar interessado na descrição ou compreensão de uma situação, é possível verificar a exatidão da informação através da comparação com descrições feitas por vários indivíduos. Se os

respondentes ocupam posições diferentes e concordam com uma determinada afirmação, a base para aceitá-la, como verdadeira, é muito maior. Mas, se há contradições entre as descrições, isto pode levar a orientações para outros estudos, que podem confirmá-las ou não.

Quando o conteúdo do questionário ou entrevista é dirigido para a verificação de crenças quanto aos fatos, deve-se encaminhá-lo aos indivíduos para saber o que eles acreditam que sejam os fatos. O que se pretende é obter uma imagem de suas crenças e não verificar o que é verdade objetivamente. É importante saber, em primeiro lugar, se o respondente tem qualquer crença relativa ao fato e só depois arguir sobre a natureza de sua crença.

No caso do conteúdo estar relacionado à verificação de sentimentos, a melhor maneira é formular perguntas que dêem ao indivíduo, grande liberdade para resposta, pois, as reações emocionais são, frequentemente, complexas. As perguntas que pedem respostas simples, supõem que os indivíduos reajam com a mesma emoção diante de elementos de um grupo independente da situação em si e de sua relação com os elementos do grupo. Mas, isto é discutível, pois, é necessário saber não só os seus sentimento, mas também as circunstâncias em que estes aparecem.

As respostas de uma pessoa quanto ao comportamento adequado, em várias situações sociais, é de interesse para o conhecimento de padrões de ação, tanto como reflexo do clima predominante de opinião, como para prever o seu comportamento em tais situações. As perguntas podem referir-se a dois componentes:

- padrões éticos do que deve ser feito;
- considerações práticas relativas ao que é possível fazer.

Em se tratando de conteúdo dirigido a comportamento atual ou passado, a melhor maneira de obtê-los é através de perguntas específicas. O comportamento de uma pessoa em determinada situação é uma indicação de seu comportamento futuro, em situações semelhantes.

Quando o conteúdo está relacionado a razões conscientes de crenças, sentimentos ou comportamento, as respostas não são simples. A própria pergunta deve considerar os vários fatores que possam influir na crença, comportamento.

TIPOS DE ENTREVISTAS E QUESTIONÁRIOS

A forma de entrevistas e questionários pode variar muito. As entrevistas podem ser rigidamente padronizadas e completamente assistemáticas.

a) Tipo padronizado

No caso de entrevista e questionário padronizado, as perguntas são dirigidas, a todos os indivíduos, exatamente com as mesmas palavras, obedecendo a mesma sequência, de modo que todos respondam à mesma pergunta, podendo, as respostas, ser comparáveis.

Quanto à sistematização das perguntas, pode-se apresentar respostas de alternativas fixas ou abertas. As alternativas fixas limitam a resposta às alternativas apresentadas. As perguntas abertas permitem que o indivíduo dê uma resposta livre de acordo com o seu quadro de referências.

1) Vantagens e desvantagens de perguntas de alternativas fixas:

Vantagens

- . questões padronizadas;
- . questões facilmente aplicáveis;
- . a análise é rápida e pouco dispendiosa;
- . as alternativas ajudam a esclarecer o sentido da pergunta.

Desvantagens

- . pode haver imposição de apresentação de opinião, em questões que o individuo não tem nenhuma opinião;
- . pode haver omissão de uma alternativa possível.

2) Vantagens e desvantagens de perguntas abertas

Vantagens

- . as questões dão abertura para uma amplitude de respostas.

Desvantagens

- . pode haver dificuldade de entendimento de conceitos envolvidos na questão;
- . há necessidade de treinamento aos codificadores;
- . a análise é difícil e dispendiosa.

Seria possível sanar o problema de omissão de alternativa (no caso de alternativa fixa) se se aplicasse o instrumento em uma amostra da população e com os resultados, colocar alternativas mais significativas.

A entrevista ou o questionário podem ter, exclusivamente, perguntas abertas ou fechadas ou ainda uma combinação delas.

b) Tipo assistemático

Não existe questionário assistemático por causa da rigidez de sua estrutura.

A entrevista assistemática é usada para assuntos em que se exige uma abordagem mais flexível, como em estudos mais intensivos de percepção, atitudes. A entrevista assistemática exige mais habilidade do entrevistador, e alcança seu objetivo se as respostas dos indivíduos forem espontâneas, concretas e específicas.

Como o entrevistador tem a liberdade de acrescentar perguntas ou explicá-las, os resultados da entrevista não são comparáveis com os de outros correspondentes, e ainda tem a desvantagem de a análise ser mais difícil e demorada.

FORMATOS DE QUESTÕES PARA ENTREVISTAS E QUESTIONÁRIOS

Há diversos formatos que podem ser utilizados na elaboração de questões para entrevistas e questionários:

- . questões diretas X indiretas;
- . questões específicas X inespecíficas;
- . fatos X opiniões
- . questões X afirmativas.

A diferença entre questões diretas e indiretas está em quão óbvia é a questão, na solicitação de informação. A abordagem indireta é mais promissora para produzir respostas francas e abertas. Exemplos:

- Você gosta de seu trabalho? (direta)
- O que você acha de seu trabalho? (indireta).

As questões específicas tratam de um objeto, pessoa, ou idéia que está sendo solicitada através de uma atitude, crença ou conceito. As questões inespecíficas tratam de uma área mais geral. As questões específicas podem levar o indivíduo a tornar-se mais cauteloso e dar respostas menos honestas. Exemplos:

- Você gosta das obras de José de Alencar? (específica).
- Você gosta das obras do movimento romântico? (inespecífica).

Nas questões que se exige que o indivíduo forneça informações sobre fatos, tem-se a desvantagem de se contar com falhas de memória ou desejo consciente de criar uma impressão particular. As questões relativas a opiniões, pode levar o respondente a distorcê-las, em função de conveniência social. Exemplos:

- Qual é a marca de seu carro? (fato)
- Você prefere Fiat ou Volkswagen? (opinião).

Em muitos casos, pode-se construir questões diretas aos respondentes, ou fornecer-lhes afirmações e perguntar-lhes se concordam ou discordam (ou se acham se são falsas ou verdadeiras). Em medidas de atitude é mais comum ver afirmativas do que questões. Não existe uma base para distinguir, entre estes dois formatos, o potencial em provocar respostas honestas. Exemplos:

- Você acha que o treinamento em serviço forma bons profissionais? (questão)
- O treinamento em serviço forma bons profissionais (afirmativa).

MODOS DE RESPOSTA DE QUESTIONÁRIO E ENTREVISTA

Há uma variedade de modos de resposta de questionários e entrevistas:

- a - resposta não estruturada - permite que o indivíduo dê a sua própria resposta, da forma que escolher.

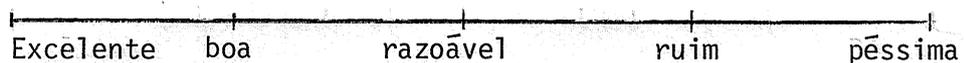
Exemplo: Descreva o que você acha do professor X

- b - resposta de preencher - exige que o indivíduo construa uma resposta, mas limitada a uma palavra ou frase.

Exemplo: Qual é a sua profissão? _____

- c - resposta escalada - pede que o indivíduo expresse acordo ou rejeição de uma afirmativa de atitude ou algum aspecto de si

Exemplo: Sua juventude foi:



- d - resposta de classificação - exige que o sujeito ordene uma série de afirmativas, em termos de um critério particular.

Exemplo: Ordene as afirmativas segundo sua utilidade na apren
dizagem de levantamento de hipóteses de pesquisa. Use
números de 1 a 5 indicando da menos a mais útil. As
afirmativas que não tiverem nenhuma utilidade corres
pondem a zero.

- _____ aula expositiva pelo professor;
- _____ discussões em pequenos grupos;
- _____ estudos individuais em livros, revistas e
apostilas;
- _____ atividades semanais;
- _____ estudos orientados pelo professor.

e - resposta do tipo múltipla escolha, isto é, respostas com vã
rias alternativas - o indivíduo deve selecionar uma das alter
nativas que lhe são apresentadas.

Exemplo: Assinale uma das alternativas:

Eu tenho recebido estímulo profissional de:

- _____ chefe;
- _____ outros profissionais do próprio local;
- _____ profissionais de outros locais;
- _____ colegas;
- _____ livros.

f - resposta categórica - o indivíduo deve marcar uma entre duas
alternativas, para cada item.

Exemplo: Você é católico? Sim _____ Não _____

VANTAGENS E DESVANTAGENS DE QUESTIONÁRIO E ENTREVISTA

A Tabela F.8 visualiza as vantagens e desvantagens da
utilização de questionários e entrevistas.

VANTAGENS E DESVANTAGENS DE QUESTIONÁRIO E ENTREVISTA

	Q U E S T I O N Á R I O	E N T R E V I S T A
VANTAGENS	<ul style="list-style-type: none"> . não exige muita habilidade para sua aplicação, podendo ser entregue pessoalmente ou pelo correio, com poucas explicações; . pode ser aplicado a um grande número de pessoas ao mesmo tempo; . processo pouco dispendioso; . permite o anonimato do indivíduo testado; . permite mais tempo de reflexão para resposta. 	<ul style="list-style-type: none"> . permite que o entrevistador repita as questões ou as apresente de maneira diferente, ou ainda que faça ou tras perguntas a fim de esclarecer o sentido de uma resposta; . atinge quase todos os segmentos da população; . oferece condições de notar a validade de descrições, pois o entrevistador pode observar o que o entrevistado diz e como o diz, podendo verificar afirmações com tradicionais.
DESVANTAGENS	<ul style="list-style-type: none"> . é aplicado somente em indivíduos alfabetizados; . há poucas devoluções (quando enviado pelo correio) podendo enviezar a amostra inicialmente planejada; . permite interpretações diferentes pelos indivíduos testados, levando, conseqüentemente, a respostas confusas. 	<ul style="list-style-type: none"> . processo dispendioso; . exige mais habilidade para a aplicação; . deve ser feita a cada indivíduo separadamente; . pressiona para uma resposta imediata; . a personalidade de diferentes entrevistadores influencia de forma diversa, na situação de mensuração; . embora o entrevistador possa assegurar anonimato, o entrevistado pode duvidar disto.

ELABORAÇÃO DE QUESTIONÁRIO E ENTREVISTA

O primeiro passo na elaboração desses instrumentais é a definição das variáveis envolvidas no estudo, seguido da decisão do uso ou de questionário ou entrevista, levando-se em conta suas vantagens e desvantagens.

A escolha do formato de questão depende do que se está tentando medir, ou seja, fatos, atitudes, preferências.

Quanto a escolha do modo de resposta, não há regras específicas. Em muitos casos, o tipo de informação que se procura, determina o modo de resposta mais adequado. A decisão deve basear-se:

- . tipos de dados desejados, isto é, a maneira com que os dados serão tratados;
- . flexibilidade de resposta;
- . facilidade de atribuição de pontos.

VALIDADE E FIDEDIGNIDADE DE QUESTIONÁRIO E ENTREVISTA

A validade refere-se à extensão em que se mede o que se pretende mensurar.

A confiabilidade está ligada à precisão com que se mede o que se estiver mensurando. O instrumento realiza sua função de medida, consistentemente.

Embora deva-se interessar pela validade e confiabilidade de questionário e entrevista, poucos as verificam, principalmente porque estes instrumentos são elaborados para uma única situação de decisão.

Para garantir a validade destes instrumentos, é importante pré-testá-los em uma sub amostra representativa da população alvo, a fim de verificar o entendimento de direções, linguagem, tempo de aplicação. O resultado da pré-testagem leva ao aperfeiçoamento dos instrumentos em

Atitudes, referem-se a sentimentos sobre objetos sociais particulares, isto é, objetos físicos, tipos de pessoas, Instituições Sociais.

Exemplo: A utilização de Tecnologia Educacional facilita a aprendizagem.

As medidas de atitudes, em particular, são as mais usadas. Pode-se medir atitudes com relação ao projeto, professores, materiais e outros.

Há várias abordagens para a medida de atitude, mas a mais usada é chamada escala de atitude, em que se pergunta diretamente ao indivíduo. É mais comum empregar método de resposta absoluta (escala de classificação), isto é, a resposta do indivíduo é relacionada diretamente com o objeto em estudo (atitude em questão) e não em comparação com outros.

O que se obtém com as escalas de atitude é limitado ao que o indivíduo sabe sobre suas atitudes e que deseja relatar.

As escalas de classificação podem ser gráficas, numéricas ou gráfico-numéricas.

Na escala gráfica, o indivíduo deve colocar um sinal no ponto adequado, em uma linha que vai de um extremo ao outro.

Exemplo de um item: A duração de cada aula deve ser de 45'

discordo _____ concordo
plenamente plenamente

Na elaboração dos itens com escalas gráficas, deve-se:

- . evitar afirmações finais muito extremas;
- . colocar as afirmações de modo que correspondam diretamente aos pontos na escala.

Na escala numérica, o indivíduo recebe uma série de afirmativas (itens) em que deve colocar um número. Os números são definidos anteriormente.

Exemplo: 1 - discordo plenamente
2 - discordo a maioria das vezes
3 - discordo ligeiramente
4 - concordo ligeiramente
5 - concordo a maioria das vezes
6 - concordo plenamente.

Exemplo de um item: Enumere cada afirmativa conforme o seu grau de acordo: As tarefas escolares devem ser sempre individuais —

São usadas, também escalas gráficas com números.
Exemplo: A retroalimentação do professor deve ser constante.

discordo _____ concordo
plenamente 1 2 3 4 5 6 plenamente

É melhor usar a escala gráfica ou gráfica numérica em vez de escala numérica por muitas razões:

- . frequentemente, as pessoas acham que as quantidades são representadas por graus de extensões físicas. A escala gráfica-numérica dá uma idéia de uma classificação contínua;
- . a escala numérica, quando coloca o significado dos números uma única vez, no início da escala, tem a desvantagem de o indivíduo precisar lembrar do significado dos mesmos. E ainda, a escrita do número, pelo indivíduo, no espaço em branco, pode causar a necessidade de se decifrar o número. A escala gráfica não permite tais problemas, pois é ancorada em seus extremos.

Com relação as escalas, em muitos casos, deve-se decidir qual o seu número de passos. Por exemplo: Para a afirmativa "Os estudos devem ser sempre em grupo" pode-se usar escalas a partir de dois passos, como:

discordo _____ concordo
plenamente plenamente

ou:

discordo _____ concordo
plenamente 1 2 3 4 5 6 plenamente

Muitas pesquisas já verificaram que quanto maior o número de passos, na escala, mais confiável ela é. A única exceção ocorre quando se emprega mais de vinte passos, pois pode confundir o indivíduo.

É preferível usar número de passos pares, na escala. O número de passos ímpares permite o uso do passo do meio, que significa posição neutra, ou nem concorda, nem discorda. O indivíduo que assinala a posição neutra pode realmente ser neutro com relação à atitude em questão ou pode não querer participar do estudo.

Existem vários tipos de escalas de atitudes:

- a) Escalas diferenciais (Thurstone)
- b) Escalas somatórias (Likert)
- c) Escalas cumulativas
- d) Escolha Q
- e) Diferencial semântico

a) Escalas diferenciais

A escala consiste de várias afirmativas (itens), geralmente vinte, em que o indivíduo deve assinalar cada afirmativa que concorda, ou os dois ou três itens que mais se aproximam de sua opinião.

A escala diferencial que é a mais usada é a do tipo Thurstone, que utiliza a escala de intervalo (distâncias numéricas aparentemente iguais).

As escalas diferenciais preocupam-se com a posição das pessoas na escala.

A escala tipo Thurstone é elaborada da seguinte maneira:

- . reúne-se várias afirmativas (itens) relativas à atitude que se quer medir;
- . um grande número de juizes (50 a 300), trabalhando independentemente, classificam as afirmativas em 11 grupos (Selltiz et al., 1974), desde as mais favoráveis à atitude (1º grupo) até as mais desfavoráveis (11º grupo);
- . calcula-se, para cada item, o valor mediano, na escala. As afirmativas que têm grande dispersão, pelo julgamento dos juizes, são eliminadas ou por serem ambíguas ou não significativas. A seleção final dos itens é feita considerando-se a distribuição uniforme dos itens pela escala.

Ao dispor os itens no questionário, os valores dos itens, atribuídos pelos juizes, não aparecem. A disposição dos itens é casual.

Exemplo (Selltiz et al 1974) referente a atitude com relação aos nativos da África do Sul, retirado do estudo de MacCrone:

Valor na escala	Item	
3,8	22	Acho que a comunidade branca desta país tem uma dívida real de gratidão para com os missionários, pela maneira com que tentaram erguer o nativo.
0,8	11	Prefiriria ver o povo branco perder sua posição neste País a mantê-la à custa da injustiça para com o nativo.

Não existe razão para se esperar, "a priori", que todos os indivíduos tenham a mesma atitude com relação às mesmas coisas, ou que as dimensões das atitudes sejam as mesmas para todos os indivíduos. Mas, se as respostas de uma pessoa se dispersam muito, o resultado da atitude tende

a não ter o mesmo significado de um resultado com pouca dispersão. A dispersão é indicadora ou que o indivíduo não tem atitude, ou que sua atitude não se organiza da forma que a escala supõe.

Tem-se feito críticas com relação às escalas tipo Thurstone:

- . quantidade de trabalho necessário para a sua elaboração;
- . o escore de um indivíduo é a média ou mediana dos valores, na escala, dos itens que ele assinala. Então, atitudes totalmente diversas podem exprimir o mesmo escore;
- . até que ponto os valores na escala são livres das atitudes dos próprios juizes que os atribuem?

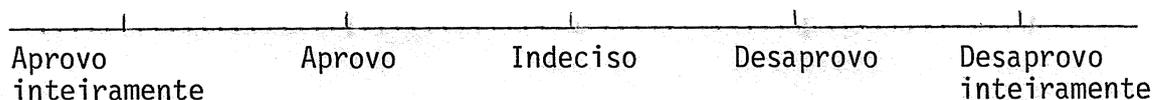
b) Escalas somatórias

Consiste de uma série de itens que parecem nitidamente favoráveis ou desfavoráveis ao objeto, devendo cada indivíduo assinalar seu acordo ou desacordo para cada um.

A soma dos escores obtidos em cada item representa o escore resultante do indivíduo, na escala total, por isso o nome escala somatória.

A escala somatória mais usada é a do tipo Likert que consiste de uma série de itens com vários graus de acordo ou desacordo, em que se supõe que o intervalo, entre os pontos, é igual. Apesar de Likert ter usado cinco graus de acordo-desacordo, alguns pesquisadores usam número menor enquanto que outros, maior.

Exemplo (Selltiz et al 1974) retirado de Murphy e Likert. Uma pessoa que ama seus semelhantes deve recusar-se a participar de qualquer guerra, por mais sérias que sejam as consequências disso para o seu País.



verificada por especialistas. A validade de conteúdo é conceituada detalhadamente no tópico "Validade", página F.179.

A precisão é ligada com os erros introduzidos durante o processo de testagem.

Um dos erros mais comuns, introduzidos pelos julgadores, é o efeito de halo, isto é, a tendência que se tem, no caso de se julgar mais de uma característica de um objeto, de transferir, de uma Avaliação para outra, uma impressão generalizada (positiva ou negativa) dele.

A generosidade e a severidade também introduzem erros, pois o julgador pode superestimar ou subestimar características do que está sendo julgado.

A tendência também de assinalar o ponto do meio introduz um outro tipo de erro.

d) Escolha Q

A escolha Q foi desenvolvida, embasada em uma metodologia elaborada por Stephenson (Nunnally, 1967), partindo do princípio que é mais importante basear-se em comparações entre diferentes respostas entre pessoas do que entre pessoas.

O indivíduo recebe um grande número de afirmações significativas ao estudo, em cartões isolados. O indivíduo deve compará-las para colocar sua resposta. O método é, portanto, de respostas comparativas. O número de cartões é determinado por conveniência de necessidades estatísticas: a distribuição deve ser aproximadamente normal, forçando que todos os indivíduos tenham a mesma média, o mesmo desvio padrão e a mesma forma de curva.

O indivíduo deve separar os cartões em um certo número especificado de pilhas (9 ou 11), segundo algum critério. O objetivo da se

O respondente indica sua opinião fazendo uma marca, na es
cãla, acima da palavra escolhida.

O processo de construção da escala tipo Likert consiste do
seguinte:

- . reúne-se um grande número de itens significativos e relacionados
ã atitude em questão, e que sejam explicitamente favoráveis ou
desfavoráveis;
- . aplica-se os itens em uma amostra representativa de indivíduos da
população-alvo;
- . avalia-se as respostas de cada item. A resposta que indica a ati
tude mais favorável recebe o resultado mais elevado;
- . calcula-se o resultado total, de cada indivíduo, nos itens;
- . analisa-se as respostas a fim de verificar quais itens discrimi
nam mais claramente entre os indivíduos que recebem escores al
tos e baixos, na escala total. Os itens que não apresentam uma
alta correlação (grau de relação entre duas variáveis, no caso,
item e escore total) são eliminados.

A escala Likert permite o uso de itens que não estão liga
dos explicitamente ã atitude em questão, porque o importante é a verifi
cação empírica da coerência entre qualquer item e o escore total da esca
la.

A escala ordena os indivíduos através da favorabilidade de
sua atitude com relação a determinado objeto, mas não fornece uma base
para se dizer quanto uma pessoa é mais favorável que a outra, e isto se
apresenta como desvantagem.

Como a escala convoca uma abrangência de resposta, o esco
re total de uma pessoa pode ter pouco sentido, uma vez que muitos pa
drões de resposta, aos vários itens, apresentam o mesmo resultado. Isto
dificulta a interpretação dos resultados.

c) Escalas cumulativas

Consiste de uma sêrie de itens em que o indivíduo deve assinalar acordo ou desacordo. Algumas vezes, os itens vêm ordenados de acordo com a favorabilidade e outras vezes a ordenação é casual.

Os itens são relacionados entre si de uma maneira tal que os indivíduos que respondem favoravelmente a determinado item, devem ter escores mais elevados na escala total. O escore do indivíduo é calculado pela contagem dos itens a que responde favoravelmente.

Exemplo (Selltiz et al, 1974) retirado da escala de Bogardus - escala de distância social (mensuração de atitudes quanto aos tipos étnicos): Para cada raça ou nacionalidade indicadas abaixo, assinale cada uma das classificações em que você estaria disposto a admitir a pessoa média dessa raça ou nacionalidade:

RAÇA OU NACIONALIDADE	ITENS						
	A parentesco por casamento.	A meu clube como amigos.	A minha rua como vizinhos.	A emprego, em minha profissão.	A cidadania, em meu País.	São como visitantes do meu País	Excluiria do meu País
Inglês	1	2	3	4	5	6	7
Negro	1	2	3	4	5	6	7
Francês	1	2	3	4	5	6	7
Chinês	1	2	3	4	5	6	7
etc.							

O interesse pelas escalas cumulativas esteve ligado ao desenvolvimento de escalas unidimensionais, isto é, escalas constituídas por itens que medem apenas uma variável.

Alguns pesquisadores indicaram que estas escalas, apesar de pretender medir apenas uma atitude, continham elementos a respeito de vários aspectos do objeto considerado. Isto explica a dispersão de resposta.

A técnica criada por Guttman - análise de escala ou método de escalograma - tem como objetivo enfrentar este problema, isto é, verificar se a característica ou atitude em estudo, realmente, inclui uma única dimensão. A característica ou atitude em estudo é denominada universo de conteúdo ou universo de atributos (Selltitz et al 1974).

O processo de Guttman considera que o universo de conteúdo é unidimensional, somente se permite uma escala cumulativa perfeita ou quase perfeita.

Na prática, é raro encontrar escalas cumulativas perfeitas ou unidimensionais, mas é possível criar escalas que se aproximam disso. A análise de escalograma usa alguns critérios para decidir se determinada série de itens se aproxima ou não de uma escala unidimensional perfeita. Um dos critérios é a reprodutividade das respostas, isto é, a ocorrência de uma grande proporção de respostas anteriores.

A técnica de Guttman oferece pouca orientação para a seleção de itens que tendem a formar uma escala cumulativa-unidimensional.

Existe uma técnica denominada discriminação de escala (Selltitz et al 1974) que fornece orientação para selecionar um conjunto de afirmações que tendem a formar uma escala unidimensional. Esta técnica combina aspectos de elaboração de escalas de Thurstone e Likert, através dos passos:

- . reúne-se um grande número de itens relacionados à característica em estudo;
- . examina-se os itens, sendo eliminados os ambíguos, neutros ou extremos;
- . os itens restantes são julgados por um grande número de juízes, que os coloca em 11 grupos, conforme sejam favoráveis ou desfavoráveis ao objeto em estudo. Isto resulta na eliminação dos itens em que existe grande dispersão de julgamento pelos juízes; e na atribuição do valor mediano aos itens restantes;
- . os itens que vão compor a escala total são transformados numa escala acordo-desacordo em 5 ou 6 graus;
- . aplica-se esta escala em um grande número de indivíduos e depois suas respostas são analisadas em termos de coeficiente de discriminação, isto é, quais itens que discriminam mais entre os indivíduos, que, pelo resultado total, se encontram no grupo alto e baixo;
- . seleciona-se o dobro de itens desejados para a escala final, em termos dos maiores coeficientes de discriminação;
- . dispõe-se os itens de acordo com a ordem de seu valor na escala;
- . divide-se os itens em duas formas similares da escala, colocando os itens de número par em uma e os de número ímpar em outra.

Existem duas limitações relativas à utilização de escalas unidimensionais:

- . pode não ser a base mais eficiente para medir a atitude relativa a objetos complexos, ou para fazer predições quanto ao comportamento em relação a tais objetos;
- . pode ser unidimensional para determinado grupo de indivíduos, e não ser para outro.

VALIDADE E PRECISÃO DE ESCALAS DE ATITUDE

A ênfase da validade das escalas de atitude é a validade de conteúdo, isto é, a congruência entre os itens e a atitude em questão,

Forte	___:	___:	___:	___:	___:	___:	___:	fraco
	1	2	3	4	5	6	7	
Agradável	___:	___:	___:	___:	___:	___:	___:	desagradável
	1	2	3	4	5	6	7	

A lógica do diferencial semântico está no fato de que as coisas reais e idéias são comunicadas, em grande parte, através de adjetivos. Então, é razoável supor que os adjetivos possam ser usados para medir conceitos.

Embora possa parecer que muitos dos adjetivos não estejam relacionados ao conceito a ser julgado, esta falta aparente de significado dos adjetivos, frequentemente, leva o sujeito a dar respostas mais honestas. As críticas, desta abordagem reclamam que os adjetivos são irrelevantes aos conceitos a serem julgados, mas é esta irrelevância aparente que é a fortaleza desta abordagem, desde que ela frequentemente se limita à tendência de produzir respostas aceitáveis socialmente.

O diferencial semântico mede três dimensões:

- . dimensão avaliativa - com os adjetivos do tipo: bom-mau/ agradável-desagradável/ limpo-sujo;
- . dimensão potencial - com os adjetivos do tipo: forte-fraco/ pequeno-grande/ pesado-leve;
- . dimensão atividade - com os adjetivos do tipo: passivo-ativo/ vagaroso-rápido/ impetuoso-calmo.

O fator avaliativo é o mais forte, pois, quase todos os adjetivos implicam em características positivas ou negativas.

É possível usar as respostas do indivíduo para verificar se, para ele, dois conceitos são semelhantes ou diferentes, através da comparação de suas respostas em ambos os conceitos. É possível, também, comparar os conceitos que dois indivíduos têm a respeito de um objeto, através da medida de semelhança apresentada por suas escolhas, nas escalas.

Cada tipo de escala discutido envolve análises específicas. Estas análises podem ser encontradas em livros de psicometria.

3.3 - SUB-UNIDADE 3 : TESTES DE APRENDIZAGEM

O enfoque, nesta sub unidade, será, em particular, os testes de aprendizagem. Para tanto, é necessário situá-los dentro de uma perspectiva maior do conceito de testes em geral. Assim, a primeira parte desta sub-unidade tratará de testes em seu sentido amplo.

O que é um teste?

O teste (Cronbach, 1970) pode ser definido como uma coleção de itens desenvolvidos para prover alguma(s) medida(s) de um(s) atributo(s) do comportamento humano. O teste é, pois, um procedimento sistemático para observar o comportamento de uma pessoa e descrevê-lo com o auxílio de uma escala numérica ou um sistema de categoria.

Um item de teste é um conjunto de instruções orientando como evocar, detectar e corrigir uma unidade de informação específica do comportamento (Sanders, et. al., 1976). O item deve incluir direções para:

- . apresentar uma situação ou estímulo;
- . registrar a resposta.

Os testes podem se focalizar em uma única pessoa ou em um grupo. A diferença entre um e outro é que os testes de grupo permitem que muitos sujeitos sejam testados de uma só vez. Os testes de grupo podem ser administrados a um único indivíduo, se for o caso. Os testes de grupo exigem menos treinamento para sua administração do que os testes individuais.

CLASSIFICAÇÃO DE TESTES

Os testes podem ser classificados de muitas maneiras. Um dos modos, segundo Cronbach (1970), Lyman (1963), Popham (1975):

- a) testes de desempenho máximo
- b) testes de desempenho típico.

a) Testes de Desempenho Máximo

Os testes de desempenho máximo são utilizados quando se quer saber quão bem um indivíduo pode se desempenhar. O que está sendo me di do é a habilidade do indivíduo, por isso ele deve mostrar seu trabalho, ganhando o melhor escore possível. Os testes de desempenho máximo incluem os testes de:

- . inteligência
- . aptidão
- . aprendizagem.

. Testes de Inteligência

Inteligência é um conceito abstrato. Tem-se idéias sobre seus significado, mas há pouco acordo geral sobre seu significado preciso.

O conceito de inteligência tem sido usado como sinônimo de maturidade mental, classificação geral, habilidade mental e outros.

Alguns testes de inteligência resultam em um único escore total (QI) enquanto que outros em muitos escores. Esta variação leva a diversos resultados.

. Testes de Aptidão

Os testes de aptidão implicam em predição, isto é, fornecem uma base para predizer níveis de desempenho futuro. Eles são utilizados, geralmente, na seleção de indivíduos para trabalhos, admissão em programas de treinamento.

. Testes de Aprendizagem

São usados para medir níveis atuais de habilidades.

Os testes de aprendizagem e de aptidão diferem no que se refere aos seus propósitos:

- . teste de aprendizagem está relacionado ao desempenho atual
- . teste de aptidão está relacionado à predição do nível de desempenho.

Os testes de aprendizagem são planejados para se ajustar aos usos particulares a serem feitos dos resultados. Podem ser:

- teste de comportamento de entrada;
- pré-teste;
- testes intermediários;
- pós-teste;
- reteste;

O teste de comportamento de entrada tem como objetivo de detetar se os indivíduos realmente dominam os pré-requisitos do curso, ou seja, se já dominam o que se supõe que deveriam dominar antes do início da instrução propriamente dita. Se os comportamentos de entrada não forem dominados pelas pessoas, deve-se ministrar ou uma instrução remedial para supri-los, ou planejar uma estratégia instrucional individualizada.

O pré-teste é aplicado no início de um curso e tem como objetivo medir o nível de domínio das habilidades propostas para o curso.

Os testes intermediários são aplicados durante o curso com o objetivo de detetar o domínio dos objetivos instrucionais das unidades em que o curso foi dividido.

O pós-teste é dado no final de um curso e é usado para detetar se os indivíduos dominaram os objetivos instrucionais propostos para o curso. O pós-teste consiste em uma forma paralela do pré-teste.

Com os dados do pré e pós-teste, faz-se a Avaliação Formativa de alunos e materiais (Unidade 3.4 - pág. F.223 e calcula-se a efetividade da instrução (tópico 4.5.3 - pág. F.355).

O reteste consiste na reaplicação de um mesmo teste aos mesmos indivíduos com o propósito de verificar a aprendizagem.

O pré e pós-teste e o teste-reteste são utilizados para calcular a confiabilidade de testes, conforme será descrito no tópico "Erro de Medida", página F.190.

b) Testes de Desempenho Típico

Os testes de desempenho típico são utilizados quando se quer saber o que o indivíduo é realmente, ou o que faz em uma dada situação ou uma classe de situações.

Os testes de desempenho típico são divididos em:

- . observações de comportamento;
- . auto-relatório;
- . observação de comportamento.

São tentativas de estudar o sujeito quando este está agindo naturalmente. Pode ser através de situações de testes padronizados ou em condições naturais.

As situações padronizadas exigem que cada sujeito seja colocado exatamente na mesma situação, o que permite uma comparação precisa das pessoas que não são vistas, normalmente, em circunstâncias similares.

Nas situações naturais o comportamento é observado em atividades diárias.

. Auto-Relatório

No auto-relatório, o indivíduo relata o seu próprio comportamento. Os questionários são usados para obter tais relatos.

O auto-relatório tem o problema de a pessoa não fornecer uma imagem honesta de si.

ELABORAÇÃO E USO DE TESTES DE APRENDIZAGEM

. Formato de Itens

Collins et al (1969) discutem sobre formatos de itens. A discussão deles é específica aos testes feitos por professores, o que não impede a utilização por avaliadores.

Os vários formatos de itens discutidos por estes autores são:

- a) Questão aberta
- b) Questão de completar
- c) Questão de múltipla escolha
- d) Questão do tipo falso-verdadeiro
- e) Questão de acasalamento
- f) Combinação de formatos

a) Questão aberta

É um dos mais antigos formatos de item de teste. É aquele onde se coloca uma questão geral para o aluno, que para respondê-la, discute livremente. A questão aberta pode variar de uma simples afirmativa (ex.: escreva uma sentença com a palavra abacaxi, no espaço que se segue), até proposições mais complexas (ex.: Responda no espaço que se segue: Quais os problemas atuais da economia brasileira? Explique-os sucintamente).

Algumas vantagens da questão aberta são:

- . permite a máxima liberdade para respondê-la;
- . é excelente forma de avaliar a escrita;

- . é ótima para testar a habilidade de solução de problema;
- . incentiva o estudo de informação verbal de forma mais aprofundada;
- . é relativamente mais fácil de ser construída;
- . é praticamente impossível ao estudante "chutar" ou adivinhar a resposta correta.

Algumas desvantagens da questão aberta são:

- . apresenta dificuldade quanto a atribuição objetiva de notas. Vários professores corrigindo um mesmo item de um mesmo aluno podem atribuir notas diferentes, ou o mesmo professor em ocasiões diferentes pode atribuir notas diferentes. Isto, apesar de ter sido fixado um critério para correção;
- . cobre somente um número limitado de tópicos;
- . exige muito tempo do aluno para respondê-la;
- . exige muito tempo do professor para corrigi-la;
- . favorece o estudante que tem facilidade de expressão verbal.

Exemplo de uma questão aberta:

"Responda a questão abaixo no espaço que se segue: Quais as consequências econômicas da cultura do café para o processo de industrialização do Estado de São Paulo?"

Alguns cuidados devem ser tomados na elaboração desse formato de item:

- . cada item deve ser escrito de forma clara e precisa de tal maneira que o aluno não tenha dúvidas a respeito do significado da questão. Por exemplo: ao invés de "Relacione a cultura do café no Estado de São Paulo com o processo de industrialização" usar por exemplo: "Quais as consequências econômicas da cultura de café para o processo de industrialização do Estado de São Paulo?"

- . não é boa prática deixar o estudante escolher os itens sobre os quais escrever. Ex.: "Escreva sobre duas das três questões seguintes ..." Isto leva o aluno a fazer um estudo limitado ao se preparar para o teste.
- . deve-se dirigir o aluno para os detalhes que precisam ser abordados na sua resposta;
- . um teste não deve ser composto só de questões abertas. Devido a subjetividade na atribuição de notas, é aconselhável usar outros tipos de item em combinação com este.

Alguns cuidados devem ser tomados na correção desse formato de item:

- . deve-se estabelecer um critério de correção para cada questão, envolvendo os pontos importantes, e corrigir todas as provas segundo este mesmo critério;
- . para dar nota mais objetivamente, deve-se corrigir uma determinada questão de todos os alunos antes de passar à seguinte;
- . o nome do aluno não deve ser visto, antes da correção da prova, para que não haja influência de impressões passadas do aluno, na sua nota;
- . a nota não deve ser influenciada pela caligrafia, soletração, sintaxe, ou pobreza de expressão do aluno.

b) Questão de completar

Este formato de item requer que o aluno coloque uma palavra ou frase que falta para completar uma afirmativa.

Algumas vantagens da questão de completar são:

- . exige que o estudante relembra informações ao invés de só reconhecê-las;
- . é relativamente fácil de ser construído;

- . é relativamente difícil para o estudante adivinhar a resposta certa;
- . pode ser corrigido objetivamente, quando bem construído;
- . exige pouco tempo para correção;
- . exige pouco tempo para ser respondido;
- . permite uma cobertura mais ampla do campo de estudo.

Algumas desvantagens da questão de completar:

- . se tal tipo de questão é usado sistematicamente, os estudantes podem limitar seus hábitos de estudos. O estudante pode, por exemplo, estudar apenas palavras chaves, datas, nomes, locais, etc...
- . permite correção mais objetiva que a questão aberta, porém sua correção é menos objetiva que a dos outros tipos de questões.

Exemplo de uma questão de completar:

"Complete a frase abaixo com o antônimo da palavra grifada:

- . tenho um sapato feio e um _____".

Alguns cuidados devem ser tomados na elaboração deste formato de item:

- . o item deve ser formulado de modo que o aluno não tenha dúvidas do que lhe é pedido. Ao invés de "O Governador do Estado do Sergipe é _____" em que o aluno poderia dar as seguintes respostas: de tal partido político, de tal religião, ou fulano de tal, todas tecnicamente corretas, usar por exemplo: "O nome do atual Governador do Estado do Sergipe é _____".
- . a questão deve ser formulada de tal maneira que o espaço a ser completado fique no fim ou próximo do fim da proposição. Ao invés de "_____ é o nome do atual Presidente do Brasil", usar por exemplo: "O nome do atual Presidente do Brasil

é _____". Isto permite que o aluno responda a questão numa primeira leitura, sem precisar voltar ao começo novamente.

. o item de teste deve ser construído com 1 ou no máximo 2 espaços, numa mesma afirmativa, para serem completados. Um item de completar, com muitas lacunas, é mais um quebra-cabeça que um item de teste. Neste caso, apresentar mais itens é uma alternativa desejável.

. algumas vezes é melhor formular uma frase interrogativa e fornecer espaço em branco para o aluno escrever a resposta. Ex.: "Quais são as cores primárias"?

1 - _____

2 - _____

3 - _____

. Não devem ser dadas "pistas" ao estudante, como por exemplo a primeira letra da resposta, ou apresentar a lacuna com traços ou em número correspondente ao número de letras da resposta;

. todas as lacunas devem ser apresentadas num mesmo tamanho, evitando que os alunos tentem estabelecer uma relação entre tamanhos da lacuna e da resposta.

OBS.: Quando num item existe mais de uma lacuna para ser preenchida, deve ser atribuído ponto para cada lacuna.

c) Questão de múltipla escolha

Este formato de item requer que o aluno escolha uma resposta dentre as várias listadas.

Algumas vantagens do item de múltipla escolha são:

. é corrigido objetivamente, pois o professor não precisa fazer julgamentos subjetivos;

- . é rapidamente corrigido, podendo inclusive ser corrigido por computador;
- . admite uma grande variedade de situações e objetivos;
- . permite variedade de escolha da elaboração das alternativas;
- . não facilita o êxito casual ou adivinhação. A probabilidade de acerto casual é igual a $1/n$, onde n é igual ao número de alternativas;
- . permite uma cobertura mais ampla do campo de estudo.

Algumas desvantagens do item de múltipla escolha são:

- . é difícil construir bons itens de múltipla escolha. Não é fácil elaborar boas alternativas;
- . é de construção demorada;
- . quando utilizados sistematicamente para cobrar informação verbal, limita o hábito de estudo dos alunos;
- . acontece às vezes, que o aluno por um processo de eliminação, chegue a resposta correta sem grande esforço e preparação.

Exemplo de uma questão de múltipla-escolha:

"Leia o objetivo abaixo. Assinale a alternativa que apresenta o tipo de aprendizagem envolvida no mesmo:

OBJETIVO: Dadas as representações gráficas de 4 sólidos geométricos, o aluno deverá identificar aquela que representa um prisma de base triangular, marcando com um "X" esta figura.

- a) solução de problema;
- b) discriminação;
- c) conceito;
- d) associação verbal".

Alguns cuidados devem ser tomados na elaboração deste formato de item.

- . quatro ou cinco alternativas é o número mais adequado. Num teste o número de alternativas deve ser constante para todos os itens.
- . evitar dar indícios gramaticais para a resposta correta. Ex.: "Assinale com um "X" a alternativa correta. Um exemplo de animal doméstico é o:
 - a) gato;
 - b) girafa;
 - c) aranha;
 - d) vaca
- . uma das alternativas deve ser a melhor, mas todas as outras devem parecer possíveis a pessoa não informada;
- . cada item deve ter apenas uma resposta correta. Ao invés de: "Assinale a(s) alternativa(s) correta(s):

Quando fomos ao cinema ontem, _____ encontramos Tereza chorando:

- () nós
- () tu
- () tu e ele
- () eu e ele

Deve-se usar dois itens em separado:

a - "Assinale a alternativa correta:

Quando fomos ao cinema ontem, _____ encontramos Tereza chorando".

- () nós
- () vós
- () eles
- () tu

b - "Assinale a alternativa correta:

Quando fomos ao cinema ontem, _____ encontramos Tereza chorando".

- () tu e ele
- () eu e ele
- () ele e ela
- () tu e ela

- . evitar o uso de padrões de resposta correta. Usar cada posição de resposta correta, aproximadamente o mesmo número de vezes;
- . evitar sempre que possível o uso de itens com formulação negativa. Ex.: "Qual não é ...
- . não utilizar as expressões "nenhuma acima" ou "todas acima" simplesmente para completar o item. A expressão "todas acima" pode ser empregada quando em caso legítimo.

d) Questão falso-verdadeiro

Este formato de item consiste em declarações que o estudante deve marcar como falsas ou verdadeiras.

Algumas vantagens da questão falso ou verdadeiro são:

- . é relativamente fácil de ser construída;
- . um estudante pode responder a um grande número de questões num pequeno espaço de tempo. É possível portanto testar um amplo campo de estudo;
- . é corrigido objetivamente;
- . é adequado para cobrar informação verbal.

Algumas desvantagens da questão falso ou verdadeiro são:

- . há uma probabilidade de 50% do estudante "adivinhar" a resposta correta;
- . exige do estudante que apenas reconheça informação;

- . é muito fácil construir-se itens falso-verdadeiros pobres.

Alguns cuidados devem ser tomados na elaboração desse formato de item:

- . não construir itens que contenham verdade parcial. Deve haver a penas uma resposta, ou falsa, ou verdadeira.

Ex.: "Leia a afirmativa abaixo, marque um "X" se ela for falsa ou verdadeira.

O Rio Amazonas, o maior do mundo em extensão, desemboca no Oceano Atlântico".

Falsa

Verdadeira

Itens deste tipo devem ser desmembrados em dois. Assim:

- 1 - "Leia a afirmativa abaixo, marque um "X" se ela for falsa ou verdadeira.

O Rio Amazonas é o maior rio do mundo em extensão".

Falsa

Verdadeira

- 2 - "Leia a afirmativa abaixo, e marque um "X" no espaço indicado se ela é falsa ou verdadeira".

"O Rio Amazonas desemboca no Oceano Atlântico".

Falsa

Verdadeira

- . evitar o uso de palavras como "nenhum", "nenhuma", "todo", "sempre", "nunca" ou "algumas vezes", "pode", "geralmente", por serem tendenciosas;
- . se itens falso-verdadeiro são estabelecidos negativamente, as palavras negativas, devem ser destacadas e grifadas;

- . construir itens falso-verdadeiro de tal modo que num teste, aproximadamente metade destes, sejam de respostas falsas e metade de respostas verdadeiras. Cuidar para que não haja um padrão na sequência de resposta, isto é, uma sequência de itens falsos e depois uma sequência de itens verdadeiros.

e) Questão de acasalamento

Este formato de item consiste na apresentação de duas listas ao aluno, uma delas contendo questões e a outra contendo respostas, que deve fazer a correspondência entre ambas.

Algumas vantagens da questão de acasalamento são:

- . é útil para medir aprendizagem em áreas que podem ser reduzidas a listas, tais como: autores e nomes de livros, capitais e Estados, datas e acontecimentos históricos, substantivos e seus sinônimos, etc.;
- . é relativamente fácil construir um bom item deste formato;
- . permite ampla cobertura de um campo de estudo;
- . é corrigido rapidamente;
- . é corrigido objetivamente;
- . há baixa probabilidade de adivinhar as respostas corretas. Para isso, é conveniente apresentar mais respostas que questões.

Algumas desvantagens desse formato são:

- . exige do aluno apenas que relembre informação;
- . se usado sistematicamente, pode limitar os hábitos de estudo do aluno.

Exemplo de questão de acasalamento:

"Enumere a 2a. coluna de acordo com a 1a., relacionando as obras aos seus respectivos autores".

- | | |
|-----------------------------|---------------------------------|
| 1 - Martin Cererê | () - Machado de Assis |
| 2 - Gabriela Cravo e Canela | () - José de Alencar |
| 3 - A Mão e a Luva | () - Aluísio de Azevedo |
| 4 - Iracema | () - José Mauro de Vasconcelos |
| 5 - A Carne | () - Jorge Amado |
| 6 - Meu Pé de Laranja Lima | () - Júlio Ribeiro |
| | () - Cassiano Ricardo |

Alguns cuidados devem ser tomados na elaboração desse formato de item:

- . dar ao aluno uma orientação clara, fixando em que bases deve ser feita a correspondência;
- . a lista de questões deve ser de conteúdo homogêneo, assim como a lista de respostas. Cada resposta deve ser plausível a qualquer questão;
- . o número de questões e respostas em cada item deve idealmente variar de 6 a 8;
- . usar mais respostas que questões;
- . colocar as colunas de questões e respostas lado a lado, economizando tempo para o aluno responder;
- . evitar dar mais de uma resposta correta a cada questão.

f) Os vários formatos de itens discutidos nas páginas anteriores podem ser combinados de várias maneiras, resultando em outros formatos como:

- . itens de interpretação
- . itens de arranjo
- . itens de analogia
- . questões orais

. Itens de Interpretação

Certas situações de testagem exigem o uso de figuras, mapas e modelos de vários tipos em que se pede ao indivíduo que explique ou interprete o que vê.

Exemplo:

- . Projeta-se um mapa e aponta-se, por exemplo, determinado rio e pede-se para o indivíduo identificá-lo.

. Itens de Arranjo

Fornece-se ao indivíduo uma lista em que se pede que a ordene de forma correta. Este tipo de item é útil, em áreas que têm uma natureza sequencial, cronológica ou alfabética.

Exemplo:

- . Arranje os meses de acordo com sua ordem correta:

dezembro	novembro	outubro	maio
abril	fevereiro	junho	agosto
julho	março	janeiro	setembro

Uma desvantagem deste tipo de item é que um erro no arranjo, causa, automaticamente, erros na ordenação.

. Item de Analogia

O indivíduo deve reconhecer a relação ou analogia entre 2 (duas) palavras. Fornece-se-lhe também uma 3a. em que ele deve completar com um 4o termo estabelecendo a mesma analogia do 1o e 2o termo.

Exemplo:

- . 10 está para 100, assim como 8 está para _____.

. Questões Orais

Este formato de item é muito útil em situações que não permitem questões escritas, como por exemplo, para testar indivíduos com

insuficiência visual bastante acentuada, ou para retestar indivíduos que fraudam testes escritos, etc..

Uma grande desvantagem de questões orais é o tempo gasto, uma vez que deve ser administrado individualmente. Outra desvantagem, que é consequência desta última é a divulgação das questões para os indivíduos ainda não examinados.

Exemplo:

. Pedir para o aluno ler determinado texto de leitura.

Existem ainda outras variações e combinações de itens de testes.

É importante que se conheça as vantagens e desvantagens dos tipos de itens, em medir precisamente o progresso do indivíduo em relação aos objetivos estabelecidos.

GERAÇÃO DE ITENS

Naturalmente, um teste não pode ser melhor do que os itens que o compõem. Um bom plano representa uma intenção de construir um bom teste, mas pode ocorrer que o plano não se concretize, se os itens não forem escritos habilidosamente. Escrever itens de teste é uma arte que poucas pessoas parecem dominar.

Frequentemente os itens falham por 2 motivos:

- . ambiguidade - falham em relação ao tipo de resposta exigida;
- . abordam aspectos triviais do assunto como datas, nomes.

A elaboração de cada item deve estar relacionada a um objetivo instrucional e para cada objetivo é aconselhável elaborar de 2 a 3 itens de teste. No caso de testes paralelos, constrói-se itens para as duas formas. É importante examinar, cuidadosamente, o objetivo para construir itens de teste adequados para medi-lo. O item não deve pedir nada além do que está estabelecido no objetivo.

Quando o teste é construído para medir a aprendizagem de uma unidade (poucos objetivos) é desejável construir itens para cada objetivo. No caso de um teste para medir a aprendizagem em um curso (grande número de objetivos) uma opção é construir itens a partir dos objetivos terminais das unidades ou amostrar randomicamente os objetivos específicos e construir os itens. Ambas as abordagens estão ainda em estudo (Stone, 1974).

Para a geração de itens, precisa-se elaborar a direção e o corpo do item.

A direção consiste de informações breves de como responder ao mesmo. Existem itens que se restringem apenas à direção. É o caso de itens que exigem, como respostas, a formulação do indivíduo.

O corpo do item é o item propriamente dito, isto é, o que o indivíduo deve responder.

Exemplo: Leia a oração: "A colheita iniciou ontem". Assinale com um (X) o sujeito desta oração.

- () colheita
- () A colheita
- () A colheita iniciou
- () iniciou ontem

ABORDAGEM DE MEDIDA

Existem duas abordagens de medida que são:

- a)"Norm-referenced" (NR) - abordagem referente a norma
- b)"Criterion-referenced"(CR) - abordagem referente a critério.

Apenas olhando um teste, não é possível dizer se é NR ou CR. De fato, um teste CR pode ser usado como NR mas o contrário não é possível.

A distinção entre NR e CR (Popham, 1973), é dada por:

- . propósito do teste
- . modo de construção
- . especificidade da informação dada sobre o domínio das tarefas relevantes à instrução
- . generalização da informação do desempenho, no teste, ao domínio
- . uso a ser feito das informações obtidas no teste.

a) Abordagem NR

1 - Características de um teste NR

Até duas décadas atrás, os propósitos básicos de quase todas as atividades de medidas educacionais era avaliar o desempenho do indivíduo de modo que o desempenho pudesse ser interpretado de acordo com o desempenho de outros indivíduos, na mesma medida, isto porque o desempenho de um indivíduo só tinha sentido se fosse comparado ao desempenho de um grupo normativo de indivíduos. O grupo normativo consiste de uma amostra representativa de sujeitos que fez o teste anteriormente.

A ênfase chave do teste NR são as diferenças individuais. Para permitir comparações entre o desempenho do indivíduo e o do grupo normativo, o teste deve ter considerável variância entre os escores individuais. Se a variância é pequena, torna-se difícil comparar os indivíduos, como é o caso, se todos os indivíduos obtiverem escores similares, em teste fácil ou difícil. É exatamente por esta razão, que os elaboradores de teste NR tentam, dispersar, ao máximo, os escores, para que seja possível fazer as comparações.

O teste NR é empregado em situações de decisão sobre os indivíduos, em questões em que há um grau de seletividade. Por exemplo: Quais alunos devem ganhar bolsas de estudo?

Como o escore do indivíduo é comparado com normas de grupo, não é possível saber o grau de proficiência exibido pelos comportamentos testados, em termos do que cada indivíduo pode fazer. O teste NR diz que uma pessoa é mais ou menos proficiente que a outra, mas não diz como uma das duas é proficiente em relação às tarefas envolvidas.

Os itens do teste NR são construídos para discriminar, por isso, despreza-se itens muito fáceis ou muito difíceis.

Os itens são frequentemente desenvolvidos a partir de habilidades comuns ou geralmente aceitas por muitos programas educacionais. Porém, eles podem não ser representativos do conteúdo comum desses programas, apesar de ter tal pretensão.

O tipo de escore relatado no teste NR é em forma de percentil e escore-padrão, uma vez que o significado do escore do indivíduo emerge da comparação. O tópico "Atribuição de Pontos", pág. F.199 trata rá destes escores.

Há procedimentos sistemáticos para a determinação da validade e confiabilidade dos testes NR, conforme será descrito nos tópicos "Validade" pág. F.179, e "Confiabilidade" pág. F.190.

2) Deficiências do teste NR para a Avaliação Instrucional

As medidas NR são elaboradas, principalmente, para permitir comparações entre os indivíduos e o grupo normativo. Elas são mais apropriadas, portanto, quando se toma decisões sobre indivíduos, referentes à seleção. Quando a função da Avaliação é avaliar o mérito de tratamento instrucional, em termos absolutos, o teste NR não é aplicável.

Um dos maiores defeitos do teste NR, para propósitos de Avaliação da instrução é que há, frequentemente, uma falta de congruência entre o que o teste mede e os currículos locais, no caso de testes que servem a muitas localidades. Os currículos variam de um lugar para

outro e por isso há rejeição, por parte dos professores, uma vez que o teste não enfatiza o seu programa instrucional. Por causa disto, deve-se ter prudência em usar testes de aprendizagem NR, a menos que haja congruência satisfatória teste-currículo.

Outro problema relativo ao teste NR para fins de avaliação da instrução está na sua necessidade de produzir variância. Os itens que não contribuem para tal podem não ser incluídos no teste. E, talvez estes itens tratem de habilidades que os professores acham importante. Então, a significância curricular do instrumento, pode ser prejudicada como consequência da preocupação de rigor técnico.

O que se quer, na realidade, é a melhoria das sequências instrucionais, para que todos indivíduos alcancem o domínio de habilidades-chave. Por esta razão, é impossível conciliar o uso de teste NR e o propósito de melhorar a efetividade absoluta da instrução.

3) Procedimentos de desenvolvimento de um teste NR

Os procedimentos para desenvolver um teste NR, que se encontram na literatura, podem ser resumidos nos seguintes passos básicos:

- . planejamento do teste
- . geração e integração de itens que vão compor o teste
- . revisão técnica do teste
- . Avaliação Formativa do teste
- . montagem do teste
- . determinação da validade do teste
- . estabelecimento de normas para o teste
- . preparação do manual do teste.

Planejamento do teste

Antes de se construir o teste é necessário preparar o seu plano. Se pessoas representativas, das que vão usar o teste, concordam, adiantadamente com a adequação do plano, o fato de se chegar a um instrumento aceitável é principalmente uma questão de habilidade técnica e Pesquisa Aplicada.

O plano deve contar o esquema das habilidades, número a proximado de itens que vai conter cada seção (se o teste tiver mais de uma), número de formas, os formatos de itens a serem usados, incluindo exemplos, procedimentos de administração e sua duração, atribuição de pontos, normas a serem computadas.

Quando o plano está completo, deve ser revisto por muitas pessoas, inclusive professores, especialistas de conteúdo, especialista em educação e medida psicológica. Muitas sugestões podem ser dadas para a melhoria do plano, que, quando revisado, deve ser submetido novamente aos revisores.

No caso de se construir um teste para ser aplicado somente para poucos indivíduos, não vale a pena desenvolver tal plano, mas o instrutor deve fazer um esquema da cobertura das habilidades.

Geração e integração de itens

É necessário que se elabore uma coleção de itens duas vezes maior da que vai compor o teste final. Deve-se:

- . elaborar a chave e os critérios de correção
- . verificar a congruência entre os itens e as habilidades
- . verificar se os itens são livres de ambiguidade e deficiência estrutural
- . verificar se os itens têm algumas indicações não intencionais
- . verificar se as direções estão claras e completas

- . verificar se os itens das várias formas são paralelos (quando houver)
- . amostrar os itens para as formas de teste.

Revisão Técnica

Consiste em um controle de qualidade preliminar, isto é, na revisão do teste por especialistas em testagem e em conteúdo. Após esta revisão, faz-se as modificações necessárias quanto a erros técnicos e de conteúdo.

Avaliação Formativa do teste

Trata-se de:

- testagem um-a-um
- testagem em uma amostra da população
- análise dos resultados para a seleção de itens

A testagem um-a-um consiste na aplicação do teste em indivíduos representativos da população-alvo. O número de pessoas e os detalhes para a aplicação e observação das mesmas devem ser estabelecidos antes da aplicação. Com base nas informações dos indivíduos testados, deve-se melhorar o teste.

Para a testagem do teste em uma amostra da população deve-se, primeiro, elaborar um plano de amostragem dos indivíduos que vão compor esta amostra. Deve-se, também, indicar a ordem de aplicação das formas (se houver) e a aplicação. Por causa de haver muitas oportunidades de se levar vantagem do acaso, os resultados podem ser enviesados, a menos que haja cinco (5) vezes mais indivíduos do que os itens (Nunnally, 1967).

A análise dos resultados para a seleção de itens consiste em:

O registro dos escores em cada item consiste em identificar o acerto no item (1) ou erro no item (zero) e o escore total no teste, para todos os indivíduos. Depois deve-se registrá-los na tabela de escores. A Tabela F.10 visualiza um exemplo do registro de escores em cada item.

TABELA F.10

REGISTRO DOS ESCORES EM CADA ITEM

FORMA

INDIVÍDUOS ↓	ITENS →							ESCORE TOTAL
	1	2	3	4	5	6	n	
1	0	1	0	1	1	0		13
2	1	1	1	0	1	1		15
3	1	0	0	1	0	0		12
4	0	1	0	1	1	0		12
5	0	0	1	0	0	0		10
n	0	0	1	0	0	0		

Para o cálculo dos índices de dificuldade, discriminação e correlação item-escore total, deve-se ordenar os escores totais dos indivíduos do mais alto ao mais baixo e seus respectivos escores individuais. A Tabela F.11 apresenta um exemplo dos escores ordenados, baseado no exemplo da tabela anterior.

TABELA: F.11^{ef}

ESCORES ORDENADOS

F O R M A		ITENS							ESCORE TOTAL	
INDIVÍDUOS		1	2	3	4	5	6	7		
2		1	1	1	0	1	1		15	} Grupo Alto
1		0	1	0	1	1	0		13	
3		1	0	0	1	0	0		12	
4		0	1	0	1	1	0		12	} Grupo Baixo
5		0	0	1	0	0	0		10	

. Correlação

O tipo mais importante de análise de item é feito pela correlação de cada item com o escore total do teste. Este índice diz até que ponto o item mede o que o teste quer mensurar. No caso de teste ter diferentes partes ou seções, cada item deve ser correlacionado com o subescore total de cada seção do teste. O coeficiente de correlação apropriado é o ponto-biserial, pois, no caso, trata-se de correlação entre um item dicotômico e a distribuição multiponto de escores (Nunnally, 1967). A fórmula é a que segue:

$$\gamma_{Pb} = \frac{MS - Mi}{\sigma} \sqrt{pq} \quad \text{em que:}$$

γ_{Pb} = coeficiente de correlação ponto-biserial;

MS = média dos escores do teste total dos indivíduos que acertaram o item;

Mi = média dos escores do teste total dos indivíduos que erraram o item;

P = valor p = $\frac{\text{número de indivíduos que acertaram o item}}{\text{número total de indivíduos}}$;

$$q = 1 - p;$$

σ = desvio padrão do grupo.

Os índices de correlação próximo a zero devem ser cuidadosamente inspecionados. Normalmente, estes itens são descartados. Entre os itens restantes, aqueles que têm correlação mais alta com o escore total, geralmente, são os melhores itens, pois, provavelmente, são menos ambíguos e não são muito extremos em dificuldade. Os índices a partir de 0.20 são considerados bons (Nunnally, 1967). Se há poucos itens a partir deste nível, não há escolha, a não ser empregá-los. Se o número de itens é menor do que o exigido para o teste final, o único recurso é recomeçar com uma coleção maior de itens.

. Índice de discriminação e dificuldade

O índice de discriminação refere-se a extensão na qual um item de teste é respondido corretamente pelos indivíduos que possuem, em maior grau, a qualidade que está sendo medida e respondido erroneamente por aqueles que possuem, em menor grau, a qualidade em questão. O objetivo é separar indivíduos com maior e menor grau de habilidade.

O índice de dificuldade refere-se a extensão em que cada item do teste pode ser respondido, corretamente, por qualquer indivíduo.

Com os resultados da Tabela F.11 (escores ordenados - calcula-se 27% do total dos indivíduos (Nunnally, 1967). Isto tem como finalidade, trabalhar apenas com os dados de indivíduos de grupo alto e baixo. Assim, se se tem dados de 120 pessoas, 27% corresponde a 32 indivíduos de cada grupo. Para calcular os índices de discriminação e dificuldade, elabora-se uma tabela com todos os itens e número de acertos no grupo alto e baixo. Exemplo de resultados para o cálculo de índices de discriminação e dificuldade, conforme o exemplo anterior.

TABELA F.12

NÚMERO DE ACERTOS EM CADA ITEM

ITEM	GRUPO ALTO	GRUPO BAIXO
1	1	0
2	2	1
3	1	1
n		

Depois calcula-se os índices segundo as fórmulas:

$$\text{Índice de discriminação} = \frac{AL - BA}{n}$$

$$\text{Índice de dificuldade} = \frac{AL + BA}{n} \quad \text{em que:}$$

AL = número de acertos no grupo alto

BA = número de acertos no grupo baixo

n = total de indivíduos nos grupos alto e baixo.

O índice de discriminação varia de -0,5 a 0,5. Quanto maior o índice, mais o item discrimina. Idealmente, os índices de discriminação devem ser positivos, ou seja, uma percentagem maior dos indivíduos com escore alto, acerta o item e uma percentagem mais baixa dos indivíduos com escore baixo, erra o item. Os índices negativos significam que uma proporção maior de indivíduos com escore baixo, no teste, acerta o item e os sujeitos com escore alto, no teste, perdem o item. Isto significa que há algum defeito no item. Os índices com discriminação negativa devem ser eliminados do teste.

O índice de dificuldade varia de zero a 1,00. Quanto menor o índice, mais difícil é o item. Idealmente, o nível de dificuldade do item deve ser 0.50.

Após os cálculos dos índices de discriminação, dificuldade e correlação item-escore total deve-se tabelá-los. Exemplo de uma tabela com os resultados dos índices.

TABELA F.13

ÍNDICES DE DISCRIMINAÇÃO, DIFICULDADE E CORRELAÇÃO ITEM-TOTAL

ITENS	DISCRIMINAÇÃO	DIFICULDADE	CORRELAÇÃO ITEM-TOTAL
1	0,53	0,42	0,68
2	0,34	0,25	0,61
3	0,66	0,27	0,52
4	0,58	0,66	0,51
n			

Montagem do teste

A montagem inicial do teste consiste na seleção dos melhores itens, em termos dos índices de correlação, dificuldade e discriminação, para a composição do teste final.

Para a montagem final, escolhe-se os melhores itens em termos dos índices citados e computa-se a fórmula da consistência interna (discutida na pág. F.195, do conjunto escolhido).

A determinação da fidedignidade do teste, validade, estabelecimento de normas e preparação do manual do teste, estão detalhados nos tópicos "Validade" pág. F.179, "Erro de Medida" pág. F.190, "Atribuição de pontos", pág. F.199.

b) Abordagem CR de Medida

1) Características de um Teste CR

Por volta de 1950 um grupo de psicólogos e educadores voltou suas atenções para as possibilidades de uma abordagem de ensino conhecida como Instrução Programada (IP), desenvolvida por Skinner (Popham, 1975).

O movimento da instrução programada pregava que, com as sequências instrucionais desenvolvidas, todos os indivíduos alcançariam 100% do domínio de qualquer objetivo em 50% do tempo exigido pela instrução convencional.

Skinner aplicava os princípios de aprendizagem derivados de laboratórios, aos esquemas instrucionais. Ele experimentou seus esquemas instrucionais, em alunos, e revisava a instrução até que funcionasse.

Robert Glaser (Popham, 1975), que esteve envolvido no início do movimento de instrução programada, chamou a atenção de educadores e psicólogos da inapropriedade de aplicar as noções clássicas de medida às situações envolvendo técnicas instrucionais mais novas. À medida que os elaboradores da instrução programada começaram a medir os efeitos das intervenções instrucionais, eles se conscientizaram que os testes "padronizados" e as técnicas tradicionais do NR, realmente, não serviam bem ao seu propósito.

O próprio termo CR (Popham, 1975) foi introduzido por Glaser, para significar uma avaliação contra um padrão absoluto ou critério. O teste CR é usado para assegurar o "status" de um indivíduo com relação ao critério.

O significado do termo critério é uma questão chave. Há várias interpretações sobre o critério ligado ao CR, (Popham, 1975) como:

- . critério = domínio do comportamento, ou seja, classes de comportamentos baseados em conteúdos particulares. O domínio especifica a situação-estímulo para a qual os indivíduos dão respostas. É a definição mais defendida.
- . critério = objetivo instrucional. Por causa da preocupação com objetivos educacionais a partir de 1960, houve uma tendência de referenciar, nessa nova abordagem de construção de testes, a ligação de itens com objetivos instrucionais. Porém, quando os objetivos não eram bem definidos, permitiam várias interpretações. Várias pessoas, com um mesmo adjetivo, podiam gerar itens completamente diferentes. Os objetivos instrucionais devem ser definidos em nível de detalhes suficientes, de modo a se constituírem em descrições bem definidas de classes de comportamentos.
- . critério = nível de proficiência. O termo critério tem sido usado para indicar o desempenho de cada indivíduo, em algum ponto de um continuum de aquisição de habilidades. Este continuum vai desde a nenhuma proficiência até o desempenho perfeito. Um escore de um indivíduo fornece informação explícita do que ele pode ou não fazer.

Os testes CR são planejados para se tomar decisões sobre indivíduos e tratamentos (por exemplo, programas instrucionais).

No caso de decisões sobre indivíduos, pode-se usar um teste CR para determinar se um indivíduo dominou um critério cujas habilidades são consideradas pré-requisitos para começar um novo programa de treinamento.

Para decisões relativas a tratamento, pode-se usar um teste CR para testar, por exemplo, uma sequência instrucional. A fim de que o sistema instrucional opere eficientemente é importante que o programa instrucional e a avaliação do programa funcionem perfeitamente, isto é, o que é ensinado é derivado dos resultados desejados que por sua vez determina o que vai ser testado (Stone, 1974). O conteúdo do currículo determina o conteúdo do teste CR.

Como o objetivo do teste CR é verificar o desempenho absoluto do indivíduo, a variabilidade nos escores é irrelevante, pois não interessa comparação. O significado do escore vem diretamente da conexão item-critério. É verdade que se obtêm escores variados em qualquer teste psicológico, mas a variabilidade não é condição necessária para um bom teste CR. Então, um item de teste CR pode ser bom, sem a qualidade de discriminação, à medida que ele estiver diretamente relacionado ao critério. Segundo Gagné (Stone, 1974) são características desejadas de um item CR:

- . distintividade - cada item deve medir uma classe particular de um desempenho
- . livre de distorção - deve-se eliminar os efeitos de fontes viés como extinção, adaptação, interferência e outros.
- . confiabilidade de detectar a aprendizagem - deve ser designada, para testar se um indivíduo é ou não capaz de exibir o desempenho. Os escores são basicamente binários.

No caso do teste CR ter considerável variância é possível relatar os escores como o é nos testes NR. No caso do teste CR se focalizar na individualização da instrução e avaliação absoluta, é mais útil atribuir pontos por dicotomização, em termos de alcance de objetivos. O tópico "Atribuição de Pontos", pág. F.199, tratará deste aspecto.

A aplicação dos coeficientes de validade e fidedignidade dos testes CR será discutida nos tópicos "Validade", pág. F.179, e "Confiabilidade", pág. F.190.

2) Vantagens do Teste CR para a Avaliação de Instrução

O papel do teste CR é ser o mais descritivo possível, com pouca ou nenhuma esperança de variância nos escores. Por causa disto, os elaboradores de testes CR podem reter itens particulares aos objetivos, ou mesmo o teste todo. Um item que é respondido corretamente por todos continua no teste CR, apesar de para o NR não ter valor.

Em geral, por causa dos testes CR reterem os itens de teste, que são reflexos do critério, independente da contribuição para a variância, trabalha-se com índices mais descritivos.

Os testes CR são mais indicados para detetar os efeitos da instrução, em termos das mudanças comportamentais planejadas, como também para avaliar o aluno nos mesmos termos.

3) Procedimentos de Desenvolvimento de Teste CR

Atualmente, não existem procedimentos sistemáticos estabelecidos para a construção de um teste CR. Uma das maneiras possíveis é:

- . Planejamento do Teste
- . Geração e integração de itens
- . Revisão técnica do teste
- . Avaliação formativa do teste

- . Planejamento do Teste

Basicamente o plano do teste deve ser feito como já descrito para o planejamento do teste NR, acrescido da definição dos limites dos comportamentos que os itens de teste devem medir, com os quais os desempenhos dos indivíduos vão ser referenciados. É a definição do domínio.

A definição do domínio deve ser o suficiente:

- concisa de modo que possa ser usada tanto pelos elaboradores de itens como pelos educadores que vão agir com base nos dados de medida. A definição serve como um conjunto de especificações aos elaboradores de itens. A tarefa de elaborar é, portanto, produzir itens que sejam congruentes com as especificações;
- restrito à classe de comportamentos em consideração, de modo que avaliadores, independentes, registrem alto acordo em termos de que

os itens do teste medem, realmente, os comportamentos descritos no domínio.

. Geração e Integração de Itens

Deve-se gerar o número necessário de itens que vão compor o teste, uma vez que não é necessário descartar itens.

Deve-se elaborar a chave e os critérios de correção, verificar a congruência entre os itens e o objetivo e outros (já descritos no tópico "Geração de Itens" pág. F.161).

. Revisão Técnica

O procedimento é o mesmo da abordagem NR. É bom que os especialistas não estejam envolvidos nem na definição do domínio, nem na elaboração de itens. O papel do especialista é verificar a congruência entre os itens e a definição do domínio. Isto permite a descoberta de ambiguidade e outros defeitos tanto nos itens como no domínio, que precisa ser reformulado.

. Avaliação Formativa do Teste

Consiste em:

- testagem um a um
- testagem em uma amostra
- análise dos resultados para avaliar os itens.

. Testagem Um a Um

É igual à testagem um a um do teste NR

. Testagem em uma Amostra

Após as revisões feitas com base no controle preliminar dos especialistas e testagem um a um, deve-se administrar os itens a um grupo

de indivíduos representativos de população-alvo e que possuem o domínio das habilidades que se está tentando medir. Se, muitos itens, em termos gerais, são respondidos da mesma maneira pelos indivíduos, tem-se, então, alguma evidência da clareza de definição do domínio. Mas, se acontecer o contrário, isto deve ser revisto cuidadosamente.

O número de indivíduos, a ser incluído nesta amostra, depende do tempo e recursos financeiros disponíveis.

. Análise dos Resultados para Avaliar os Itens

O ponto crítico na abordagem CR é a interpretação dos dados, pois não há um guia definido.

Alguns autores defendem a utilização dos índices de discriminação e dificuldade ao teste CR (Stone, 1974), no caso de haver algum grau de variância. Em se tratando do índice de discriminação, havendo muitos índices negativos, o problema está na definição do domínio. O índice de dificuldade foi estudado, por exemplo, em situação de pré e pós-teste. No pré, o índice seria próximo a zero (difícil antes da instrução) e um valor próximo a um pós-teste (fácil após instrução).

O que acontece, na verdade, é que o avaliador deve empregar o seu bom senso para inspecionar os resultados do ensaio dos itens. Os avaliadores devem acompanhar o desenvolvimento de novos procedimentos para analisar dados de teste CR.

Uma opção para melhorar a qualidade dos itens é verificar as características desejáveis de um teste CR segundo Gagné (Stone, 1974), já citado na pág. F.176.

VALIDADE

Este tópico trata de validade especificamente para teste de aprendizagem, embora seja aplicável a todos os instrumentos.

Um teste é válido na extensão em que mede a variável que pretende mensurar.

Testes diferentes têm virtudes diferentes. Nenhum teste em algum campo é o melhor para todos os propósitos. Por exemplo, se um teste é usado para selecionar calouros, deve ser válido para este propósito, mas não necessariamente válido para outros propósitos.

O que se valida não é o instrumento de medida, mas sim o uso no qual o instrumento é colocado.

Não existe instrumento válido ou não válido em termos absolutos. O que existe é validade em determinado grau. A pergunta que se deve fazer a um teste é: Quanto válido é o teste para a decisão que se quer tomar? (Cronbach, 1970).

O processo de validação exige sempre investigações empíricas. As medidas psicológicas servem aos propósitos maiores de:

- a) representar um universo de conteúdo especificado - validade de conteúdo;
- b) estabelecer uma relação funcional com uma variável particular - validade preditiva e concorrente;
- c) medir características psicológicas - validade de construto.

a) Validade de Conteúdo

Um teste é uma tentativa de determinar como um indivíduo se desempenha em um conjunto de situações reais. Em vez de colocar a pessoa, em cada situação real, o teste é usado como um atalho para determinar seu comportamento em um conjunto de situações. Então, o teste exige uma amostragem de um universo de situações do conjunto total. Com base no desempenho do indivíduo nas situações amostradas, pode-se generalizar para o conjunto completo de situações.

O julgamento da adequação do conteúdo do teste como evidência do desempenho de um universo maior de situações é chamado validação de conteúdo. A adequação do conteúdo é especialmente importante para testes que medem resultados de educação ou treinamento.

A validade de conteúdo é assegurada pelo plano de procedimentos de construção do teste, através de uma definição apropriada do universo que é representado no teste, através do processo de amostragem.

A definição do universo deve englobar:

- . os tipos de tarefas, estímulos ou situações;
- . os tipos de respostas com que o observador ou atribuidor de pontos deve contar.

Se a definição do universo é clara, o usuário pode ter confiança que o teste reflete o universo de interesse.

Frequentemente, é logicamente impossível ou impraticável amostrar realmente o conteúdo. O que se faz comumente é:

- . a formulação de uma coleção de itens que representa amplamente a unidade da instrução. Para assegurar que os itens representam realmente a unidade da instrução é necessário ter um esquema detalhado dos tipos de questões e problemas que serão incluídos;
- . a seleção do conteúdo geralmente envolve questões de valores, isto é, amostra-se apenas o que os professores acham mais importantes, ficando o teste restrito a isto.

Às vezes examina-se a correlação de cada item com o escore total do teste, e os itens com baixo coeficiente são retirados. Se se tirar um item particular, provavelmente não afetará a validade de conteúdo do teste. O perigo é que se retire muitos dos itens pobremente construídos de uma mesma área de conteúdo, perdendo assim sua representatividade (Nunnally, 1967).

O elaborador de teste que examina um item com correlação insatisfatória pode explicar os erros de bons alunos: uma dupla negação, duas respostas plausíveis e outras, devendo o item ser reformulado.

Talvez a maneira máxima de assegurar validade de conteúdo seja através da não introdução de dificuldade irrelevante, por exemplo, às vezes erra-se um item por dificuldades verbais. Alguns elaboradores de teste tentam fazer itens difíceis, oferecendo alternativas traiadoras ou exigindo boas discriminações, porque acreditam que um "bom" teste dispersa as pessoas. Onde a validade de conteúdo é o alvo, a discriminação entre pessoas não é.

A validade de conteúdo envolve a inspeção dos itens por alguém que decide se eles estão o suficiente consoante com o conteúdo ou comportamento que vai ser medido. Há, portanto, uma forte confiança no julgamento humano. Às vezes a expressão validade curricular é usada para descrever esta abordagem julgamental.

A taxionomia de Benjamin Bloom, et. al (1973) e os tipos de aprendizagem de Robert Gagné (1973) são bons guias para definir as habilidades do Universo. Assim, pode-se verificar a validade de conteúdo de testes de aprendizagem, pois é possível comparar os itens de teste com as referidas classificações que os originam.

b) Validade Preditiva e Concorrente

. Validade Preditiva

A validade preditiva é de interesse quando o propósito é usar um instrumento para estimar alguma forma importante de comportamento, sendo este referente a um critério, que é, normalmente, um instrumento que serve como um padrão de comparação.

O termo predição é usado para referir-se a relações funcionais entre um instrumento e evento que ocorrem antes, durante e depois que um instrumento é aplicado.

A validade preditiva consiste na correlação entre escores, dos indivíduos, no teste preditivo com escores de uma variável (critério) dos mesmos indivíduos. A validade é determinada pelo grau de correspondência entre as duas medidas envolvidas.

Examina-se, por exemplo, a validade preditiva entre um teste para selecionar um candidato a um emprego e o desempenho dele, no emprego.

A validade preditiva é somente usada em certos tipos de problemas aplicados em psicologia e educação, como por exemplo, separar alunos em diferentes níveis de habilidades.

Não se pode confundir o instrumento preditor com o critério que se pretende predizer. É mais difícil obter um bom critério do que um teste preditor. Um tipo comum de critério é a escala ou grau.

Os testes preditivos são válidos somente na extensão que servem bem as funções de predição.

As técnicas para empregar testes preditivos e para validar tais usos são parte da psicologia aplicada para seleção de pessoal. Na seleção de pessoal, há muitas estratégias de decisão e muitos "designs" de Estatística que relacionam-se ao emprego de diferentes tipos de medidas preditivas, para diferentes tipos de trabalhos, com relação aos diferentes tipos de critérios.

A validade de instrumentos preditivos é determinada pelo grau de correlação entre o instrumento e o critério. É errado pensar em termos de correlações perfeitas ou mesmo correlações altas; é razoável ter correlações entre 0.30 e 0.40 (Nunnally, 1967):

Empregando-se um instrumento preditivo, podem ser feitos muitos erros na predição do desempenho dos indivíduos, isoladamente. A maneira apropriada de interpretar um coeficiente positivo de validade

preditiva, por exemplo, é em termos da extensão em que ele indica uma melhoria possível na qualidade média que seria obtida das pessoas, empre-gando-se tal instrumento. Na média, as pessoas que têm escore alto no teste, desempenham-se consideravelmente melhor do que as pessoas que têm escore baixo no teste (Nunnaly, 1967).

. Validade Concorrente

A validade é concorrente quando o teste é validado em relação a um critério, na mesma época, isto é, sem lapso de tempo entre a testagem e os valores obtidos do critério.

O que é comumente feito é correlacionar o desempenho num teste com o desempenho em outro, ou ainda comparar qualidades ou desempenhos avaliados, pelo teste, com qualidades ou desempenhos avaliados por outro procedimento qualquer.

c) Validação de Construto

O grau no qual é necessário validar medidas de variáveis psicológicas é proporcional ao grau em que a variável é concreta ou abstrata. Na extensão em que a variável é abstrata ela é chamada de construto.

A validação de construto é uma análise do significado dos escores do teste em termos de conceitos psicológicos ou construtos. Todo teste é impuro em algum grau e, muito raramente, mede exatamente o que seu nome implica.

Na determinação da validade de construto, o objetivo é identificar todos os fatores que influenciam o desempenho no teste e determinar o grau de influência de cada um. O processo inclui três partes:

- . identificar os construtos que poderiam, possivelmente, explicar o desempenho no teste. Isto é um ato de imaginação, baseado na observação ou estudo da lógica do teste;

- . derivar da teoria, hipóteses testáveis em torno do construto;
- . coletar dados para testar uma hipótese após a outra.

A validade de construto consiste no processo de tomar uma palavra abstrata, explícita em termos de variáveis observáveis (indicadores). Esquematizar um construto consiste, essencialmente, em estabelecer o que se quer dizer pelo uso de palavras particulares, como inteligência, coragem, ansiedade. Nenhum método preciso pode ser estabelecido para esquematizar apropriadamente o domínio de variáveis para um construto. O esquema constitui, essencialmente, uma teoria cuidando de como as variáveis se relacionarão uma com a outra. Em tentativas anteriores de esquematizar um domínio, o esquema comumente consiste de somente uma definição na qual a palavra que denota o construto é relacionada a palavras em um nível de abstração mais baixo.

Na validade de construto tanto o teste como a própria teoria do construto estão sendo validadas, ao mesmo tempo. Se a(s) hipótese(s) for(em) confirmada(s), a teoria relativa ao construto é sustentada e a medida é validada.

Os métodos (Cronbach, 1970) mais comuns de obter evidência do uso da abordagem da validação de construto são:

- . comparação dos escores antes e depois de um tratamento particular;
- . comparação de escores de grupos conhecidos;
- . correlações com outros testes.

O coeficiente de correlação resultante será evidência de validade de construto. Se a evidência falha em sustentar a hipótese, há algo errado ou:

- . no modo em que o construto foi conceitualizado;
- . no teste;
- . em ambos - teoria e teste.

A Tabela F.14 apresenta as questões levantadas pelos tipos de validade, seus procedimentos, aplicações e exemplos, baseadas na classificação de Cronbach (1970).

Não há dúvida que os tipos de validade discutidos são aplicáveis à abordagem NR.

Como a abordagem CR é recente, ainda estão em discussão, entre os teóricos, a questão da validade aplicada a esta abordagem.

Há, pelo menos, acordo entre muitos, sobre a aplicação da validade de conteúdo à abordagem, sendo as medidas CR validadas primariamente em termos da adequação com que representam o critério. O procedimento geral para a validação consiste em um julgamento cuidadosamente feito, baseado na relevância aparente do teste aos comportamentos inferidos daqueles delimitados pelo critério.

Alguns autores recomendam (Stone, 1974), além das estratégias de validade de conteúdo, as abordagens de validade de construto, para se sustentar a confiança que se tem no instrumento.

Popham (1975) usa, para a abordagem CR, 3 tipos de validade:

- . validade descritiva
- . validade funcional
- . validade de seleção de domínio

. Validade Descritiva

Um teste é válido na extensão em que mede acuradamente a classe de comportamentos do aprendiz, descritos na definição do domínio.

Os testes CR que não possuem validade descritiva são de pouca utilidade, porque o avaliador educacional não pode esclarecer sobre o que o teste está medindo.

QUESTÕES, PROCEDIMENTOS, APLICAÇÃO E EXEMPLOS DOS TIPOS DE VALIDADE

QUESTÃO	TIPO DE VALIDAÇÃO E SEUS PROCEDIMENTOS	APLICAÇÃO	EXEMPLO
As observações realmente a mostram o universo de tarefas ou situações que pretendem representar?	<u>Validação de conteúdo</u> Examinam-se os itens de teste e suas respostas, comparando-os com uma descrição completa do universo que se supõe que o teste representa.	Teste usado para avaliar programas educacionais e para avaliar a aprendizagem do indivíduo.	Os itens de um teste sobre redação de um ofício é comparado com o conteúdo de um manual sobre redação oficial.
Como as medidas de algum desempenho (critérios) relacionam-se aos escores do teste?	<u>Validação preditiva</u> Aplica-se o teste preditivo. Depois de algum tempo de treinamento ou de trabalho, coleta-se dados sobre o critério. Examina-se a correspondência dos escores do critério e os do teste.	Teste usado na seleção de empregados e estudantes.	Aplica-se um teste para admissão de estudantes em uma universidade. No final do curso, correlaciona-se os escores do teste com as notas do curso (critério).
Como os escores do teste podem ser explicados psicologicamente? O teste mede o atributo que diz medir?	<u>Validação do construto</u> Levantam-se hipóteses relativas ao significado dos escores do teste, estabelecendo-se como se espera diferenças entre os escores altos e baixos. Testam-se as hipóteses uma a uma.	Teste que mede processos mentais ou características pessoais	Estuda-se um teste de arte para determinar o quanto grandemente os escores dependem de treinamento em arte.

Segundo este autor, a validade descritiva se assemelha à validade de conteúdo da abordagem NR. Ele afirma, que nas situações de medida NR, o julgamento dos revisores é, algumas vezes, obscuro, porque a descrição do conteúdo, que fornece limites para o julgamento do item, é pouco detalhada. No CR, contudo, a existência de uma descrição do domínio, satisfatoriamente explicada, torna a tarefa do revisor mais rigorosa. O trabalho do revisor é identificar a proporção de itens congruentes com a descrição do domínio. Se os itens estão fora do alvo, ou se a descrição do domínio não é precisa, o revisor pode tomar uma decisão negativa quanto a validade descritiva do instrumento. Sem validade descritiva, não se pode entender muito sobre o desempenho de um indivíduo em um teste.

Como as abordagens usadas para se estabelecer validade descritiva (CR) parecem as mesmas da validade de conteúdo (NR), o autor acha que o termo conteúdo não é preciso para refletir dimensões não cognitivas, por isso, considera o termo descritiva mais geral. Como o propósito total do teste CR é referenciar acuradamente o desempenho de um indivíduo em relação a um domínio de comportamento, o uso do termo validade descritiva parece ser mais adequado.

A validade descritiva é considerada um pré-requisito das outras formas de validade.

. Validade Funcional

Corresponde à validade preditiva, em que se tenta prever o desempenho de um indivíduo em uma situação subsequente.

A razão de escolher um termo mais geral, validade funcional, é que há muitas funções em que o avaliador coloca medidas CR. Algumas funções envolvem um critério, outras não.

Ocasionalmente, os avaliadores encontram situações em que desejam que os testes não só descrevam acuradamente o desempenho

dos indivíduos em termos do que eles podem ou não fazer (validade funcional), mas também desempenhem outra função. Neste caso, pode-se inspecionar a validade funcional.

Para o teste CR ter validade funcional é necessário que os indivíduos que se desempenhem bem no teste, depois da instrução, desempenhem-se bem quando em situação real.

Quando se usa um teste CR, deve-se reconhecer que deve ser abandonado qualquer movimento que tenda ao aumento da validade funcional, em prejuízo da validade descritiva.

. Validade de Seleção do Domínio

Refere-se a acuracidade com que o domínio foi selecionado. A primeira decisão que o elaborador de teste CR toma é que domínio (classes de comportamento do aprendiz) indica melhor a dimensão maior que se quer medir, uma vez que existem diferentes domínios. Em algumas situações, a escolha do domínio é clara, em outras é mais difícil. Na verdade, não se empenha numa seleção de um domínio entre 12 ou mais alternativas. O que se faz é definir um domínio de uma maneira, em vez de muitas maneiras alternativas.

Um teste CR possui validade de seleção do domínio quando pode ser determinado que o domínio particular escolhido pelo elaborador de teste, serve, adequadamente, como indicador da dimensão maior de interesse que foi planejado para medir. O conceito de "generalizabilidade" é o núcleo da validação de domínio, porque o que se quer saber é quão generalizável o domínio selecionado é para os outros domínios.

Há uma similaridade entre as técnicas de validação de seleção do domínio - CR - e técnicas de validação de construto - NR. Contudo, na validação de seleção do domínio não se está interessado em estabelecer a existência de algum construto hipotético. A tarefa do teste CR é simplesmente decidir se se fez um bom trabalho quando se resolveu limitar a classe de comportamentos que se constitui na definição do domínio.

A abordagem de validade ao teste CR segundo Popham (1975) ainda está em estudo.

ERRO DE MEDIDA

Nenhuma simples observação é inteiramente representativa da habilidade de um indivíduo. O escore obtido indica grosseiramente o nível da habilidade ou comportamento típico. Para saber a confiança que se tem no teste, examina-se a consistência entre medidas. A consistência pode ser verificada aplicando-se o mesmo teste repetidamente, ou usando-se formas similares.

Os escores de um teste variam de uma medida para outra. Vários fatores podem provocar essa variância. Todas as causas que não estão ligadas ao comportamento que está sendo medido, pelo teste, são interferências. Tais variâncias produzidas são "erros".

A discussão básica do erro de medida é que cada pessoa tem um escore verdadeiro ou escore-universo que seria obtido se o escore do indivíduo não fosse sujeito a influências não desejadas.

Uma observação particular é somente uma das muitas observações que podem ser feitas. Uma única observação não é representativa do escore do indivíduo.

Os escores de um indivíduo, em um teste, depende de condições como: testador, dia, hora, itens de teste, tensão emocional, fadiga, flutuações de memória, atribuidor de pontos e outros. Cada uma destas variações é uma fonte de variância, porque se a pessoa for retetada o escore muda.

O escore-universo de um indivíduo é aquele que se obteria se um grande número de medidas fosse coletado e se extraísse a média.

Como não é possível obter o escore-universo de um indivíduo, o escore observado tem de substituí-lo. A medida que há desacordo entre ambos, há erro de medida. Qualquer escore observado é uma amostra da distribuição e fornece uma estimativa relativamente inexata.

Se os escores observados aproximam-se do escore-universo, pode-se dizer que a observação é precisa ou confiável. E desde que as observações concordem uma com as outras, diz-se que são consistentes e têm pouco erro de variância. O termo mais usado na literatura é confiabilidade e em anos recentes "generalizabilidade" (Cronbach, 1970).

O erro de medida é a diferença entre o escore-universo de um indivíduo e o seu escore em uma observação. Se, por exemplo, o escore observado é 62 e o escore-universo 58, a discrepância é de 4 pontos, isto é, o erro de medida é de 4 pontos.

O fato de se adicionar observações, no teste, aumenta a probabilidade de o escore observado se aproximar do escore-universo.

Como não se sabe o escore-universo de um indivíduo, não se pode determinar o erro de qualquer escore particular observado. Mas, pode-se estimar o tamanho que tais erros tendem a ter, através do desvio-padrão da distribuição de uma única pessoa. Este desvio padrão dos escores observados é chamado erro padrão de medida que é típico, isto é, é igual para todos os indivíduos. O erro padrão de medida é simbolizado por e.p.m.

A Figura F.14 (Cronbach, 1970) mostra a relação entre o escore observado e o escore-universo.

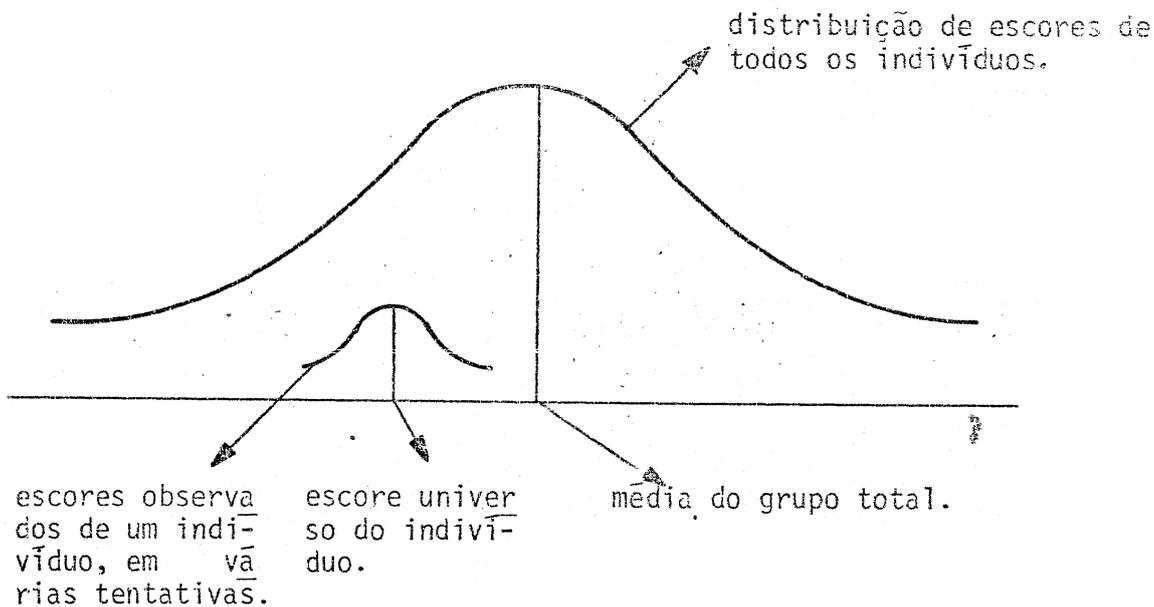


Fig. F.14 -- Relação entre o Escore-Universo e o Escore Observado

COEFICIENTES DE CONFIABILIDADE OU "GENERALIZABILIDADE"

Se os escores dispersam em uma escala de 100 pontos, um erro de 5 pontos tem pouco efeito. Se é de somente 10 pontos, um erro de 5 pontos é enorme. Compara-se, então, o tamanho dos erros com a distribuição de escores. Tal comparação (Cronbach, 1970) é a razão de duas variâncias que é chamado coeficiente de "generalizabilidade" ou confiabilidade.

$$\frac{\text{Variância dos escores-universo}}{\text{Variância dos escores observados}}$$

Esta razão é geralmente denotada pelo símbolo r_{xx} (ou similar). Uma razão de 0,45 significa que 45% da variância observada pode ser explicada pelas diferenças reais dos escores-universo (Cronbach, 1970).

Um simples escore observado não diz nada sobre o erro de medida. Para investigar a magnitude dos erros, devemos ter duas ou mais observações, por pessoa. O estudo do erro é parte essencial do desenvolvimento do instrumental. O elaborador de instrumentos coleta a informação e talvez a usa para melhorar o instrumento e relata-a no manual do teste, de modo que se pode decidir se o teste é o suficientemente preciso para seus propósitos, isto é, se mede consistentemente o que quer mensurar.

Hã 4 formas gerais de confiabilidade de testes estimadas:

- a) Estabilidade
- b) Equivalência
- c) Estabilidade e equivalência
- d) Consistência interna.

- a) Estabilidade

Consiste em administrar a mesma forma de testes, aos mesmos indivíduos, em ocasiões diferentes e depois comparar o desempenho dos indivíduos em ambas as testagens. O método é conhecido como teste-reteste. A estabilidade é indicada pela correlação dos escores dos sujeitos na 1a. ocasião (t_1) e 2a. ocasião (t_2).

O coeficiente pode variar de zero (não existe relação entre eles) até 1.00 (relação perfeita). Os coeficientes próximos de zero são raros. É desejável obter coeficientes altos. O coeficiente de correlação resultante frequentemente fica entre 0,80 a 0,95 (Popham, 1975).

O teste-reteste tem a vantagem de exigir apenas uma forma de teste, mas tem como desvantagem ser influenciado pelo hábito e memória ou, ainda, por quaisquer eventos que possam ocorrer entre as duas testagens. O intervalo entre as duas aplicações é importante.

O método teste-reteste é o único que fornece informação sobre a consistência do teste no tempo, isto é, diz se o teste é estável. Esta qualidade de um teste é importante principalmente quando se pretende pré e pós testagem.

b) Equivalência

Envolve a administração de duas formas de um teste, aos mesmos indivíduos, em ocasiões diferentes. As duas formas de teste devem ser paralelas no conteúdo e nas operações mentais exigidas, isto é, os dois testes devem ter itens de uma forma condizente com os itens de outra forma.

A equivalência é indicada pela correlação entre os resultados obtidos nas duas formas de teste. Este coeficiente é uma medida tanto de estabilidade temporal, quanto de consistência de resposta a formas similares do teste.

Esta abordagem pode ser usada para:

- . avaliar a confiabilidade de ambas as formas, pela comparação de uma com a outra;
- . determinar o grau no qual as duas formas são paralelas. Esta é particularmente importante, por exemplo, se uma forma vai ser usada como pré-teste e outra como pós-teste.

Frequentemente o coeficiente resultante excede a 0,80 (Popham, 1975).

Para a elaboração de formas paralelas é preciso:

- . construir testes independentes a fim de satisfazer às mesmas especificações;
- . elaborar o mesmo número de itens, sob a mesma forma e conteúdo;
- . manter a mesma amplitude e nível de dificuldade.

c) Estabilidade e equivalência

Este método envolve um esforço simultâneo de estabelecer o grau de consistência de uma medida com relação à estabilidade e equivalência.

Consiste na administração de duas formas equivalentes de um mesmo teste, aos mesmos indivíduos, com um intervalo de tempo entre ambas.

A equivalência e estabilidade é indicada pela correlação entre os resultados obtidos em ambas as formas, pelos mesmos indivíduos. O coeficiente resultante é, comumente, mais baixo do que com os métodos usados separadamente.

d) Consistência interna

A consistência interna consiste na verificação se todos os itens de um mesmo teste medem a mesma qualidade.

Envolve o cálculo de confiabilidade com dados de uma única administração do teste. Existem dois procedimentos:

- 1 - confiabilidade das metades
- 2 - confiabilidade de Kuder-Richardson

1 - Confiabilidade das metades

Consiste na divisão do teste em duas metades, separando-se os itens de número par e os de número ímpar, como se constituíssem testes separados. Administra-se o teste total a um grupo de indivíduos e relaciona-se os escores obtidos de cada indivíduo, em ambas as metades.

O coeficiente de correlação resultante é considerado uma estimativa do grau em que duas metades de um teste estão medindo a mesma qualidade ou característica, consistentemente. Com a relação obtida

(r_1) calcula-se a confiabilidade do teste total (r_2) pela aplicação da fórmula de Spearman Brown (Popham, 1975):

$$r_2 = \frac{2 r_1}{1 + r_1}$$

2 - Confiabilidade Kuder-Richardson

Consiste na administração de um teste a um grupo de indivíduos e depois verifica-se o escore de cada item individual para ver o grau em que os itens estão medindo a mesma característica ou qualidade. Depois, aplica-se a forma de Kuder-Richardson KR-20 ou KR-21. A fórmula KR-20 (Popham, 1975) é apropriada para itens dicotômicos e a KR-21 para itens multipontos. As fórmulas são as que seguem:

$$r_{KR20} = \left(\frac{K}{K - 1} \right) \left(\frac{\sigma^2 - p_i q_i}{\sigma^2} \right) \quad \text{onde}$$

r_{KR-20} = coeficiente de Confiabilidade Kuder-Richardson

K = número de itens do teste

σ^2 = variância do teste

p_i = proporção de indivíduos que acerta o item i

q_i = proporção de indivíduos que erra o item i

$$r_{KR-21} = \frac{K}{K - 1} \left(1 - \frac{M (K - M)}{K s} \right) \quad \text{onde}$$

K = número de itens do teste

M = média do conjunto de escores do teste

s = desvio padrão do conjunto de escores do teste

RELAÇÃO ENTRE OS PROCEDIMENTOS DE ESTIMAR A CONFIABILIDADE ÀS ABOR-
DAGENS NR E CR

Não há dúvida de que os quatro procedimentos discutidos para estimar a confiabilidade de testes são aplicáveis às medidas NR.

Quanto às medidas CR, ainda não existe um acordo sobre a aplicabilidade dos procedimentos clássicos, uma vez que estes são baseados em variância. Ainda não há acordo sobre o que deve substituí-los.

Há diferentes interpretações da confiabilidade na abordagem CR. Para Robert Gagné (Stone, 1974), a confiabilidade está relacionada a questão: quantas amostras de desempenho devem ser tomadas de uma simples classe, a fim de que se saiba se o indivíduo pode ou não desempenhar a classe. Gagné (Stone, 1974), sugere o sucesso em 2 a 3 itens de uma classe, como o mínimo razoável para uma avaliação confiável do domínio.

Ivens (Stone, 1974) indica o uso de dois tipos de índices de confiabilidade ao teste CR:

- . confiabilidade de item - significando a proporção de indivíduos com o mesmo escore, no item, no pós teste e reteste (segunda aplicação do pós-teste). Seria uma medida de estabilidade de item (no caso de se utilizar reteste) ou equivalência de item (no caso de se usar uma forma paralela do pós-teste);
- . confiabilidade de escore - é calculada através de dois métodos:
 - computação da média da confiabilidade de escore. Isto permite detetar os itens que causam baixa confiabilidade de escore;
 - computação da diferença entre a percentagem de resposta em uma administração de teste e a percentagem em uma segunda administração; devendo ser, tal computação, para cada indivíduo.

Existem ainda muitas pesquisas relativas ao problema da confiabilidade do teste CR.

Cronbach, Rajaratnan e Glaser (Stone, 1974) apresentaram uma reinterpretação do conceito de confiabilidade ao teste CR, através do coeficiente de "generalizabilidade" (G). O estudo deste conceito responde a questão: quão bem o domínio das habilidades do teste pode ser generalizado ao domínio de todas as habilidades do universo de desempenho, isto é, quão adequadamente pode-se generalizar a observação a uma classe de observações a que pertence.

O coeficiente G é definido como o quadrado da correlação entre o escore observado e o escore-universo. O procedimento básico, que compõe o estudo G, utiliza a técnica de análise de variância ao computar este coeficiente.

O coeficiente G foi aplicado, no teste CR, em um estudo feito por Stone (1974).

O propósito do estudo G é avaliar a técnica de medida e não avaliar indivíduos. Porém, como o G depende da variabilidade de escores, pode não ser significativo se for usado com um grupo de indivíduos, homogêneo, no seu nível de domínio. O problema pode ser resolvido (Stone, 1974), no estudo G, tendo-se uma amostra da população, razoavelmente grande, em que seja possível incluir pessoas de diferentes níveis de domínio, como por exemplo, diferentes graus de aprendizagem.

O coeficiente G também responde parcialmente a questão da validade, em termos de quão validamente pode-se interpretar uma medida como representativa de um certo conjunto de medidas possíveis.

Popham (1975) discute que os quatro tipos de estimativas de confiabilidade do teste NR podem ser usados para medidas CR, mas com algumas cautelas, uma vez que estas abordagens contam com a presença de considerável variância nos escores do teste.

Quando a variância é pequena, como seria o caso de um tratamento instrucional, pode-se usar métodos menos sofisticados

ticados como, por exemplo, simplesmente calcular a percentagem de escores dos indivíduos em duas ocasiões diferentes.

O autor acrescenta que o significado das estimativas de estabilidade e equivalência é essencialmente o mesmo para as abordagens NR e CR, mas que a estimativa de consistência interna apresenta um problema diferente.

Como as estimativas de consistência interna estão focalizadas, não na consistência de um teste em diferentes ocasiões ou tempo, mas a consistência dos itens individuais dentro do teste, faz mais sentido, para o teste CR, concentrar-se na confiabilidade como teste total, portanto as estimativas de equivalência e estabilidade parecem ser mais apropriadas (Popham, 1975).

É importante ressaltar que não há, ainda, nada definitivo com relação a aplicação dos coeficientes de confiabilidade clássicos aos testes CR, pois ainda estão sendo estudados.

RELAÇÃO ENTRE VALIDADE E CONFIABILIDADE

A confiabilidade é uma condição necessária, mas não suficiente para a validade do teste. A existência de um teste confiável torna possível a validade. A confiabilidade não garante validade, pois um teste pode ser confiável, mas não ser válido. O contrário não pode ocorrer, isto é, o teste válido é confiável.

ATRIBUIÇÃO DE PONTOS

Os procedimentos de atribuição de pontos variam conforme o teste seja de questões abertas ou não.

Na discussão deste tópico, será suposto que o indivíduo tenha todos os pré-requisitos colocados anteriormente nas páginas 100-101, vol. 1. Os conceitos estatísticos referidos, em particular, podem ser encontrados em qualquer livro de estatística básica.

Os testes de questões abertas que exigem resposta cons
truída precisam ter regras para julgamento para que o(s) atribuidor(es)
de pontos siga(m).

Nos outros formatos, o indivíduo deve escolher uma res
posta. A chave de atribuição de pontos é fixada anteriormente. Neste ca
so os testes podem ser corrigidos à mão ou à máquina.

O processo de atribuição de pontos das abordagens NR e
CR pode ser idêntico, quando, mesmo no CR houver variabilidade nos esco
res. Mas, como a variabilidade é irrelevante ao CR, no caso de ins
trução individualizada e Avaliação absoluta é mais útil atribuir pontos
por dicotomização em termos de alcance de objetivos.

INTERPRETAÇÃO DE ESCORES CONSIDERANDO:

- a) Avaliação para decisão comparativa
 - b) Avaliação para decisão absoluta.
-
- a) Avaliação para Decisão Comparativa

1) Escores Brutos

O relato numérico direto de um teste de desempenho é cha
mado escore bruto.

Os escores brutos não podem ser interpretados em testes
psicológicos como em medidas físicas. As escalas de medida física, geral
mente, têm um zero verdadeiro e unidades iguais ao longo da escala. Isto
permite dizer, por exemplo, que uma menina é duas vezes mais pesada que
outra. O mesmo não pode ser feito em medidas psicológicas. Um escore de
25% obtido por um indivíduo em um teste de soletração, não significa
que ele sabe 1/25 d
as palavras que devia saber. Provavelmente, o profes
sor não perguntou as palavras que ele sabia. Mesmo um escore zero num
teste não significa zero na habilidade medida. As diferenças em escore
bruto não representam ordinariamente distâncias reais entre os indivíduos.

Os escores brutos de indivíduos pertencentes a grupos diferentes não podem ser comparados. Os escores brutos não dizem nada, por si sós, a não ser que se conheça o tamanho do grupo, média e desvio padrão, ou seja, que sejam transformados em escores derivados.

2) Escores Percentis

O modo mais fácil de fazer comparações é ordenar os escores do mais alto ao mais baixo. A ordenação depende do número de indivíduos no grupo, pois este tamanho varia. Os escores brutos são transformados em escores percentis. O escore percentil é expresso em termos de percentagem.

Um percentil diz que proporção do grupo fica abaixo dele. O percentil 50 é também chamado mediana.

Um dos procedimentos usados para o cálculo de percentis (Cronbach, 1970) é dado pela construção de gráfico de frequência que consiste em:

- . ordenar os escores brutos do mais alto ao mais baixo
- . agrupá-los dentro de um intervalo de classe escolhido
- . colocar as frequências para cada classe
- . computar a frequência acumulada
- . computar a frequência relativa (%)
- . traçar o gráfico, tendo, na abcissa, os escores brutos (sendo o limite inferior) e na ordenada a percentagem
- . determinar a percentagem equivalente ao escore bruto pela leitura da curva.

Exemplo:

A partir dos escores desordenados:

63	56	47	57
62	48	57	62
60	46	61	55
58	45	49	51
45	43	62	44
52	54	68	52

ordená-los do mais alto ao mais baixo:

63	61	57	54	49	45
62	60	57	52	48	45
62	58	56	52	47	44
62	58	55	51	46	43

escores agrupados, frequências, frequências acumuladas, frequências relativas.

Exemplo baseado nos escores ordenados, agrupados.

ESCORES	FREQUÊN CIA	FREQUÊNCIA ACUMULADA	FREQUÊNCIA RELATIVA	FREQUÊNCIA RELATIVA ACUMULADA %
61 - 63	5	24	0,21	1,00
58 - 60	3	19	0,13	0,79
55 - 57	4	16	0,16	0,66
52 - 54	3	12	0,13	0,50
49 - 51	2	9	0,08	0,37
46 - 48	3	7	0,13	0,29
43 - 45	4	4	0,16	0,16
	<u>4</u> N : 24			

A Figura F.15 apresenta o gráfico relativo aos escores e a percentagem para determinação do percentil.

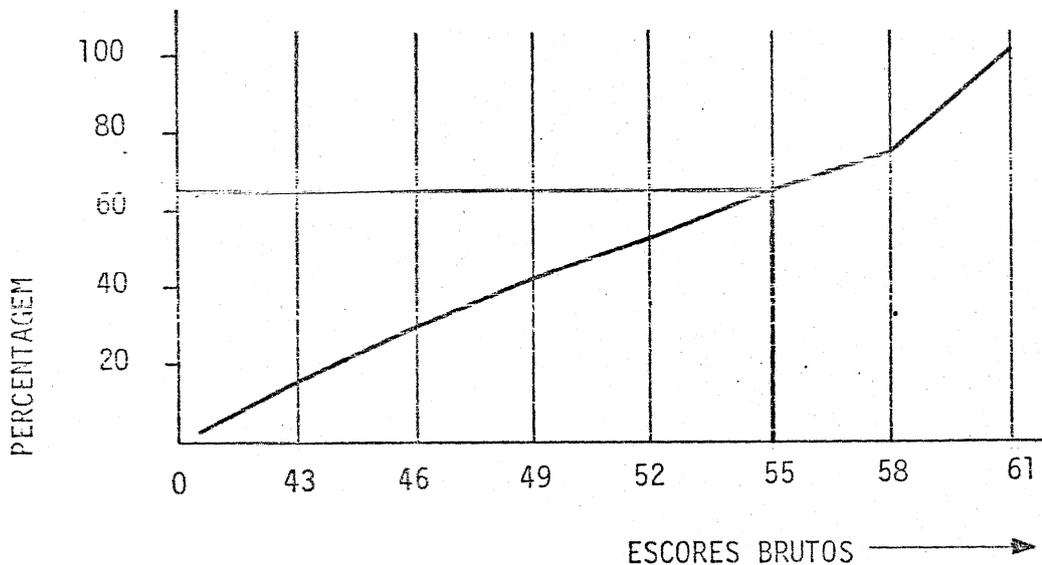


Fig. F.15 - Gráfico com as Percentagens e Escores Brutos para a Determinação do Percentil.

- determinar o percentil equivalente ao escore pela leitura da curva. O percentil equivalente ao escore bruto 55 é 56.

Os escores percentis não podem ser comparados a menos que os grupos, em que as conversões foram feitas, sejam similares. Sempre que se usa os percentis, o grupo de norma envolvido precisa ser lembrado.

O escore percentil tem como vantagens:

- . é facilmente entendido
- . é computado com facilidade
- . pode ser interpretado exatamente, mesmo quando a distribuição dos escores do teste não é normal.

Suas desvantagens são:

- . magnificar diferenças pequenas que são próximas à média.
- . reduzir o tamanho aparente de diferenças grandes e importantes na distribuição.
- . ser menos apropriados para análises estatísticas.

3) Escores-Padrão

O escore padrão fornece a distância que separa um escore bruto da média em termos de desvio padrão. Ele diz quantos desvios-padrão um indivíduo está abaixo ou acima da média.

Deve-se portanto transformar os escores brutos em escores padrão para que se possa fazer comparações.

Os escores padrão recebem nomes diferentes, conforme os valores respectivos da média e o desvio padrão:

<u>Escore Padrão</u>	<u>Média</u>	<u>Desvio Padrão</u>
escore Z	0	1
escore T	50	10
e outros	-	-

A transformação do escore bruto em escore Z obedece a fórmula:

$$Z = \frac{X - \bar{X}}{s} \quad \text{em que:}$$

X = escore bruto

\bar{X} = média do grupo

s = desvio padrão do grupo

Para a transformação do escore Z em T utiliza-se a fórmula:

$$T = 50 + (10) Z$$

É preferível adotar o escore T devido a duas inconveniências do escore Z:

- . muitos escores são negativos por causa da média ser igual a zero;
- . seu cálculo é muito trabalhoso, indo, muitas vezes, além da primeira casa decimal.

4) Escores Padrão Normalizados

Os escores-padrão só podem ser comparados se vierem de distribuições similares. Para ser possível proceder à comparação de escores vindos de distribuições com formas diversas, faz-se necessário a normalização de escores. Os escores normalizados são ajustados à curva normal.

A curva normal é uma curva de frequência que tem importantes propriedades matemáticas:

- . simétrica - isto é, a média, moda e mediana coincidem, e as percentagens de casos entre a média e os desvios-padrão positivos e negativos são os mesmos;
- . assintótica - ou seja, vai de menos infinito a mais infinito sem nunca tocar no eixo horizontal.
- . a área total sob a curva é 1,00, assim distribuída:

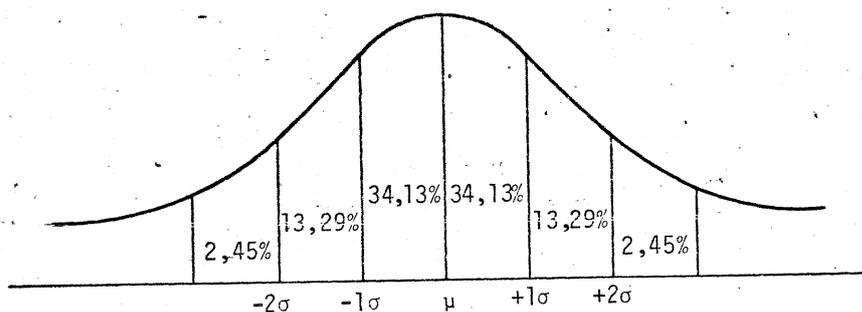


Fig.F.16 - Área sob a Curva Normal

Por exemplo, entre a média e $+1\sigma$ tem-se 34,13% dos casos; entre a média e $+2\sigma$ tem-se 47,42% dos casos, e assim por diante.

Em testes psicológicos e de aprendizagem é possível obter distribuição de escores aproximadamente normal. A forma de distribuição depende tanto da escala de medida como das experiências dos indivíduos.

Sempre que se tem uma distribuição normal, pode-se converter os escores padrão em percentis ou vice-versa. Abaixo da média há 50% dos casos. Abaixo $\mu + 1\sigma$ estão 84% dos casos.

Os escores são interpretados mais facilmente se são reduzidos a uma distribuição comum, isto é, se são normalizados. Os escores normalizados são obtidos através da extensão da distribuição para torná-la aproximadamente normal e depois transformá-la em uma escala de escore padrão.

O processo de normalização consiste em:

- . transformar os escores brutos em percentil;
- . transformar os percentis em escore Z com base na tabela da curva normal;
- . transformar os escores obtidos em uma distribuição normal, tomando-se como base uma média e um desvio padrão.

A vantagem do escore padrão é que:

- . as diferenças no escore padrão são proporcionais às diferenças do escore bruto;

Os escores padrão tem como desvantagens:

- . não podem ser interpretados facilmente, no caso de distribuição assimétrica;
- . as pessoas precisam ser treinadas em como interpretá-lo.

5) Perfis

Os testes variam tanto em tamanho como em dificuldade, por isso, o escore bruto sozinho não pode identificar o melhor desempenho dos indivíduos.

Os perfis (Cronbach, 1970) são elaborados a partir de escores percentis ou normalizados. Eles permitem comparar a posição de cada pessoa em vários testes.

A Figura F.17 apresenta um perfil de um aluno nos exames finais.

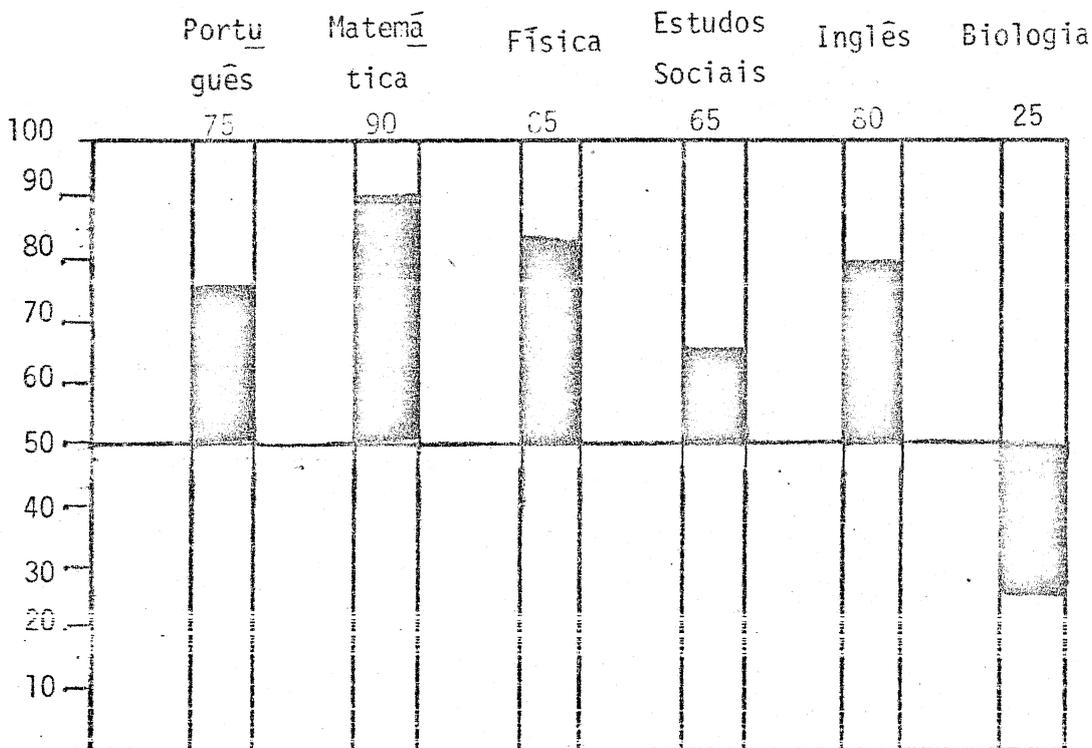


Fig. F.17 - Perfil de um Aluno nos Exames Finais.

Deve-se ter cuidado com as interpretações de perfis, pois, percentis iguais não significam igualmente bom ou mal desempenho.

NORMAS

O propósito das normas é descrever os escores do teste em termos de localização relativa a um grande corpo de escores coletados através de aplicação anterior.

As normas podem ser expressas em percentis ou escores padrão. Às vezes, apresentam-se tabelas de percentis equivalentes por grupos, sexo, grau, etc.

Para a elaboração de normas é necessário uma amostra aleatória da população. Se se pretende uma mistura de todos os segmentos da população, deve-se representar todos eles, sem amostrar mais um deles, pois pode-se enviesar as normas.

Em muitos testes, é necessário que se desenvolvam normas locais, além das nacionais.

Não se deve confundir normas com padrões, pois, por exemplo, indivíduos da 6a. série não alcançam, necessariamente, normas da 6a. série, isto porque os testes da escola representam somente o que a escola está habilitando, no momento, e o teste "padronizado" pode não ser representativo das habilidades do currículo.

As normas tornam-se obsoletas. Quando um teste é modificado, as normas devem ser alteradas.

O MANUAL DO TESTE

O manual é a principal fonte de informação sobre a qualidade de um teste publicado. O manual provê direções para a testagem, procedimentos de atribuição de pontos e resultados de pesquisa.

Os manuais não são sempre tão úteis como deveriam ser, pois podem ter seções inadequadas. Proeminência e popularidade não são, necessariamente, sinais de qualidade nos testes.

Um bom manual é difícil de ser preparado. Ele precisa ser claro, o suficiente, para que qualquer usuário qualificado possa compreendê-lo.

O manual deve conter informações sobre:

- . disseminação de informações sobre o teste
- . interpretação dos escores
- . validade do teste
- . confiabilidade do teste
- . administração do teste
- . atribuição de pontos do teste
- . escalas e normas do teste.

b) Avaliação para decisão absoluta

Os testes CR são mais compatíveis com os requisitos de Avaliação Instrucional, em termos absolutos. Toda educação é um esforço consciente de levar à modificação do comportamento, através da aprendizagem. O papel das técnicas instrucionais é causar as mudanças e o das técnicas de Avaliação é detetá-las a fim de tomar decisões referentes ao:

- . indivíduo
- . instrução.

Com base na elaboração de 2 a 3 itens por objetivo, este pode ser considerado alcançado quando o indivíduo acerta todos os itens (no caso de 2) e pelo menos 2 (no caso de 3 itens). Quando o objetivo exige a elaboração de um item apenas, o indivíduo deve acertá-lo para atingir o objetivo.

Deve-se corrigir o teste CR dicotomicamente, isto é, indicando sucesso (1) ou falha (zero) do indivíduo em cada item e transformando, depois, em sucesso ou falha nos objetivos correspondentes. As Tabelas F.15 e F.16 apresentam um exemplo de correção dos itens e respectivo alcance de objetivo.

TABELA F.16

ACERTO OU ERRO NOS ITENS

ITENS →	OBJETIVO 1			OBJETIVO 2		OBJETIVO 3			OBJETIVO 4		OBJETIVO n
	1	2	3	4	5	6	7	8	9	10	
INDIVÍDUOS											
1	1	1	1	0	1	1	0	1	0	0	
2	1	0	0	1	1	1	1	0	1	1	
3	1	1	1	0	0	1	0	0	1	1	
4	1	0	1	1	0	0	0	0	1	0	
n											

TABELA F.16

ALCANCE DOS OBJETIVOS

INDIVÍDUOS \ OBJETIVOS	OBJETIVOS				
	1	2	3	4	n
1	1	0	1	0	
2	0	1	1	1	
3	1	0	0	1	
4	1	1	0	0	
n					

Trabalhando-se com alcance em objetivos é possível detectar pontos falhos no processo instrucional, ou seja, pontos da instrução que precisam ser reformulados.

Com os resultados de alcance de objetivos pode-se calcular tanto o nível de aprendizagem do aluno e a efetividade da instrução. Estes tópicos são detalhados na unidade 5.0, pág. F.355.

ADMINISTRAÇÃO DE TESTES

O processo de administração de testes é independente das abordagens de medida.

Alguns testes são suficientemente simples para qualquer adulto administrar; outros precisam de meses de treinamento especial. Exige-se habilidade quando é necessário questionar, oralmente, o sujeito.

Antes da administração de um teste, o administrador deve familiarizar-se com o mesmo.

O administrador deve manter uma atitude imparcial e científica. Não se deve dar indicações aos sujeitos através de gestos faciais ou palavras de encorajamento. O que o indivíduo testado fizer, deve ser feito pelo seu próprio esforço.

a) Condições de testagem

Certos problemas gerais de administração (Cronbach, 1970) são comuns a todos os testes. As sugestões que seguem devem ser mantidas durante a situação de testagem:

- . assegurar que o indivíduo esteja fisicamente confortável para tomar o teste, em sala de tamanho adequado, com boa iluminação, espaço suficiente para a colocação dos materiais;

- . cuidar para que o indivíduo esteja livre de distrações. Não se pode esperar um bom desempenho se se está distraído;
- . colocar as pessoas testadas em uma posição de modo que possam ou vir claramente as direções do teste;
- . assegurar que as direções orais, quando houver, sejam todas da das antes do início do teste;
- . não dar mais atenção a alguns indivíduos do que outros. Se o teste for bem construído, sem questões ambíguas e contiver direções completas, claras e concisas não haverá necessidade de esclarecer questões individualmente;
- . terminado o tempo regular de testagem, deve-se recolher todos os testes. Recolher os testes de alguns e permitir que outros o continuem é prejudicial.

O estado do indivíduo testado afeta os resultados do teste. Se na ocasião do teste a pessoa está cansada, abalada, distraída, ou com outros problemas, os resultados não serão amostra de seu comportamento.

b) Controle de grupo

O administrador (Cronbach, 1970) deve ficar em uma posição, em relação ao grupo, de modo que enxergue todos os sujeitos. Ele sõ deve iniciar a testagem quando toda a atenção estiver voltada para si, para que não seja necessário repetir nada.

Para a eficiência de testagem, os sujeitos devem seguir, prontamente, as instruções e todos devem fazer a mesma coisa, ao mesmo tempo. Esta atitude deve ser mantida sem interferir com a oportunidade de os indivíduos fazerem perguntas.

As direções devem ser dadas simples e claramente. Sõ se deve passar para a próxima direção depois de se ter certeza que todos a cumpriram.

Em teste padronizado, se acontecer de um indivíduo sentir-se mal durante a testagem e precisar retirar-se da sala, o teste deste indivíduo fica invalidado, devendo o sujeito ser testado em outra ocasião com uma forma diferente do teste, uma vez que a meta do administrador do teste é obter informação útil.

c) Direções aos sujeitos

A responsabilidade mais importante na administração do teste é dar direções. Estas fazem parte da situação de teste.

É importante que o administrador dê as direções exatamente como fornecidas pelo manual do teste. O propósito de um teste padronizado é obter uma medida que possa ser comparada com medidas feitas em outras épocas, por isso o administrador não deve acrescentar idéias às direções-padrão para não prejudicar os sujeitos testados em outras ocasiões.

As direções do teste devem ser claras, livres de ambiguidade de modo que todos os testados entendam exatamente a mesma coisa.

d) Motivação para tomar o teste

Não há problemas de motivação quando se trata de medida física. Por exemplo, não importa como uma pessoa está se sentindo quando se quer saber a sua altura. Mas, no caso de teste psicológico, o sujeito é colocado em uma escala e então é importante que o sujeito esteja motivado.

e) Incentivos que aumentam os escores

Uma grande fonte de motivação é o desejo de manter auto-respeito em relação aos outros. O esforço também é estimulado pelo interesse na tarefa, hábito de obediência a autoridade, e amizade com o administrador do teste.

Os escores do teste tendem a aumentar quando o administrador intensifica, o indivíduo testado, a fazer o melhor possível.

f) Preparação do sujeito para o teste

A motivação mais proveitosa para validar uma testagem é o desejo do indivíduo testado que o escore seja válido.

A cooperação entre o administrador e o sujeito é perfeitamente possível. Está aumentando, consideravelmente, a necessidade de preparar o sujeito, através de informações prévias, apropriadas.

Considerando o papel técnico do avaliador, ou seja, a obtenção de informações, esta unidade (4.3) enfocou o aspecto referente ao desenvolvimento de instrumentais de medida. A próxima unidade (4.4) tratará da Avaliação Formativa de materiais instrucionais impressos e alunos utilizando os testes de aprendizagem.

4.0 - UNIDADE 4 : OBTENÇÃO DE INFORMAÇÕES: AVALIAÇÃO FORMATIVA DE MATERIAIS INSTRUCIONAIS IMPRESSOS (M.I.I.) E DE ALUNOS

Antes de iniciar suas atividades, o avaliador precisa saber, exatamente, "o quê" será avaliado, que mudanças são pretendidas e quais são as suas características. O "o quê" poderá ser um sistema ou apenas um componente instrucional. Por exemplo, o "o quê" poderá referir-se a aprendizagem de um grupo de alunos (componente instrucional) em matemática, durante o período de um ano. Estes alunos devem ter entre 14 e 16 anos de idade, possuírem conhecimentos básicos de equações de primeiro grau e maturidade intelectual correspondente a sua idade cronológica. O "o quê" poderá significar, também, um texto programado, uma aula de TV, uma estratégia instrucional, um conjunto de componentes em interação. O conhecimento claro e preciso do "o quê" será avaliado é condição essencial para a atividade de avaliação.

Pode-se avaliar com propósito formativo ou somativo qualquer componente educacional, quer seja ele um currículo, um método de en

sino, um material instrucional. Se apenas um destes componentes é avaliado, isoladamente, temos a Avaliação de componentes instrucionais. Porém, se o que será avaliado é um conjunto de componentes (materiais impressos, aulas de TV, método de ensino, alunos) em interação, temos a Avaliação de sistemas instrucionais.

Esta unidade trata da Avaliação de componentes instrucionais com propósito formativo, ou seja, descreve a Avaliação Formativa de M.I.I. (materiais instrucionais impressos) e de alunos, pois, sendo ela de grande utilidade e de baixo custo, poderá ser realizada pela maioria das Instituições Educacionais brasileiras. Será feita, também, alguma referência à Avaliação Formativa de materiais instrucionais audio-visuais, pois, embora não seja este o enfoque deste trabalho, algumas considerações a respeito serão apresentadas.

Antes de se falar de Avaliação Formativa de M.I.I. e de alunos, faz-se necessário descrever alguns conceitos e algumas fontes de informações que servem aos propósitos deste trabalho. Em seguida, serão apresentados alguns procedimentos para avaliar, formativamente, os componentes educacionais (M.I.I. e alunos), objetos deste estudo.

4.1 - SUB UNIDADE 1: PAPEL E FONTES DE INFORMAÇÕES DA AVALIAÇÃO FORMATIVA

4.1.1 - O PAPEL FORMATIVO DA AVALIAÇÃO

Questiona-se sobre currículos escolares, materiais instrucionais, estratégias de ensino, aprendizagem dos alunos, programas educacionais em geral. Pergunta-se o que deve ser feito para torná-los mais adequados às metas da Educação, reconhecendo-se a necessidade de melhorá-los, sistematicamente. Para cumprir tal necessidade, surgiu, recentemente, uma importante atividade de Avaliação que ocorre durante o desenvolvimento de componentes instrucionais. Este tipo de Avaliação relaciona-se com a melhoria do Sistema Educacional ou de seus componentes e é chamada Avaliação Formativa. Este termo foi primeiramente utilizado por

Scriven (1976) que atribuiu à Avaliação dois papéis distintos: formativo e somativo.

No seu papel formativo, a Avaliação procura auxiliar o elaborador de componentes educacionais (currículos, métodos, materiais, programas instrucionais), fornecendo informações que possibilitam revisar e aprimorar estes componentes, enquanto estão sendo desenvolvidos.

No seu papel somativo, a Avaliação procura auxiliar o consumidor de componentes educacionais, fornecendo informações que possibilitam identificar se um determinado produto é bom ou ruim, se é melhor do que outros.

As atividades de Avaliação Formativa ocorrem durante o desenvolvimento de um componente educacional, quando modificações, no mesmo, podem ser feitas. As atividades de Avaliação Somativa ocorrem no final do processo de desenvolvimento, quando o componente está pronto para ser utilizado pelo consumidor.

A Avaliação Formativa assim, no processo de desenvolvimento de componentes instrucionais precede a Avaliação Somativa.

Para melhorar um sistema ou seus componentes são necessárias informações a eles relativas. Algumas fontes de informações úteis para este propósito, são descritas a seguir.

4.1.2 FONTES DE INFORMAÇÕES PARA A AVALIAÇÃO FORMATIVA

Cunningham (1972) identifica três fontes de informações nas quais o avaliador poderá encontrar várias espécies de evidências:

- a) informações internas
- b) informações externas
- c) informações contextuais

a) Informações Internas

Segundo Cunningham, informações internas, para servirem ao propósito formativo, são obtidas através da análise do próprio componente. Tais informações referem-se:

- às características físicas do componente;
- ao conteúdo;
- aos objetivos;
- aos procedimentos.

Inclui também a análise criteriosa do componente por professores, estudantes e especialistas no assunto.

Por exemplo, analisando um material instrucional qualquer (o componente em questão) poder-se-á obter informações sobre a estrutura e sequência do seu conteúdo. Estas informações podem revelar que algumas de suas partes não estão estruturadas e sequenciadas da melhor maneira. A indicação destas falhas levará à revisão e melhoria das partes do conteúdo que estão apresentando problemas. Ou ainda, analisando procedimentos e objetivos propostos para a instrução, o avaliador poderá coletar evidências mostrando que:

- . os objetivos propostos não estão clara e precisamente definidos;
- . os procedimentos instrucionais são inadequados para o alcance dos objetivos selecionados.

b) Informações Externas

Informações externas, segundo o mesmo autor, são obtidas através da análise dos efeitos de partes do componente sobre o comportamento de estudantes, professores, pais e outras pessoas envolvidas no processo. Estes efeitos podem incluir não só o desempenho do estudante como também a atitude dos pais, professores em relação a um certo componente instrucional que não foi, ainda, totalmente desenvolvido. Por exemplo, algumas informações obtidas dos pais dos estudantes podem evi

denciar uma reação desfavorável dos mesmos, quanto ao uso de um texto programado. Esta informação poderá levar à revisão e melhoria deste material, na tentativa de eliminar ou diminuir a reação negativa dos pais. Ou ainda, informações obtidas dos professores podem mostrar que eles atribuem pouco valor a alguns objetivos, propostos para o texto programado. Neste caso, tais objetivos deverão ser revistos e melhorados, na tentativa de conseguir uma mudança positiva na atitude dos professores.

Para reformular e melhorar os componentes em desenvolvimento, o avaliador não utiliza apenas um tipo de informação, mas procura relacionar informações internas, externas e contextuais para propor maneiras de solucionar um determinado problema.

c) Informações Contextuais

Estas informações são relativas às condições nas quais o componente deveria funcionar. Frequentemente, elas são supostas pelos e laboradores, mas raramente são testadas. Algumas informações contextuais referem-se mais especificamente às características dos alunos, professores e currículo. Por exemplo, analisando as condições nas quais um material instrucional, em desenvolvimento, está sendo utilizado, podem ser obtidas informações, mostrando que os alunos que estão passando por ele, não têm os pré-requisitos necessários para utilizá-lo. Neste caso, uma instrução remedial poderá ser sugerida aos estudantes, como uma solução para o problema.

As informações internas, externas e contextuais são muito úteis para o avaliador formativo, desde que sua tarefa é identificar falhas nos componentes instrucionais e o porquê destas falhas. Se ele tiver informações sobre existência de um problema qualquer no componente, ele irá perguntar: porque existe este problema? Por exemplo, um material instrucional impresso falhou no alcance dos objetivos 3 e 5 e conduziu ao domínio dos objetivos 1, 2 e 4. O avaliador, informado da existência deste problema, vai procurar identificar sua causa, coletando evidências nas várias fontes de informações (internas, externas e contextuais). Assim, ele poderá descobrir que:

- a sequência de aprendizagem deveria ser: objetivo 1, 2, 4, 5 e 3 e não objetivo 1, 2, 3, 4 e 5 (informação interna);
- os professores não possuem a qualificação necessária para utilizar os materiais instrucionais, em sala de aula, na orientação da aprendizagem dos alunos (informação contextual);
- os estudantes têm uma atitude negativa em relação aos objetivos não dominados (informação externa);
- os objetivos não dominados estão mal definidos (informação interna).

Relacionando todas as informações obtidas, o avaliador chegará à identificação da causa do problema e proporá maneiras de solucioná-lo. Portanto, ao avaliador formativo não basta apenas uma espécie de informação. Ele procura diferentes informações que sejam relevantes e que possibilitem ajustamento apropriados no contexto, no próprio componente ou nos resultados pretendidos. Ele procura evidências dentro do "porquê dos resultados", dentro das relações entre informações contextuais, internas e externas.

Para coletar informações internas, externas e contextuais, vários procedimentos, formais ou informais, podem ser usados. A utilização de um ou outro depende do que se deseja. Se um certo grau de generalização de resultados é pretendido, devem ser utilizados procedimentos formais para coletar informações. Por exemplo, a análise hierárquica pode ser utilizada para coletar evidências sobre estrutura e sequência de conteúdo; técnicas de observação sistemática, "designs" experimentais e quase-experimentais podem servir para identificar efeitos de certos materiais sobre o comportamento humano. Também podem ser usadas outras técnicas e instrumentais, como:

- questionários para serem respondidos por especialistas no assunto;
- observações em sala de aula, estruturadas ou não;
- exames, exercícios escritos, cadernos dos alunos;
- testes de rendimentos;

- escala de atitudes e opiniões;
- entrevistas com professores e especialistas.

E ainda, recomendações para modificações em um componente qualquer, podem resultar de discussões entre os membros da equipe de elaboração.

Concluindo: o avaliador formativo, utilizando vários tipos de procedimentos, procura assegurar-se de que estão sendo mantidas as relações previstas entre informações contextuais, internas e externas. Estuda o componente enquanto está sendo desenvolvido, propondo revisões, modificações e melhoria no mesmo.

4.2 - SUB-UNIDADE 2: AVALIAÇÃO FORMATIVA DE MATERIAIS INSTRUACIONAIS IMPRESSOS

4.2.1 - PROPÓSITO DA AVALIAÇÃO FORMATIVA DE MATERIAIS INSTRUACIONAIS IMPRESSOS (M.I.I.)

O avaliador formativo de M.I.I. se propõe a coletar evidências para serem utilizadas na tomada de decisão a respeito de como melhorar estes materiais, enquanto estão sendo desenvolvidos. Por exemplo, um texto programado que se propõe a ensinar análise sintática, é utilizado por um grupo de alunos. Algumas evidências, coletadas através de testes de aprendizagem aplicados com propósito formativo, revelam deficiências no domínio deste objetivo. Uma análise destas evidências mostra que os erros dos estudantes relacionam-se com a identificação do objeto direto. Esta informação é então utilizada para tomar decisões acerca do material, elaborado para ensinar objeto direto, de como revisá-lo e melhorá-lo.

4.2.2 - ALGUMAS EVIDÊNCIAS PARA TOMADA DE DECISÃO RELATIVA À MELHORIA DE M.I.I.

Segundo Briggs (1974) as avaliações formativas são frequentemente caracterizadas por procedimentos informais, porém a informalidade destes procedimentos não deve afetar a precisão dos dados coletados. Evidências convincentes são desejadas. Por exemplo, se alguém pro

cura evidências sobre o domínio dos objetivos de uma certa instrução, nenhuma evidência é tão convincente como o desempenho dos alunos em um teste planejado adequadamente para medir o alcance destes objetivos.

Muitas evidências são necessárias para tomada de decisão a respeito de como revisar e melhorar M.I.I. em desenvolvimento. Briggs (1974) propõe que a primeira delas é relativa a acuracidade do conteúdo da instrução. Esta pode ser obtida através de uma revisão do material feita por especialista no conteúdo em estudo. Outras evidências são relativas à praticabilidade e eficácia dos materiais. A questão da praticabilidade pode ser decidida, considerando-se o registro das dificuldades experienciadas pelos professores e estudantes, na condução da instrução. A questão da eficácia é um julgamento mais complexo; pode depender em parte de informações revelando:

- as atitudes dos alunos, acidentalmente escolhidos, para receber uma determinada instrução;
- o desempenho do professor durante o processo instrucional;
- a extensão na qual os materiais foram utilizados como pretendido;
- a extensão na qual os estudantes apresentaram o desempenho esperado.

Estas evidências podem ser coletadas através da utilização de:

- . registro de observações;
- . questionários;
- . testes de aprendizagem;
- . observações das atitudes e interesses do alunos;
- . críticas de especialistas no assunto.

e devem ser obtidas para todos os objetivos. Devem também ser procuradas pelos estudantes, professores e ainda por um observador que pode ser um dos membros da equipe de planejamento. Briggs apresenta uma lista de questões pretendidas para identificar as várias espécies de evidências

procuradas (diz, porém, que elas não representam o conteúdo dos instrumentos). Cada uma delas pode exigir várias outras questões específicas.

A lista proposta por Briggs (1974) é apresentada a seguir.

a) A partir do Observador

1. Até que ponto os materiais e meios estão sendo utilizados da maneira pretendida pelo planejador?
2. Até que ponto o professor tem utilizado os procedimentos e tomado as decisões pretendidas?
3. Até que ponto os estudantes têm seguido os procedimentos gerais especificados?

b) A partir do Professor

1. Quais as dificuldades práticas na condução da lição? (Exemplo: tempo, equipamentos, etc.).
2. Qual é a estimativa do grau de interesse ou de absorção dos estudantes na lição?
3. Que dificuldades metodológicas foram encontradas pelo professor na condução dos procedimentos pretendidos?

c) A partir do Estudante

1. Com que prazer você escolheria fazer as coisas que aprendeu nesta lição?
2. Com que prazer você recomendaria esta lição para seus amigos?
3. Quais os resultados do teste de aprendizagem relativo aos objetivos desta lição?

Estas várias espécies de evidências coletadas através de vários meios são então utilizadas para tirar conclusões sobre a necessidade de conservar uma instrução como está, modificá-la ou rejeitá-la.

4.2.3 - COMO OCORRE A AVALIAÇÃO FORMATIVA DE M.I.I.

a) Algumas Variáveis do Ambiente de Sala de Aula

Os materiais instrucionais não operam no vácuo, mas funcionam como parte de um sistema instrucional constituído de muitos componentes interdependentes. Todos estes componentes contribuem em maior ou menor grau para a aprendizagem dos alunos. Assim sendo, o avaliador deverá identificá-los como variáveis (ou fatores) que influenciam o desempenho dos estudantes durante o processo de instrução. Se ele deseja uma situação na qual possa, com alguma segurança, atribuir o desempenho dos alunos aos materiais instrucionais utilizados, deve, de alguma forma, procurar eliminar a influência destas variáveis.

Se a Avaliação Formativa de M.I.I. deve ocorrer em sala de aula com o processo de ensino/aprendizagem em andamento, muitas variáveis que interferem no desempenho dos alunos, podem ser identificadas. Light e Reynolds (1972) identificaram algumas categorias destas variáveis. Entre elas, temos:

- 1) Testagem;
- 2) Papel instrucional do professor;
- 3) Motivação dos estudantes;
- 4) Habilidades de estudo.

1) Testagem

Os instrumentais de medida a serem utilizados devem ser válidos e confiáveis (conforme especificado na unidade 3.0, páginas F.179 e F.190).

Os procedimentos para a aplicação dos instrumentais, quer seja em relação ao administrador ou aos alunos, devem ser clara e preci

samente definidos. Testes inválidos ou procedimentos inadequados, do administrador ou dos alunos, durante a testagem podem interferir no desempenho dos estudantes e tornar pouco significativos os esforços da Avaliação Formativa.

2) Papel Instrucional do Professor

O papel do professor no processo instrucional, deve ser claramente identificado. Seus procedimentos não devem suplantar a atuação dos materiais instrucionais. Se o professor apresenta comportamentos não previstos, estes comportamentos poderão interferir nos resultados da aprendizagem dos estudantes.

3) Motivação dos Estudantes

A falta de motivação dos estudantes para utilizar os materiais instrucionais pode causar problemas de aprendizagem. Se alguns alunos estão desmotivados, as deficiências que apresentarem no domínio de certos objetivos poderiam ser atribuídas à falta de motivação e não aos materiais instrucionais.

4) Habilidades de Estudo

Se são dados, aos estudantes, materiais muito abaixo ou acima do seu nível de habilidades, a Avaliação Formativa não teria sentido desde que:

- um estudante adiantado pode ser capaz de demonstrar domínio em uma habilidade, mesmo que os materiais sejam defeituosos.
- estudantes encontrando dificuldades com materiais avançados podem estar tendo dificuldades muito mais porque os materiais estão tão acima do seu nível de habilidades, do que por causa de defeitos nos mesmos. O sucesso dos materiais instrucionais está relacionado com a habilidade do estudante para aprender a partir daqueles materiais. (Estas habilidades de estudo constituem pré-requisitos para a instrução).

b) Um Modelo de Avaliação Formativa de M.I.I.

Na literatura pertinente podem ser encontrados alguns procedimentos de avaliação formativa de materiais instrucionais.

Abedor (1972) e Light Reynolds (1972) apresentam alguns procedimentos que, tendo sido testados experimentalmente, mostraram-se eficazes para o propósito de melhoria de materiais instrucionais. Um estudo destes procedimentos conduziu à seleção de alguns deles, que posteriormente vieram a constituir o modelo que é proposto a seguir, para a Avaliação Formativa de M.I.I. A utilidade do modelo é evidente, considerando que:

- Constantes mudanças estão ocorrendo no Sistema Educacional e que estas mudanças precisam ser testadas;
- Em situações de mudanças, a ênfase principal da Avaliação é de caráter formativo.

O modelo poderá ser utilizado para melhorar M.I.I. que estão sendo desenvolvidos como parte de um novo método de ensino, de um novo currículo escolar em desenvolvimento, ou de uma grande mudança que está sendo introduzida no sistema.

Toda mudança introduzida no sistema ou em qualquer um dos seus componentes, tem o propósito de melhorar uma situação existente. Assim sendo, as modificações pretendidas devem ser previamente testadas, com propósito formativo e somativo, antes de serem implantadas em todo o sistema. Esta testagem se caracteriza como uma fase de ensaio na qual a mudança desejada é obtida e executada em pequena escala e desde que sejam alcançados os resultados pretendidos, as modificações são introduzidas em todo o sistema. Por exemplo, se o sistema escolar pretende introduzir materiais impressos de instrução individualizada, no ensino de 2º grau, eles devem ser inicialmente testados em uma amostra destas escolas. Tais materiais são obtidos ou elaborados e aplicados em um grupo de alunos representativos da população alvo. Após este ensaio, quando os materiais foram avaliados e comprovada sua validade, serão introduzidos em todas as escolas do 2º grau.

Portanto, qualquer mudança no sistema total ou em apenas um dos seus componentes, deve ser primeiramente testado, ou seja, deve ser obtida e executada em pequena escala.

O modelo que será descrito a seguir, adequa-se às situações de mudança, e é utilizado quando se pretende avaliar as modificações pretendidas, com propósito formativo. O modelo especifica passos para a obtenção da mudança (por exemplo, a introdução de materiais impressos de instrução individualizada) e, para a sua execução na fase de ensaio da mesma. Implica assim em duas etapas:

1a. Etapa: Durante a obtenção da mudança planejada, na sua fase de ensaio.

2a. Etapa: Durante a execução da mudança planejada e em pequena escala.

A 1a. etapa ocorre enquanto estão sendo executadas as atividades e obtidos os recursos necessários à execução do ensaio da mudança planejada, ou seja, enquanto todos os recursos estão sendo conseguidos ou desenvolvidos em sua forma inicial (por exemplo, a elaboração de materiais instrucionais em sua primeira forma).

A 2a. etapa ocorre enquanto os recursos obtidos, para o ensaio da mudança estão sendo processados (por exemplo, os materiais instrucionais sendo utilizados por uma amostra da população alvo).

Na 1a. etapa, enquanto os materiais estão sendo desenvolvidos ou obtidos, estes serão avaliados através dos seguintes passos:

- 1 - Controle de qualidade por um especialista no assunto;
- 2 - Reformulação dos M.I.I.;
- 3 - Testagem um a um dos M.I.I. - Pré-Teste - M.I.I. - Pós-Teste;
- 4 - Análise dos dados coletados;
- 5 - Reformulação dos M.I.I.

Na 2a. etapa, os M.I.I. (em sua primeira forma) enquanto estão sendo processados na fase de ensaio da mudança, serão avaliados a través dos seguintes passos:

- 6 - Testagem dos M.I.I. revisados, com um grupo maior: Pré-teste - M.I.I. - Pós-teste;
- 7 - Análise dos dados coletados;
- 8 - Reformulação dos M.I.I. (se houver maiores problemas nos mes mos);
- 9 - Avaliação Somativa dos M.I.I. (se não forem constatados maio res problemas no mesmo).

Os passos referentes a estas duas etapas serão descritos a seguir.

Assim, na 1a. etapa, temos:

- 1 - Controle de Qualidade dos M.I.I. por um Especialista no Assunto.

Os M.I.I., em sua primeira forma, serão submetidos a um especialista no assunto para a identificação de possíveis erros conceituais e para um julgamento a respeito da qualidade do conteúdo apresentado (clareza, precisão, estrutura física, etc...). Se for desejado, um membro da equipe de Planejamento poderá também rever o material para opinar a respeito de sua validade.

- 2 - Reformulação dos M.I.I.

Os M.I.I. serão reformulados em função das informações obtidas através do controle de qualidade. Estas informações serão fornecidas ao elaborador dos M.I.I., que as utilizará, para modificar partes do material que estão apresentando problemas. Estas reformulações objetivam sempre a melhoria dos M.I.I.

3 - Testagem Um a Um: Prê-Teste - M.I.I. - Pós-Teste

Para cada unidade de ensino, será utilizado o seguinte "design".

Prê-Teste - M.I.I. - Pós-Teste

O pré-teste proverá informações sobre conhecimentos e habilidades já dominados pelos alunos. O pós-teste proverá informações sobre conhecimentos e habilidades dominados pelos alunos, após terem passado pelos materiais instrucionais.

Um grupo de aproximadamente seis alunos serão selecionados a partir da população para quem os M.I.I. são pretendidos (para esta seleção são úteis as técnicas de amostragem descritas na unidade 5.0, página F.270, de uma vez, será submetido ao pré-teste e após usar o material instrucional tomará o pós-teste. Assim, a aplicação dos testes de aprendizagem e a utilização dos materiais instrucionais, por este pequeno grupo, deverá ocorrer individualmente com cada aluno selecionado. Aqui, algumas informações sobre problemas nos testes ou nos materiais instrucionais poderão ser obtidas diretamente de cada aluno pelo contato face a face entre ele e o avaliador. Outras informações sobre possíveis deficiências nos M.I.I. serão obtidas através da análise dos dados coletados com a aplicação dos testes de aprendizagem.

4 - Análise dos Dados Coletados

Os dados obtidos serão, aqui, analisados para a obtenção de informações que possibilitem a melhoria dos M.I.I. Uma forma de obter estas informações é através da determinação do ganho por objetivo, obtido pelo grupo, após ter passado pelos materiais instrucionais e completado os testes de aprendizagem. A diferença entre a percentagem de pessoas dominando um mesmo objetivo no pré e pós-teste pode ser considerada como um indicador da qualidade do material. O ganho pode ser computado pelo fórmula:

$$\text{ganho por objetivo} = \frac{\text{ganho real por objetivo}}{\text{ganho possível por objetivo}}$$

onde

$$\text{ganho real} = \left(\begin{array}{l} \text{número de indivíduos que} \\ \text{atingiram o objetivo no} \\ \text{pós-teste} \end{array} \right) - \left(\begin{array}{l} \text{número de indivíduos que} \\ \text{atingiram o objetivo no} \\ \text{pré-teste} \end{array} \right)$$

$$\text{ganho possível} = \left(\begin{array}{l} \text{número máximo de alunos} \\ \text{que poderiam ter atingi-} \\ \text{do o objetivo no pós-} \\ \text{teste.} \end{array} \right) - \left(\begin{array}{l} \text{número de indivi-} \\ \text{duos que atingi-} \\ \text{rã o objetivo no} \\ \text{pré-teste.} \end{array} \right)$$

Considerando o seguinte exemplo:

TABELA F.17

RESULTADOS DA TESTAGEM UM A UM: PRÉ E PÓS-TESTE

OBJETIVOS ALUNOS ↓	1		2		3		4		5		6		7	
	PRÉ	PÓS												
1		X		X	X	X		X		X		X	X	X
2	X			X		X		X						
3		X			X		X	X				X		
4						X		X						X
5		X				X								
6						X								X
Nº DE ALUNOS POR OBJETIVO	1	3	0	2	2	5	1	4	0	1	0	2	1	3

X = Objetivos Dominados

E, aplicando a fórmula de ganho para os objetivos propostos, temos os seguintes resultados:

TABELA F.18

GANHO POR OBJETIVO, OBTIDO PELO GRUPO

OBJETIVOS	ÍNDICES DE GANHO
1	40%
2	33%
3	75%
4	60%
5	17%
6	33%
7	40%

A análise dos índices de ganho por objetivo, obtidos pelo pequeno grupo, fornecerão indicações de quais materiais precisam ser melhorados. Em situações de ensino individualizado, o ganho ideal seria 100%. Entretanto, não existe um índice padrão de ganho que possa ser utilizado nas várias situações de aprendizagem. Assim, em cada situação, poderão ser utilizados índices julgados mais adequados para aquela situação particular. No exemplo acima, se o índice de ganho julgado mais adequado foi, por exemplo, 70%, os materiais relativos aos objetivos que não atingiram este índice, devem estar apresentando problemas. Estes materiais devem ser revisados, e, desde que sejam identificadas deficiências em algumas de suas partes, estas devem ser reformuladas, tendo em vista sua melhoria (por exemplo, melhoria da estrutura e sequência, do formato de apresentação, das informações).

Uma outra abordagem para análise dos dados é sugerida por Bloom (1972). Esta análise é feita considerando-se a percentagem de atingimento dos objetivos no pós-teste. No exemplo citado anteriormente, temos:

TABELA F.19

PERCENTAGEM DE ALUNOS POR OBJETIVO

OBJETIVOS	PERCENTAGEM DE ALUNOS QUE ATINGIRAM O OBJETIVO
1	50%
2	33%
3	83%
4	66%
5	17%
6	33%
7	50%

Bloom sugere que os erros cometidos por uma grande parte dos alunos (60% ou mais), podem ser atribuídos aos materiais instrucionais. Assim sendo, analisando as percentagens de alcance de objetivos, indicadas acima, podemos supor que os materiais instrucionais que pretendem conduzir ao domínio dos objetivos 2, 5 e 6 estão provavelmente apresentando problemas. Temos que 67% dos alunos não dominaram o objetivo 2, 83% o objetivo 5 e 67% o objetivo 6. O avaliador deve identificar estes problemas, suas causas e indicar partes dos materiais que devem ser melhoradas. Porém, antes de se tomar uma decisão final, deve-se considerar não só os resultados da análise sugerida por Bloom, que considera os dados do pós-teste, como também o índice de ganho por objetivo que considera dados do pré e pós-teste.

Para se identificar possíveis causas de deficiências no alcance dos objetivos propostos, o avaliador poderá analisar os erros cometidos pelos alunos e em seguida localizar, nos materiais, o conteúdo correspondente aos itens errados. As respostas dadas as questões que se seguem poderão ajudá-lo nesta tarefa. (Estas questões foram propostas por Light e Abedor, 1972).

a) O que existe de semelhante nos itens errados no teste?

Por exemplo, após a aplicação de um teste medindo a aprendizagem dos alunos em subtração, serão identificados os itens errados. Uma análise destes itens pode mostrar que todos eles referem-se a subtração de números inteiros iguais ou maiores do que uma dezena.

b) Como os itens errados diferem dos itens acertados?

Por exemplo, todos os itens acertados pelos alunos, referem-se a subtrações com números menores que uma dezena.

c) Onde está, nos materiais, o conteúdo referente aos itens errados?

Dos exercícios relacionados para a aprendizagem de subtração de números inteiros, apenas dois deles referem-se a subtração de números iguais ou maiores do que uma dezena.

d) O que, nos materiais instrucionais pode ser a causa dos erros cometidos no teste?

Pode-se supor que se os materiais tivessem relacionado um maior número de exercícios com números iguais ou maiores do que uma dezena, os alunos teriam dominado este objetivo.

e) Como pode a causa hipotetizada dos erros, ser experimentalmente testada?

Outros exercícios de subtração, semelhante aqueles errados nos testes, poderão ser acrescentados nos materiais.

Todas as informações obtidas através da comunicação direta entre avaliador e alunos e dos dados coletados a partir dos testes de aprendizagem, serão fornecidas para o elaborador que as utilizará para reformular seus materiais nas partes indicadas como apresentando deficiências.

5 - Reformulação dos M.I.I.

Aqui, o elaborador dos materiais fará as reformulações indicadas para melhorar os materiais instrucionais de tal forma que eles possam conduzir os alunos ao domínio dos objetivos que não foram ainda alcançados.

Os materiais instrucionais reformulados serão, então, avaliados em uma 2a. etapa. Como visto anteriormente, esta etapa ocorre durante o processamento da fase de ensaio da mudança, ou seja, durante a execução da mudança planejada e em pequena escala. Assim, tem-se:

6 - Testagem dos M.I.I. revisados, com um grupo maior:Pré-teste - Pós-Teste

Aqui é selecionado um grupo maior de indivíduos que tenha as mesmas características da população para a qual os materiais são destinados. (Técnicas de amostragem descritas na Unidade 5.0, pág. F.270, serão ser utilizadas para esta seleção). Os mesmos procedimentos descritos no passo 3 serão aqui utilizados, exceto que a testagem, neste momento, é realizada com o grupo todo, simultaneamente, e não individualmente, com cada um dos estudantes.

7 - Análise dos dados coletados

Os mesmos procedimentos utilizados no passo 4.

8 - Reformulação dos M.I.I.

Se os materiais instrucionais ainda apresentarem problemas, serão reformulados de acordo com as informações obtidas pela avaliação, através da análise dos dados coletados a partir dos testes de aprendizagem.

9 - Avaliação Somativa dos M.I.I.

Se não forem constatados outros problemas nos materiais instrucionais, eles serão então avaliados do ponto de vista somativo, ou seja, serão obtidas informações que permitam um julgamento de valor dos mesmos quanto a sua eficiência e/ou efetividade. (Procedimentos para Avaliação Somativa encontram-se especificados na Unidade 5.0.

A Figura F.18 apresenta um esquema do modelo de Avaliação Formativa de M.I.I. que foi descrito.

Como foi visto, o modelo descrito, incluindo duas etapas, obtenção e execução da mudança planejada, é utilizado na fase de ensaio, quando os M.I.I. estão sendo avaliados com propósito formativo. Enquanto informações avaliativas revelarem falhas nestes materiais e reformulações se fizerem necessárias, os mesmos procedimentos de Avaliação Formativa, indicados no modelo, devem ser retomados. Desde que os M.I.I. sejam considerados OK, é feita, então, a Avaliação Somativa dos mesmos. Se as informações, fornecidas por esta Avaliação, indicarem que os resultados obtidos são os desejados, os materiais instrucionais serão introduzidos em todo o sistema. Alguns procedimentos para obtenção de informações que possibilitam chegar a uma decisão final a respeito dos resultados obtidos na fase de ensaio, são apresentados na Unidade 5.0, pág. F.270, que descreve dois esquemas de Avaliação Somativa dos produtos da mudança em termos de aprendizagem.

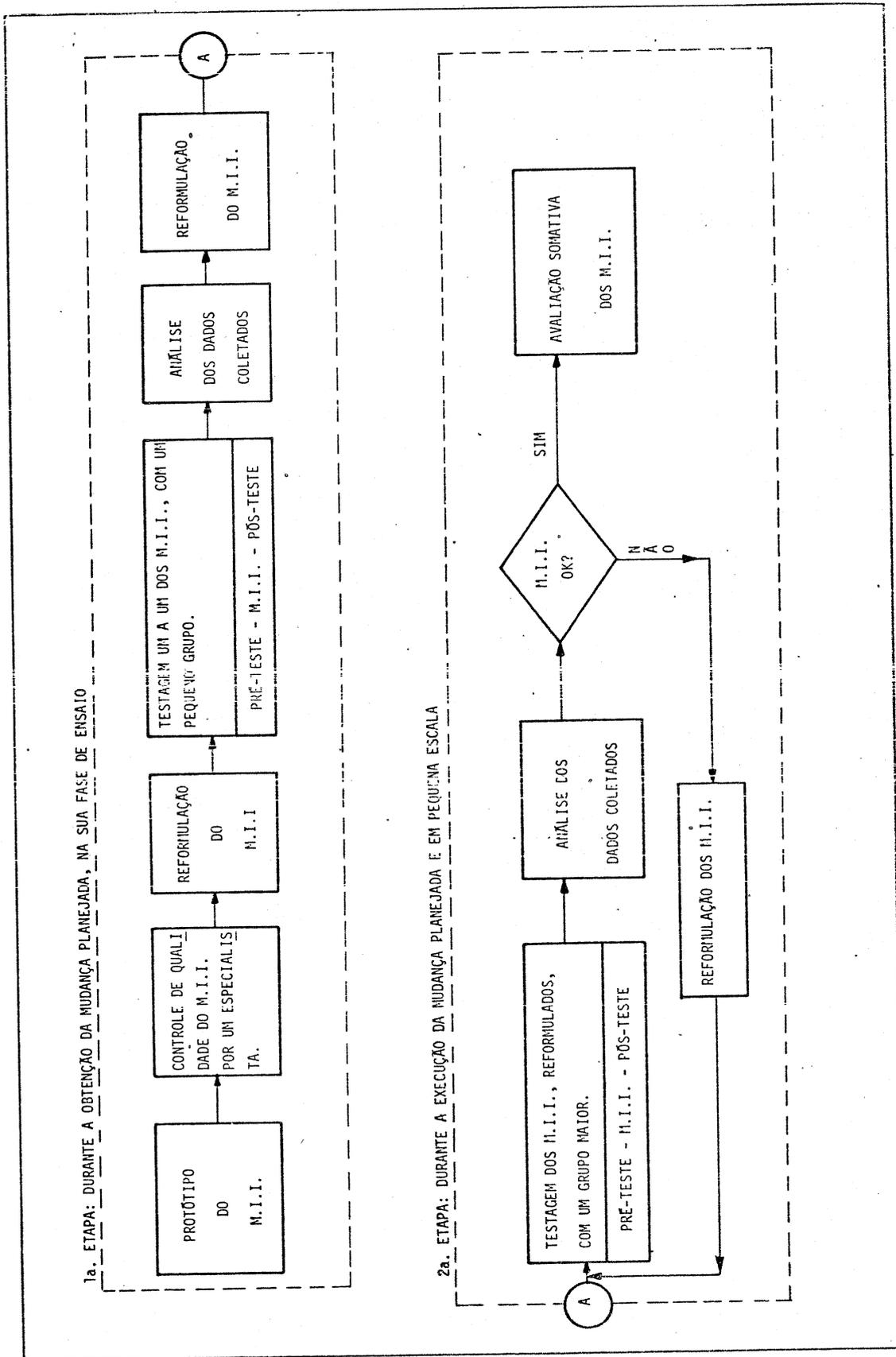


Fig. F.18 - Modelo de Avaliação Formativa de Materiais Instrucionais Impressos

4.2.4 - ALGUMAS CONSIDERAÇÕES SOBRE AVALIAÇÃO FORMATIVA DE MATERIAIS INSTRUCIONAIS AUDIO-VISUAIS

Os procedimentos para a Avaliação Formativa de materiais instrucionais audio-visuais, em comparação com outros, apresentam uma certa complexidade, implicando em maior dispêndio de tempo, de recursos materiais e humanos qualificados para esta tarefa. Por exemplo, é bastante conhecido o elevado custo de programas de TV. Se um programa de TV é utilizado para o alcance de determinado(s) objetivos(s) instrucional (ais), e se dificuldades são encontradas no alcance deste(s) objetivo(s), as aulas do programa deverão ser revisadas. Esta revisão implica em modificações e, conseqüentemente, regravação das mesmas, tendo em vista a melhoria do programa. Isto iria requerer tempo, pessoal especializado em áreas de conhecimento relacionadas, recursos financeiros, etc...

As Instituições Educacionais brasileiras, na sua maioria, não dispõem de tantos recursos e, portanto, dificilmente utilizam tais materiais. Porém, isto não significa que programas de TV não devam ser utilizados, ou que nenhuma atividade de Avaliação Formativa de aulas de TV, não possa ser realizada se existem poucos recursos (materiais financeiros e humanos) e tempo disponíveis. Por exemplo, nesta situação dois procedimentos de Avaliação Formativa poderiam ser utilizados para melhorar estas aulas:

- 1 - Controle de qualidade do "script".
- 2 - Testagem do "script" com dois ou mais indivíduos.

1 - Controle de Qualidade - o "script" é analisado por um especialista em conteúdo e um em planejamento para detetar falhas no mesmo e verificar sua adequação aos objetivos propostos;

2 - Testagem do "script" com dois ou mais indivíduos - o script será lido por dois ou mais indivíduos que tenham as mesmas características da população alvo. Esta testagem objetiva a identificação de dificuldades, ambigüidades ou qualquer confusão que possa apresentar-se para o aluno.

As informações obtidas através do especialista em conteúdo, planejador e alunos, poderão mostrar pontos que, se reformulados, podem melhorar a qualidade da aula, antes que ela seja realizada. Estes únicos procedimentos não conduzem, porém, à eficiência e efetividade de uma aula ou programa de TV. Porém, tem algum significado dentro das várias limitações impostas ao Sistema Educacional Brasileiro.

Na literatura pertinente podem ser encontrados alguns procedimentos específicos para a Avaliação Formativa de materiais instrucionais audio-visuais. Estes procedimentos não serão, aqui, apresentados, visto que fogem ao objetivo desta unidade.

A seguir serão descritos alguns procedimentos de Avaliação formativa de alunos.

4.3 - SUB-UNIDADE 3: AVALIAÇÃO FORMATIVA DE ALUNOS

A Avaliação Formativa de alunos refere-se a avaliação de componentes instrucionais, considerando que o aluno é apenas um dos vários elementos de um sistema de instrução. Os procedimentos descritos a seguir referem-se a Avaliação do desempenho dos mesmos, no processo de ensino/aprendizagem.

4.3.1 - PROPÓSITO DA AVALIAÇÃO FORMATIVA DE ALUNOS

O propósito da Avaliação Formativa de alunos é determinar o grau de domínio de uma tarefa de aprendizagem e apontar a parte da tarefa não dominada. Ajuda o aluno a focalizar-se sobre uma habilidade particular que precisa ser melhor dominada. Ela ocorre durante o processo instrucional e tem uma função retroalimentadora, possibilitando melhoria na aprendizagem dos alunos.

4.3.2 - COMO OCORRE A AVALIAÇÃO FORMATIVA DE ALUNOS

Especificamente, o processo de ensino/aprendizagem tem como meta provocar mudanças duradouras no comportamento dos indivíduos.

Estas ocorrem como consequência do domínio de habilidades e informações particulares. Assim sendo, relativamente a esse processo, devem ser claramente especificados:

- . os objetivos de mudanças comportamentais a serem alcançados pelos alunos;
- . as unidades instrucionais identificadas para o alcance de cada um dos objetivos propostos;
- . os procedimentos e materiais a serem utilizados durante o processo.

Desde que as mudanças pretendidas estejam muito bem definidas, o avaliador formativo poderá ajudar a conduzir o aluno em direção ao domínio das habilidades especificadas e, conseqüentemente, alcançar os objetivos de mudanças comportamentais duradouras.

a) Procedimentos para Avaliação Formativa de Alunos

Para auxiliar o aluno, no alcance dos objetivos de um determinado processo instrucional, o avaliador poderá:

- 1) Identificar objetivos não dominados pelos alunos e o porquê do não domínio dos mesmos;
- 2) Informar o estudante acerca do domínio ou não dos objetivos propostos, e fazer indicações de caminhos alternativos a serem requeridos para o alcance dos objetivos não dominados.

Se existe um planejador instrucional, o segundo passo poderá ser executado pelo avaliador ou pelo próprio planejador. Porém, em nossa realidade educacional, na maioria das vezes o professor assume os dois papéis, de planejador e de avaliador, e assim sendo, os dois passos serão executados por ele.

1) Identificar Objetivos não Dominados pelos Alunos e o Porquê do não Domínio dos Mesmos

Para identificar objetivos não dominados pelos alunos, o avaliador irá procurar evidências que indicam deficiências no alcance dos mesmos. Para coleta destas evidências, ele poderá utilizar vários instrumentais como: questionários, testes objetivos, fichas de observação. Estes instrumentais devem ser elaborados, especificamente para medir o alcance de objetivos propostos para uma instrução particular. Os dados obtidos com sua aplicação irão auxiliar o avaliador na identificação dos erros cometidos pelos estudantes e das causas destes erros; ou seja, os dados analisados fornecerão informações sobre objetivos não dominados e a causa do não domínio dos mesmos. Por exemplo, em um teste os itens 5 e 6 estão medindo o objetivo correspondente a adição de números inteiros. Nos resultados do teste foi verificado que alguns alunos não dominaram este objetivo, assumindo que o critério para o seu alcance era acertar os dois itens (Gagné, 1967). Como apenas alguns poucos alunos não o dominaram a causa foi então atribuída a eles próprios. Algumas evidências obtidas pelo avaliador mostraram que estes estudantes não possuíam os pré-requisitos necessários à instrução.

O aluno é constantemente avaliado com relação ao seu progresso nas unidades de ensino. Para cada unidade são definidos os objetivos que devem ser alcançados e elaborados testes de aprendizagem para medir o domínio destes objetivos. Para cada item dos testes, são utilizados padrões de respostas. Nos testes com propósito formativo, os padrões são os próprios objetivos das unidades. As respostas dadas a cada item são comparadas com os objetivos que estão sendo medidos pelo teste e, através desta comparação é identificado o grau de alcance dos mesmos e confirmado ou não o seu domínio. O domínio destes objetivos implica no domínio das unidades de ensino.

Se a instrução é individualizada, os alunos a ela submetidos e, avaliados com base nesta instrução individual, deverão dominar 100% dos objetivos propostos. Porém, se a instrução não é individualizada, não se espera, igualmente, que todos os alunos dominem

todos os objetivos. Para a avaliação do desempenho do aluno em situações de aprendizagem grupal Bloom (1975) propõe alguns procedimentos úteis. Segundo este autor, tem-se utilizado níveis de alcance de objetivos de 80 a 85 por cento em cada uma das provas com propósito formativo de pequenas unidades de ensino como indicação do domínio das mesmas. Por exemplo, se um estudante alcançou 80% dos objetivos de uma certa unidade, poder-se-á concluir que ela foi dominada por ele. Diz ainda que cada professor talvez queira variar este nível de precisão mas, que se ele for demasiadamente alto (95 a 100%) é provável que seja alcançado apenas por poucos alunos. Porém, se for também muito baixo (50 a 60%), a maioria dos alunos terá a ilusão de que dominou a unidade, quando na realidade cometeu muitos erros.

Identificar objetivos ou unidades de ensino não dominados pelos estudantes é uma tarefa mais fácil do que descobrir o porquê do não domínio dos mesmos, considerando que os objetivos devem expressar as mudanças, pretendidas no comportamento dos alunos. Estas mudanças são resultantes de todo um processo de ensino/aprendizagem que inclui um ou mais componentes em interação, tais como:

- materiais instrucionais (materiais impressos, audio-visuais, transparências);
- procedimentos de ensino;
- informações;
- professor;
- etc...

Assim sendo, pode-se supor que cada um deles, individualmente ou combinados entre si, contribuem ou não para a aprendizagem de habilidades específicas, ou seja, para o domínio ou não de objetivos particulares. Qualquer um dos componentes do processo pode ser a causa das deficiências de aprendizagem dos alunos, desde que não tenha sido avaliado e comprovada a sua eficácia para uma dada situação. Outros fatores, tais como:

- problemas de interação professor/aluno;
- problemas na administração dos instrumentais de medida;
- falta de motivação dos alunos, para o alcance dos objetivos propostos;
- incapacidade do professor para transmitir um certo conteúdo ou para executar o método em questão.

poderiam, também, ser identificados como outras possíveis causas. Por exemplo, tem-se a seguinte situação de ensino/aprendizagem: alguns materiais instrucionais elaborados para o domínio de conceitos básicos de ciências biológicas, foram utilizados por um certo grupo de alunos. Em seguida, estes alunos foram testados no alcance dos objetivos propostos. Os resultados do teste com propósito formativo revelaram que alguns dos objetivos não foram dominados pelo grupo. Sabendo que os materiais instrucionais e os testes de aprendizagem não foram pré-avaliados, assim como os alunos não foram pré-testados, o avaliador formativo poderá supor que a causa do não domínio está em qualquer um destes componentes ou em todos eles, ao mesmo tempo. Tem-se que considerar que:

- . os materiais não tendo sido previamente avaliados e nem aprimorados, poderão estar apresentando problemas e, assim, impedindo os alunos de alcançarem certos objetivos. (Pode-se observar que embora a testagem dos materiais seja o ideal, na realidade não é muito fácil realizar, anteriormente, a avaliação dos mesmos; e quase sempre estão sendo avaliados ao mesmo tempo que os alunos);
- . os alunos, não tendo sido pré-testados, podem não possuir os pré-requisitos necessários à instrução, não tendo, portanto, condições para uma aprendizagem efetiva. (No caso de instrução individualizada, tal problema, provavelmente, não ocorrerá desde que nesta situação leva-se em consideração os pré-requisitos dos alunos);
- . o instrumental de medida não tendo sido anteriormente avaliado pode não ser um teste válido ou confiável. (É importante notar, aqui, que sendo os testes e os materiais instrucionais ambos ligados diretamente com os objetivos, torna-se impossível isolar um componente do outro, ao avaliá-los. Assim, avaliar os testes antes de validar os materiais implica no uso dos testes sem os materiais, o que invalidaria o próprio processo de validação dos testes. O avaliador deve conscientizar-se de que, provavelmente,

ele sempre estará trabalhando dentro desta restrição. Talvez a utilização sucessiva de várias avaliações formativas de materiais instrucionais e testes ligados aos mesmos objetivos, viesse minimizar tal problema).

O avaliador poderia ainda identificar outras possíveis causas, além destas, desde que numa situação de instrução muitas variáveis existem, influenciando todo o processo.

Se o avaliador enfrenta uma situação como a descrita acima, sua tarefa torna-se mais complexa. Ele deverá procurar muitas evidências de tal modo que seja possível localizar exatamente a causa das falhas de aprendizagem dos alunos. As evidências coletadas, além de serem úteis na identificação da causa destas falhas, poderão, também, ser utilizadas para melhorar todos os componentes do processo, simultaneamente. Por exemplo, tem-se os seguintes resultados obtidos através da aplicação de um teste com propósito formativo, conforme a Tabela F.20.

Analisando os dados da Tabela F.20, pode-se obter algumas informações, como:

- aproximadamente 30% dos alunos não dominou 60% dos objetivos;
- os objetivos 6 e 7 não foram dominados por 70% dos alunos;

No primeiro caso, tem-se que uma percentagem de alunos (30%), relativamente pequena, não dominou grande parte dos objetivos. Bloom (1971) sugere que erros cometidos por uma minoria de alunos podem ser atribuídos a eles próprios e que os objetivos não dominados devem ser revistos individualmente. Esta revisão é feita com o objetivo de sugerir instrução remedial a estes alunos e conduzi-los ao domínio dos objetivos não alcançados (avaliação formativa de alunos). No segundo caso, tem-se que uma percentagem de alunos (70%) relativamente grande não dominou dois objetivos da unidade. Bloom (1971) sugere que erros cometidos por uma parte considerável da classe (60% ou mais) podem ser atribuídos aos testes, aos materiais instrucionais ou ao processo de ensino. Assim sendo, os materiais devem ser revistos e melhorados nos pontos que estão apresentando dificuldades (avaliação formativa de materiais instrucionais). Tem-se ainda que a causa do não domínio dos objetivos 6 e 7 pode

ria estar nos testes de aprendizagem. Uma análise dos mesmos poderia mostrar que os itens medindo estes objetivos apresentam alguma tendenciosidade (por exemplo: item mal formulado, direcionando a resposta para uma das alternativas que não é a correta). Assim, eles devem ser revistos e reformulados (Ver unidade 3.0, pág. F.168).

Em resumo, tem-se que qualquer um dos componentes de um processo instrucional, desde que não tenha sido avaliado e aprimorado, poderá ser causa do não domínio dos objetivos propostos para aquela instrução. A tarefa de identificar as causas do não domínio de certos objetivos, torna-se mais complexa e difícil se vários materiais são utilizados de forma combinada para o alcance de um conjunto de objetivos. Nesta situação, é realmente difícil localizar exatamente a causa do não domínio, desde que vários materiais (por exemplo: textos impressos, transparências e aulas de TV) interferem na aprendizagem dos alunos. Todos eles devem ser revistos e, provavelmente, modificações em um deles implicará em modificações em todos os outros.

Considerando que, em situações como as descritas acima, a tarefa do avaliador formativo torna-se mais complexa e difícil; pode-se dizer que para maior sucesso da Avaliação Formativa de alunos, deve-se ter assegurada a eficácia de todos os componentes do processo. Esta seria uma situação ideal e iria significar que:

- . todos os materiais instrucionais a serem utilizados não apresentam problemas ou dificuldades;
- . o professor está suficientemente qualificado para desempenhar suas tarefas;
- . os procedimentos de ensino são adequados para aquela situação específica;
- . os testes de aprendizagem são válidos e confiáveis;
- . os alunos possuem todos os pré-requisitos necessários à instrução. (No caso de instrução individualizada, este problema é eliminado, considerando que os objetivos instrucionais são elaborados em função do comportamento de entrada de cada aluno).

E ainda, mesmo nesta situação, o avaliador formativo não deve desconsiderar outros possíveis fatores, capazes de influenciar o desempenho do aluno. Deve procurar identificá-los e exercer um certo controle sobre os mesmos. Por exemplo, se um novo método de ensino (incluindo procedimentos, materiais instrucionais, testes de aprendizagem, alunos e professor) foi avaliado e aperfeiçoado antes de ser definitivamente instalado no sistema, a tarefa do avaliador formativo será então simplificada. Nesta situação, com maior segurança, ele poderá atribuir a causa do não domínio de objetivos a alguns outros fatores (ou variáveis) tais como:

- os alunos não leram todas as páginas indicadas nos materiais;
- não fizeram um número suficiente de exercícios práticos;
- trocaram idéias entre si enquanto respondiam os itens do teste;
- os alunos estavam ansiosos enquanto respondiam o teste;
- não estavam motivados para receber a instrução.

É importante relembrar, aqui, que se está falando de uma situação ideal. Na realidade, é praticamente impossível obter uma situação de ensino/aprendizagem que não envolva qualquer tipo de problema. Mesmo que cada um dos componentes instrucionais tenha sido previamente avaliado, dentro do contexto de sala de aula, em interação com outros componentes, poderão apresentar dificuldades que, embora pequenas, possam causar deficiências na aprendizagem dos alunos. Para se conseguir um certo grau de perfeição destes componentes, é preciso que eles sejam testados e revisados várias vezes durante um período relativamente longo. Assim sendo, o avaliador deve atentar para todos os componentes instrucionais, mesmo que já tenham sido avaliados e aperfeiçoados.

2) Informar o Aluno Acerca do Domínio ou não dos Objetivos Propostos a Fazer Indicações de Caminhos Alternativos a serem Seguidos para o Alcance dos Objetivos não Dominados.

Se ao aluno é indicado quais itens acertou ou errou, ou quais foram exatamente os objetivos não dominados, ele poderá ter informações acerca do conteúdo e habilidades que tem de aprender e revisar.

Conhecerá os pontos em que está tendo dificuldades, o que lhe será muito útil para efetuar as revisões necessárias. Para quem alcançou ou aproximou-se do domínio dos objetivos, os resultados da Avaliação Formativa podem ser uma recompensa ou reforço eficaz. À medida em que o estudante vai sendo comunicado do domínio de pequenas unidades, ele será reforçado positivamente. É provável que, neste caso, ele consiga maior motivação e segurança para prosseguir nas unidades seguintes e se antecipe às provas, preparando-se adequadamente para as mesmas. Por outro lado, os estudantes que não alcançaram o domínio de uma pequena unidade de ensino também poderão ter segurança com relação ao que aprendeu e ao que terão que aprender. Com relação aos objetivos não dominados, os alunos deverão voltar à aprendizagem das habilidades, neles identificadas, através de caminhos alternativos indicados. Em seguida, uma forma paralela de prova com propósito formativo será proporcionada a estes alunos. Os resultados desta 2a. testagem poderão reforçar positivamente, aqueles que a ela se submeteram.

Várias formas alternativas de aprendizagem podem ser proporcionadas aos estudantes. Algumas delas podem ser:

- ajuda de monitores;
- cooperação de grupos especiais;
- explicações mais claras e detalhadas, pelo professor em sala de aula;
- novos materiais impressos individualizados;
- ilustrações concretas sobre o assunto;
- exercícios sobre problemas específicos.

E muitas outras podem ser pensadas. Como existem diferenças individuais, certas alternativas podem ser mais eficazes para uns do que para outros. É importante que o estudante compreenda que se não aprendeu por um método de instrução ou conjunto de materiais, pode aprender através de outros. Assim, os objetivos não dominados devem ser tratados sob um novo enfoque, ou seja, novos procedimentos ou caminhos alternativos devem ser seguidos para alcançar o domínio dos mesmos.

Para informar ao aluno sobre seu progresso em uma pequena unidade de ensino, é útil o uso de uma folha contendo: (1) o registro dos erros cometidos no teste e (2) a prescrição detalhada dos materiais instrucionais que poderão ser utilizados para superar as dificuldades e alcançar os objetivos não dominados.

São apresentadas, a seguir, como sugestão, duas tabelas (Tabela F.21 e Tabela F.22) que poderão ser utilizadas na Avaliação Formativa dos alunos.

A Tabela F.21 poderá ser fornecida aos alunos com todos os itens já preenchidos.

Uma outra tabela é apresentada a seguir. Esta será preenchida pelo próprio aluno que deverá:

- . marcar com 0 (zero) os itens que respondeu incorretamente e com 1 (um) os itens respondidos corretamente;
- . assinalar com X os objetivos dominados ou não;
- . escolher (sozinho ou em consulta com o professor) as alternativas de aprendizagem indicadas, aquela que julgar mais eficaz para si, para revisão e alcance dos objetivos não dominados.

Além das alternativas de aprendizagem identificadas nas Tabelas F.21 e F.22 existem, como foi visto, algumas outras que podem também ser utilizadas, dependendo da situação. Por exemplo, se grande parte dos alunos não dominou um certo objetivo, o professor pode achar mais apropriado reuni-los em uma hora de aula e fornecer-lhes, ele próprio, explicações mais detalhadas sobre alguns conhecimentos básicos. O professor poderia, também, achar apropriado, para esta situação, a utilização da técnica de cooperação de grupos especiais. Estes grupos são organizados de tal forma que alguns alunos que os compõem, tendo dominado o objetivo, podem ajudar aqueles que não adquiriram, ainda, a habilidade. Ou ainda, a solução mais eficiente poderia ser a realização de exercícios práticos sob a orientação de monitores. E muitas outras alternativas de aprendizagem podem ser identificadas para cada situação particular.

TABELA F.21

UM EXEMPLO DE FICHA DE AVALIAÇÃO FORMATIVA DE ALUNO, A SER PREENCHIDA PELO PROFESSOR

ALUNO:
 DISCIPLINA:
 UNIDADE DE ENSINO:
 TESTE Nº:
 DATA:

OBJETIVOS PROPOSTOS		ITENS DE TESTE CORRESPONDENTES	OBJETIVOS DOMINADOS	OBJETIVOS NÃO DOMINADOS	CAMINHOS ALTERNATIVOS DE APRENDIZAGEM	
1 - O aluno deverá identificar equações matemáticas de 2º grau.		1,2,3	X		LIVRO TEXTO: MATEMÁTICA FUNDAMENTAL - 6ª. SÉRIE	pgs. 35 a 39
2 - O aluno deverá encontrar a solução de equações matemáticas de 2º grau		4,5	X		TEXTO - INSTRUÇÃO PROGRAMADA: EQUAÇÕES MATEMÁTICAS DE 2º GRAU	pg 1 a 18
3 - O aluno deverá solucionar problemas utilizando equações matemáticas de 2º grau		6,7,8		X		pg 10 a 20
n -						pg 20 a 30

X = OBJETIVO DOMINADO

TABELA F.22

UM EXEMPLO DE FICHA DE AVALIAÇÃO FORMATIVA DE ALUNO, A SER PREENCHIDA
PELO PROFESSOR E PELO ALUNO

NOME:
DISCIPLINA:
DATA:

Escreva o número 1 nos itens acertados e 0 (zero) naqueles que respondeu incorretamente. Assinale, também, com um X os objetivos dominados, se você acertou os dois itens, correspondentes a cada objetivo. Caso contrário, assinale em objetivo não dominado. Para revisão dos itens errados, escolha uma ou mais das alternativas indicadas. Estas explicam corretamente as idéias contidas nos itens

OBJETIVOS PROPOSTOS		ITENS DE TESTE	RESPOSTAS CORRETAS	OBJETIVOS DOMINADOS	OBJETIVOS NÃO DOMINADOS	CAMINHOS ALTERNATIVOS DE APRENDIZAGEM		
1 - 0 aluno deverá identificar operações de adição	1	1		X		LIVRO TEXTO: MATEMÁTICA FUNDAMENTAL - 6a. SÉRIE	TEXTO Nº 5: EXERCÍCIOS PRÁTICOS	
	2	1			pgs 15 a 19			pg 5 a 8
2 - 0 aluno deverá solucionar operações de adição	3	1		X		pgs 20 a 27	pg 7 a 9	
	4	1						
3 - 0 aluno deverá solucionar problemas de adição	5	0			X	pg 33 a 41	pg 10 a 15	
	6	1						
4 - 0 aluno deverá solucionar problemas com operações de adição e de subtração	7	0			X	pg 33 a 38	pg 12 a 18	
	8	0						
	9	1						

1 = RESPOSTAS CORRETAS

0 = RESPOSTAS INCORRETAS

Os procedimentos de avaliação formativa de alunos, que foram descritos, são apenas sugestões que podem ser acrescidas de outras idéias ou adaptadas para situações específicas. (Não são portanto, procedimentos rígidos, mas devem ser adaptados para melhor atender a necessidades particulares).

4.4 - ALGUMAS CONCEPÇÕES ERRÔNEAS EM ESTUDOS DE AVALIAÇÃO FORMATIVA

Antes de concluir esta unidade, julgou-se útil apresentar, aqui, algumas concepções errôneas em estudos de Avaliação Formativa, identificados por Sanders (1972).

A primeira concepção errônea diz respeito a variáveis que devem ser consideradas pelo avaliador formativo. Em muitos estudos desta natureza, persiste a idéia de que somente os estudantes deveriam ser observados. Esta atitude limita o objetivo da Avaliação Formativa. No sistema instrucional, existem muitas outras variáveis importantes que devem ser consideradas pelo avaliador, além do desempenho do estudante. Por exemplo, a atitude dos pais e professores em relação a um novo método de ensino não deixam de ser, também, informações importantes para o avaliador, quer seja ela a Avaliação Formativa ou Somativa.

Hammond (Sanders, 1972) produziu a Figura F.19, identificando variáveis importantes dentro do processo de avaliação.

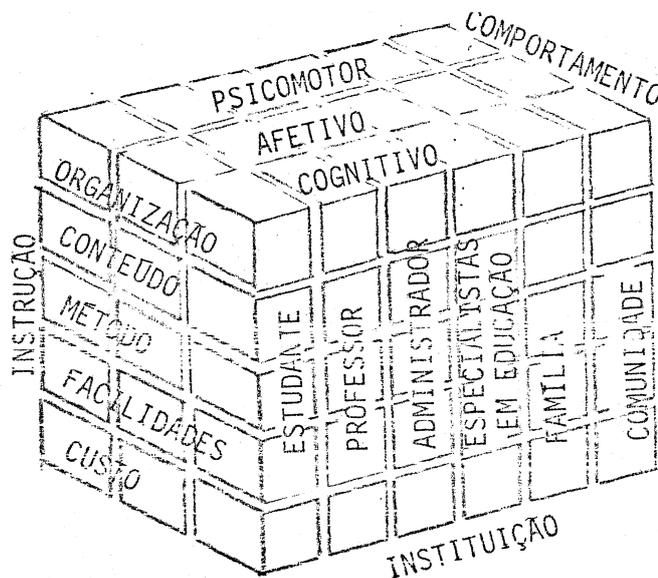


Fig. F.19 - Estrutura para Avaliação

A segunda concepção errônea refere-se à elaboração de objetivos. Tem-se acreditado que o elaborador ou o professor deve chegar até o avaliador com um conjunto de objetivos já precisamente descritos, claros e sem ambiguidade de tal forma que possam ser utilizados pelo avaliador, para desenvolver instrumentais válidos. É de responsabilidade do avaliador auxiliar na elaboração de objetivos comportamentais. O avaliador não tem obrigação de escrever objetivos para o professor ou para o elaborador, mas sim de ajudá-los a definir claramente os seus objetivos.

Uma terceira concepção errônea refere-se ao papel do elaborador do componente e do avaliador. Uma pessoa pode ser ao mesmo tempo elaborador e avaliador, porém, é necessário que estes papéis sejam bem discriminados. O elaborador deve ser capaz de identificar claramente sua atuação como um ou outro. Assim, deve ser capaz de dizer: "Eu, agora, estou atuando como avaliador", ou "eu estou atuando, agora, como elaborador".

Uma outra concepção errônea refere-se a explicação de fenômenos. A Avaliação não tem, necessariamente, obrigação de explicar por que as coisas acontecem. Isto não significa que o avaliador evite procurar por explicações, mas a explicação não é necessária para uma boa Avaliação. A explicação do porquê de um fenômeno é função da pesquisa.

Concluindo os estudos de Avaliação Formativa de alunos e de materiais instrucionais impressos é necessário dizer que os procedimentos utilizados para avaliar formativamente simples componentes instrucionais podem, também, ser parte de um conjunto de procedimentos identificados para avaliar formativamente um sistema inteiro. Assim, tem-se em mente que um modelo de Avaliação Formativa de sistema abrange vários outros procedimentos capazes de fornecer informações mais complexas e mais diversas daquelas requeridas para a Avaliação de um simples componente instrucional.

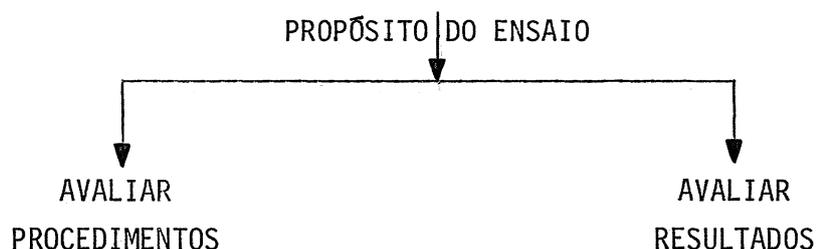
Esta unidade tratou de alguns procedimentos de Avaliação Formativa. A unidade seguinte descreverá alguns procedimentos de Avaliação Somativa dos produtos da mudança em termos de aprendizagem.

5.0 - UNIDADE 5 - OBTENÇÃO DE INFORMAÇÕES: AVALIAÇÃO SOMATIVA DOS PRODUTOS DA MUDANÇA, EM TERMOS DE APRENDIZAGEM

5.1 - INTRODUÇÃO: DOIS ESQUEMAS PARA AVALIAÇÃO SOMATIVA DOS PRODUTOS DA MUDANÇA

Antes que mudanças inovativas sejam instaladas definitivamente em sistemas mais amplos, elas devem passar por uma fase de ensaio na qual são testadas. Estas mudanças referem-se mais frequentemente a adoção de novos programas educacionais, novos materiais instrucionais, métodos de ensino, utilização de novos equipamentos, e outros.

A finalidade desta testagem é avaliar a execução dos procedimentos planejados e os resultados obtidos com tais procedimentos, relacionando ambos a fim de controlar o funcionamento do protótipo e explicar os resultados.



Com relação aos resultados do ensaio, informações a eles relativas podem servir aos propósitos formativo e somativo. Na Avaliação Formativa, Avaliação interna de um esforço de mudança, através da comparação do desempenho dos estudantes submetidos, ao mesmo, com seus objetivos, capacitam-se os seus elaboradores a dizer se ele está funcionando, e a sugerir modificações para melhorá-lo.

Na Avaliação Somativa, tenta-se, através de procedimentos mais sistemáticos, chegar-se a uma decisão final, indubitável, acerca do valor da mudança.

A sistematização dos procedimentos adequados para conduzir a Avaliação Somativa depende do tipo de informação que será útil ao decisor para concluir acerca do valor da mudança em ensaio.

Ao decisor pode ser útil informação do tipo:

- a) Absoluta X Comparativa
- b) Generalizável X Não Generalizável
- c) Envolvendo relações Causa-Efeito X Não envolvendo relações Causa-Efeito.

a) Absoluta X Comparativa

A Informação Absoluta envolve a comparação dos resultados produzidos por uma intervenção, com padrões de desempenho pré-estabelecidos. Um exemplo de questão que, para sua resposta, exige informação absoluta é o seguinte: "a mudança introduzida está atingindo seus objetivos a nível de um padrão absoluto desejado?".

A Informação Comparativa envolve a comparação dos resultados produzidos por uma intervenção, com aqueles produzidos por intervenções alternativas. A questão "a mudança introduzida está atingindo seus objetivos com maior êxito do que mudanças alternativas?" implica na utilização de informação comparativa. Neste caso, a Avaliação deve proporcionar uma comparação entre as saídas (ex.: aprendizagem dos alunos) ocasionadas por diferentes entradas (programas, materiais instrucionais, métodos de ensino e outros).

b) Generalizável X Não Generalizável

A Informação Generalizável envolve, através da utilização de procedimentos de Estatística Inferencial, conclusões acerca da população-alvo, com base em dados de uma amostra*.

* Os conceitos de amostra e população serão tratados mais adiante, na página F.268.

A informação Não Generalizável não envolve inferências a cerca da população alvo a partir de dados amostrais. Ela é descritiva de uma realidade.

c) Envolvendo relação causa e efeito X Não envolvendo relação causa-efeito.

Ao decisor pode ser útil a informação referente aos efeitos causados pelos esforços de mudança nas saídas do sistema, ou então, a informação somativa não relativa a tais efeitos. Assim, em função da informação requerida pelo decisor: absoluta ou comparativa, generalizável ou não generalizável, envolvendo relações causa e efeito ou não envolvendo relações causa efeito; é possível conceber vários modos de conduzir a Avaliação Somativa dos produtos da mudança. O avaliador deve ser eclético na seleção dos conceitos, técnicas, etc. disponíveis para elaboração de esquemas de Avaliação.

Serão apresentados, nesta unidade, dois esquemas para quando o produto da mudança a ser avaliado somativamente for a aprendizagem dos alunos. Tais esquemas são resultado da sistematização de procedimentos para conduzir a Avaliação que fornecerá, ao decisor, informação que lhe será útil em duas situações específicas:

SITUAÇÃO I - Em que:

- Existe a necessidade de escolher entre alternativas de ação. Numa fase de testagem, a inovação que se pretende introduzir no sistema é implementada paralelamente à alternativa vigente, para que, ao final do ensaio, seja decidido se a mudança deve ser instalada definitivamente no sistema, em substituição à alternativa tradicional, ou se é esta última que permanece (informação comparativa);
- Quer assegurar-se que o "efeito líquido" da instrução em si mesma seja revelado, ou melhor, quer saber-se acerca dos efeitos que um esforço de mudança causa sobre as saídas do sistema. O objetivo maior, sendo, portanto, chegar-se a relações causa-efeito en

tre especificação de procedimentos da mudança, e obtenção do projeto (informação envolvendo relação causa-efeito);

- Existe necessidade, por questões econômicas e práticas, de trabalhar-se na fase de ensaio de uma inovação, com amostras. E, é preciso que a informação obtida com a amostra seja generalizável para a população, através da utilização de processos científicos (informação generalizável).

Esta situação caracteriza-se, portanto, por envolver si multaneamente informação comparativa, referente a relação causa-efeito e generalizável.

SITUAÇÃO 2 - Em que:

- Existe necessidade de testar uma inovação com o propósito de ve rificar sua VALIDADE, isto é, verificar se a inovação conduz de terminada clientela especificada, ao alcance dos objetivos propostos (informação absoluta);
- Não existe preocupação com relações causa-efeito (informação não envolvendo relação causa-efeito);
- Não existe preocupação de utilizar processos científicos para in ferir da amostra para a população. Ao invés disso, a ênfase é da da na utilização de índices de efetividade, que indicam em que me dida, ou grau, os objetivos propostos foram alcançados. Tais ín dices descritivos possibilitam aos decisores informações mais de talhadas a respeito dos resultados da mudança, fazendo transpare cer, através deles, a VALIDADE ou não da mesma (informação não ge neralizável).

Esta situação caracteriza-se, portanto, por envolver si multaneamente informação absoluta, não referente a relação causa-efeito e não generalizável.

Serão fornecidos esquemas úteis nestas duas situa ções específicas, por serem elas as situações de Avaliação Soma

tiva que mais se adequam ao contexto de avaliação de Produto proposto por Stufflebeam et al (1971), cuja abordagem foi tratada neste curso.

O primeiro dos esquemas é baseado no processo e na base lógica da Pesquisa Experimental, e vem sendo tradicionalmente utilizado em Avaliação Somativa dos resultados.

O segundo esquema envolve a sistematização de procedimentos para levantamento de dados para descrição do valor de uma mudança, em termos de sua efetividade. É uma abordagem mais recente que se caracteriza por avaliar os resultados produzidos por uma mudança, em comparação com padrões absolutos pré-estabelecidos.

A seleção, pelo avaliador, do esquema de Avaliação Somativa a ser utilizado, deve basear-se nas questões que o decisor quer respondidas. Ambos os esquemas são instrumentais úteis ao avaliador para o exercício de sua atividade de delinear, obter e prover informações que orientarão aos decisores no seu julgamento final acerca de um esforço de mudança. O que distingue a utilidade de um e outro esquema é o tipo de informação que cada um permite fornecer.

5.2 - SUB UNIDADE 1 - UM ESQUEMA DE AVALIAÇÃO SOMATIVA DOS PRODUTOS DA MUDANÇA (NO CASO APRENDIZAGEM) BASEADO NOS PROCEDIMENTOS DA PESQUISA EXPERIMENTAL

Dentro da metodologia abrangente da Avaliação, definida como o processo de delinear, obter e prover informações úteis para o julgamento de alternativas de decisão, está incluída a metodologia da Pesquisa Experimental. Mais especificamente, é na Avaliação de Produto, verificando os resultados de uma mudança introduzida num sistema, e nesta, ainda mais especificamente, é na Avaliação Somativa para Decisão Comparativa que o uso do "design" experimental de coleta e análise de dados tem sua principal utilidade para o processo avaliatório.

Através da utilização de um esquema de Avaliação Somativa dos produtos da mudança, baseado nos procedimentos da Pesquisa Experimental*, o que o avaliador pretende é fornecer aos decisores informações úteis acerca dos efeitos que certas(s) mudança(s) introduzida(s) (variável de entrada, também denominada variável independente) ocasiona(m) sobre variável(eis) de saída (também denominadas variáveis dependentes).

O efeito de qual variável independente, sobre qual variável dependente, será objeto de estudo, é identificado a partir de um problema de Avaliação. Assim, por exemplo, se existe necessidade de decidir sobre a adoção de um método de ensino entre dois possíveis (A=sintético, B=analítico) para a alfabetização de adultos - PROBLEMA - é necessário que se faça uma Avaliação Somativa dos dois métodos, ou seja, é necessário que se estude os efeitos dos métodos de ensino - VARIÁVEL INDEPENDENTE - na alfabetização de adultos - VARIÁVEL DEPENDENTE. Existindo diferença significativa entre os efeitos de cada método, deverá ser adotado aquele de maior efeito.

O esquema geral dentro deste processo de Avaliação, baseado no modelo da Pesquisa Experimental é propor para o problema identificado uma resposta provável do mesmo (hipótese), e coletar dados da realidade que comprovarão ou não tal hipótese sugerida.

Este estudo deve ocorrer em escala reduzida, numa fase de testagem anterior à utilização por todo o sistema, do método já então apontado como o melhor.

É claro que, não havendo uma identificação precisa do avaliador acerca de quais dados coletar, sob que condições fazê-lo, e que procedimentos utilizar para coletá-los, poder-se-á chegar a interpretações errôneas ou não significativas acerca destes efeitos. Uma série

* Neste curso serão tratados apenas os procedimentos de análise dos dados relativos a testagem de hipóteses. Não serão envolvidas as técnicas de estimativa.

de fatores reduzem a interpretabilidade dos dados avaliativos. É exata mente para controlar tais fatores e aumentar a confiança nos resultados de uma Avaliação Somativa, que são utilizados esquemas que orientam a es truturação e o planejamento de tais esforços avaliativos.

4.5.2.1 - PASSOS DO ESQUEMA

- P
L
A
N
E
J
A
M
E
N
T
O

E
X
E
C
U
Ç
Ã
O
- 1 - Identificar e definir o problema da Avaliação
 - 2 - Formular as hipóteses do problema
 - 3 - Identificar e classificar variáveis
 - 4 - Definir operacionalmente as variáveis
 - 5 - Elaborar um plano experimental para coleta e análise de dados, envolvendo, possivelmente:
 - elaboração de um plano de amostragem de sujeitos para repre sentar a população dada;
 - identificação das variáveis não experimentais que possam conta minar o experimento, e o como controlá-las;
 - seleção de um "design" de pesquisa para coleta de dados;
 - elaboração de um plano para designação dos sujeitos para os grupos e de determinação dos tratamentos experimentais para os grupos;
 - esquematização de procedimentos para coleta de dados;
 - elaboração de um plano de construção dos instrumentais para me dir os resultados do experimento;
 - elaboração de um plano de testagem de hipótese
 - 6 - Executar o experimento
 - 7 - Reduzir os dados brutos
 - 8 - Computar um teste estatístico apropriado para fazer decisões acer ca da(s) hipótese(s) que está(ão) sendo testada(s).

Os itens de 1 a 5 deste esquema são relativos a fase de planejamento do experimento. Os itens de 6 a 8 são relativos à fase de execução da coleta e análise de dados.

A seguir será tratado, em detalhes, cada um destes passos que compõem o esquema de Avaliação Somativa, baseado nos procedimentos da Pesquisa Experimental.

1) IDENTIFICAR E DEFINIR O PROBLEMA DE AVALIAÇÃO

Em Avaliação, o problema é determinado pela própria situação. Os objetivos do próprio projeto o definem. Onde houver necessidade de determinar-se o valor de novos programas educacionais, materiais instrucionais, etc., estará identificado um problema de Avaliação Somativa. O avaliador em interface com os decisores do projeto deverá identificar as questões avaliativas que os mesmos necessitam respondidas.

A definição de um problema deve ter as seguintes características:

- deve envolver relação entre duas ou mais variáveis. (O avaliador manipula, no mínimo, uma variável para determinar seu efeito sobre outras variáveis. Assim, a definição de um problema requererá a inclusão de pelo menos duas variáveis e sua relação);
- deve ser clara e não ambígua, usualmente apresentada sob a forma de uma questão;
- deve ser possível coletar dados para resolver a(s) questão(ões) feita(s).

Exemplos:

Problema 1: "Existe efeito de método de ensino, na aprendizagem dos alunos da 1ª. série do 1º grau, no Projeto Y?"

Nesta afirmativa do problema:

- existem duas variáveis: método de ensino e aprendizagem dos alunos;
- existe relação entre estas variáveis (efeito da primeira sobre a segunda);

- o problema está claro e é apresentado sob a forma de uma questão;
- é possível coletar dados para responder a questão. Podem ser aplicados testes de aprendizagem em dois grupos de alunos, um submetido ao método A, outro ao método B, por exemplo.

Um outro exemplo seria:

Problema 2: "Existe efeito dos procedimentos relativos à intervenção instrucional X, na aprendizagem dos alunos da 1ª série do 1º grau, no Projeto Y?".

Nesta afirmativa do problema:

- existem duas variáveis: procedimentos relativos a intervenção instrucional X e aprendizagem;
- existe relação entre estas variáveis (efeito da primeira sobre a segunda);
- o problema está claro e é apresentado sob a forma de uma questão;
- é possível coletar dados para responder a questão.

2) FORMULAR A HIPÓTESE DO PROBLEMA

Uma hipótese é uma resposta sugerida para o problema, acerca de relações entre os conceitos identificados no mesmo. Ela é afirmada antecipadamente à observação dos dados, representando uma expectativa acerca de eventos. Tem as seguintes características:

- deve supor uma relação entre duas ou mais variáveis;
- deve ser clara e não ambígua, em forma de uma sentença afirmativa;
- deve ser testável, isto é, deve ser possível de ser reescrita em forma operacional, que pode então ser avaliada como base nos dados.

Hipótese 1: "O método de ensino tem efeito na aprendizagem dos alunos da 1ª série do 1º grau, do Projeto Y".

Hipótese 2: "Os procedimentos relativos a intervenção instrucional X têm efeito na aprendizagem dos alunos da 1ª. série do 1º grau do Projeto Y".

3) IDENTIFICAR E CLASSIFICAR VARIÁVEIS

Uma hipótese pode envolver as seguintes variáveis:(Tuckman, 1972):

- variável independente;
- variável dependente;
- variável moderadora;
- variável controladora;
- variável interveniente.

Variável Independente: Segundo Tuckman (1972), é aquela variável (ou FATOR) que é medida, manipulada, ou selecionada pelo experimentador para determinar sua relação com um fenômeno observado. É a variável que ele irá manipular ou variar, para causar uma mudança em alguma outra variável. Ela é denominada independente porque existe o interesse em saber como ela afeta outra variável e não no que ela é afetada. Para Ferguson (1971), uma variável independente pode tanto ser uma variável de tratamento, como uma variável de classificação.

. Variável de tratamento: envolve uma modificação nos sujeitos, uma modificação que é controlada pelo avaliador. Ex.: os alunos poderão ser submetidos a um novo método de ensino (variável de tratamento), ou poderão receber instrução via TV através de um novo formato de programa (variável de tratamento).

. Variável de classificação: envolve características dos sujeitos que estavam presentes já anteriormente ao tratamento, e à parte deste, e que não resultam de manipulação do avaliador. Ex.: Sexo, idade, status sócio-econômico. Apesar das categorias de uma variável de classificação não serem criadas pelo avaliador, é ele que seleciona as variáveis de classificação que serão incluídas no estudo.

As diferentes variações de categorias das variáveis independentes são denominadas níveis. Ex.: Se estiver em teste o efeito de três métodos de ensino: a) instrução programada, b) aula expositiva, c) leitura e discussão em grupo, na aprendizagem de Estatística, tem-se, neste caso, a variável independente: Método de ensino, com três níveis: a, b, c. Se estiver sendo testado o efeito do tipo de escola na aprendizagem dos alunos, tem-se uma variável independente: tipo de escola, que poderá ter os três níveis: grupo escolar, escola reunida, escola isolada.

Os níveis da(s) variável(is) independente(s) são selecionados pelo avaliador. Tais níveis podem ser fixos ou randômicos, qualitativos ou quantitativos.

Nível fixo: determinados pelo avaliador. Ex.: o avaliador fixa os dois níveis, masculino e feminino, para um estudo ou os três níveis QI - 90, 110, 140 para outro estudo.

Nível Randômico: selecionado aleatoriamente dentre todos os possíveis. Ex.: trabalhar com os níveis 70 - 90 - 110, quando a variável é quociente de inteligência, e tais níveis foram selecionados ao acaso.

Nível Qualitativo: Ex.: urbano, rural ou masculino, feminino.

Nível Quantitativo: exprimem quantidades. Ex.: QI: 70, 80 e 120 (três níveis quantitativos).

Variável Dependente: Segundo Tuckman (1972), é aquele fator que é observado e medido, para determinar o efeito da variável independente. É a variável que irá mudar, como um resultado da alteração da variável independente. É denominada variável dependente porque seu valor é função do valor da variável independente. Mudando o valor desta última muda o valor da primeira.

Nas hipóteses seguintes, são identificadas suas variáveis independente e dependente.

Hipótese 1: "Existe efeito do método de ensino na aprendizagem dos alunos do Projeto Y".

Variável independente: método de ensino

Variável dependente: aprendizagem dos alunos

É esperado que, variando-se o método de ensino, varie o grau de aprendizagem dos alunos.

Hipótese 2: "Existe efeito dos procedimentos relativos à intervenção instrucional X, na aprendizagem dos alunos da 1ª série do 1º grau do Projeto X".

Variável independente: procedimentos relativos à intervenção instrucional.

Variável dependente: aprendizagem dos alunos.

É esperado que, variando-se os procedimentos instrucionais, varia o grau de aprendizagem dos alunos.

Comumente, em se tratando de Avaliação Somativa de um sistema escolar, trabalha-se com a variável dependente - aprendizagem dos alunos. Mais especificamente, trabalha-se com os objetivos educacionais que, com a mudança introduzida, pretende-se alcançar. O que será testado serão as expectativas acerca do que os estudantes poderão fazer depois de terem passado por tal mudança.

Variável Moderadora: De acordo com Tuckman (1972), é um tipo especial de variável independente. Ela é medida, manipulada ou selecionada para estudo, a fim de que se determine se afeta a relação entre a variável independente primária, e a variável dependente. Por exemplo, se para a hipótese 1, já estabelecida, o avaliador tiver interesse em examinar se o método de ensino tecnológico afeta mais a aprendizagem dos alunos do sistema municipal, então tal hipótese deverá ser reformulada, incluindo-se a dependência administrativa como uma variável moderadora, ou seja, uma variável que afeta a relação entre método de ensino e aprendizagem dos alunos.

Hipótese 3: "No Projeto Y:

Os alunos submetidos ao método tecnológico, nas escolas municipais, aprendem mais que os alunos submetidos

tidos ao método tecnológico, nas escolas estaduais, enquanto os alunos submetidos ao método tradicional nas escolas estaduais aprendem mais que os alunos submetidos ao método tradicional, nas escolas municipais".

Se todos os alunos fossem reunidos em 2 grupos (método tecnológico e método tradicional), o efeito do método tecnológico poderia ter sido maior que o do método tradicional; porém se os alunos são separados segundo a dependência administrativa (municipal-estadual) os dois métodos podem ter efeito diferente em cada grupo. Quando isto ocorre, diz-se que existe efeito da INTERAÇÃO das duas variáveis independentes na aprendizagem dos alunos.

Hipótese 4: "No Projeto Y:

O método de ensino tecnológico tem mais efeito do que o método de ensino tradicional na aprendizagem dos alunos tanto da zona urbana quanto da zona rural, mas ele tem singularmente mais efeito na aprendizagem dos alunos da zona rural".

Variável independente: método de ensino

Variável moderadora: zona de localização da escola.

Variável dependente: aprendizagem dos alunos.

Graficamente, esta relação de moderação seria do tipo apresentado na Figura F.20.

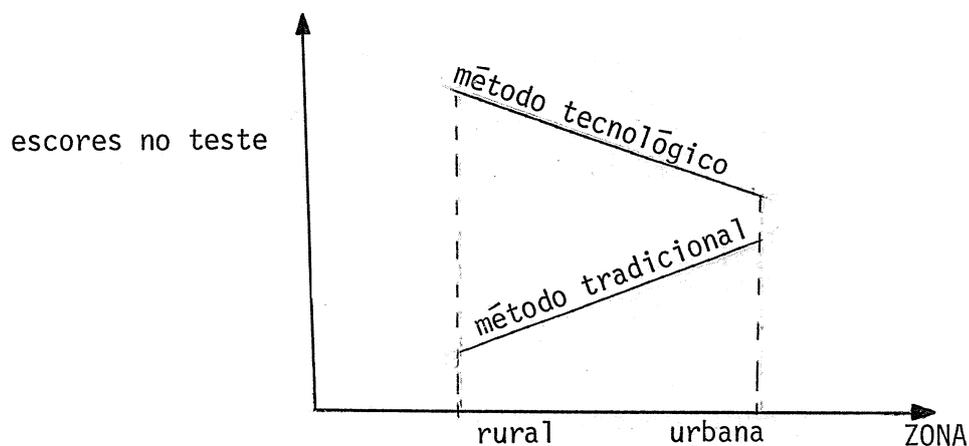


Fig. IV.20 - Efeito da Interação da Variável Independente primária e da Variável Moderadora.

Variáveis de Controle: Ainda segundo Tuckman (1972), são aquelas que são controladas pelo avaliador para cancelar ou neutralizar qualquer efeito que elas poderiam ter sobre o fenômeno observado. Todas as variáveis não podem ser estudadas ao mesmo tempo; algumas precisam ser neutralizadas para garantir que não terão efeito diferencial ou moderador na relação entre a variável independente e a variável dependente, em estudo.

Num estudo, o efeito da variável de controle é neutralizado, enquanto o efeito da variável moderadora é realmente estudado.

Efeitos das variáveis de controle podem ser neutralizados através de procedimentos que serão estudados no tópico b.1, página

F.283

Hipótese 5: "No projeto Y:

Na zona urbana o método de ensino tecnológico tem mais efeito na aprendizagem dos alunos do que o método de ensino tradicional".

Variável independente: método de ensino

Variável dependente: aprendizagem dos alunos

Variável de controle: zona.

Todas as variáveis descritas anteriormente são concretas. Cada variável independente, moderadora e de controle pode ser manipulada pelo avaliador, e cada variação pode ser observada por ele, na maneira que elas afetam a variável dependente.

Variável Interveniente: Para Tuckman (1972), é aquele fator que teoricamente afeta o fenômeno observado, mas não pode ser visto, medido, ou manipulado; seu efeito precisa ser inferido a partir dos efeitos das variáveis independente e moderadora sobre o fenômeno.

Hipótese 6: "O modo de apresentação de uma aula tem efeito na aprendizagem dos alunos".

Variável independente: modo de apresentação de uma aula.

Variável interveniente: personalidade do professor.

Variável dependente: aprendizagem.

Realmente, se dois modos de apresentação de uma aula estão sendo comparados (aula por TV X aula expositiva), pode ocorrer que um modo seja mais efetivo que outro, não por que apresente maiores ou menores informações sobre o assunto, mas porque a personalidade do professor (quer da aquele que transmite a aula via TV, quer daquele que transmite a aula diretamente em sala de aula) influi na aprendizagem dos alunos. Neste caso, a variável "personalidade do professor" é interveniente.

As relações entre os 5 tipos de variáveis descritas podem ser visualizadas na Figura IV.21, embasado numa diagramação apresentada por Tuckman (1972).

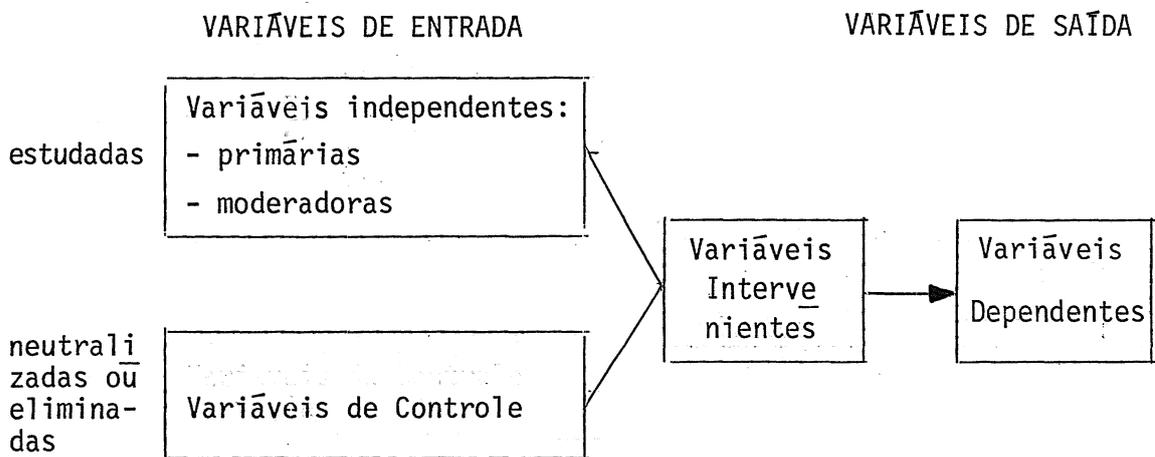


Fig. F.21 - Relação entre os 5 Tipos de Variáveis.

4) DEFINIR OPERACIONALMENTE AS VARIÁVEIS

Consiste na conversão de variáveis de uma forma abstrata ou conceitual para uma forma operacional.

A característica principal de uma definição operacional é que ela envolve um critério observável e mensurável daquilo que está sendo definido. Embora possa haver mais que uma definição operacional para uma variável, na situação na qual vai ser usada, ela deve ser única.

A possibilidade de se testar uma hipótese depende da possibilidade de serem construídas definições operacionais adequadas, para suas variáveis.

Neste passo, sendo a variável dependente, a aprendizagem dos alunos ou, mais especificamente, os objetivos educacionais que a mudança introduzida pretende alcançar, estes são operacionalizados, ou seja, são escritos em termos comportamentais. Tais objetivos comportamentais devem explicar o desempenho que o aluno demonstrará se os objetivos da mudança foram alcançados.

Exemplos de definições operacionais de variáveis são as seguintes:

aprendizagem: - escores individuais dos alunos num teste com referência ao critério (C.R.T).

- ou, escores médios das classes num teste com referência a normas (N.R.T).

efetividade: - percentagem de objetivos instrucionais atingidos por aluno, num teste com referência ao critério.

zona rural: - local cuja densidade demográfica é igual ou inferior a $\frac{x \text{ habitantes}}{\text{Km}^2}$

zona urbana: - local cuja densidade demográfica é superior a $\frac{X \text{ habitantes}}{\text{Km}^2}$

A definição operacional das variáveis tem, também, a vantagem de facilitar a comunicação. Para uma comunicação precisa, há necessidade de operacionalização da mesma, para que pessoas diferentes compreendam, exatamente, o seu significado.

5) ELABORAR UM PLANO EXPERIMENTAL PARA COLETA E ANÁLISE DE DADOS.

Neste passo, é elaborado um plano para condução do experimento, ou seja, é criado um plano para investigar o problema na realidade. Este plano envolverá possivelmente:

a) ELABORAÇÃO DE UM PLANO DE AMOSTRAGEM DE SUJEITOS PARA REPRESENTAR A POPULAÇÃO DADA.

Uma POPULAÇÃO é o conjunto de todos os casos que são conformes a algum conjunto indicado de especificações. Pelas especificações, alunos e matriculados nas escolas do município X, define-se uma população formada por todos os alunos matriculados nas escolas do município X. Pelas especificações, alunos do 1º grau e matriculados nas escolas do município X, define-se uma população formada por todos os alunos do 1º grau matriculados nas escolas do município X. Observa-se que esta segunda população está incluída na primeira, ou seja, é uma sub-população ou um estrato da população. Um ESTRATO é definido por uma ou mais especificações, que dividem uma população em segmentos mutuamente exclusivos.

Para estudar-se as características dos elementos de uma população, pode-se obter informação para cada um destes elementos, ou então obter a informação desejada apenas para alguns elementos e não para todos. Quando se selecionam alguns elementos com a finalidade de conhecer algo a respeito da população de que fazem parte, denomina-se este grupo de elementos de AMOSTRA. O que se espera é que aquilo se verifica para a amostra, seja verdadeiro a respeito da população como um todo. O nível com que a informação colhida da amostra corresponde àquela que seria obtida, caso se tivesse trabalhado com todos os elementos da população, depende, em grande parte, da maneira de escolher esta amostra. Existem situações em que se pode, até certo ponto, controlar as características da amostra. Nestas situações, a maneira de obter a amostra pode aumentar a probabilidade de que os resultados desta, não estarão muito longe dos números reais da população.

O uso de amostras para o estudo de populações se constitui num instrumento valioso à disposição de investigadores, dentre eles dos avaliadores, devido a três características básicas da amostragem: confiabilidade, economia e velocidade. Estas três características são apresentadas por King (1964).

- confiabilidade, porque uma amostra, apropriadamente planejada, conduz a resultados que são precisos dentro de limites especificados;
- economia, porque tais resultados podem ser obtidos com somente uma fração dos esforços requeridos para fazer um estudo completo;
- velocidade, porque o menor esforço requerido para a amostragem pode transformar-se na mais rápida condução dos estudos.

Estas três características básicas da amostragem tornam seu uso extremamente adequado para a Avaliação, uma vez que a Avaliação consiste no fornecimento de informação útil para a tomada de decisão, e que a utilidade da informação se prende também a sua ocasião - fornecimento da informação em tempo hábil, pois a melhor informação é inútil quando chega muito tarde - e a sua eficiência - relação entre quantidade de informação e tempo, custo e pessoal necessário para sua obtenção.

Assim, por questões de economia de tempo e recursos, avaliador se propõe a fornecer informações acerca dos resultados obtidos com uma mudança introduzida num sistema educacional, não precisará realizar mensurações individuais com todos os alunos submetidos a tal mudança. Deste conjunto maior (população), ele seleciona um sub conjunto (amostra), que utilizará para realizar seu experimento. Uma única concessão deve ser feita para se usufruir destas três vantagens possibilitadas pela amostragem - aceitar resultados aproximados. E, realmente, a necessidade de resultados absolutamente exatos não é frequente em Avaliação. O grau de precisão exigido dos resultados varia com o propósito particular do estudo e, quando um resultado mais preciso é necessário, este pode ser obtido, pelo aumento do tamanho da amostra, ou através da utilização de um método de amostragem mais eficiente.

O plano de amostragem deve incluir a determinação do tamanho da amostra. Para isto, são envolvidos procedimentos que exigem conhecimentos estatísticos mais complexos que aqueles que pretende-se apresentar aqui. Neste item, em geral o Avaliador recorre aos serviços de um Estatístico.

Neste passo, incluir-se-ão apenas os métodos de amostragem.

a.1) MÉTODOS DE AMOSTRAGEM

Basicamente, existe uma distinção entre 2 grandes grupos de amostras - amostras não probabilísticas e amostras probabilísticas.

Na amostragem não probabilística, não existe forma de estimar a probabilidade que cada elemento tem de ser incluído na amostra, nem mesmo existe segurança de que todos os elementos tenham alguma oportunidade de serem incluídos.

Na amostragem probabilística, é conhecida a probabilidade que cada elemento tem de entrar para a amostra. A amostragem probabilística permite ao investigador que estime o quanto os resultados baseados em sua amostra tendem a diferir dos que seriam encontrados, caso trabalhasse com toda a população, ou inversamente, permite ao investigador que especifique o tamanho da amostra com a qual deverá trabalhar, para que os resultados que obtiver, não difiram, por mais de uma quantidade especificada, dos que seriam obtidos se trabalhasse com a população total.

O primeiro passo, no processo de amostragem é identificar a população. Se o avaliador quer, por exemplo, obter informações acerca do efeito de um método de ensino na aprendizagem dos alunos de 1ª série do 1º grau do Estado X, então a população será constituída de todos os alunos da 1ª série do 1º grau deste Estado, submetidos a este método. Se está interessado em obter informações acerca do efeito de 2 métodos de ensino na aprendizagem dos alunos da 2ª série do 1º grau da zona rural, a população, no caso, serão os alunos da 2ª série do 1º grau da zona rural, submetidos a estes métodos.

A definição da população alvo é baseada na identificação das variáveis independente, moderadora e de controle, envolvidas no estudo.

Ex.: Hipótese: Na zona urbana, o método de ensino tecnológico tem mais efeito do que o método de ensino tradicional na aprendizagem dos alunos, tanto nas escolas municipais quanto nas estaduais, mas ele tem singularmente mais efeito na aprendizagem dos alunos das escolas estaduais.

Variável dependente: aprendizagem dos alunos.

Variável independente: método de ensino (tecnológico-tradicional).

Variável moderadora: sistema administrativo (municipal-estadual).

Variável de controle: zona.

A população definida a partir destas variáveis será composta de todos os alunos das escolas municipais e estaduais, da zona urbana, submetidos ao método tecnológico ou ao método tradicional. Serão excluídos do estudo, alunos da zona rural, e os da zona urbana de escolas federais, por exemplo.

Definida a população, o passo seguinte é a escolha de um esquema de amostragem que será utilizado quando da extração da amostra propriamente dita. Seguem-se vários destes esquemas. Alguns são relativos à amostragem não probabilística, outros à amostragem probabilística.

- ESQUEMAS DE AMOSTRAGEM NÃO PROBABILÍSTICA -

Os principais esquemas de amostragem não probabilística são:

Amostragem acidental: Neste esquema são considerados apenas os casos que acidentalmente aparecem, e continua-se o processo até que a amostra atinja determinado tamanho. Assim, dentro deste esquema, quando um avaliador deseja fazer generalizações acerca de um método de ensino, por exemplo, ele pode considerar os resultados de alunos em classes, que ele tem disponíveis.

Não existe forma (a não ser através de um estudo paralelo com uma amostra probabilística ou com a população) de avaliar o viés introduzido em tais amostras.

Amostragem de quotas: Acrescenta certa segurança ao esquema anterior de amostragem, pois garante a inclusão, na amostra, de diversos elementos da população, e a consideração destes, nas proporções em que ocorrem na mesma. Assim, como se deseja uma amostra de uma população de alunos da zona urbana e rural, considerando-se que existe uma diferença entre ambos os grupos de alunos, quanto às características que se deseja medir, deve-se garantir a inclusão na amostra, de alunos de ambas as zonas. Caso contrãrio, os resultados refletirão uma imagem enganadora da população total. Na amostragem por quotas, o objetivo é a obtenção de uma amostra que seja uma réplica da população para o qual deseja-se generalizar. Esta preocupação não existe na amostragem acidental.

É importante salientar que a amostragem por quotas deve ser utilizada, apenas, quando se sabe de antemão que, com relação às características que se pretende medir, há uma diferença entre os elementos considerando-se um nível ou outro de determinado fator.

Determinados os estratos, é retirada uma amostra casual dentro de cada um deles.

O importante na amostragem por quotas não é que os vãrios estratos da população sejam amostrados em sua proporção correta, mas que exista número suficiente de casos de cada estrato para que seja possível uma estimativa do valor do estrato da população e que se saiba a proporção que cada estrato constitui na população total. Se tais condições forem satisfeitas, as estimativas dos valores para os vãrios estratos podem ser combinadas, a fim de dar uma estimativa do valor para a população total.

Amostragem intencional: A suposição básica deste esquema de amostragem é que, com um bom julgamento e uma estratégia adequada, pode-se escolher os casos que devem ser incluídos na amostra, e assim chegar a amostras que sejam satisfatórias para as necessidades.

Uma estratégia comum da amostragem intencional é escolher casos julgados como típicos da população de interesse, supondo-se que os erros de julgamento na seleção tenderão a contrabalançar-se. É necessário, no entanto, que exista uma base objetiva para se fazer tais julgamentos.

- ESQUEMAS DE AMOSTRAGEM PROBABILÍSTICA -

É essencial para a validade de um estudo que envolva amostragem, que a amostra seja representativa da população. Uma maneira de assegurar que uma amostra será representativa da população é utilizar, para a sua extração, os procedimentos de amostragem probabilística.

Os esquemas mais comumente utilizados são:

Amostragem casual simples: Conduz à obtenção de AMOSTRA CASUAL SIMPLES, em que todos os elementos da população têm a mesma probabilidade de serem nela incluídos.

Procedimentos:

- produzir um sistema de referência, isto é, uma listagem completa e numerada de todos os elementos que compõem a população. A listagem é completa se nenhum elemento da população fica fora dela, nem entra mais de uma vez. A listagem é numerada se ao seu 1º elemento é atribuído o número 1, ao 2º elemento o número 2, e assim por diante até o enésimo elemento que recebe o número n;
- consultar uma tabela de números aleatórios. Estas tabelas são construídas de modo que os algarismos 0, 1, 2, 3, 4, 5, 6, 7, 8 e 9 são distribuídos casualmente nas linhas e colunas da Tabela com a mesma probabilidade de ocupar uma posição qualquer. Consiste numa série extensiva de números gerados aleatoriamente por computador. A Tabela F.23 é um exemplo de tabela de números aleatórios. Tabelas deste tipo encontram-se publicadas na maioria dos textos estatísticos.

A sequência de leitura da tabela pode ser qualquer uma, de cima para baixo, de baixo para cima, da esquerda para a direita ou vice-versa.

TABELA F.23

TABELA DE NÚMEROS ALEATÓRIOS

86515	90795	66155	66434	56558	12332	94377	57802
69186	03393	42505	99224	88955	53758	91641	18867
41686	42163	85181	38967	33181	72664	53807	00607
86522	47171	88059	89342	67248	09082	12311	90316
72587	93000	89688	78416	27589	99528	14480	50961
52452	42499	33346	83935	79130	90410	45420	77757
76773	97526	27256	66447	25731	37525	16287	66181
04825	82134	80317	75120	45904	75601	70492	10274
87113	84778	45863	24520	19976	04925	07824	76044
84754	57616	38132	64294	15218	49286	89571	42903
75593	51435	73189	64448	31276	70795	33071	96929
73244	61870	28709	38238	76208	76575	53163	58481
23974	14783	17932	66686	64254	57598	26623	91730
32373	05312	94590	22561	70177	03569	21302	17381
59598	56774	08749	43448	28484	16325	62766	31466
91682	12904	29142	65877	64517	31466	02555	52905
87653	98088	75162	97496	59297	79636	74364	16796
79429	66186	59157	95114	16021	30890	21656	93662
85444	39453	67981	49687	36801	38666	50055	11244
85739	44326	91641	40837	93030	03675	18788	91332
84637	76154	14150	07876	41899	69207	66785	87225
59575	32764	91090	66515	05498	51512	16107	52141
81305	58846	69558	41675	88898	23775	30649	86545
29835	35801	23472	22700	39976	21279	36694	85970
32795	54313	39072	16809	22148	60102	18465	87650
37837	12507	54594	30814	23277	99497	11037	63718
58394	96952	12181	11641	83373	14726	23541	25774
74543	46849	95714	70358	95873	94136	83991	77299
77338	59570	29277	82041	06923	01795	77022	17443
21157	50634	16432	44292	20030	38547	67134	95995

Se a população tem até 10 elementos, utiliza-se, para de terminar a constituição da amostra, apenas uma coluna. Nela encontrar-se-ão números de 0 até 9. Neste caso, zero corresponde ao número 10.

Se a população tem até 100 elementos, utilizam-se 2 colu nas. Nelas são encontrados números de 00 até 99. Neste caso, 00 corres ponde ao número 100. Utilizando 3 colunas encontram-se números de 000 até 999. 000 correspondendo ao número 1.000. E assim por diante.

Suponha-se, por exemplo, que se queira extrair uma amostra de 5 elementos de uma população de 100 elementos, e que foi resolvido que a leitura da tabela seria feita de cima para baixo e a partir da esquerda. Então, os números selecionados seriam: 86, 69, 41, 72, 52. Consultando-se um sistema de referência, fica a amostra constituída com aqueles elementos cuja numeração corresponde a estes números.

Amostragem Sistemática: Neste caso, o procedimento é o seguinte:

- produzir uma listagem completa e numerada de todos os elementos que compõem a população;
- fixar o tamanho da amostra;
- definir a constante K, dada pela relação tamanho da população/tamanho da amostra;
- escolher, ao acaso, um número entre 0 e K. Este número definirá o 1º elemento da amostra;
- somar a este 1º número sorteado a constante K, definindo o 2º elemento da amostra. Somar K a este 2º número, definindo o 3º elemento da amostra. E assim por diante.

Se, por exemplo, a amostra fosse 10% da população, a constante K seria 10. Se sorteado, ao acaso, um número entre 0 e 10 e este fosse o 5, então a amostra seria constituída dos elementos que, no sistema de referência, tivessem a seguinte numeração: 5, 15, 25, 35, 45...

Amostragem por Conglomerados: Muitas vezes a construção de um sistema de referência para uma certa unidade de amostragem é impossível. Exemplo: Quer-se estudar o efeito de um método de ensino na aprendizagem dos alunos da 1ª série do 1º grau da cidade X, mas não se tem meios de conseguir uma relação completa destas alunos. Tem-se, porém, a relação completa das classes de 1ª série do 1º grau desta cidade. A classe é uma unidade de amostragem maior do que o aluno, englobando um certo número de alunos. O procedimento a ser seguido é, então, o seguinte:

- obter uma amostra casual simples das classes;
- considerar, como constituinte da amostra, todos os alunos destas classes escolhidas ao acaso.

Observação: Unidade de Análise, Unidade de Amostragem e Amostras.

Após a coleta de dados, estes deverão ser analisados. A análise destes dados envolve a menor unidade independente de informação (Popham, 1975). Em estudos avaliativos é comum utilizar-se, como unidade de análise, não o escore individual do aluno, mas sim o escore médio da classe. Quando o tratamento instrucional envolve trabalho em equipe, a menor unidade independente de informação, ou seja, a unidade de análise, será o escore representativo do grupo. Assim, a unidade de análise não é necessariamente o escore do aluno. Tais unidades de análise já deverão estar especificadas na operacionalização das variáveis envolvidas na hipótese.

Quando a unidade de análise é o escore individual, e a unidade de amostragem não é o aluno, mas sim uma classe intacta (ou natural), utiliza-se a amostragem por conglomerados. Tal procedimento, portanto, se adequa à seleção de grupos pré-formados de indivíduos de uma população.

Ex.: Suponha-se uma população de 5.000 alunos da qual deverá ser extraída uma amostra de 500, e que a unidade de análise é o escore individual. Ao invés de selecionar randomicamente cada um dos 500 alunos, o que implicaria em aplicação de instrumentais em numerosas salas, são selecionados, randomicamente, 20 classes com 25 alunos cada uma delas (Unidade de amostragem: a classe).

Se a unidade de análise é o escore individual, e os mesmos são selecionados aleatoriamente, tem-se uma amostra casual simples. Assim também, quando a unidade de análise é o escore das classes e estas são selecionadas aleatoriamente.

Resumindo:

- . Quando a unidade de amostragem é igual à unidade de análise, tem-se a amostragem casual simples;
- . Quando a unidade de amostragem é uma unidade maior que a unidade de análise, tem-se a amostragem por conglomerados.

Amostragem por estágios múltiplos: Imagine que um avaliador está interessado em estudar o efeito de um método de ensino na aprendizagem dos alunos da 1ª série do 1º grau do Estado X, e que sua unidade de análise será a escora média da classe. Imagine, ainda, que ele não dispõe de um sistema de referências das mesmas. Então o procedimento a ser seguido é extrair uma amostra por estágios múltiplos. Assim:

- escolher uma 1ª unidade de amostragem (no caso, poderia ser o Município);
- construir um sistema de referência para os Municípios do Estado;
- extrair uma amostra casual deles;
- escolher uma 2ª unidade de amostragem (no caso, as classes);
- construir para cada Município sorteado, um sistema de referência para as classes, isto é, uma listagem completa e numerada das mesmas;
- extrair uma amostra casual das classes em cada Município;
- considerar constituinte da amostra final, todas as classes amostradas de todos Municípios amostrados.

OBS.: Se sua unidade de análise fosse o escore individual, o avaliador poderia incluir, na sua amostra, todos os alunos das classes amostradas e trabalhar com os escores individuais, ao invés de com os escores médios das classes.

Amostragem Estratificada: Suponha que um avaliador está interessado em estudar o efeito de um método de ensino na aprendizagem dos alunos do 1º grau do Município X. Neste caso, ele poderia construir um sistema de referência dos alunos de 1º grau do Município, e extrair uma amostra casual dos mesmos. Porém, o avaliador sabe, de antemão, que os alunos da zona rural apresentam certas características diferentes dos alunos da zona urbana, que podem influenciar na aprendizagem. Então, para garantir a inclusão, na amostra, tanto de alunos de escolas rurais quanto de alunos das escolas urbanas, ele divide a população em 2 estratos (alunos da zona urbana e alunos da zona rural) e extrai uma amostra casual simples para cada estrato. Estas sub amostras são, então, reunidas para formar a amostra total.

A divisão da população em estratos pode ser baseada num único critério (por exemplo zona - que daria dois estratos: rural e urbano) ou numa combinação de dois ou mais critérios (exemplo - zona e sexo - que daria 4 estratos: alunos homens da zona rural, alunos homens da zona urbana, alunos mulheres da zona rural, alunos mulheres da zona urbana).

Os procedimentos gerais para extração de amostra estratificada são:

- identificar os critérios para estratificação da população (no exemplo acima: zona e sexo);
- identificar os níveis destes critérios (ex.: urbana e rural - masculino e feminino);
- identificar as combinações destes níveis (ex.: alunos homens da zona rural, alunos homens da zona urbana, alunos mulheres da zona rural, alunos mulheres da zona urbana);
- construir um sistema de referência para cada combinação destes níveis;
- extrair uma amostra casual simples para cada combinação de nível;
- reunir as sub amostras formando a amostra total.

O uso da amostra estratificada permite incluir no estudo variáveis de interesse, e permite controlar melhor a validade interna deste, através do uso de variáveis moderadoras e de controle. É importante salientar que a divisão da população em estratos só deve ser feita quando os critérios de divisão forem relevantes para o que se está medindo. Somente dispo de informações adicionais acerca da população, será possível decidir pela estratificação, que resultará em amostras mais representativas da mesma.

Embora os conceitos básicos de amostragem sejam simples, sua aplicação, geralmente, resulta em procedimentos bastante complexos. Os detalhes de planejamento e execução de um estudo de amostragem seriam mais convenientemente conduzidos, se contassem com assessoria de estatísticos.

b) IDENTIFICAÇÃO DAS VARIÁVEIS NÃO EXPERIMENTAIS QUE POSSAM CONTAMINAR O EXPERIMENTO

Uma afirmativa válida do efeito de um tratamento instrucional particular é impossível de ser feita, a menos que outros fatores, que também influenciem o efeito obtido, tenham sido eliminados ou limitados. Ou seja, só é possível fazer uma afirmativa válida acerca do efeito de uma variável independente sobre uma variável dependente, quando as demais variáveis, que possam influenciar a variável dependente, tenham sido eliminadas ou limitadas. Quando num estudo seus resultados são uma função do programa ou abordagem que está sendo avaliada, ao invés de resultados de outros fatores estranhos, diz-se que o estudo tem VALIDADE INTERNA. Quando num estudo seus resultados são generalizáveis, para um contexto maior do que aquele estudado, diz-se que o estudo tem VALIDADE EXTERNA.

VALIDADE INTERNA tem relação com a legitimidade da interpretação dos resultados de um experimento. Campbell e Stanley (1963) identificaram 8 tipos de variáveis estranhas que, se não controladas, podem ser fonte de invalidade interna.

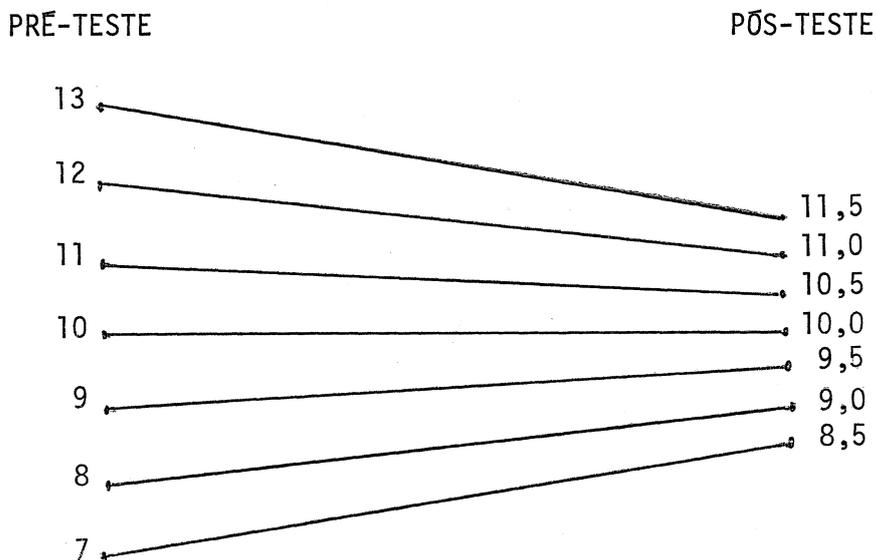
- História: efeitos de eventos específicos que ocorrem durante o tratamento instrucional, em adição aos efeitos do próprio tratamento. Isto, quando este se estende além de um período de tempo. Ex.: se numa escola elementar inicia-se um curso de conversação inglesa e, paralelamente, a rádio local emite um curso da língua, ficará obscuro se a causa de qualquer melhoria no domínio da mesma, pelos alunos, foi devida ao curso iniciado na escola, ou a programação de rádio.

- Maturação: efeito de processo interior dos respondentes, operando como uma função da passagem do tempo (incluindo crescimento biológico, psicológico ou sociológico), e que pode conduzir aos resultados desejados tanto quanto o tratamento instrucional. Ex.: alguém que 6 meses mais tarde tem um desempenho melhor num teste de interpretação de textos, o teve devido ao tratamento instrucional que recebeu, ou porque está 6 meses mais velho?

- Testagem: efeito sobre os escores num 2º teste, devido ao fato de ter-se submetido a uma primeira testagem. Estudantes que realizam o mesmo teste duas vezes (pré e pós-teste), algumas vezes desempenham-se diferentemente no pós-teste como resultado da pré-testagem.

- Instrumentação: efeito modificador sobre as medidas devido a mudanças nos instrumentais, nos observadores, ou naqueles que corrigirão os testes. Se, durante um estudo de Avaliação são mudados os instrumentais de medida, qualquer mudança aparente no desempenho dos alunos deverá ser mais diretamente atribuída a esta modificação instrumental, do que ao fenômeno educacional em Avaliação. Ex.: imagine que o pré-teste é mais difícil que o pós-teste. Então, qualquer diferença no desempenho dos alunos, nestas duas ocasiões, deverá incluir o efeito desta diferença de dificuldade dos dois instrumentais.

- Regressão Estatística: efeito da correlação imperfeita entre o pré-teste e o pós-teste, que ocorre quando grupos são selecionados com base em seus escores extremos. Suponha que foram aplicados um pré-teste e um pós teste num grupo de alunos e que tais instrumentais não tinham correlação perfeita. Então, os resultados que seriam obtidos seriam do tipo:



Ou seja, os escores extremos (tanto os mais altos quanto os mais baixos) tenderiam para a média do grupo. Quanto menor a correlação entre pré e pós-teste, maior a regressão em direção a média. Efeitos de regressão são consequências inevitáveis de correlação imperfeita teste-reteste para grupos selecionados por sua extremidade. Como correlações per

feitas são ideais, dificilmente atingíveis, a solução não é selecionar grupos extremos, pois quando estes são utilizados, os escores no pós-teste ao invés de evidenciarem efeito, apenas, do tratamento instrucional, evidenciarão, também, efeito da regressão estatística, em direção à média do grupo. Popham (1975) ilustra este efeito, dizendo que no passado muitos programas educacionais foram sustentados por ele. Era identificado um grupo de alunos a pelo menos um ou dois desvios padrão abaixo da média, num teste de aprendizagem. Os mesmos eram submetidos ao tratamento, aplicava-se um teste subsequente no qual o desempenho dos alunos, obviamente, era mais alto. A conclusão a que se chegava era de que o programa fora efetivo. No entanto, devido ao efeito da regressão estatística, os alunos poderiam ter demonstrado o mesmo progresso sem nenhum tratamento.

- Seleção: efeito que aparece em estudos de Avaliação, quando dois ou mais grupos estão sendo utilizados, devido ao recrutamento diferencial das pessoas constituindo os grupos. Ex.: Se estudantes mais motivados são designados para o tratamento X (voluntários), enquanto estudantes menos motivados são designados para o tratamento Y, o tratamento X poderá parecer mais efetivo, simplesmente devido à diferenças iniciais dentre os grupos.

- Mortalidade Experimental: efeito provocado pela perda diferente de respondentes dentre os diferentes grupos de comparação, utilizados no experimento. Dois ou mais grupos que no início do experimento poderiam ser considerados equivalentes, já não o seriam no final do mesmo, se houvesse, por exemplo, num grupo a desistência de seus melhores alunos. Medições ao final do experimento expressariam, também, o efeito desta desigualdade.

- Interação de Fatores: efeito provocado pela combinação de 2 fontes de invalidade. Por exemplo, um efeito poderia ser o provocado por uma interação Seleção X Maturação, ou Seleção X História. Com relação ao efeito da interação Seleção X Maturação: se os grupos utilizados, no experimento, não são equalizados com relação à variável idade, poderá haver, em adição a este problema de seleção, um problema de maturação, uma vez que, em algumas idades os indivíduos têm uma "taxa" de maturação maior que, em outra. Assim, os dados colhidos de observações em ambos os grupos, poderiam refletir mais uma diferença do efeito da interação Seleção X Maturação, por exemplo, do que uma diferença de efeito dos tratamentos.

VALIDADE EXTERNA: Em pesquisa o objetivo maior é chegar a GENERALIZAÇÕES. Já em Avaliação, este é um aspecto mais secundário. O avaliador busca, apenas, de modo limitado a generalização dos resultados dos seus estudos. Ele preocupa-se com generalizações em escalas muito mais modestas, Enquanto para a Pesquisa o ideal é a busca de resultados generalizáveis, para qualquer tempo e qualquer lugar, já para a Avaliação a busca pela generalização, se ela existe, é mais comumente em função de populações de estudantes em períodos subsequentes, considerando aquele contexto restrito em que se verificou o estudo. Porém, sempre que possível, as seguintes fontes de validade externa devem ser consideradas num estudo de Avaliação.

- Efeito Reativo da Testagem: efeito em que um pré-teste poderia aumentar ou diminuir a sensibilidade ou conformidade de um respondente, com relação à variável experimental, e então fazer com que os resultados obtidos por uma população pré-testada não fossem representativos dos efeitos da variável experimental, para uma população não pré-testada da qual os respondentes foram selecionados. Isto é, o tratamento só teria tal efeito em grupos pré-testados. Quando o pré-teste fosse suprimido, os efeitos não seriam os mesmos.

- Efeitos de Interação do Viés de Seleção e Tratamento: se características particulares dos grupos em estudo causaram os resultados observados com relação ao efeito de X, então ele não poderá ser generalizado para grupos com características diferentes. Ex.: um estudo feito com alunos da zona urbana pode ter seus resultados não generalizáveis para alunos da zona rural se existir algum aspecto único à população urbana que possa ter contribuído para os efeitos do experimento.

- Efeitos Reativos dos Arranjos Experimentais: os arranjos para o experimento podem torná-lo tão artificial, que os resultados obtidos, com ele, dificilmente serão generalizáveis para situações não experimentais. Além disso, o conhecimento dos alunos de que estão participando de um experimento, pode afetar os resultados. E, em situações experimentais onde é patente o artificialismo, é muito difícil evitar tal conhecimento.

- Interferência de Múltiplos Tratamentos: quando os mesmos participantes são submetidos a um número de tratamentos, simultaneamente, alguns deles experimental e outros não, os tratamentos sempre interagem e o efeito causado por esta combinação de tratamentos, pode produzir efeitos que não serão reproduzidos em situações nas quais somente o tratamento experimental for aplicado.

b.1) TÉCNICAS PARA CONTROLE DE VARIÁVEIS NÃO EXPERIMENTAIS QUE POSSAM CONTAMINAR O EXPERIMENTO.

Duas são as técnicas mais utilizadas em "designs" de Avaliação para controlar o efeito de variáveis estranhas, que podem contaminar o experimento:

- Inclusão de um grupo de controle ou de um grupo de comparação.
- Equalização dos Grupos-Controle da Seleção.

- Inclusão de um grupo de controle ou de um grupo de comparação -

Grupo de controle é um grupo de sujeitos cuja seleção e história são tanto quanto possível idênticas ao do grupo experimental, de modo que a única diferença entre ambos seja que o grupo experimental recebe o tratamento instrucional e o grupo de controle não o recebe. Se dois grupos existem em condições tais que a única diferença entre ambos é o tratamento que um deles recebe, parece lógico que qualquer diferença dentre os resultados obtidos para cada um deles, ao final do experimento, seja devido a este tratamento. Empregando apropriadamente grupos de controle, alguém pode eliminar a maior parte das 8 fontes de invalidade interna.

Tais estudos envolvendo Grupo de Controle X Grupo de Tratamento são muito pouco frequentes para Avaliadores Educacionais. Isto porque Avaliação Educacional diz respeito ao fornecimento de informações úteis para decisões acerca de qual seria a melhor alternativa dentre as possíveis. E, geralmente, não existe necessidade de escolha entre um tratamento e nenhum tratamento, mas sim o que existe é sempre a necessidade de escolha entre tratamentos alternativos. Desse modo, os estudos de Avaliação, em geral, devem envolver grupos de comparação. Grupos de com

paração são grupos de sujeitos cuja seleção e experiência são tanto quanto possível idênticos entre si, de modo que a única diferença seja que enquanto um recebe um tratamento instrucional determinado, os outros recebem tratamentos alternativos a este.

Utilizando-se um grupo de controle ou um grupo de comparação, a fim de que sejam neutralizados os efeitos de variáveis estranhas, todas as condições que não as variáveis independentes e moderadoras, devem ser mantidas constantes para ambos os grupos.

- Equalização dos Grupos - Controle da Seleção -

Segundo Tuckman (1972) a equalização dos grupos pode ser feita através de:

- a) designação aleatória dos sujeitos para os grupos de comparação;
- b) acasalamento aos pares;
- a) designação aleatória dos sujeitos para os grupos de comparação

A fim de constituir grupos de comparação essencialmente equivalentes (controlando, então, a seleção), a melhor técnica é a designação randômica dos sujeitos, ou seja, é selecionar ao acaso, a partir de um conjunto de sujeitos, aqueles que indistintamente constituirão um ou outro grupo de comparação. A randomização é a técnica mais efetiva para equalizar grupos. Mesmo se os grupos parecem ser diferentes em um ou outros fatores, haverá tipicamente outros fatores não discernidos que cancelam aquela diferença aparente. A randomização controla as variáveis de seleção sem antes ter que identificá-las. Através do uso desta técnica, o avaliador pode assumir a equalização dos grupos para qualquer propósito prático, pois, é suposto que a chance reduz a probabilidade que um tipo de pessoa mais do que outro, esteja num grupo de comparação ou em outro.

O exemplo que segue implica na utilização de dois grupos de comparação designados aleatoriamente.

Um avaliador vai testar dois métodos (A e B) para o ensino de Matemática.

variável independente: método de ensino, com 2 níveis A e B.

variável dependente: aprendizagem dos alunos.

Cada método será, então, empregado em um grupo de alunos. Suponha que existam 50 alunos, e que serão formados 2 grupos de 25. A equalização dos grupos ficará garantida, se o processo de designação dos sujeitos aos grupos for aleatória. Para isto, bastará enumerar cada aluno com um número de 01 a 50, e em seguida, sortear aqueles que constituirão um ou outro grupo, ou então escalar os sujeitos para os mesmos, através do uso de uma tabela de números aleatórios. Formados os grupos, cada um deles passará por um período de instrução sob um ou outro método de ensino. Após este período, o desempenho dos alunos será medido através da aplicação de um teste de aprendizagem. O avaliador analisará, então, os dados para informar acerca de qual método deverá ser usado para o ensino de Matemática.

Haverá, é claro, numerosas situações nas quais o avaliador não poderá designar aleatoriamente os sujeitos para os grupos. É muito comum, em Educação, o uso de classes intactas, ou seja, classes formadas pela própria administração da escola, anteriormente ao planejamento do experimento. Neste caso, é preferível trabalhar com grupos de algum modo equivalentes do que com aqueles muito desiguais. Dispondo-se de classes intactas para serem utilizadas, deve-se tentar pelo menos a designação aleatória de cada classe, para cada tratamento instrucional que estará sendo comparado.

b) Acasalamento aos pares

É utilizado para se obter grupos equivalentes. É uma técnica alternativa para a designação aleatória, descrita anteriormente.

O uso desta técnica consiste em identificar, no conjunto de sujeitos, os pares de pessoas que são mais equivalentes com respeito a(s) variável (eis) especificada(s) (estas variáveis são aquelas que poderão ser fontes de invalidez interna para o estudo. Em geral são: sexo, ida

de, Q.I., status s̄ocio-econ̄mico, ou escores no pr̄e-teste que mede a variável dependente), e em seguida, designar aleatoriamente cada elemento de cada par para um dos grupos de comparaç̄o. Os grupos resultantes podem ser considerados razoavelmente iguais, obtendo-se, assim, o controle das variáveis de seleç̄o.

Tal acasalamento mais designaç̄o aleat̄ria geralmente produz grande precis̄o nos resultados.

Os passos gerais desta t̄cnica s̄o, portanto:

- identificaç̄o da(s) variável(eis) que serã(ão) controlada(s);
- acasalamento dos pares a partir de mensuraç̄es destas variáveis;
- designaç̄o aleat̄ria de cada sujeito de cada par, para cada um dos grupos de comparaç̄o.

Al̄m destas duas t̄nicas - inclus̄o de um grupo de comparaç̄o ou de um grupo de controle - equalizaç̄o dos grupos, que s̄o as mais utilizadas em "designs" de Avaliaç̄o, existem outras que permitem, tam b̄m, o controle de variáveis estranhas que podem contaminar o experimento.

O controle da variável hist̄ria, possibilitado, principalmente, pela inclus̄o do grupo de controle (ou de comparaç̄o), ẽ umentado atrav̄s de:

a) M̄todo da Remoç̄o

Atrav̄s deste, qualquer fator estranho deverã ser removido do experimento, tanto na condiç̄o de controle quanto na condiç̄o experimental. Características f̄sicas estranhas do ambiente (relativas a som, luz, ventilaç̄o, etc.) deverã ser evitadas.

b) Método da Constância

As experiências, que não aquelas resultantes da manipulação da variável independente, deverão ser constantes para os diferentes grupos. As tarefas, procedimentos, experiências não exclusivas do tratamento, deverão ser idênticas para os grupos. Tempo de exposição ao experimento, número de testes e retestes aplicados, etc., são exemplos de variáveis que devem ser mantidas constantes.

O controle da instrumentação pode ser conseguido através de:

- utilização de instrumentos cuja precisão é assegurada;
- uso do mesmo instrumental ou instrumental paralelo para o pré e pós-teste;
- utilização de instrumentos com validade estabelecida;
- estabelecimento de um sistema de correção a ser utilizado.

Além das técnicas já descritas, existem outras que podem auxiliar os avaliadores no planejamento de estudos de Avaliação.

- a pequena duração de experimentos controla a maturação e história;
- a utilização de alunos médios, numa distribuição, controla a regressão estatística;
- o emprego de medidas não obstrusivas controla a testagem.

c) SELEÇÃO DE UM "DESIGN" DE COLETA DE DADOS

No plano experimental para coleta de dados deverá ser apresentado o "DESIGN" DE COLETA selecionado. Os designs característicos da experimentação tradicional em Pesquisa, vem sendo utilizados em Avaliação. Os avaliadores precisam conhecer os mais comumente utilizados no seu campo, e ter consciência de como cada um deles atua em relação ao controle das fontes de validade de um estudo.

Os "designs" que seguem são alguns daqueles apresentados e analisados por Campbell e Stanley (1963), que identificam para cada um deles as suas vantagens e desvantagens, no controle das fontes de validade interna e externa de um experimento, além de características especiais que garantem seu uso em diferentes tipos de situações:

- uns são mais fáceis e baratos de aplicar;
- uns são grosseiros e obscuros (não controlando várias fontes de validade) apropriados apenas para certas avaliações formativas, mas impróprias para avaliações somativas;
- uns são adequados quando é impossível a designação randômica dos sujeitos;
- uns são adequados quando grupos de controle não podem ser constituídos.

Na esquematização dos "designs" de coleta de dados as seguintes convenções serão utilizadas:

- R: Designação aleatória dos sujeitos aos grupos, a fim de obter a equalização dos mesmos;
- X: Tratamento Instrucional (efeito do qual será medido);
- X': Tratamento Alternativo;
- T₁: Algum processo de medida anterior a X;
- T₂: Algum processo de medida posterior a X.

A disposição destes elementos, em linha, da esquerda para a direita, indica a sequência dos eventos no tempo. Os elementos mais à esquerda ocorrerão primeiro e assim por diante. Dois elementos em duas linhas paralelas serão simultâneos, quando localizados na mesma vertical.

Os "designs" de coleta de dados, em função do maior ou menor controle das fontes de validade interna de um experimento, são classificados em 3 grandes categorias:

- . "designs" pré-experimentais ——— nenhum ou mínimo controle sobre as fontes de validade interna;
- . "designs" experimentais ——— controle total das fontes de validade interna;
- . "designs" quase-experimentais ——— controle parcial sobre as fontes de validade interna.

Uma tabela que apresenta para cada "design" as fontes de validade que ele seguramente controla (+), as fontes de validade que ele não controla (-), aquelas para as quais existe a possibilidade de controle, mas esta não é assegurada (?), fatores que não lhe são relevantes (), e utilidade do "design" para a Avaliação é mostrada ao final da apresentação de todos os "designs", num resumo de suas vantagens e desvantagens no controle das fontes de validade.

C.1) "DESIGNS" PRÉ EXPERIMENTAIS

Tais "designs" não têm nenhum ou tem o mínimo controle sobre as fontes de validade interna. Os tipos, aqui, descritos, não usam grupo de comparação, não envolvendo, portanto, comparação entre tratamentos alternativos. São mais adequados para Avaliação Formativa do que Avaliação Somativa.

= "design" um: um grupo pós-teste

X	T ₂
---	----------------

a) Procedimentos:

- . expor um grupo de alunos ao tratamento por um período de tempo;
- . efetuar mensuração após o tratamento;
- . comparar os resultados com os de outros eventos ocasionalmente lembrados ou observados.

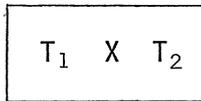
b) Controle das fontes de validade

Não há nenhum controle, não há validade interna. Este design tem a maioria das fraquezas dos "designs" subsequentes. Variáveis estranhas como história, maturação, mortalidade, seleção (VALIDADE INTERNA), podem obscurecer a interpretabilidade dos dados obtidos com este "design", bem como interação seleção X tratamento (VALIDADE EXTERNA).

c) Utilidade

Com todas as suas fraquezas, este "design" pode ser utilizado nos 1ºs estágios de Avaliação Formativa de um Projeto, assegurando grosseiramente o que acontece aos aprendizes depois que experienciam um certo tipo de tratamento. Detetando-se objetivos não atingidos pelos alunos, identificam-se pontos falhos no programa a serem reformulados. Não deve ser utilizado para decidir sobre a adoção de um programa (Avaliação Somativa).

- "design" dois: um grupo pré-teste pós-teste



a) Procedimentos

- . efetuar mensuração antes do tratamento;
- . expor um grupo de alunos ao tratamento instrucional por um período de tempo;
- . efetuar mensuração após o tratamento;
- . realizar a análise estatística adequada.

b) Controle das fontes de validade

Os defeitos de tal "design" são numerosos, mas ele é particularmente suscetível às seguintes variáveis estranhas: história, matu

ração, testagem, instrumentação, interação seleção com outros fatores (VALIDADE INTERNA); interação seleção e tratamento, interação pré-teste e tratamento (VALIDADE EXTERNA).

VALIDADE INTERNA

- História:
- (-) . Se T_1 e T_2 ocorrem em dias diferentes, então os eventos ocorridos paralelamente ao tratamento instrucional, podem ter causado a mudança;
 - . Quanto maior o intervalo de tempo entre T_1 e T_2 , mais plausível é o efeito desta variável;
 - . Em estudos acerca de métodos de ensino é impossível ocorrer o "isolamento experimental", tão relevante ao controle da variável história.
- Maturação:
- (-) . Se o intervalo de tempo entre T_1 e T_2 é longo, a diferença dos resultados obtidos, nestas duas ocasiões, pode ser efeito do processo de maturação dos alunos (maturação biológica ou psicológica), mais do que do tratamento.
- Testagem:
- (-) . Se o mesmo teste for usado duas vezes, pode ocorrer o efeito da testagem, ou seja, a medida em T_2 refletirá, também, o fato do aluno ter experienciado T_1 , anteriormente.
- Instrumentação:
- (-) . Se for usado um teste em T_1 e outro em T_2 , pode ocorrer o efeito da instrumentação;
 - . Se a medida é obtida por observação (menos frequentes em estudos de Avaliação), diferentes observadores, nas duas ocasiões, poderão produzir a diferença dos resultados. Ou, então, um mesmo observador trabalhando nas duas ocasiões, poderá, pela prática, ser mais hábil na segunda ocasião, produzindo então diferença nos resultados.

Interação Seleção e outros fatores: . Embora a seleção seja controlada pelo pré-teste, o efeito da interação seleção e outro fator não controlado pode ocorrer, uma vez que não houve seleção randômica dos sujeitos, e que não existe grupo de comparação.

(-)

Regressão Estatística: . Uma vez que não houve seleção aleatória dos sujeitos, pode ocorrer que um grupo de escores extremos esteja sendo comparado no pré-teste e no pós-teste, o que ocasionaria o efeito da regressão estatística.

(?)

Seleção e Moralidade: . São controladas se o mesmo grupo passa por T_1 e T_2 .

(+) (+)

VALIDADE EXTERNA

Interação seleção e tratamento: . Não tendo havido seleção aleatória dos alunos, pode ocorrer que os resultados observados, com relação ao efeito de X, sejam devidos às características particulares do grupo, e, portanto, não generalizáveis.

(-)

Interação pré-teste e tratamento: . Os resultados podem não ser generalizáveis para grupos de alunos não pré-testados.

(-)

Reação aos procedimentos experimentais: . Este "design" representa um esquema de uma situação natural em Avaliação. Porém, ainda assim, o artificialismo pode ser criado, produzindo resultados dificilmente generalizáveis para outras ocasiões.

(?)

c) Utilidade

O design é particularmente útil para Avaliação Formativa. Para a Avaliação Somativa, entretanto, embora variações deste, sejam frequentemente usadas (Avaliação Absoluta), os "designs" com grupos de comparação são preferíveis.

Observa-se que, com este "design":

- . se o intervalo de tempo entre pré-teste e pós-teste é pequeno, controla-se história e maturação;
- . se o pré-teste é de um tipo tal que não influencia o desempenho nos pós-teste, então a ameaça de testagem pode ser evitada;
- . usando-se dois testes idênticos, a instrumentação é controlada;
- . se a inclusão dos alunos, no estudo, é feita de tal modo a ser o grupo representativo da população, então, o efeito de seleção não existe.

Então, para alguns propósitos e com alguns cuidados, tal "design" pode ser utilizado com suficiente respeito.

C.2) "DESIGNS" EXPERIMENTAIS

São aqueles que têm controle total das fontes de validade interna.

Tais "designs" são, teoricamente, perfeitos. Apresentam, entretanto, restrições quanto à prática. Não é fácil encontrar situações de Avaliação em que seu uso seja possível.

Todos eles envolvem grupo de controle, bem como designação aleatória dos sujeitos aos grupos.

- "design" três: grupo de controle pré-teste pós-teste.

R	T ₁	X	T ₂
R	T ₃	X'	T ₄

a) Procedimentos

- . designar aleatoriamente os sujeitos para os grupos;
- . efetuar a mensuração antes do tratamento;
- . expor o grupo experimental ao tratamento e o grupo de comparação ao tratamento alternativo, por um período de tempo;

- . efetuar a mensuração após o tratamento;
- . realizar a análise estatística apropriada.

b) Controle das fontes de validade

Este é um "design" forte que, quando cuidadosamente utilizado, controla todas as fontes de validade interna.

VALIDADE INTERNA

História: É controlada uma vez que os eventos históricos que possam produzir uma diferença entre T_2 e T_1 , também podem produzi-la entre T_4 e T_3 . No entanto, este "design" pede por simultaneidade entre as seções dos grupos de comparação. E, se realmente elas ocorrem em paralelo, diferentes professores, salas de aula, horário, etc. precisam ser utilizados. A solução é designar aleatoriamente para cada grupo, cada um destes elementos.
(+)

Maturação e Controlados pelo uso dos grupos alternativos. Se houver teste: efeito de maturação ou teste, estes serão idênticos para os diferentes grupos, o que não enviesará os resultados.
(+) (+)

Instrumentação: É controlada facilmente, principalmente, se a resposta do aluno é obtida por um instrumento fixo, tal como um teste de aprendizagem impresso. Sendo utilizados observadores, estes deverão ser designados aleatoriamente para os grupos. Havendo apenas um observador para dois grupos, ele deverá desconhecer qual é o grupo de tratamento e qual o de comparação.
(+)

Regressão Estatística: É controlada não importa quão extremos (com relação aos escores da população) sejam os escores do pré-teste, uma
(+)

vez que ambos os grupos serão designados aleatoriamente do mesmo conjunto extremo. Neste caso, se houver regressão em um deles, haverá também no outro.

Seleção: É controlada uma vez que a designação aleatória assegura a equalização dos grupos. A equalização é melhor assegurada quanto maior for o número de alunos designados aleatoriamente.

(+)

Mortalidade: A diferença de presenças nas três ocasiões (pré-teste, tratamento, pós-teste) é controlada pela existência do pré e pós-testes. Deverão ser utilizados para análise, dados de todos os alunos que fizerem ambos pré e pós-testes. Os escores dos alunos que fizerem apenas um destes, deverão ser excluídos. Este procedimento, embora atenuo o efeito do tratamento, evita que seja introduzido um viés na amostra. Um exemplo pode demonstrar a inadequação de não proceder de tal modo. No exemplo abaixo, dois grupos de alunos foram submetidos aos tratamentos X e X' respectivamente. Em cada grupo 5 alunos fizeram o pré-teste e 4 fizeram o pós-teste (uma ausência em cada grupo).

		GRUPO 1	Pré-teste	Pós-teste
CASO 1	X		10	<input type="text"/>
			9	9
			8 $\bar{x} = 8$	8 $\bar{x} = 7,5$
			7	7
			6	6
	X'	GRUPO 2	10	10
			9	9
			8 $\bar{x} = 8$	8 $\bar{x} = 8,5$
			7	8
			6	<input type="text"/>

CASO 2	GRUPO 1	Pré-teste	Pós-Teste
		10	<input type="checkbox"/>
		9	9
		8 $\bar{x} = 7,5$	8 $\bar{x} = 7,5$
		7	7
	6	6	
	GRUPO 2	10	10
	9	9	
	8 $\bar{x} = 8,5$	8 $\bar{x} = 8,5$	
	7	7	
6	<input type="checkbox"/>		

Note que cada aluno tirou a mesma nota no pré e pós-teste, o que mostra que foi nulo o efeito de cada tratamento, em cada grupo. Porém, em função de uma análise mal estruturada, isto poderia parecer não verdadeiro. No 1º caso, quando não foram eliminados os escores dos alunos que não fizeram ambos pré e pós-testes, a diferença das médias do grupo 1 foi negativa ($7,5 - 8,0 = -0,5$), e a diferença das médias do grupo 2 foi positiva ($8,5 - 8,0 = 0,5$). Isto poderia parecer que o efeito do tratamento alternativo (X') foi melhor que o de X , o que não foi verdade. Considerando-se apenas os escores daqueles que realizaram os 2 testes, a ausência de efeito do tratamento se evidencia.

GRUPO 1 ($7,5 - 7,5 = 0,0$)

GRUPO 2 ($8,5 - 8,5 = 0,0$)

Interação seleção e outros fatores:

Controlada uma vez que a designação aleatória equaliza os grupos. Qualquer efeito de interação seria idêntico, para os dois grupos.

(+)

VALIDADE EXTERNA

Interação teste e tratamento:

(-)

Neste "design", se o teste é reativo, sua influência pode confundir os resultados. Medidas de afetividade serão mais frequentemente reativas que medidas em outros domínios. Mesmo medidas cognitivas podem facilitar a interação bem sucedida dos alunos com uma sequência instrucional. Porém, o efeito do pré-teste sobre o tratamento, restringindo a validade externa, é função da extensão na qual tais medidas são características do universo para o qual alguém quer generalizar. E, em Avaliação interessa-se em generalizar para situações nas quais a testagem é fenômeno regular. Neste caso, nenhum efeito indesejável da interação teste e tratamento estaria presente.

Interação seleção e tratamento:

(?)

Embora este "design" controle os efeitos de seleção, explicando diferenças entre os grupos de comparação, existe ainda a possibilidade de que os efeitos demonstrados se restrinjam somente àquela população única da qual os grupos foram selecionados. Em Avaliação, este efeito não é tão relevante quanto para a Pesquisa pois generalização em Avaliação é vista num sentido muito mais restrito.

Efeitos Reativos dos arranjos experimentais:

(?)

Este "design" é altamente controlador das fontes de validade interna. Apresenta um esquema que dificilmente refletirá uma situação natural. Aleatoriedade, designação aleatória dos alunos, pré-testagem, etc. podem criar expectativas do não usual.

c) Utilidade

Este é um "design" muito forte. Porém, só é possível quando a designação aleatória dos grupos é permitida, o que nem sempre ocorre na prática, em estudos de Avaliação. É próprio para Avaliação Somativa.

Variações deste "design" podem ser empregadas, ampliando-se o número de grupos de comparação.

R	T ₁	X	T ₂
R	T ₃	X'	T ₄
R	T ₅	X''	T ₆
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮

- "design" quatro: grupo de controle somente pós-teste

R	X	T ₂
R	X'	T ₄

a) Procedimentos

- . designar aleatoriamente os sujeitos para os grupos de comparação;
- . expor o grupo experimental ao tratamento experimental e o grupo de comparação ao tratamento alternativo por um período de tempo;
- . efetuar a mensuração após o tratamento;
- . realizar a análise estatística apropriada.

Este "design" pode ser considerado o design três, sem pré-teste. As técnicas de aleatoriedade garantem a equalização inicial dos grupos, o que possibilita a exclusão do pré-teste.

b) Controle das fontes de validade

Este "design" controla todas as fontes de validade interna, como o "design" três, mas, mais do que este, controla os efeitos da interação teste-tratamento, pela exclusão do pré-teste.

c) Utilidade

Por envolver aleatoriedade é um "design" que nem sempre é possível de ser utilizado. É apropriado aquelas situações nas quais o "design" trêz poderia ser usado, além daquelas que, permitindo a aleatoriedade, não permitem o pré-teste, como é o caso particularmente frequente nos primeiros anos escolares de experimentos com métodos para introdução inicial de conteúdo inteiramente novo, para os quais pré-testes, no senso comum, são impossíveis. Deve ser utilizado, também, nos casos em que os pré-testes pareçam reativos.

Variações deste "design" podem ser empregadas, ampliando-se o número dos grupos de comparação.

Observa-se que os "designs" trêz e quatro conforme esquematizados, envolvem o estudo do efeito de UM SIMPLES FATOR, com um número de níveis igual ou maior que dois. Constituem-se em "DESIGNS" COMPLETAMENTE RANDOMIZADOS, isto é, "designs" nos quais todos os tratamentos são designados às unidades experimentais de uma maneira completamente aleatória. Na prática, isto é geralmente feito do seguinte modo: a) um grupo de sujeitos é dividido, utilizando o método aleatório, em tantos grupos quanto são os níveis da variável independente; b) um tratamento diferente é aplicado a cada grupo; c) mensurações são obtidas de cada sujeito. Um esquema para orientar a coleta de dados poderia ser apresentado na Tabela F.24.

TABELA F.24
UM ESQUEMA DE COLETA DE DADOS PARA
"DESIGNS" DE UM SIMPLES FATOR

FATOR "Y"		
Nível 1	Nível 2	Nível 3
Mensuração 1	Mensuração 1	Mensuração 1
Mensuração 2	Mensuração 2	Mensuração 2
⋮	⋮	⋮
Mensuração n	Mensuração n	Mensuração n

Neste caso estar-se-ia com um experimento com um simples fator, com três níveis. Portanto, seriam constituídos aleatoriamente três grupos de alunos, cada um submetido a um dos três tratamentos. As mensurações representam os escores de cada aluno no teste. No caso do "design" três, tal esquema seria preenchido duas vezes (Pré-Teste e Pós-Teste).

Existem situações, porém, onde há uma base para classificação dos sujeitos em sub grupos (também chamados blocos) dentro dos grupos de tratamento. Isto é feito com o objetivo de se conseguir maior homogeneidade na ambientação em que ocorre o experimento. A aleatoriedade é então, restrita dentro dos blocos - "DESIGN" COM UMA RESTRIÇÃO DE ALEATORIEDADE.

Exemplo:

Se, para estudar o efeito de três métodos de ensino na aprendizagem dos alunos (TRADICIONAL - TECNOLÓGICO COM RÁDIO - TECNOLÓGICO COM TELEVISÃO) dispuser-se, para realização do Experimento, de 3 escolas, cada uma delas com 3 classes, num total portanto de 9 classes, é necessário evitar que o fato de se estar trabalhando em escolas distintas enviesasse os resultados. Um dos métodos pode parecer melhor, não devido a sua qualidade intrínseca, mas porque foi executado nas classes de uma escola melhor. Qualquer variação entre os métodos pode refletir variação entre escolas. Então, esta variação precisa ser removida. O "design" adequado para isto é o de BLOCOS RANDOMIZADOS. As classes são separadas em 3 blocos, segundo a escola que pertencem, e então cada uma delas é designada aleatoriamente para cada tratamento.

Ex.: Classes: 1 2 3 4 5 6 7 8 9
 └───┬───┘ └───┬───┘ └───┬───┘
 escola 1 escola 2 escola

	TRADICIONAL		RÁDIO		TELEVISÃO				
Classes:	7	1	4	8	3	6	9	2	5

Deste modo, cada método é aplicado em 3 classes, cada uma delas pertencente a um tipo de escola.

Existem outros "designs" com um simples fator, que se constituem em modificações, com relação ao processo de aleatoriedade dos "designs" três e quatro. Não serão estudados, aqui, por serem menos úteis à Avaliação. É o caso dos "designs" BLOCOS RANDOMIZADOS INCOMPLETOS, QUADRADO LATINO, QUADRADO GRECO-LATINO, cuja referência pode ser encontrada em Hicks (1964).

- "design" cinco: "design" fatorial

"Designs" Fatoriais são modificações dos "designs" verdadeiramente experimentais, descritos anteriormente. Tais modificações consistem no estudo simultâneo do efeito de duas ou mais variáveis independentes. "Designs" Fatoriais são, portanto, experimentos nos quais cada tratamento é uma combinação dos níveis de dois ou mais fatores cujo efeito na variável medida é de interesse. Se todas as combinações possíveis dos níveis são estudadas, o experimento é um FATORIAL COMPLETO. Se, num experimento está se estudando o efeito sobre uma variável dependente de duas variáveis independentes: a primeira delas com três níveis e a segunda com dois níveis, então teremos $3 \times 2 = 6$ tratamentos, resultantes das combinações destes níveis. Um exemplo é apresentado na Tabela F.25.

TABELA F.25

COMBINAÇÃO DE NÍVEIS NUM "DESIGN"
FATORIAL COMPLETO

		MÉTODO DE ENSINO		
		MÉTODO TRADICIONAL	MÉTODO TECNOLÓGICO (RA)	MÉTODO TECNOLÓGICO (TV)
LOCALIZAÇÃO	ZONA URBANA	1º tratamento	2º tratamento	3º tratamento
	ZONA RURAL	4º tratamento	5º tratamento	6º tratamento

Neste caso, teriam que ser constituídos 6 grupos, cada um deles recebendo um dos tratamentos abaixo:

- . método tradicional na zona urbana;
- . método tradicional na zona rural;
- . método tecnológico com RA na zona urbana;
- . método tecnológico com RA na zona rural;
- . método tecnológico com TV na zona urbana;
- . método tecnológico com TV na zona rural.

Design 3 X 2 → 2 fatores: 1º com 3 níveis, 2º com 2 níveis
= 6 tratamentos.

Design 4 X 3 → 2 fatores: 1º com 4 níveis, 2º com 3 níveis
= 12 tratamentos.

Os "designs" com dois fatores são denominados FATORIAL-"TWO-WAY".

Os "designs" com 3 fatores são denominados FATORIAL "THREE-WAY". Ex.:

Design 2 X 2 X 2 → 3 fatores: 1º com 2 níveis, 2º com dois níveis, 3º com 2 níveis = $2 \times 2 \times 2 = 8$ tratamentos.

O propósito do "design" fatorial completo é estudar:

- . o efeito isolado de cada variável independente sobre a variável dependente;
- . o efeito de interações de variáveis independentes sobre a variável dependente.

Diz-se que existe efeito da interação entre dois fatores quando uma mudança em um fator produz uma mudança diferente na variável de resposta, a um nível de outro fator, em comparação com a mudança em outros níveis, deste fator.

Nos gráficos abaixo, no primeiro deles existe representação de um efeito de interação, enquanto, no segundo, não existe.

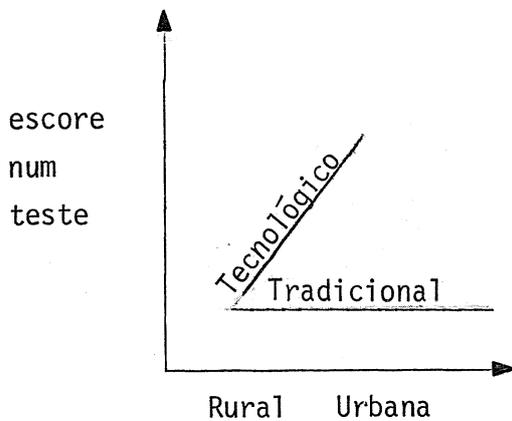


Fig. IV.22 - Efeito de Interação das Variáveis Método de Ensino e Zona.

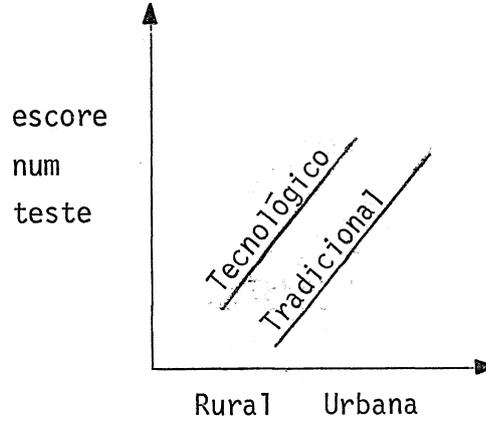


Fig. IV.23 - Ausência de Efeito de Interação das Variáveis Método de Ensino e Zona.

O paralelismo entre linhas implica em nenhuma interação. Como o método gráfico é menos preciso que o método analítico, é necessário aplicar testes estatísticos, a fim de determinar se a interação é ou não, estatisticamente significativa.

Como foi dito, "designs" fatoriais são modificações dos "designs" verdadeiramente experimentais descritos anteriormente. Uma ilustração da modificação de um "DESIGN" GRUPO DE CONTROLE SOMENTE PÓS-TESTE que estuda o efeito na aprendizagem dos alunos de uma variável independente (método de ensino) com dois níveis (convencional - tecnológico), para um "DESIGN" FATORIAL que estuda simultaneamente o efeito na aprendizagem dos alunos desta variável independente, e de mais outras (localização da escola) também com dois níveis (zona urbana e zona rural) seria esquematizado como o que segue:

R	X	Y ₁	T ₁
R	X'	Y ₁	T ₂
R	X	Y ₂	T ₃
R	X'	Y ₂	T ₄

Em que:

- X → método tecnológico
- X' → método convencional
- Y₁ → zona urbana
- Y₂ → zona rural

Neste caso, tem-se um "DESIGN" FATORIAL COMPLETO COM DOIS FATORES, cada um deles com dois níveis. Tem-se, portanto, um experimento com quatro tratamentos. O número de tratamentos é dado pelo produto dos níveis dos fatores (2 x 2 = 4), conforme a Tabela F.26.

- a) Método tecnológico na zona urbana
- b) Método tecnológico na zona rural
- c) Método convencional na zona urbana
- d) Método convencional na zona rural.

TABELA F.26

COMBINAÇÃO DE NÍVEIS NUM "DESIGN" FATORIAL COMPLETO DO TIPO 2 X 2

		MÉTODO	
		TECNOLÓGICO	CONVENCIONAL
LOCALIZAÇÃO	ZONA URBANA	escores dos alunos submetidos ao tratamento: (a)	(c)
	ZONA RURAL	(b)	(d)

a) Procedimentos

- . escolher os níveis dos fatores (QUANTITATIVOS ou QUALITATIVOS, FIXOS ou ALEATÓRIOS) que serão utilizados;
- . combinar os níveis dos fatores estabelecendo quantos grupos se serão utilizados no experimento;
- . definir o número de mensurações a serem realizadas para cada tratamento;
- . selecionar, usando o método aleatório os sujeitos da população.
- . designar, sempre que possível, aleatoriamente os sujeitos aos grupos. Quando uma das variáveis independentes for uma variável de classificação, os sujeitos não poderão ser designados aleatoriamente para os diferentes níveis desta variável. Apenas, eles deverão ser designados aleatoriamente para os diferentes níveis da variável de tratamento.
- . efetuar ou não, conforme o "design", a mensuração antes do tratamento;
- . expor os grupos aos tratamentos;
- . efetuar a mensuração após o tratamento
- . realizar a análise estatística apropriada.

b) Controle das fontes de Validade

O "design" FATORIAL é um "design" verdadeiramente experimental. Daí, tudo que foi dito a respeito do controle das fontes de validade interna e externa, para os "designs" três e quatro, vale também para este "design" cinco, se ele é uma modificação do "design" três ou quatro respectivamente.

c) Utilidade

A utilidade do "design" fatorial se prende ao fato de estudar, simultaneamente, num mesmo experimento, os efeitos de dois ou mais fatores, além de serem obtidas informações acerca das interações entre estes fatores. Porém, deve-se evitar experimentos fatoriais com muitos

fatores, pois com o número de combinações tornando-se muito grande, será muito difícil, do ponto de vista prático, conduzir a experimentação.

- "nested-designs"

Num "design" FATORIAL COMPLETO todas as combinações possíveis dos níveis dos fatores são estabelecidas.

Já num experimento "NESTED", os níveis de um fator são aninhados dentro, ou são sub-amostras, de níveis de outro fator. Neste caso, pode-se estudar o efeito de cada fator, isoladamente, na variável de resposta; mas não se pode estudar o efeito da interação entre ambos.

As Tabelas F.27 e F.28 exemplificam esquemas de coleta de dados de um "nested design".

TABELA F.27

UM ESQUEMA DE COLETA DE DADOS DE UM "NESTED DESIGN"

ESCOLA TIPO I		ESCOLA TIPO II	
JOÃO	PEDRO	MÁRIO	PAULO
-	-	-	-
-	-	-	-
-	-	-	-
-	-	-	-
-	-	-	-

TABELA F.28

UM ESQUEMA ALTERNATIVO DE COLETA DE DADOS DE UM "NESTED DESIGN"

	ESCOLA DO TIPO I	ESCOLA DO TIPO II
JOÃO		
PEDRO		
MÁRIO		
PAULO		

As Tabelas F.27 e F.28 representam os diagramas de coleta de dados, para um experimento que estuda os efeitos do:

- . tipo de escola (I e II) na aprendizagem dos alunos;
- . professor (João, Pedro, Mário, Paulo) na aprendizagem dos alunos;

Neste "design" não é possível estudar o efeito da interação entre os 2 fatores, uma vez que cada professor ensinou apenas num tipo de escola. Se o 2º esquema estivesse totalmente preenchido, isto é, se cada professor tivesse ensinado em ambos os tipos de escola, ter-se-ia um "design" fatorial completo ($2 \times 4 = 8$ tratamentos), onde seria possível o estudo dos efeitos dos fatores isolados e de sua interação. Neste caso, tem-se um "nested design" em que o fator professor foi aninhado no fator tipo de escola.

Deve-se observar que o número de níveis do fator não precisa ser o mesmo, para todos os níveis do outro fator.

- "nested factorial designs"

São aqueles em que vários fatores estão envolvidos, alguns podendo ser cruzados com outros (2 fatores são cruzados quando todas as combinações possíveis de seus níveis são estabelecidas), alguns podendo

ser aninhados aos níveis de outros. Então, experimento "nested fatorial" é aquele em que fatores cruzados e aninhados aparecem simultaneamente. A Tabela F.29 exemplifica a combinação de níveis de um "nested fatorial design".

TABELA F.29

UM ESQUEMA DE COLETA DE DADOS DE UM "NESTED FATORIAL DESIGN"

	ESCOLA TIPO I		ESCOLA TIPO II	
	JOÃO	PEDRO	MÁRIO	PAULO
MÉTODO TECNOLÓGICO	- - - -	- - - -	- - - -	- - - -
MÉTODO TRADICIONAL	- - - -	- - - -	- - - -	- - - -

A Tabela F.29 representa o diagrama de coleta de dados para um experimento que estuda o efeito:

- . do tipo de escola (I e II);
- . do professor (João, Pedro, Mário, Paulo);
- . do método de ensino;
- . da interação tipo de escola X método de ensino (uma vez que todos os níveis de ambos os fatores estão combinados);
- . da interação professor X método de ensino, na aprendizagem dos alunos.

A interação do tipo de escola X professor não é estudada, uma vez que o fator professor está aninhado ao fator tipo de escola (cada professor leciona em apenas 1 tipo de escola, e não em ambos).

Existem outros "designs" que estudam o efeito de 2 ou mais fatores e nos quais existem restrições na randomização. "DESIGNS" FATORIAIS DE BLOCOS RANDOMIZADOS, "DESIGNS" FATORIAIS QUADRADOS LATINOS, "DESIGNS" FATORIAIS "SPLIT PLOT", são exemplos destes. Não serão estudados aqui por serem menos utilizados em Avaliação. Referências a respeito podem ser encontradas em Hicks (1964).

C.3) "DESIGNS" QUASE-EXPERIMENTAIS

São aqueles que têm controle parcial sobre as fontes de validade interna. São não envolvem designação aleatória dos elementos que constituirão a unidade de análise. São por isso mesmo, muito úteis em Avaliação, onde é frequente a necessidade de se trabalhar com grupos intactos, isto é, com classes pré-constituídas. O avaliador que conheça as variáveis específicas que tais "designs" deixam de controlar, e as situações nas quais eles podem ser aplicados, poderá sentir-se encorajado para usá-las.

- "Design seis" grupo de controle não equivalente

T ₁	X	T ₂
T ₃	X'	T ₄

a) Procedimentos

- . selecionar grupos já formados;
- . designar aleatoriamente o grupo que receberá o tratamento experimental;
- . efetuar a mensuração antes do tratamento;
- . expor o grupo experimental ao tratamento e o grupo de comparação ao tratamento alternativo, por um período de tempo;
- . efetuar a mensuração após o tratamento;
- . realizar a análise estatística adequada.

b) Controle das fontes de Validade

VALIDADE INTERNA

A Validade interna do estudo é satisfatória se os grupos tiverem média e desvio padrão similares no pré-teste.

História

(+)

Maturação:

(+)

Testagem:

(+)

Instrumentação:

(+)

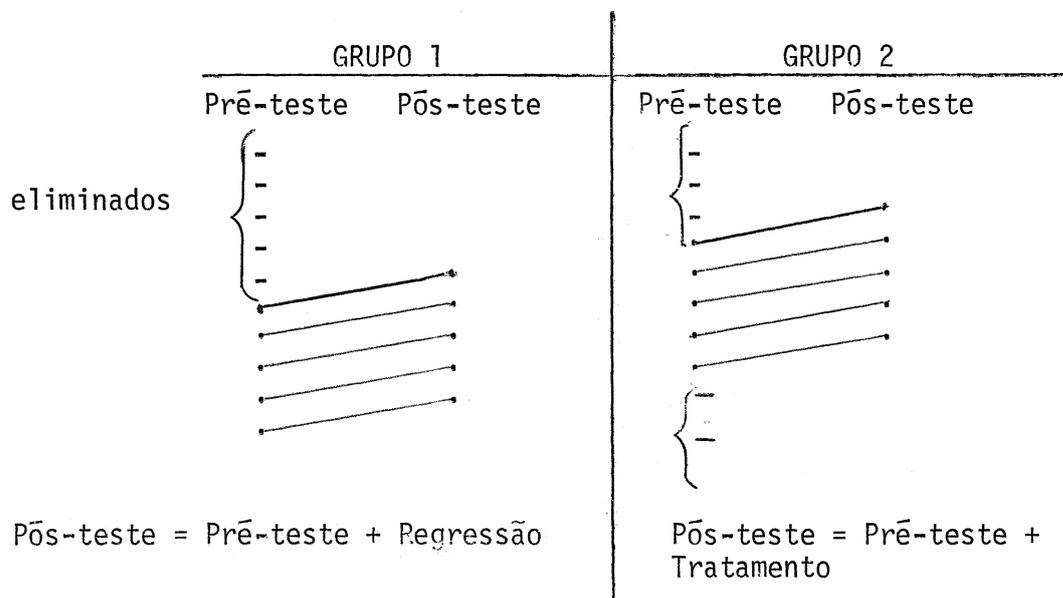
controladas pela presença do grupo de comparação, pois efeitos destas variáveis seriam encontradas afetando ambos os grupos

Regressão Estatística:

(?)

É uma das ameaças à validade interna deste "design". Deve ser evitado o uso de sujeitos selecionados por seus escores extremos, o que, sem dúvida, ocasionaria o efeito de regressão. Ao se tentar equalizar os grupos, eliminando através do acasalamento os alunos com escores discordantes no pré-teste, pode ocorrer o efeito da regressão estatística.

Ex.: Suponha que foram os seguintes os resultados de um experimento:



No caso, os 2 grupos foram "equalizados", usando o acasalamento aos pares. Permaneceram participando do grupo 1, os elementos com escore extremo no grupo, enquanto no grupo 2 os elementos com escore médio. Pelos resultados do pós-teste, poder-se-ia inferir que ambos os tratamentos tiveram o mesmo efeito. Porém, o que verdadeiramente ocorreu foi que, para o grupo 1, o pós-teste reflete o efeito de Regressão, enquanto para o grupo 2, aí sim, está refletido o efeito do tratamento. Acasalamento é desaconselhável neste "design" para equalizar os grupos. Este só é útil quando passível de ser utilizado junto com randomização.

Seleção:

(+)

É controlada pela aplicação do pré-teste. As diferenças iniciais dos grupos (medida através do pré-teste) podem ser anuladas através do uso de técnica de análise estatística (ANACOVA).

Interação Seleção X Maturação
Seleção X História, etc...

(-)

A ameaça maior à validade interna do experimento é devida ao efeito seleção X maturação, seleção X história, etc... Com a ausência da randomização, há sempre a possibilidade de que alguma diferença entre os grupos, não refletida no pré-teste, contamine os resultados do pós-teste.

Ex.: Se um dos grupos é constituído de voluntários, estes podem ser mais motivados.

Mortalidade:

(+)

Controlada pela existência do pré e pós-teste.

VALIDADE EXTERNA

Interação teste X tratamento:

(-)

Idem ao "design" três.

Interação seleção X tratamento: (?)

Idem ao "design" três.

Efeitos Reativos dos arranjos experimentais: (?)

A ameaça à validade externa devido a arranjos reativos do experimento estão presentes, porém, provavelmente num grau menor do que em verdadeiros designs experimentais. O uso de classes intactas constitui em Educação uma situação mais natural do que o uso de grupos formados por designação aleatória de seus elementos.

c) Utilidade

Este é um dos "designs" mais amplamente utilizados em Educação. A necessidade do uso deste "design" ocorre, sempre, no mundo prático da escola, onde a designação aleatória dos sujeitos aos grupos nem sempre pode ser empregada. Muitas vezes o avaliador necessita iniciar seu trabalho com o uso de grupos pré-estabelecidos de sujeitos. Neste caso, ele deve esforçar-se em utilizar grupos os mais similares possíveis em todos os aspectos individuais, potencialmente relevantes ao estudo. Este é um "design" valioso que deve ser usado em muitas ocasiões em que os "designs" três, quatro ou cinco não são possíveis.

Variações deste "design" podem ser empregadas, ampliando-se o número dos grupos de comparação.

- "design" sete: um grupo série no tempo.

T ₁	T ₂	T ₃	T ₄	X	T ₅	T ₆	T ₇	T ₈
----------------	----------------	----------------	----------------	---	----------------	----------------	----------------	----------------

a) Procedimentos

- efetuar uma série periódica de mensuração (quanto mais melhor) antes do tratamento;
- expor o grupo de sujeitos ao tratamento por um período de tempo;

- efetuar uma s̄erie perīodica de mensuraç̄ao ap̄os o tratamento;
- realizar an̄alise estat̄istica dos dados.

OBS.: *Frequentemente as medidas ser̄ao obtidas de registros existentes, tais como: testes de realizaç̄ao regularmente administrados, rela- t̄orios de freqūencia, etc... Tais dados de arquivo s̄ao particular- mente ajustados a este tipo de "design".*

Embora a an̄alise estat̄istica requerida para este design seja complexa, em muitas ocasīoes, um julgamento racional, acerca do efei- to do tratamento pode ser feito atrav̄es da an̄alise de uma representaç̄ao gr̄afica dos dados, como a apresentada a seguir na Figura F.24.

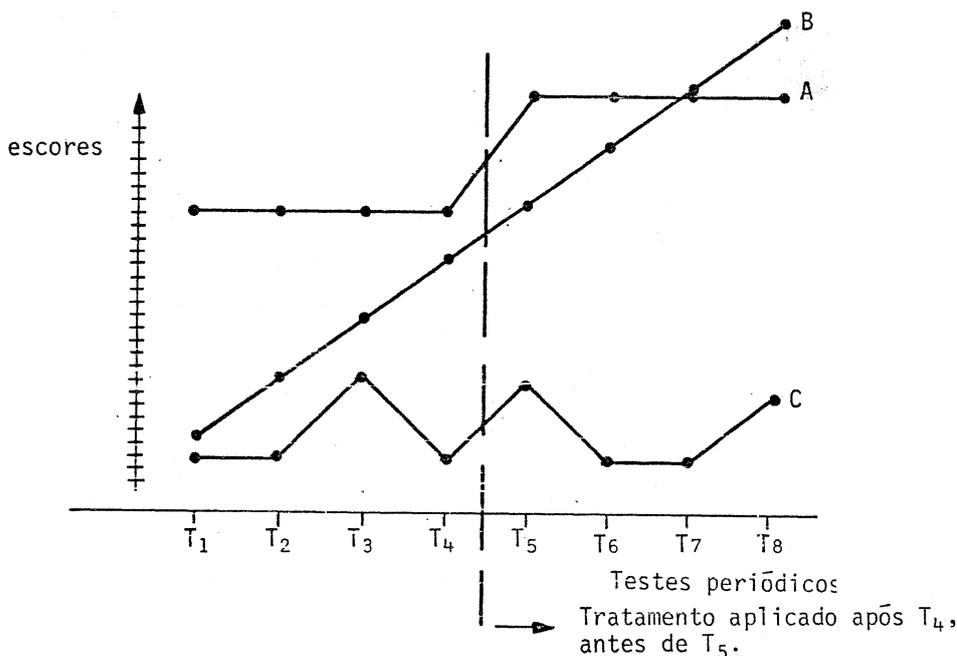


Fig. F.24 - Representaç̄ao Gr̄afica de S̄eries Temporais.

Atrav̄es deste gr̄afico, embora se tenha um acr̄escimo nos escores de T₄ para T₅ nas tr̄es linhas que representam resultados de um experimento, nota-se que, somente em A, este acr̄escimo pode ser tido co- mo devido ao tratamento.

b) Controle das fontes de validade:

VALIDADE INTERNA

<u>História:</u> (-)	É a maior ameaça a validade deste "design". Existe a hipótese de que não o tratamento, mas algum evento paralelo produziu a mudança observada no escore. Experimentos deste tipo se estendem em período de tempo mais ou menos longos. Assim, variáveis como efeitos do clima, realização de jogos escolares podem ter efeito na variável medida. Mas, ainda assim, o controle desta ameaça é maior que nos "designs" um e dois.
<u>Maturação:</u> (+)	Maturação, usualmente, não produz efeito paralelo ao tratamento se, por exemplo, não existe diferença nos primeiros quatro testes, existe diferença entre T_5 e T_4 , e novamente não existe diferença nos quatro últimos testes. Na linha A, por exemplo, não transparece efeito de maturação.
<u>Testagem:</u> (+)	O efeito da testagem não se refletiria apenas de T_4 para T_5 .
<u>Instrumentação:</u> (?)	Seria difícil esperar que o efeito da instrumentação ocorresse juntamente numa ocasião particular (entre T_4 e T_5). No entanto, o ponto de interrogação aparece para chamar a atenção às situações em que uma mudança no calibre do instrumento poderia ser erroneamente interpretada como um efeito de X.
<u>Interação seleção X maturação, etc.</u> (+)	Controlada pela testagem seriada.

<u>Regressão:</u> (+)	Ameaça não plausível como explanação de um efeito em T_5 maior do que em T_2 , T_3 , e T_4 .
<u>Seleção:</u> (?)	Como fonte de efeitos principais é excluída se as mesmas pessoas estão envolvidas em todas as observações.
<u>Mortalidade:</u> (+)	Se os dados são coletados por indivíduo, há controle de mortalidade. Se as observações são coletadas por grupo, deve ser feito um registro da evasão, recolocação de alunos, para assegurar que mudança de pessoal (composição do grupo), não tenha efeito paralelo ao tratamento.

VALIDADE EXTERNA

<u>Interação Seleção X tratamento:</u> (?)	Uma vez que se trabalhou com um grupo intacto, o efeito do tratamento pode ter sido restrito a este grupo.
<u>Interação pré-teste X tratamento:</u> (-)	Os resultados obtidos podem ser restritos a situações em que exista esta testagem seriada. Isto, quando esta modifica os sujeitos, de modo que eles não sejam mais representativos da população.
<u>Efeitos reativos dos arranjos experimentais:</u> (?)	Se os dados são obtidos através de registros escolares, não haverá efeito reativo do experimento, uma vez que o estudo é posterior ao fato. Porém, se as testagens são realizadas excepcionalmente para o estudo, isto pode dar ao experimento um aspecto pouco natural.

c) Utilidade

Existem numerosas ocasiões, em Educação, em que um grupo de comparação ou de controle não pode ser incluído num experimento. É o caso de introdução de uma mudança num sistema, que não encontre mudança

alternativa disponível para comparação. Nestes casos, este "design" é mais forte que os "DESIGNS": UM GRUPO-PÓS TESTE e UM GRUPO - PRÉ - TESTE PÓS TESTE.

Também, tal "design" provê uma perspectiva diferente, para avaliar os efeitos de uma mudança educacional. Muitas intervenções produzem seus efeitos através do tempo, não num simples instante como o medido por um único teste. É, portanto este, um "design" adequado para estudar efeitos a longo prazo.

Constitui-se ainda num instrumento "após o fato" para avaliar mudanças educacionais para as quais não foi planejado nenhum grupo de comparação.

Sua forma mais simples pode ser modificada de modo a envolver dois ou mais grupos, conforme segue:

T ₁	T ₂	T ₃	T ₄	X	T ₅	T ₆	T ₇	T ₈
T ₉	T ₁₀	T ₁₁	T ₁₂	X'	T ₁₃	T ₁₄	T ₁₅	T ₁₆

Esta última forma é, obviamente, mais forte que a forma básica. A presença do grupo de controle manipula o efeito da variável história que é a maior ameaça à validade do "design" sete. Esta forma é também superior ao "design" seis. GRUPO DE CONTROLE NÃO EQUIVALENTE - pois controla a interação entre seleção e maturação, seleção e história, etc.

A Tabela F.30 que segue, baseada em Campbell e Stanley (1963), apresenta para cada "design":

- fontes que ele seguramente controla (+)
- fontes que ele não controla (-)
- fontes com controle não assegurado (?)
- fontes que não são relevantes aos designs ()

TABELA F.30

CONTROLE DAS FONTES DE VALIDADE, E UTILIDADE PARA A AVALIAÇÃO
DE CADA "DESIGN" DE COLETA DE DADOS

FONTES DE INVALIDADE "DESIGN"	INTERNA							EXTERNA				UTILIDADE DO "DESIGN" PARA A AVALIAÇÃO	
	HISTÓRIA	MATURAÇÃO	TESTAGEM	INSTRUMENTAÇÃO	REGRESSÃO ESTATÍSTICA	SELEÇÃO	MORTALIDADE	INTERAÇÃO DE FATORES	EFEITO REATIVO DA TESTAGEM	INTERAÇÃO SELEÇÃO X TRATAMENTO	REACÇÃO AOS ARRANJOS EXPERIMENTAIS		MÚLTIPLOS TRATAMENTOS
UM GRUPO PÓS-TESTE	-	-				-	-			-			Pode ser utilizado nos primeiros estágios de Avaliação Formativa
UM GRUPO PRÉ-TESTE PÓS-TESTE	-	-	-	-	?	+	+	-	-	-	?		Extremamente adequado para Avaliação Formativa
GRUPO DE CONTROLE PRÉ-TESTE - PÓS TESTE	+	+	+	+	+	+	+	+	-	?	?		"Design" muito forte no controle das fontes de validade, porém, com restrições práticas
GRUPO DE CONTROLE SOMENTE PÓS-TESTE	+	+	+	+	+	+	+	+	+	?	?		"Design" muito forte no controle das fontes de validade, porém com restrições práticas.
FATORIAL COMPLETO	Idêntico ao "design" 3 ou 4, uma vez que é modificação de um ou de outro, respectivamente.											- Estuda o efeito de dois ou mais fatores. - Estuda o efeito da interação destes fatores.	
"NESTED DESIGN"	Idêntico ao "design" 3 ou 4, uma vez que é modificação de um ou de outro, respectivamente.											- Estuda o efeito de dois ou mais fatores.	
"NESTED FATORIAL DESIGN"	Idêntico ao "design" 3 ou 4, uma vez que é modificação de um ou de outro, respectivamente.											- Estuda o efeito de mais de dois fatores. - Estuda o efeito de interação de alguns fatores.	
GRUPO DE CONTROLE NÃO EQUIVALENTE	+	+	+	+	?	+	+	-	-	?	?		"Design" relativamente forte no controle das fontes de validade e, muito adequado às situações reais.
UM GRUPO SÉRIE NO TEMPO	-	+	+	?	+	+	+	+	-	?	?		É o mais forte dos "designs" dentre os que têm só um grupo. Adequado para estudar efeitos a longo prazo. Constitui num instrumento "após o fato".

d) ELABORAÇÃO DE UM PLANO PARA DESIGNAÇÃO DOS SUJEITOS AOS GRUPOS E DE DETERMINAÇÃO DOS TRATAMENTOS AOS GRUPOS.

Enquanto a seleção de uma amostra, para representar a população se prende à VALIDADE EXTERNA do experimento, isto é, a probabilidade de generalizar, corretamente, para a população, os resultados obtidos através do estudo da amostra, a designação randômica dos sujeitos se prende à VALIDADE INTERNA do experimento.

Esta técnica, como já foi visto, é característica dos "designs" verdadeiramente experimentais.

Escolhido o "design" de coleta de dados, sendo ele um design verdadeiramente experimental, isto implicará que, após a seleção da amostra, o passo seguinte será designar aleatoriamente os sujeitos para os diferentes grupos de tratamento. Constituir-se-ão tantos grupos quantos forem os tratamentos. Tais grupos podem ser considerados, então, equivalentes.

Observa-se que quando a unidade de análise for a classe, se após a seleção aleatória daquelas que constituirão a amostra, possuir-se a designação aleatória destas para constituírem os diferentes grupos de tratamento, estar-se-á, ainda, dentro do esquema proposto como "design" verdadeiramente experimental.

Constituídos os grupos, o passo seguinte será determinar, ainda aleatoriamente que grupo receberá qual tratamento. Este último passo deverá ocorrer mesmo em "designs" cuja designação aleatória dos sujeitos (ou das classes) nos grupos não seja possível. Grupos intactos deverão ser submetidos a um ou outro tratamento, aleatoriamente.

e) ESQUEMATIZAÇÃO DE PROCEDIMENTOS PARA COLETA DE DADOS

O plano experimental deverá incluir, também, os procedimentos que serão utilizados para a coleta de dados. O quando e o como serão coletados os dados se constituem em pontos importantes da metodologia de

O processo experimental envolve a coleta de dados numa amostra, e a partir destes dados, a realização de inferências para um universo maior de observações potenciais.

Ao concluir-se acerca da população, com base em dados de uma amostra, estar-se-á passando por alguns riscos de tomar decisões in corretas.

Existem processos rigorosos para assegurar que a partir do estudo da amostra possa-se inferir acerca da população. O teste esta tístico de hipótese é um processo científico para generalizar da amostra para a população. Através do teste estatístico, pode-se determinar a pro bab ilidade de que os resultados de um estudo sejam devidos ao acaso, pro ven do, portanto, evidências para o julgamento da validade de uma hipótese.

Exemplo:

Imagine que por existir a necessidade de se decidir sobre a adoção de um método de ensino entre dois possíveis (A = sintético, B = analítico), para a alfabetização de adultos, foi realizado um estudo com dois grupos de comparação. No caso, o que se espera é uma resposta à questão.: "Os dois métodos têm o mesmo efeito na aprendizagem dos alu nos, ou um é superior ao outro"?

Imagine, também, que através do escore médio de cada gru po no teste realizado após os tratamentos, estes tratamentos "pareceram" ser diferentes (isto porque os grupos apresentaram escores médios dife ren tes). Neste caso, como saber se a diferença observada para estas duas amostras reflete, realmente, uma diferença entre os tratamentos e não é simplesmente devida a um erro de amostragem (sendo, portanto, ao acaso)? Isto é, como saber, se a partir da diferença observada através dos dados amostrais, pode-se inferir que a diferença seria também observada, caso se estivesse trabalhando com a população de observações?

Um teste estatístico pode determinar a probabilidade de que a diferença obtida, entre duas médias, seja devida ao acaso, isto é, seja devida a um erro de amostragem, orientando, então, que decisão tomar acerca da hipótese em teste. Quando a probabilidade de uma diferença en tre mé dias se dever ao acaso, for pequena, aceita-se a afirmativa de que a diferença entre elas se deve realmente aos tratamentos.

coleta. A determinação do grupo deverá resultar num cronograma de coleta de dados, ou numa rede PERT, enquanto as especificações do grupo deverão resultar na identificação das condições, sob as quais as mensurações serão realizadas, os métodos que serão utilizados para estas mensurações, e os instrumentais que deverão ser elaborados.

Todos os procedimentos para coleta de dados deverão ser cuidadosamente planejados, pois, caso os dados não sejam coletados de forma cuidadosa e científica, serão insuficientes as melhores definições operacionais das variáveis, a melhor seleção do "design" de coleta de dados, a utilização da mais representativa das amostras, etc.. pois os resultados estarão, infalivelmente, prejudicados.

f) ELABORAÇÃO DE UM PLANO DE CONSTRUÇÃO DOS INSTRUMENTAIS PARA MEDIR OS RESULTADOS DO EXPERIMENTO.

Neste passo, será planejada a construção dos instrumentais para medir os resultados do experimento. Estes serão constituídos com base na operacionalização da variável dependente. Em estudos avaliativos cuja variável dependente é a aprendizagem dos alunos, a operacionalização desta variável resulta na especificação de objetivos comportamentais, que se constituem, então, a base para elaboração dos instrumentais de medida. Estes serão utilizados para medir a extensão na qual aqueles objetivos do programa foram atingidos, o que refletirá o efeito da exposição dos alunos à mudança introduzida.

Os procedimentos utilizados para o planejamento e construção de instrumentais de medida - testes de aprendizagem - são apresentados detalhadamente na Unidade 3.0, página F.149.

g) ELABORAR UM PLANO DE TESTAGEM DE HIPÓTESES

A realização de um experimento se dá com o objetivo de que sejam coletados dados da realidade, que comprovem ou não, a hipótese definida para o problema em estudo.

O processo de testagem de hipótese apresenta os seguintes itens:

- | | |
|--|--|
| P
L
A
N
E
J
A
M
E
N
T
O | - estabelecer a hipótese nula; |
| | - selecionar um teste estatístico apropriado; |
| | - especificar um nível de significância que implicará no critério para rejeitar a hipótese nula; |
| | - determinar a região crítica. |
| E
X
E
C
U
C
Ã
O | - computar o valor da estatística empregada no teste |
| | - tomar a decisão estatística de rejeitar ou não a hipótese nula. |

Salienta-se que o processo total de testagem de hipótese é realizado apenas parcialmente neste passo 5 - "Elaborar um plano experimental para coleta de dados do ESQUEMA de Avaliação Somativa, baseado na Pesquisa Experimental, pág. F.258. Realmente, é neste passo, mais especificamente neste tópico G - "Elaborar um plano de testagem de hipótese" que são realizados os primeiros quatro itens do processo de testagem de hipótese. Os dois últimos são realizados após a Execução do Experimento - passo 6 - correspondendo aos passos 7 e 8, respectivamente, do citado ESQUEMA.

- estabelecer a hipótese nula

A fim de que a hipótese do problema - resposta sugerida para o mesmo e que será testada através dos dados coletados da realidade - seja aceita ou não, é necessário que ela, após operacionalizada, seja transcrita para as duas formas de hipóteses estatísticas, que se constituem

em suposições acerca da população que está sendo amostrada. Estas duas formas de hipóteses estatísticas são a hipótese alternativa e a hipótese nula.

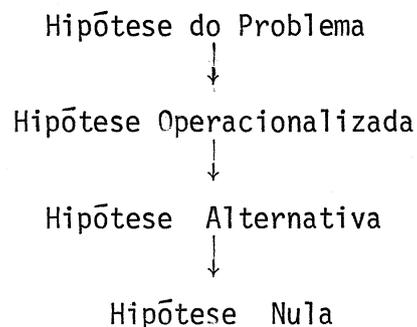
- A hipótese alternativa (H_1) é resultado da transcrição da hipótese operacional que se apresenta sob a forma de afirmativa de uma proposição, em uma hipótese sob a forma de uma afirmativa matemática (Hill, 1967).

OBS.: Alguns autores, entre estes Siegel (1956), ao invés de considerarem que existe a hipótese operacional e a hipótese alternativa, consideram que o que existe é apenas a hipótese alternativa, e que esta pode ser apresentada sob a forma de uma afirmativa de uma proposição, ou, sob a forma de uma afirmativa matemática.

- A hipótese nula (H_0) é uma afirmativa matemática desenvolvida a partir da hipótese alternativa, e é oposta a esta. A hipótese nula é, portanto, contrária e complementar a hipótese alternativa (Hill, 1967).

A hipótese nula é uma hipótese acerca de "nenhuma diferença" e é usualmente formulada com o propósito de ser rejeitada (Siegel, 1956). No processo de testagem de hipótese, só se aceita a hipótese alternativa, através da rejeição do seu contrário, que é a hipótese nula.

A sequência no estabelecimento das hipóteses segundo Hill (1967) é:



Ex.:

Hipótese do Problema: Os alunos das escolas da zona urbana aprendem mais que os alunos das escolas da zona rural

Hipótese Operacionalizada Num teste de desempenho, os escores individuais dos alunos da zona urbana são maiores do que dos alunos da zona rural.

Hipótese Alternativa (H_1) $\mu_u > \mu_r$

Onde

H_1 indica a hipótese estatística alternativa

μ_u indica o escore médio da população de alunos da zona urbana, no teste de desempenho

μ_r indica o escore médio da população de alunos da zona rural, no teste de desempenho

> indica que a primeira média é maior que a segunda.

Observa-se que, nesta hipótese, está implícita uma direção. Tem-se uma afirmativa de que o escore médio da população de alunos da zona urbana é maior do que o escore médio dos alunos da zona rural. Se $H_1 = \mu_u < \mu_r$, ainda, ter-se-ia implícita uma direção. Ter-se-ia que o escore médio da população de alunos da zona urbana seria menor do que o escore médio dos alunos da zona rural.

Nenhuma direção estaria implícita na $H_1 = \mu_u \neq \mu_r$. Neste caso, não estaria afirmado que um grupo tem escore médio maior (ou menor) do que outro, apenas que têm escore médio diferente.

Hipótese Nula: $H_0: \mu_u \leq \mu_r$

Nula:

Para $H_1: \mu_u < \mu_r$ ter-se-ia $H_0: \mu_u \geq \mu_r$

Para $H_1: \mu_u \neq \mu_r$ ter-se-ia $H_0: \mu_u = \mu_r$

Hipóteses estatísticas são afirmativas matemáticas que, dependendo de sua forma, podem ser classificadas em 3 categorias (Hill, 1967):

- a) Hipóteses de magnitude;
- b) Hipóteses de diferença;
- c) Hipóteses de relação.

a) Uma hipótese de magnitude é aquela na qual é estipulada uma afirmativa do número, valor ou tamanho de uma entidade.

Ex.:

Hipótese Operacional: O escore médio num dado teste, dos alunos submetidos ao método tradicional é menor do que 60.

Hipótese Alternativa: $H_1: \mu_{tr} < 60$

Hipótese Nula: $H_0: \mu_{tr} \geq 60$

b) Uma hipótese de diferença é aquela que trabalha com a diferença entre quantidades sendo comparadas.

Ex.:

Hipótese Operacional: Os estudantes do grupo experimental terão um escore médio maior, num dado teste do que os estudantes do grupo de comparação.

Hipótese Alternativa: $H_1: \mu_{ex} > \mu_c$

$H_0: \mu_{ex} \leq \mu_c$

c) Uma hipótese de relação é aquela na qual é estabelecido o grau de associação entre duas variáveis assumidas como sendo relacionadas causalmente e formando uma população bivariada.

Ex.:

Hipótese Operacional: Existe uma relação direta entre escores dos indivíduos num teste de inteligência

e escores dos indivíduos num dado teste de desempenho em matemática, para determinada população.

Hipótese Alternativa: $H_1: \rho > 0$ O coeficiente de correlação " ρ ", a medida de relação empregada neste caso, é positiva e significativamente maior do que zero.

Hipótese Nula: $H_0: \rho \leq 0$

O propósito do teste de hipótese é que os dados apresentem evidência suficiente para rejeitar H_0 . Se os dados rejeitam H_0 é possível então aceitar H_1 correspondente, uma vez que ambas as hipóteses são complementares. Assim, o processo de testar a hipótese nula, em essência, é o processo de testá-la contra sua hipótese alternativa. Se os dados rejeitam a hipótese nula (H_0), o investigador pode aceitar a hipótese alternativa complementar (H_1). Se os dados não podem rejeitar a hipótese nula, o investigador não pode aceitar a hipótese alternativa. Nota-se que a hipótese nula nunca é aceita. As decisões a seu respeito são sempre rejeitar H_0 , ou não se pode rejeitar H_0 . Assim, a hipótese alternativa correspondente nunca é rejeitada. As decisões a seu respeito são sempre aceitar H_1 ou não se pode aceitar H_1 .

- especificar um nível de significância que implicará no critério para rejeitar a hipótese nula.

Uma vez que os resultados de testes estatísticos são sempre em termos de "existe uma probabilidade X de que os resultados sejam devidos ao acaso, e não ao tratamento", as decisões acerca de rejeitar H_0 ou não se pode rejeitar H_0 são sempre decisões que envolvem incertezas. Realmente, uma hipótese nula, colocada em teste, pode ser VERDADEIRA OU FALSA. Uma decisão, baseada em dados amostrais, acerca desta hipótese, pode ser REJEITAR H_0 ou NÃO SE PODE REJEITAR H_0 . Disto resulta que ao ser tomada uma decisão acerca da hipótese nula, poder-se-á estar numa das quatro situações diagramadas na Tabela F.31.

TABELA F.31

QUATRO SITUAÇÕES POSSÍVEIS DE OCORRER QUANDO É TOMADA UMA
DECISÃO ACERCA DA HIPÓTESE NULA

	H_0 VERDADEIRA	H_0 FALSA
REJEITAR H_0	decisão incorreta ERRO DO TIPO I (α)	DECISÃO CORRETA
NÃO SE PODE REJEI TAR H_0	DECISÃO CORRETA	decisão incorreta ERRO DO TIPO II (β)

Ao tomar-se uma decisão estatística, portanto, existe a possibilidade de que esteja incorrendo num erro:

- pode-se rejeitar H_0 , sendo H_0 verdadeira (erro do Tipo I);
- pode-se não rejeitar H_0 , sendo H_0 falsa (erro do Tipo II).

A probabilidade (α) de que ao testar-se uma hipótese nula, seja cometido um erro do tipo I, é denominada nível de significância de um teste. Isto é, nível de significância de um teste é a probabilidade de de rejeitar-se, por exemplo, a hipótese de que a diferença entre as médias de dois grupos experimentais seja devido ao acaso, quando realmente esta diferença é devida ao acaso ($H_0: \mu_1 = \mu_2$) indicando que não existe diferença quanto ao tratamento).

Como já afirmado, em geral em Ciências Humanas trabalha-se com o nível de significância 0,05 a 0,01. Ou seja, trabalha-se com a probabilidade de 5 vezes em 100 ou 1 vez em 100 de rejeitar-se H_0 , sendo H_0 verdadeira.

A especificação de um nível de significância implica na determinação de um critério para rejeitar a hipótese nula. Fixado o nível de significância, é determinado em correspondência, através do uso adequado de uma tabela apropriada ao teste estatístico empregado, um valor X_C , denominado VALOR CRÍTICO. O valor crítico define a REGIÃO CRÍTICA da estatística do teste, que contém todos os valores desta estatística, que implicam na decisão de rejeitar H_0 .

Computado com os dados da amostra, o teste estatístico, fica determinado o valor observado da estatística do teste. Este valor é, então, confrontado com o valor teórico, X_C , para que seja verificado se este se encontra fora ou dentro da região crítica. Encontrando-se dentro desta, a decisão estatística é REJEITAR H_0 . Encontrando-se fora, a decisão estatística é que não se pode rejeitar H_0 .

Se H_0 verdadeira, e a decisão estatística: NÃO SE PODE REJEITAR H_0 , estar-se-á tomando uma decisão correta. Porém, sendo H_0 verdadeira, e a decisão estatística REJEITAR H_0 , estar-se-á cometendo um erro do tipo I. A probabilidade de cometer-se tal erro é igual ao nível de significância assumido. Assim, se o nível de significância é 0,05, a probabilidade de cometer-se um erro do tipo I é de 5 vezes em 100. Se o nível de significância é 0,01, a probabilidade de cometer-se um erro do tipo I é de 1 vez em 100. Percebe-se, que à medida que diminui o nível de significância, diminui a probabilidade de se cometer um erro do tipo I. Então, porque não trabalhar com nível de significância zero, que implicaria em NUNCA cometer um erro do tipo I? Simplesmente, porque existe a possibilidade de ocorrência do outro tipo de erro - o erro do tipo II, que, como foi visto, consiste em não rejeitar H_0 , sendo H_0 falsa. Assim, caso se trabalhasse com o nível de significância zero, nunca seria rejeitada a hipótese nula. E se a hipótese nula fosse falsa? Não existiria a possibilidade de rejeitar uma hipótese falsa. O que acontece é que a medida que se diminui a probabilidade de cometer um erro do tipo I, aumenta-se a probabilidade (β) de cometer um erro do tipo II e vice-versa. Como nunca se sabe se H_0 é verdadeira ou falsa, pois é isto justamente que está sendo colocado em teste, deve-se planejar um teste em que os riscos α e β sejam ambos pequenos. O procedimento geral seguido é fixar um valor α (0,05 ou 0,01), tornar mínimo o valor de β e determinar o tamanho da amostra em função dos valores α e β .

Apesar das decisões, de rejeitar ou não a hipótese nula, envolverem, explicitamente, a definição do nível de significância (α), e não a definição de β , deve-se ressaltar que existe uma ligação íntima entre estes valores e mais o tamanho da amostra utilizada. Portanto, estes três elementos devem ser determinados simultaneamente, sendo um deles, sempre, função dos outros dois.

- determinar a região crítica

Há dois tipos de regiões críticas, em função da hipótese nula que está sendo testada. O tipo associado à hipótese, na qual não está implícita nenhuma direção (Ex.: $H_0 : \mu_1 = \mu_2$) e o tipo associado à hipótese na qual está implícita uma direção (Ex.: $H_0 : \mu_1 \geq \mu_2$ ou $H_0 : \mu_1 \leq \mu_2$).

No primeiro caso, a região crítica é composta de duas porções iguais de distribuição amostral* da estatística envolvida no teste, uma localizada no final superior da curva, e outra no final inferior da curva. A Figura F.25, que segue, apresenta graficamente esta idéia.

Ex.: $H_0 : \mu_1 = \mu_2$ ou $H_0 : \mu_1 - \mu_2 = 0$

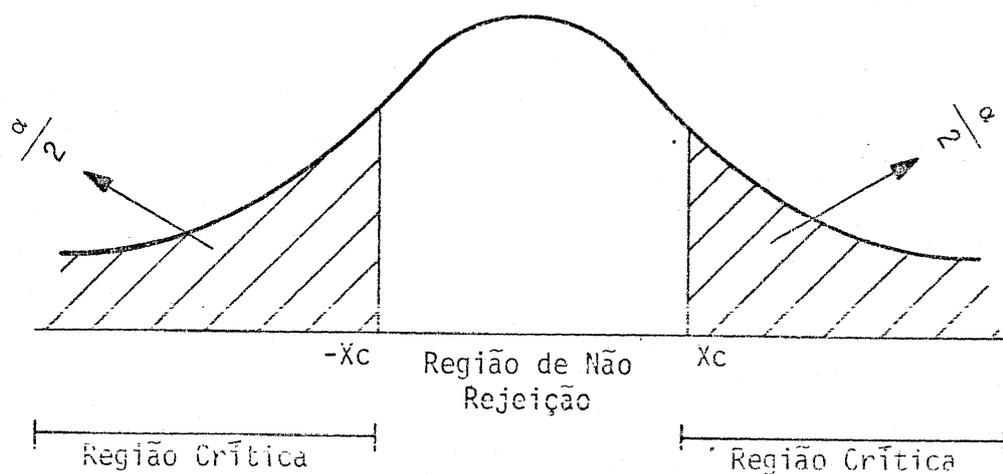


Fig. F.25 - Região Crítica para Hipótese Nula com o Sinal de Igualdade (=).

* Para revisar os conceitos relativos a distribuição amostral, pré-requisitos deste curso, recorrer a livros textos de estatística diferencial com Ferguson (1971), Guilford (1965), ou outros.

O valor crítico x_c , valor teórico do teste estatístico que define os limites ($-x_c$ e $+x_c$) das regiões críticas, é encontrado em tabelas, apresentadas na maioria dos livros de Estatística. Existem tabelas próprias dos valores teóricos de cada um dos testes.

Se o valor da estatística calculada com os dados amostrais, extraídas da população em estudo, for menor do que $-x_c$, ou maior do que x_c , então a hipótese nula será rejeitada em favor da hipótese alternativa. Observe que cada uma das porções apresenta uma área sobre a curva igual a $\alpha/2$.

No segundo caso, a região crítica é composta de uma porção da distribuição amostral da estatística envolvida no teste, localizada no final superior da curva se $H_0: \mu_1 \leq \mu_2$, ou no final inferior da curva se $H_0: \mu_1 \geq \mu_2$. As Figuras F.26 e F.27, que seguem, apresentam as regiões críticas para estas duas hipóteses nulas respectivamente.

Ex.: $H_0: \mu_1 \leq \mu_2$ ou $H_0: \mu_1 - \mu_2 \leq 0$.

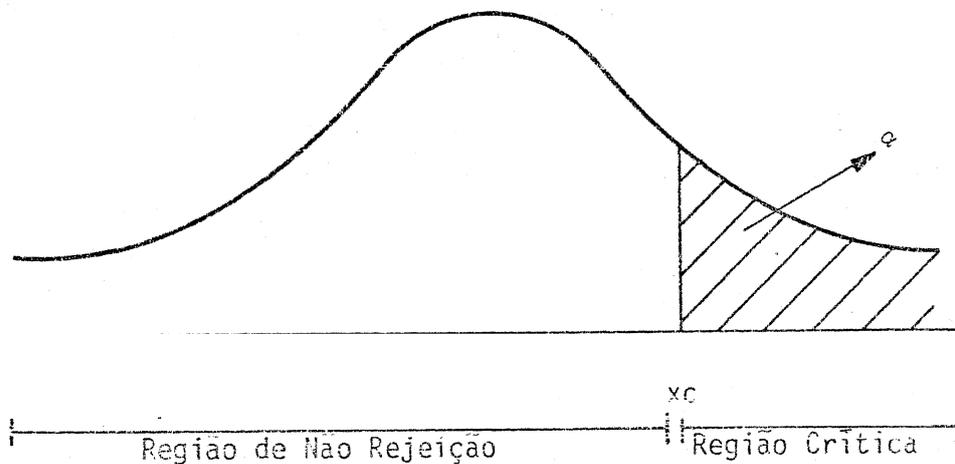


Fig. F.26 - Região Crítica para Hipótese Nula com Sinal Menor ou Igual (\leq).

Se o valor da estatística calculada com os dados amostrais extraídos da população, em estudo, for maior do que x_c , então a hipótese nula será rejeitada em favor da hipótese alternativa.

Ex.: $H_0: \mu_1 \geq \mu_2$ ou $H_0: \mu_1 - \mu_2 \geq 0$

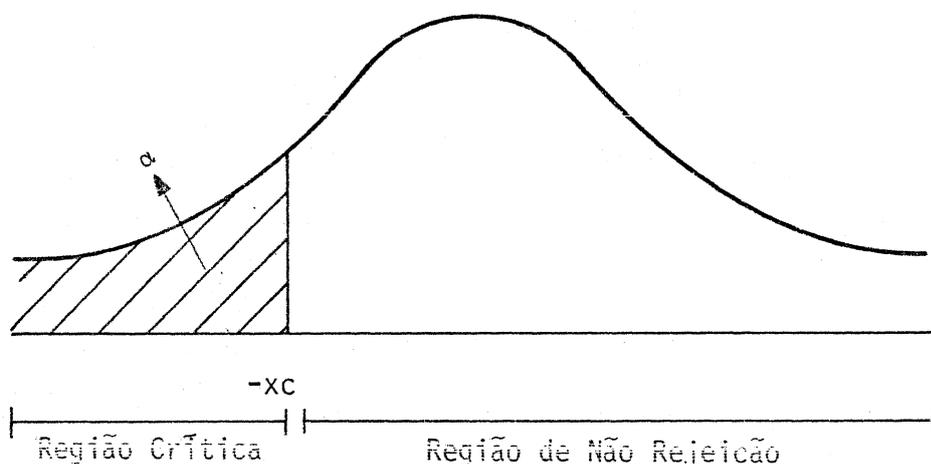


Fig. F.27 - Região Crítica para Hipótese nula com Sinal Maior ou Igual (\geq).

Se o valor da estatística calculada com os dados amostrais extraídos da população, em estudo, for menor do que $-x_c$, então a hipótese nula será rejeitada em favor da Hipótese alternativa.

- Selecionar um teste estatístico apropriado.

Em função da categoria da hipótese que está sendo testada (magnitude, diferença, relação), podem ser identificados, em correspondência, três tipos de testes de significância estatística. Destes, dois são particularmente úteis aos avaliadores, quando avaliando resultados de mudanças inovativas que se encontram em ensaio: - os testes de significância da magnitude, e os testes de significância da diferença. Estes últimos, sobretudo, são os mais úteis aos avaliadores por estarem eles envolvidos com o problema de comparação entre alternativas competitivas. Avaliadores estão sempre respondendo a questão do tipo: "Há uma diferença significativa entre estes dois (ou mais) conjuntos de dados?".

Dentro destes dois conjuntos de testes, serão tratados, aqui, os testes mais usuais em Avaliação. Não serão descritos os mecanismos computacionais destes testes, mas apenas apresentados os esquemas de análise de dados apropriados para a maioria das situações encontradas, quando da utilização dos "designs" de coleta de dados transcritos anteriormente.

Uma observação deve ser colocada: para conduzir estas análises é preciso pesquisar mais sobre o assunto, ou contratar estatísticos consultantes.

- TESTES PARAMÉTRICOS X TESTES NÃO PARAMÉTRICOS

Os testes estatísticos em função do nível de mensuração dos dados, podem ser classificados em dois grandes grupos:

- testes paramétricos;
- testes não paramétricos.

Testes paramétricos são aqueles em que o nível de medida de variável dependente é o da escala de intervalo (Popham, 1975). O uso destes testes requer que as seguintes suposições sejam satisfeitas.

- distribuição normal: os testes paramétricos são mais válidos quando aplicados a dados que seguem uma distribuição normal;
- homogeneidade da variância: os dados são mais facilmente comparados através de um teste paramétrico, quando a variância dentro dos dois grupos é igual;
- medidas de intervalos iguais e contínuos: como conceituados, os testes paramétricos, somente podem ser aplicados quando as medidas da variável dependente representam uma escala de intervalo, isto é, quando os dados são contínuos e de intervalos iguais do começo ao fim.

Testes não paramétricos são aqueles em que o nível de medida da variável dependente pode ser o da escala ordinal ou da escala nominal. O uso destes testes requer poucas suposições acerca das distribuições: não requer uma distribuição normal, nem homogeneidade da variância.

O nível de medida dos dados pode influenciar o tipo de análise estatística selecionada. Em geral, uma técnica estatística aplicável a uma escala mais fraca de medida será apropriada para dados refletindo uma escala mais forte. O reverso entretanto não ocorre.

- EXEMPLOS DE USO DE ALGUNS TESTES ESTATÍSTICOS MAIS ÚTEIS

A apresentação de cada teste seguirá um formato padrão que compreende: identificação de uma situação de testagem, a hipótese nela correspondente à situação, o teste estatístico adequado para, nesta situação, testar tal hipótese, a fórmula deste teste, a classificação do teste como paramétrico ou não paramétrico e finalmente a identificação dos "designs" de coleta de dados cuja aplicação resulta em dados adequadamente analisados pelo teste em apresentação.

SITUAÇÃO 1.

Testar a significância da magnitude da média de uma população normalmente distribuída cuja variância é conhecida. Esta situação não ocorre com frequência em Avaliação. Não é frequente não se conhecer a média de uma população e conhecer-se sua variância. Isto só acontece quando se dispõe de dados relativos a estudos anteriormente realizados.

$H_0: \mu = \mu_1$, em que μ_1 é o valor determinado proposto para μ .

Teste estatístico: teste z.

Fórmula de z: $z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$, em que:

\bar{X} = média da amostra

n = tamanho da amostra

τ = desvio padrão conhecido da população, e

$\mu = \mu_1$ para H_0 verdadeira

- substituindo-se os valores na fórmula, obtém-se o z observado.
- o z teórico, que definirá a região crítica, é obtido consultando-se o tabela "Áreas sob a curva normal" considerando-se o valor de α .

Tipo de Teste: paramétrico

Adequação: para analisar dados obtidos com utilização do "DESIGN"]
- UM GRUPO-PÓS TESTE, quando a variância da população é conhecida.

SITUAÇÃO 2

Testar a significância da magnitude da média de uma população normalmente distribuída, cuja média e variância não são conhecidas.

$H_0: \mu = \mu_1$, em que μ_1 é o valor determinado proposto para μ .

Teste estatístico: teste t

Fórmula de t: $t = \frac{\bar{X} - \mu}{S \sqrt{n}}$, em que:

\bar{X} = média da amostra

n = tamanho da amostra

S = desvio padrão da amostra

$\mu = \mu_1$, para H_0 verdadeira

- substituindo-se os valores na fórmula, obtém-se o t reservado;
- o t teórico, que definirá a região crítica, é obtido consultando-se a tabela "valores de t de Student", para (n-1) graus de liberdade, considerando-se o valor α . O grau de liberdade, (degree of freedom) é simbolizado por d.f.

OBSERVAÇÃO:

GRAUS DE LIBERDADE: Dado um conjunto de "n" observações, cuja soma dos valores é igual a Q, tem-se que o valor de (n-1), destas observações são livres para variar e o valor de uma única observação fica pré-determinado, para que a soma dos valores, destas observações, mantenha-se igual a Q. Ao número de observações, cujos valores podem variar livremente, dá-se o nome de grau de liberdade.

Ex.: sejam quatro observações (n=4), cujos valores são respectivamente 8, 7, 3 e 5. Neste caso, $Q = 8 + 7 + 3 + 5 = 23$. Aqui, o valor de 3 (n-1) observações são livres de variar, o de uma fica pré-determinada para que a soma dos quatro valores mantenha-se igual a 23. Sejam 6, 2 e 9 os valores de três observações. Para que a soma não se altere, o valor da última observação fica pré-determinada.

$$6 + 2 + 9 = 17$$

$$23 - 17 = 6 \text{ (valor pré-determinado em função dos outros três valores).}$$

Tipo de teste: paramétrico

Adequação: para analisar dados obtidos com a utilização do "DESIGN" 1 - UM GRUPO - PÓS-TESTE, quando a média e variância da população não são conhecidas.

SITUAÇÃO 3

Testar a significância da diferença entre duas médias de duas populações, normalmente distribuídas, cujas médias não são conhecidas, e cujas variâncias são conhecidas. Estar-se-á, no caso, trabalhando com amostras selecionadas aleatoriamente e independentemente. Esta é uma situação rara, pois não é frequente conhecer-se a variância da população.

$$H_0: \mu_1 - \mu_2 = 0$$

Teste estatístico: teste z

$$\text{Fórmula de z: } z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\tau_1^2}{n_1} + \frac{\tau_2^2}{n_2}}}, \text{ em que}$$

\bar{x}_1 = média da amostra 1

\bar{x}_2 = média da amostra 2

n_1 = tamanho da amostra 1

n_2 = tamanho da amostra 2

τ_1 = desvio padrão conhecido da população 1

τ_2 = desvio padrão conhecido da população 2

$\mu_1 - \mu_2 = 0$, para H_0 verdadeira.

- substituindo-se os valores na fórmula obtém-se o z observado

- o z teórico, que definirá a região crítica é obtido consultando-se a tabela "Áreas sob a curva normal", considerando-se o valor α .

Tipo de teste: paramétrico

Adequação: este teste é adequado para a análise dos dados obtidos com a aplicação do "DESIGN" 4 - GRUPO DE CONTROLE SOMENTE PÓS-TESTE, em que é testada a diferença entre T_2 e T_1 , quando se conhece as variâncias das populações 1 e 2. Nas mesmas condições, é ainda possível sua utilização para o "DESIGN" 3 - GRUPO DO CONTROLE PRÉ-TESTE PÓS-TESTE, em que é testada a significância da diferença entre T_4 e T_2 .

SITUAÇÃO 4

Testar a significância da diferença entre duas médias de populações normalmente distribuídas, cujas médias e variâncias não são conhecidas. As amostras das populações serão selecionadas aleatória e independentemente.

$H_0: \mu_1 - \mu_2 = 0$

Teste estatístico: teste t

Fórmula de t:
$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$
 em que,

\bar{x}_1 e \bar{x}_2 = média das amostras

n_1 e n_2 = tamanho das amostras

S_1 e S_2 = desvio padrão das amostras

$\mu_1 - \mu_2 = 0$, para H_0 verdadeira

- substituindo-se os valores na fórmula obtém-se o t observado
- o t teórico, que definirá a região crítica, é obtido consultando-se a tabela "valores t de student", para $(n_1 + n_2 - 2)$ graus de liberdade, considerando-se o valor α .

Tipo de teste: paramétrico

Adequação: este teste é de adequação ótima para a análise dos dados obtidos com a aplicação do "DESIGN" 4 - GRUPO DE CONTROLE E SOMENTE PÓS-TESTE, em que é testada a significância da diferença $T_2 - T_1$. Sua utilização é também possível para "DESIGN" 3 - GRUPO DE CONTROLE PRÉ-TESTE PÓS-TESTE, em que é testada a significância da diferença $T_4 - T_2$, embora para este "design", a análise de covariância seja mais adequada.

SITUAÇÃO 5

Testar a significância da diferença entre duas médias para amostras correlacionadas. Isto é, duas amostras constituídas de um único grupo de sujeitos estudado, sob duas condições experimentais separadas. Seria o caso, por exemplo, de testar a significância da diferença das médias de um mesmo grupo de alunos no pré e no pós-teste. O estudo envolve portanto PARES de observações.

$$H_0: \mu_1 - \mu_2 = 0$$

↑ ↑
médias médias
no no
pós-teste pré-teste

Teste estatístico: teste t

Fórmula de t:
$$\frac{\sum_{i=1}^n D_i}{\sqrt{\frac{N \sum_{i=1}^n D_i^2 - (\sum_{i=1}^n D_i)^2}{N-1}}}$$
 em que

N = número de pares de observação

D = diferença entre os valores de cada par de observação

- substituindo-se os valores na fórmula obtém-se o t observado
- o t teórico, que define a região crítica é obtido consultando-se a tabela "valores t de student", para N-1 graus de liberdade, considerando-se o valor α .

Tipo de teste: paramétrico

Adequação: analisar dados obtidos com a utilização do "DESIGN" 2 - UM GRUPO - PRÉ-TESTE, PÓS TESTE, em que é testada a significância da diferença entre T_2 e T_1 .

SITUAÇÃO 6: Testar a significância da diferença entre duas ou mais médias de diferentes populações, normalmente distribuídas, cujas médias e variâncias não são conhecidas. As amostras das duas ou mais populações são selecionadas aleatoriamente e independentemente.

$$H_0: \mu_1 = \mu_2 = \dots \mu_n = 0$$

Teste estatístico = teste F (análise de variância "one-way").

$$\text{F\u00f3rmula de F: } F = \frac{\sum_{j=1}^K n_j (\bar{X}_j - \bar{X})^2}{\frac{\sum_{j=1}^K \sum_{i=1}^{n_j} (X_{ij} - \bar{X})^2}{N - K}}$$

em que,

K = n\u00famero de tratamentos (ou grupos);

n_j = n\u00famero de indiv\u00edduos no j \u00e9simo grupo;

N = n\u00famero total de indiv\u00edduos em todos os K grupos;

\bar{X}_j = a m\u00e9dia de observa\u00e7\u00f5es do j \u00e9simo grupo;

\bar{X} = a m\u00e9dia de todas observa\u00e7\u00f5es tomadas juntas;

X_{ij} = o i \u00e9simo elemento do j \u00e9simo grupo.

- substituindo-se os valores na f\u00f3rmula obt\u00e9m-se o F observado;

- O F te\u00f3rico, que definir\u00e1 a regi\u00e3o cr\u00edtica, \u00e9 obtido consultando-se a tabela "distribui\u00e7\u00e3o de F de Snedecor", para $K-1$ graus de liberdade associados com o numerador e $N-K$ graus de liberdade associado com o denominador, considerando-se o valor α .

Tipo de teste: param\u00e9trico

Adequa\u00e7\u00e3o: este teste \u00e9 adequado para os "designs" de um simples fator, completamente randomizados ou de blocos randomizados, em que \u00e9 testada a signific\u00e2ncia da diferen\u00e7a entre duas ou mais m\u00e9dias. Com somente duas m\u00e9dias ($K=2$), tal teste conduz ao mesmo resultado do que aquele obtido com a aplica\u00e7\u00e3o do teste t .

Observa-se que o teste F indica se existe diferen\u00e7a estatisticamente significativa entre as m\u00e9dias dos grupos, mas no caso de existir tal diferen\u00e7a, ele n\u00e3o identifica entre quais m\u00e9dias ela existe. Assim, se o valor F observado \u00e9 significativo, s\u00e3o requeridas an\u00e1lises adi

cionais para determinar precisamente quais grupos são significativamente diferentes. Em seguida a análise de variância, poderia ser utilizado o teste de Scheffé (aplicável quando o número de elementos de cada grupo são iguais ou diferentes) ou o Duncan Range Test (aplicável quando o número de elementos de cada grupo são iguais). Através do uso destes testes fica determinado se cada média é significativamente diferente de todas as outras médias, ou se há diferenças significantes entre algumas médias e não entre outras.

Observa-se que a análise de variância "one-way" é também útil para analisar os dados obtidos com a utilização do "DESIGN" 7 - UM GRUPO SÉRIE NO TEMPO. Neste caso, considera-se cada uma das médias de cada um dos testes (de T_1 até T_n) como representativa de um tratamento, e testa-se a significância das diferenças entre médias. No caso de uma diferença significativa apenas entre as médias dos testes imediatamente anterior e imediatamente posterior ao tratamento, poder-se-á concluir pela existência do efeito deste.

Salienta-se ainda que para os dados obtidos com a utilização do "DESIGN" 7 - UM GRUPO SÉRIE NO TEMPO existem vários modelos de análise estatística. Todos eles trabalhosos. Um procedimento mais grosseiro, e por isso mesmo muitas vezes muito útil é a análise gráfica. Seu uso, cercado de cautela, é recomendado.

SITUAÇÃO 7:

Testar simultaneamente para populações que se distribuem normalmente:

- a significância da diferença entre duas ou mais médias, representativas dos níveis de uma primeira variável independente;
- a significância da diferença entre duas ou mais médias, representativas das níveis de uma segunda variável independente;

- a significância do efeito de interação das duas variáveis independentes;

As amostras das populações são selecionadas randômica e independentemente.

$H_0: \mu_1 = \mu_2 \dots \mu_n = 0$ (n, nº de níveis de A: primeira variável independente)

$H_0: \mu_1 = \mu_2 = \dots \mu_m = 0$ (m, nº de níveis de B: segunda variável independente).

$H_0: \tau^2_{A \times B} = 0$

Teste estatístico: teste F - análise de variância. "Two Way" Fórmula de F.

Neste caso, não existe uma única fórmula para o teste. Cada uma das três hipóteses nulas testadas pela análise de variância "Two-Way" implica no uso de uma fórmula. Além disso, tais fórmulas podem variar de um experimento para outro em função dos níveis das variáveis, envolvidas no mesmo, serem fixos, aleatórios, ou mistos.

Devido a maior complexidade e diversidade de tais fórmulas, elas deixam de ser aqui apresentadas por irem além das perspectivas deste curso.

Tipo de teste: paramétrico

Adequação: este teste é adequado para analisar os dados obtidos com a aplicação DO "DESIGN" 5 - "DESIGN" FATORIAL.

SITUAÇÃO 8: Testar a significância da diferença entre médias de um pós-teste, quando, embora não tenha sido possível utilizar procedimentos experimentais para assegurar a homogeneidade inicial dos grupos, foram coletados dados de um pré-teste. Neste caso, um método estatístico poderá garantir,

quando da análise, o controle de variáveis até então não controladas. O nível inicial de desempenho dos alunos pode ser diferente. Esta diferença inicial entre os grupos será removida estatisticamente, e eles poderão, então, ser considerados homogêneos.

$$H_0: \mu_1 = \mu_2 = \dots \mu_n$$

Teste estatístico: teste F - Análise de Covariância

Fórmula de F para Análise de Covariância.

Sua complexidade vai além das perspectivas do curso.

Tipo de Teste: Paramétrico.

Adequação: A análise de covariância (ANACOVA) é de adequação ótima para analisar os dados obtidos com a aplicação do "DESIGN" 6 - GRUPO DE CONTROLE NÃO EQUIVALENTE. Neste caso, não tendo sido possível a designação aleatória dos sujeitos aos grupos, é aplicado um pré-teste. Quando, através da análise de covariância testa-se a diferença entre T_4 e T_2 , os valores T_3 e T_1 são considerados; removendo-se assim, provável diferença inicial entre os grupos. A análise de covariância é também adequada para analisar os dados obtidos com o "DESIGN" 3 - GRUPO DE CONTROLE PRÉ-TESTE - PÓS TESTE. No caso, testa-se a diferença entre T_4 e T_2 , removendo-se diferenças iniciais entre T_3 e T_1 , que apesar da designação aleatória dos sujeitos, possam ainda existir.

A Tabela F.32 que segue apresenta os testes estatísticos mais adequados para cada "design" de coleta de dados apresentados anteriormente.

TABELA F.32

TESTES ESTATÍSTICOS APROPRIADOS PARA CADA "DESIGN" DE COLETA DE DADOS

"DESIGNS"	TESTES ESTATÍSTICOS
1 - X T ₂ "DESIGN"UM GRUPO PÓS-TESTE	teste z - magnitude da média (variância da população conhecida) ou teste t - magnitude da média (variância da população não conhecida)
2 - T ₁ X T ₂ "DESIGN"UM GRUPO PRÉ-TESTE PÓS-TESTE	teste t - diferença das médias para amostras correlacionadas
3 - R T ₁ X T ₂ R T ₃ X' T ₄ "DESIGN" GRUPO DE CONTROLE PRÉ-TESTE PÓS-TESTE	teste z - diferença das médias (variância da população conhecida) teste t - diferença das médias (variância da população não conhecida). teste F - "One way" ANAVA ANACOVA
4 - R X T ₁ R X' T ₂ "DESIGN"GRUPO DE CONTROLE - SOMENTE PÓS-TESTE	teste z - diferença das médias ou (variância da população conhecida) teste t - diferença das médias (variância da população não conhecida) teste F - "One-Way" ANAVA (mais de dois grupos)
5 - "DESIGN" FATORIAL "NESTED" "NESTED FATORIAL"	teste F - "Two-Way" ANAVA
6 - T ₁ X T ₂ T ₃ X' T ₄ "DESIGN"GRUPO DE CONTROLE NÃO EQUIVALENTE	ANACOVA
7 - T ₁ T ₂ T ₃ T ₄ X T ₅ T ₆ T ₇ T ₈ "DESIGN" SÉRIE NO TEMPO	- "one-way" - ANAVA - análise gráfica

Neste passo cinco, então, deve ser elaborado o plano experimental para coleta e análise de dados. Neste plano, repetindo, possivelmente estarão incluídos: o plano de amostragem, a identificação das variáveis não experimentais que possam contaminar o experimento e o como controlá-las, a seleção de um "design" de pesquisa para coleta de dados, plano para designação dos sujeitos para os grupos e determinação dos tratamentos experimentais para os grupos, esquematização de procedimentos para coleta de dados, plano de construção dos instrumentais para medir os resultados do experimento, plano de testagem de hipótese.

O passo seguinte do esquema de Avaliação Somativa, baseado nos procedimentos da pesquisa experimental, é executar o experimento.

6) EXECUTAR O EXPERIMENTO

Neste passo, é executado o plano experimental de coleta de dados. São construídos os instrumentais de medida, é executado o plano de amostragem dos sujeitos pode ocorrer a designação dos sujeitos para os diferentes grupos e a determinação dos tratamentos para os mesmos, em função do "design" de coleta de dados selecionados. Ainda, em função deste "design", e dos procedimentos de coleta de dados esquematizados, são os dados finalmente coletados.

7) REDUZIR OS DADOS BRUTOS

Neste passo são computadas as estatísticas (média, desvio padrão, etc), que serão empregadas no teste estatístico que será utilizado.

8) COMPUTAR UM TESTE ESTATÍSTICO APROPRIADO PARA ORIENTAR AS DECISÕES ACERCA DA(S) HIPÓTESE(S) QUE ESTÁ(ÃO) SENDO TESTADA(S).

Neste passo, segundo o plano de testagem de hipóteses estabelecidas, é computado o teste estatístico. Os resultados deste, orientarão as decisões acerca das hipóteses que estão sendo testadas. Os pro

cedimentos computacionais de tais testes não serão tratados neste curso, pois são mais convenientemente conduzidos por estatísticos. Inclusive, mais comumente, esta é uma tarefa executada por computadores eletrônicos que fornecem apenas as saídas, ou resultados, obtidos com a aplicação de tais testes. Importante para o avaliador é poder interpretar estes resultados.

Como já referido anteriormente, o propósito do teste de hipótese é REJEITAR H_0 , para que se possa ACEITAR H_1 , uma vez que ambas são contrárias e complementares entre si. ACEITAR H_1 , implica em ACEITAR a hipótese do problema, uma vez que a hipótese alternativa é simplesmente uma operacionalização e transcrição, em símbolos matemáticos, desta. A hipótese nula será rejeitada se o valor observado da estatística do teste (valor computado) se encontrar dentro da região crítica, definida pelo valor teórico da estatística de teste. Caso contrário, a decisão estatística será NÃO SE PODE REJEITAR H_0 , o que implicará na decisão de que NÃO SE PODE ACEITAR H_1 .

Veja-se como exemplo de interpretação de resultados da aplicação de um teste estatístico, o seguinte:

a) As hipóteses do problema eram:

- Existe efeito do método de ensino na aprendizagem dos alunos do Projeto A (método de ensino: 2 níveis);
- Existe efeito da localização da escola na aprendizagem dos alunos do Projeto A (localização da escola 2 níveis);
- Existe efeito da interação método de ensino X localização da escola na aprendizagem dos alunos do Projeto A.

OBSERVAÇÃO:

A testagem de tais hipóteses implicou no planejamento de um "design" FATORIAL "TWO-WAY", (2 x 2), com 4 grupos.

b) As hipóteses alternativas correspondentes, ou seja, estas mesmas hipóteses escritas sob símbolos matemáticos eram:

- $H_1: \mu_1 \neq \mu_2$
- $H_2: \mu_3 \neq \mu_4$
- $H_3: \tau_{AB}^2 \neq 0$

c) As hipóteses nulas correspondentes eram:

- $H_0: \mu_1 = \mu_2$
- $H_0: \mu_3 = \mu_4$
- $H_0: \tau_{AB}^2 = 0$

Significando:

- Não existe efeito do método de ensino na aprendizagem dos alunos do Projeto A;
- Não existe efeito da localização da escola na aprendizagem dos alunos do Projeto A;
- Não existe efeito da interação método de ensino X localização da escola na aprendizagem dos alunos do Projeto A.

d) Os resultados do F observado e do F teórico para $\alpha = 0,05$ df do denominador: 120 e df dos numeradores: 2, 1 e 5 pela ordem são apresentados na Tabela F.33 que segue.

TABELA F.33

RESULTADOS DE UMA APLICAÇÃO DA ANÁLISE DE VARIÂNCIA "TWO-WAY"

FONTE	F OBS	F TEORICO
Método de ensino	82,04 *	3,07
Localização	4,28 *	3,91
Método X Localização	1,58	2,45

$\alpha = 0,05$

Neste teste F, deve-se REJEITAR H_0 quando o F observado for maior que o F teórico. Assim sendo, neste exemplo, deve-se rejeitar as duas primeiras hipóteses nulas, e não se pode rejeitar a terceira. As conclusões seriam, portanto:

- Para $\alpha = 0.05$, deve-se rejeitar a hipótese nula de que "não existe efeito do método de ensino na aprendizagem dos alunos do Projeto A", ou seja, deve-se ACEITAR a hipótese de que "existe efeito do método de ensino na aprendizagem dos alunos do Projeto A".
- Para $\alpha = 0,05$, deve-se REJEITAR a hipótese nula que "não existe efeito da localização da escola na aprendizagem dos alunos do Projeto A", ou seja, deve-se aceitar a hipótese de que "existe efeito da localização da escola na aprendizagem dos alunos do Projeto A".
- Para $\alpha = 0,05$, NÃO SE PODE REJEITAR A hipótese nula que "não existe efeito da interação método de ensino X localização da escola, na aprendizagem dos alunos do Projeto A", isto é, NÃO SE PODE afirmar que "existe efeito da interação método de ensino X localização da escola, na aprendizagem dos alunos do Projeto A".

Veja-se outro exemplo, quando o teste aplicado é o teste t de diferença de médias.

a) A hipótese do problema era:

- Existe efeito do método de ensino na aprendizagem dos alunos do Projeto B (método de ensino com 2 níveis: tecnológico e tradicional).

OBSERVAÇÃO:

A testagem de tal hipótese implicou na utilização do "DESIGN" GRUPOS EQUIVALENTES-SOMENTE PÓS-TESTE

R	X	T ₁
R	X'	T ₂

X → método tecnológico
X' → método tradicional

b) A hipótese alternativa era:

- $H_1: \mu_1 \neq \mu_2$

c) A hipótese nula era:

- $H_0: \mu_1 = \mu_2$

Significando:

- Não existe efeito do método de ensino na aprendizagem dos alunos do Projeto B.

d) Os resultados t observado e t teórico para $\alpha = 0,05$ e d.f. $(n_1 + n_2 - 2 = 16)$ foram:

t teórico: 2,12

t observado: 1,67

A Figura F.28 que segue, mostra que o t observado se encontra fora da região crítica, e que NÃO SE PODE REJEITAR H_0 .

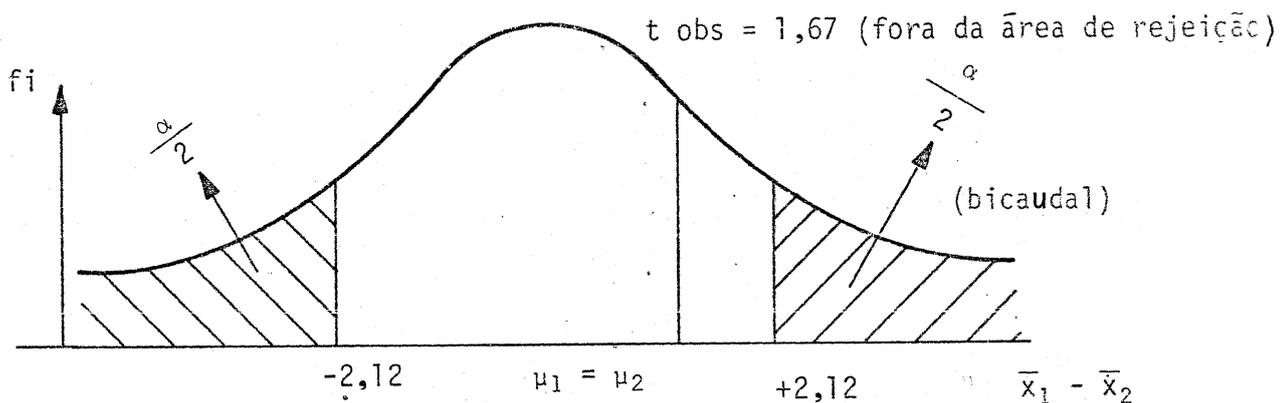


Fig. F.28 - Comparação Gráfica dos Resultados Teórico e Prático de uma Aplicação do Teste t de Student, para $H_0: \mu_1 = \mu_2$.

Assim, a conclusão seria:

Para $\alpha = 0,05$ NÃO SE PODE REJEITAR a hipótese nula de que "não existe efeito do método de ensino na aprendizagem dos alunos do Projeto B", ou seja, NÃO SE PODE AFIRMAR que "existe efeito do método de ensino na aprendizagem dos alunos do Projeto B". A diferença dos efeitos produzidos pelos dois métodos (Tecnológico e Tradicional) não foi significativa ao nível 0,05.

Veja-se outro exemplo de interpretação de resultados de um teste estatístico. No caso, o teste t de magnitude da média.

a) A hipótese do problema era:

- o tempo médio requerido para execução do teste padronizado \bar{X} é maior que 130 minutos.

b) A hipótese alternativa era:

- $H_1: \mu > 130'$

c) A hipótese nula era:

- $H_0: \mu \leq 130'$

Significando:

- o tempo médio requerido para execução do teste padronizado \bar{X} é menor ou igual a 130 minutos.

d) Os resultados t teórico e t observado para $\alpha = 0,05$ e $(26 - 1 = 25)$ graus de liberdade foram:

t teórico = 2,79 t observado = 6,59

Isto significando que o t observado se encontra na região crítica, e que deve-se REJEITAR H_0 , conforme a Figura F.27 que segue.

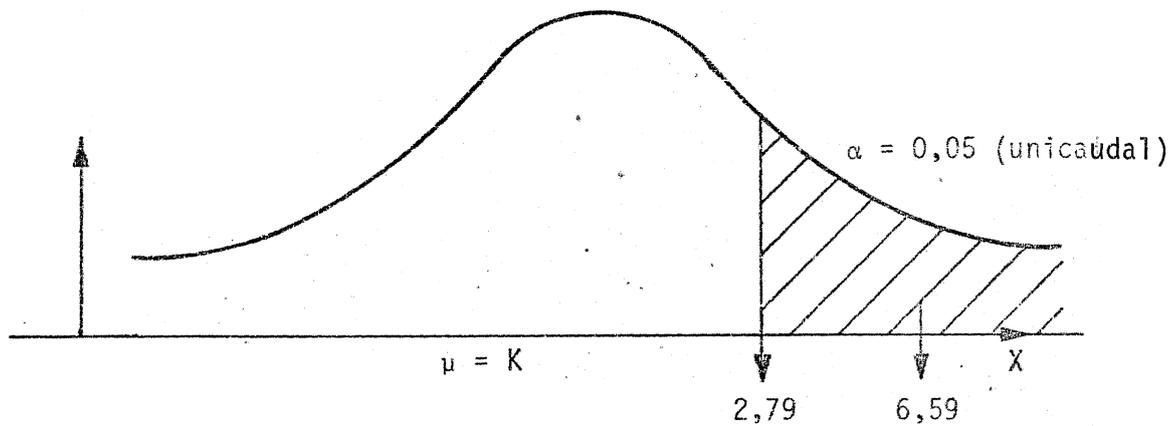


Fig. F.29 - Comparação gráfica dos resultados teórico e prático de uma aplicação do teste t de Student para $H_0: \mu \geq K$

Assim, a conclusão seria:

Para $\alpha = 0,05$, deve-se REJEITAR a hipótese nula de que o "tempo médio requerido para execução do teste padronizado X é menor ou igual a 130 minutos", ou seja, deve-se ACEITAR a hipótese de que o tempo médio requerido para execução do teste padronizado X é maior do que 130 minutos.

Veja-se mais um exemplo de interpretação de resultados de um teste estatístico. No caso, Teste F, utilizado para análise de variância "ONE-WAY", dos dados obtidos com "DESIGN" GRUPOS EQUIVALENTES SOMENTE PÓS-TESTE, com 4 grupos:

R	X	T_1
R	X'	T_2
R	X''	T_3
R	X'''	T_4

a) A hipótese do problema era:

- existe efeito do nível do professor em sua atitude com relação ao Projeto X.

OBSERVAÇÃO: A variável independente tendo quatro níveis: primário incompleto, primário completo, ginásio incompleto, ginásio completo.

b) A hipótese alternativa era

$$H_1: \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$$

c) A hipótese nula era:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$$

Significando:

- Não existe efeito do nível do professor na sua atitude com relação ao Projeto X.

d) Os resultados do F observado e F teórico para $\alpha = 0,05$ graus de liberdade do denominador = 60, e graus de liberdade do numerador = 3; são apresentados na Tabela F.34 que segue.

TABELA F.34

RESULTADOS DE UMA APLICAÇÃO DA ANAVA "ONE-WAY".

FONTE	F OBSERVADO	F TEÓRICO
Nível do Professor	69,98 *	2,76

$$\alpha = 0,05$$

De onde se conclui que:

Para $\alpha = 0,05$, deve-se REJEITAR A HIPÓTESE NULA de que "não existe efeito do nível do professor na sua atitude com relação ao Projeto X", ou seja, deve-se aceitar a hipótese de que "existe efeito do nível do professor em sua atitude com relação ao Projeto X".

5.2.2 - COMENTÁRIO ACERCA DO ESQUEMA DE AVALIAÇÃO SOMATIVA BASEADO NOS PROCEDIMENTOS DA PESQUISA EXPERIMENTAL.

O esquema aqui apresentado é de grande utilidade ao avaliador no exercício de suas atividades de delinear, obter e prover informações úteis aos decisores para o julgamento de mudanças instrucionais alternativas, quando este julgamento é baseado nos efeitos produzidos por tais mudanças na aprendizagem dos alunos. Assim, quando o decisor estiver interessado em conhecer o "efeito líquido" da instrução, isto é, seu efeito "puro", o avaliador para fornecer-lhe informação útil deverá lançar mão do uso do esquema de Avaliação Somativa baseado na Pesquisa Experimental.

Quando uma mudança instrucional é colocada em ensaio, é de se esperar que a informação requerida pelo decisor, seja de que são realmente os procedimentos instrucionais novos, introduzidos pela mudança que ocasionaram tais e tais efeitos. O estabelecimento desta relação causa (procedimentos) - efeito (resultados) só é permitido através do uso deste modelo.

Devido ao desejo de estabelecer esta relação causa-efeito, há necessidade de um certo controle da situação em que se testa a inovação. Este artificialismo intrínseco, ao modelo, é muitas vezes combatido por alguns autores que afirmam que em Avaliação é muito importante a consideração da série de fatores que ocorrem naturalmente (Worthen e Sanders, 1973). Mas, quando o decisor está interessado neste tipo de informação relativa ao efeito "puro" da instrução, nenhum conjunto de procedimentos é mais adequado que aqueles baseados na Pesquisa Experimental. Além disso, estando os avaliadores interessados em relações causa-efeito, entre procedimentos instrucionais e resultados produzidos no sistema devido a tais procedimentos, é, para tais avaliadores suficiente, fornecer aos decisores informação acerca de que alguma coisa que acompanha a implementação da intervenção instrucional A (e não uma influência estranha, não controlada, não relacionada a esta inovação), é responsável pelos resultados. Isto, não exige do estudo aquela forma de rigidez necessária à Pesquisa, em que existe a preocupação de investigar, dentro de um

complexo de fatores, aquele ou aqueles elementos fundamentais na relação causa e efeito, em estudo. A flexibilidade permitida pela utilização deste esquema depende em grande parte da flexibilidade do planejamento dos procedimentos da inovação instrucional colocada em teste.

5.3 - SUB UNIDADE 2 - UM ESQUEMA DESCRITIVO DA AVALIAÇÃO SOMATIVA DOS PRODUTOS DA MUDANÇA

Como já salientado anteriormente, ultimamente, modelos descritivos não baseados totalmente nos procedimentos de Pesquisa Experimental, também, vêm sendo utilizados com frequência para a Avaliação Somativa dos resultados de uma mudança instrucional. O aparecimento destes modelos se deu em virtude do tipo de informação exigida pelos decisores para julgar o valor destes esforços de mudança. Tais informações se referem à efetividade ou à eficiência destes esforços.

A efetividade de uma mudança é a medida do grau de alcance dos seus objetivos, e ela será tanto mais efetiva quanto maior for este grau. A efetividade da mudança relaciona-se, portanto, às SAÍDAS produzidas pela mesma. Assim, definidos os objetivos que se deseja atingir com um esforço de mudança, introduzida tal mudança e medido o alcance dos objetivos, temos a EFETIVIDADE de tal esforço. Embora a Avaliação absoluta seja mais adequada, quando se trabalha com Índices de efetividade, existe a possibilidade de realizar-se a Avaliação Comparativa, confrontando-se resultados de duas ou mais alternativas desenvolvidas para conduzir ao alcance de objetivos comuns.

A eficiência de uma mudança é a medida do seu rendimento, ou seja, da relação entre SAÍDA/ENTRADA produzida. Relaciona, por exemplo, a aprendizagem (saída) e o custo desta aprendizagem (entrada) ou ainda aprendizagem (saída) e o tempo requerido para esta aprendizagem. A mudança será mais eficiente quando exigir um número de entrada para uma mesma saída, ou quando produzir maior saída com uma mesma entrada.

Quando se dispõe de mudanças alternativas para atingir o mesmo objetivo, a eficiência de cada uma delas pode ser comparado, orientando uma escolha entre as mesmas.

O Esquema de Avaliação Somativa que aqui será apresentado, trata da Avaliação de esforços de mudança em termos de sua efetividade. Vai além das perspectivas deste curso, a determinação da eficiência dos mesmos. Salienta-se, apenas, que a efetividade é elemento componente da eficiência, e portanto parte do processo da determinação desta eficiência será envolvido.

Questões do tipo: "qual a efetividade da alternativa X?" "os objetivos da mudança estão sendo atingidos ao nível de um padrão absoluto pré-determinado?", são respondidas através do uso deste modelo.

Dentre os produtos da mudança restringir-se-á, este esquema, ao produto aprendizagem dos alunos.

Uma distinção clara deve ser feita entre esta testagem para determinação da efetividade de uma inovação, e outras testagens prévias realizadas para embasar revisões (Avaliação Formativa).

5.3.1 - PASSOS DO ESQUEMA

- 1 - Definir os objetivos educacionais da mudança;
- 2 - Redefinir tais objetivos de forma operacionalizada;
- 3 - Identificar a questão a ser respondida pela Avaliação;
- 4 - Elaborar um plano para coleta de dados;
- 5 - Identificar um padrão absoluto de efetividade;
- 6 - Executar o plano de coleta de dados;
- 7 - Determinar a efetividade da mudança;
- 8 - Comparar a efetividade da mudança com o padrão absoluto pré-determinado.

Nos tópicos seguintes serão apresentados, em detalhes, cada um dos passos que compõem o Esquema Descritivo de Avaliação Somativa dos Produtos da Mudança.

1) DEFINIR OS OBJETIVOS EDUCACIONAIS DA MUDANÇA

Quando uma mudança é introduzida num sistema escolar, ela existe com o propósito de atingir a determinados objetivos. Tais objetivos, diferentes para intervenções específicas, representam o propósito do sucesso da intervenção.

Neste passo, são identificados os objetivos da mudança, ou seja, o que é esperado dos alunos após terem experienciado a intervenção. Aqui, tais objetivos se apresentam ainda em termos conceituais, ambíguos e vagos.

Ex.: OBJETIVO: Melhorar a aprendizagem de Matemática.

Esta tarefa não é específica do Avaliador. Ela é realizada em interface com os responsáveis pelas decisões de planejamento no sistema.

2) REDEFINIR OS OBJETIVOS DE FORMA OPERACIONAL

A operacionalização dos objetivos da mudança resulta na definição dos objetivos comportamentais, que expressam os comportamentos observáveis que o aluno demonstrará, se os objetivos da mudança forem alcançados. Através do comportamento observável do aluno, infere-se a aquisição, pelo mesmo, de determinadas habilidades.

Ex.: OBJETIVO COMPORTAMENTAL: Dados 10 números, o aluno deverá identificar números fracionários, grifando-os.

Neste objetivo está explícito um comportamento observável do aluno através do qual pode-se inferir que ele possui determinada habilidade (se o aluno grifar corretamente os números fracionários, pode-se

inferir que ele tem o conceito de números fracionários). O desempenho pelo aluno de tal comportamento demonstrará que ele atingiu o objetivo.

A definição operacional dos objetivos conduz o avaliador para termos mais concretos com os quais ele pode trabalhar.

Neste passo, é obtida uma listagem dos objetivos, escritos de forma operacional, pretendidos pela mudança e que se constituirão a base para determinação da efetividade da mesma.

Também esta, não é uma tarefa específica do avaliador. Ela é realizada em interface com os responsáveis pelas decisões do planejamento instrucional.

3) IDENTIFICAR A QUESTÃO A SER RESPONDIDA PELA AVALIAÇÃO

Este passo consiste na identificação pelo avaliador ou em interface com o decisor, da questão que este último quer respondida. De modo geral, as questões que este modelo permite responder são relativas a efetividade de uma mudança instrucional que está sendo testada.

Ex.: "Qual a efetividade da mudança X?"

4) ELABORAR UM PLANO PARA COLETA DE DADOS

Neste passo, inicialmente, quando não se estiver trabalhando com toda clientela alvo, haverá necessidade da elaboração de um plano de amostragem dos alunos para representá-la. Isto deverá ser feito conforme já descrito no esquema anterior, apenas salientando que aqui é mais usual do que lá, o uso de amostragem não probabilística. A amostragem intencional é comumente utilizada em estudos destinados a avaliar a efetividade de mudanças instrucionais. Ao identificar a população alvo, e a partir desta sua amostra representativa, é importantíssima a consideração dos pré-requisitos em termos de habilidades, que tais alunos são supostos apresentar. Claro que, quando possível, seria mais desejável tirar uma amostra aleatória desta população alvo.

A tarefa seguinte é selecionar um "design" de coleta de dados. Como este esquema é basicamente um esquema de Avaliação para decisão Absoluta, o design mais adequado é o "DESIGN" UM GRUPO PRÉ-TESTE - PÓS-TESTE, apresentado no esquema anterior (vide página 475).

Neste passo será ainda planejada a construção dos instrumentais para medir os resultados. Sendo este um modelo que visa determinar a efetividade de uma mudança, obviamente os instrumentais serão construídos com base nos objetivos comportamentais, forma operacionalizada dos objetivos da mudança. A utilização dos procedimentos de Medida de Critério permitirão a evidência da mudança comportamental, refletindo a efetividade da instrução.

Os itens de teste, escritos a partir de objetivos comportamentais, são designados para medir o que o estudante pode ou não pode fazer, em termos do que está estabelecido no objetivo, revelando, deste modo, as habilidades que ele domina, refletindo os efeitos da exposição dos alunos à intervenção educacional em ensaio.

Os procedimentos específicos de medida de critérios são apresentados detalhadamente na Unidade 3.0.

Finalmente, deverão ser esquematizados os procedimentos para coleta de dados, em detalhes suficientes para que seus aspectos essenciais possam ser reproduzidos em outras ocasiões, se desejado. O quando e o como serão coletados os dados deverão ser especificados, incluindo o tamanho do grupo de estudantes, duração dos testes, tempo de instrução, disposição física dos ambientes de estudo, etc.

5) IDENTIFICAR UM PADRÃO ABSOLUTO DE EFETIVIDADE

A Avaliação Somativa de uma intervenção instrucional implica no fornecimento de informações, para o julgamento desta intervenção em termos de seus resultados finais, como produto acabado. A informação que o avaliador fornece ao decisor lhe é útil para seu julgamento quando

além de conter dados descritivos dos resultados da intervenção em teste, contém também uma comparação destes dados com um padrão absoluto de excelência, pré-determinado. Assim, é necessário que neste passo seja determinado um padrão absoluto — nível de desempenho que tem valor de referência amplamente difundido, ou, comumente aceito — em comparação com o qual será julgada a efetividade da mudança.

É bom salientar, aqui, que esta informação absoluta, relativa a comparação dos resultados de uma intervenção com padrões pré-estabelecidos, serve não só para avaliar uma única alternativa, mas pode também servir como informação comparativa, orientando a escolha entre alternativas. Quando duas ou mais intervenções alternativas estão sendo avaliadas, para que seja escolhida a melhor delas em termos de sua efetividade, este modelo pode ser utilizado, devendo ao final ser escolhida a intervenção que apresentar os maiores índices relativos ao grau de alcance dos objetivos.

Um padrão absoluto pode ser fixado com base em resultados obtidos por outros sistemas, pelo próprio sistema em tempos anteriores, ou ainda, em resultados idealizados pelos programadores da mudança. Assim, por exemplo, ao se avaliar um novo currículo de Matemática, o padrão pode ser fixado com base nos resultados obtidos pelos currículos de Matemática de outros sistemas, pelo velho currículo do sistema em que está sendo introduzida a modificação curricular, ou ainda, em opiniões que os seus programadores têm a respeito de qual deveria ser sua efetividade. Neste último caso, embora exista um consenso geral entre os educadores, que o objetivo maior, ao se introduzir uma mudança num sistema instrucional é o alcance da excelência, não existe ainda, hoje, um consenso a respeito de quais seriam os padrões de excelência (Sanders, 1973). Uma das dificuldades, quando da utilização deste modelo, é exatamente a inexistência de padrões absolutos com os quais comparar os desempenhos das mudanças em teste. É papel do avaliador, neste passo, identificar, em interface com os decisores, os padrões absolutos de efetividade, que serão utilizados como base de comparação para o julgamento, em cada caso particular em estudo.

Devido a inexistência de "valores ótimos consagrados" há o perigo de que através da adoção de um padrão de comparação inadequado sejam subestimados ou superestimados os resultados. A seleção do próprio padrão já implica num julgamento. O avaliador é responsável em conduzir os decisores a um consenso acerca de que padrões adotar. Experiência, conhecimento de outros sistemas, etc., capacitam os responsáveis pelas mudanças instrucionais a decidirem acerca de tais padrões. No entanto, muita pesquisa a respeito dos mesmos se faz necessária ainda. Um exemplo de um "padrão já consagrado" quando a instrução é completamente individualizada é o quociente 100/100 indicando que ao final da instrução 100% dos alunos deverão ter atingido 100% dos objetivos.

Salienta-se, novamente, que no caso de uso de um "design" de coleta de dados com apenas um grupo, somente através da comparação dos resultados da mudança com um padrão de desempenho aceitável dos alunos, poder-se-á concluir acerca da efetividade da mesma. Sem o padrão de comparação, não existe julgamento, e sem julgamento não existe Avaliação.

6) EXECUTAR O PLANO DE COLETA DE DADOS

Neste passo, em função do planejado no passo quatro, serão elaborados os instrumentais de medida; poderão ser selecionadas as amostras de sujeitos; sendo finalmente coletados os dados, segundo o "design" de coleta escolhido e a esquematização dos procedimentos de coleta de dados.

7) DETERMINAR A EFETIVIDADE DA MUDANÇA

Como já apresentado anteriormente, a efetividade de uma mudança é a medida do grau de alcance dos seus objetivos. Esta medida do alcance dos objetivos é possibilitada, pelas situações de testagem. Nos testes de aprendizagem, cada item de teste mede apenas um objetivo. Diz-se que um aluno atingiu um determinado objetivo quando, existindo um item para medir este objetivo, ele acerta este item, ou então existindo dois itens para medir este objetivo ele acerta ambos (dois itens por objetivo garantem maior precisão à medida). Quando se trabalha com efetivi

dade da instrução, a correção dos testes deve ser feita em termos de objetivos alcançados. O escore do aluno será diretamente proporcional ao número de objetivos atingidos, ao invés de ao número de itens acertados.

A efetividade de uma mudança instrucional pode ser computada através dos quatro métodos seguintes (Stone, 1973):

a) Usando dados médios do pós-teste

A efetividade é determinada através do quociente \bar{X}/p em que:

\bar{X} : representa o escore médio do grupo no pós-teste;

p : representa a percentagem de alunos, do grupo, que atingir com sucesso todos os objetivos;

b) Usando dados individuais do pós-teste:

Através deste método a efetividade é determinada através do quociente X/Y em que:

X : representa a percentagem de alunos do grupo que atingiu determinada percentagem dos objetivos;

Y : representa a percentagem de objetivos atingidos por tais alunos;

Assim, se $\frac{X}{Y} = \frac{60}{80}$ tem-se que 60% dos alunos atingiu 80%, ou mais, dos objetivos.

c) Usando dados individuais do pré e pós-testes (ganho individual)

A pré-testagem dos alunos, antes deles se submeterem à mudança, e a pós-testagem, depois desta, produzem dados que substanciam as asserções acerca do valor da experiência em si mesma. O ganho entre os testes indica a melhoria obtida durante o curso de instrução.

No caso, a questão essencial respondida seria:

"A mudança ocasionou acréscimo de conhecimentos e habilidades como especificados pelos objetivos desenvolvidos pela instrução?"

O ganho individual envolve, para cada aluno, o quociente entre $\frac{\text{ganho observado}}{\text{ganho possível}} \times 100\%$. Isto é operacionalizado através da fórmula:

$$\left(\begin{array}{l} \text{Índice de ga} \\ \text{nho Indivi-} \\ \text{dual:} \end{array} \right) = \frac{(\text{escore no pós-teste}) - (\text{escore no pré-teste})}{(\text{escore máximo possível no pós-teste}) - (\text{escore no pré-teste})} \times 100\%$$

Ex.: Se um aluno tira 3 no pré-teste e 8 no pós-teste, e a nota máxima possível é 10, seu índice de ganho será:

$$\left(\begin{array}{l} \text{Índice de ga} \\ \text{nho indivi-} \\ \text{dual} \end{array} \right) = \frac{8 - 3}{10 - 3} \times 100 = \frac{5}{7} \times 100\% = 71$$

Este resultado implica que o aluno ganhou 71% do que era possível, para ele ganhar.

Aplicando-se esta fórmula, para cada um dos alunos do grupo, obtêm-se o ganho de cada pessoa. Aplicando-se esta fórmula para todos os alunos e tirando-se a média destes ganhos, tal média pode expressar o índice de ganho global, ou seja, a efetividade da instrução modificada. Neste caso, um ganho de 71% implicaria que o grupo como um todo, ganhou 71% do que era possível, para ele, ganhar.

d) Usando as médias do pré e pós-testes:

Esta é uma versão modificada do método acima. Implica também na computação do ganho, apenas que, enquanto o método anterior é aplicável aos escores individuais, este método se refere aos escores médios do grupo.

$$\left(\text{Índice de ganho global} \right) = \frac{\left(\text{m\u00e9dia do grupo no p\u00f3s-teste} \right) - \left(\text{m\u00e9dia do grupo no pr\u00e9-teste} \right)}{\left(\text{m\u00e9dia m\u00e1xima poss\u00edvel no p\u00f3s-teste} \right) - \left(\text{m\u00e9dia do grupo no pr\u00e9-teste} \right)} \times 100\%$$

A utiliza\u00e7\u00e3o do \u00cdndice de ganho global tem uma vantagem sobre a utiliza\u00e7\u00e3o do \u00cdndice de ganho individual, porque enquanto para o primeiro s\u00e3o considerados todos os dados do pr\u00e9-teste e todos os dados do p\u00f3s-teste no segundo, s\u00f3 existe possibilidade de utilizar os dados dos alunos que realizarem ambos: pr\u00e9 e p\u00f3s-teste.

7) COMPARAR A EFETIVIDADE DA MUDAN\u00c7A COM O PADR\u00c3O ABSOLUTO PR\u00c9-DE-TERMINADO

Neste passo, os dados descritivos relativos \u00e0 efetividade da mudan\u00e7a s\u00e3o comparados com padr\u00f5es absolutos pr\u00e9-determinados, a fim de verificar a congru\u00eancia ou discrep\u00e2ncia entre ambos os conjuntos.

Comparando-se dados do p\u00f3s-teste, verifica-se, se um grupo teve melhor desempenho do que aquele estabelecido pelo padr\u00e3o. Comparando-se os ganhos verifica-se, se um grupo teve mais progresso do que aquele fixado pelo padr\u00e3o.

A informa\u00e7\u00e3o decorrente desta compara\u00e7\u00e3o entre resultados reais e padr\u00f5es pr\u00e9-determinados, ser\u00e1 \u00fatil para a tomada de decis\u00e3o acerca da validade da mudan\u00e7a, ou seja, sua capacidade de conduzir determinada clientela, ao alcance dos objetivos propostos. No caso de estarem sendo comparadas mudan\u00e7as alternativas poss\u00edveis, a informa\u00e7\u00e3o obtida, atrav\u00e9s do uso deste modelo, possibilitar\u00e1 a tomada de decis\u00e3o acerca de qual a alternativa mais v\u00e1lida, considerando-se os objetivos propostos que se pretende alcan\u00e7ar.

5.3.2 - COMENTÁRIO ACERCA DO ESQUEMA DESCRITIVO DE AVALIAÇÃO SOMATIVA DOS PRODUTOS DA MUDANÇA

Este é um esquema muito mais flexível que o primeiro. En quanto aquele tem como propósito maior, a verificação de hipóteses cau sais, e portanto exige procedimentos de coletas e análise de dados que não somente reduzam o viés e aumentam a precisão dos resultados, mas tam bém possibilitem inferência a respeito de causalidade, este segundo mode lo exige procedimentos menos rígidos, suficientes, apenas, para a redução do viés e o aumento da precisão. Com isto, pode haver um controle muito menor das condições em que é testada a inovação. A situação de ensaio torna-se mais natural, isto é, identifica-se mais com a situação real de sala de aula.

Ambos os esquemas são utilizados pelos avaliadores como instrumentos que os auxiliam no processo de fornecer informações úteis para a tomada de decisão. O primeiro esquema através do teste de hipóte se estatística, que lhe é inerente traz implícito o próprio critério de decisão. O segundo esquema enfatiza a definição de padrões e o julgamen to do valor ou mérito de uma alternativa instrucional, a partir da infor mação comparativa entre padrões e resultados reais obtidos, o que orien ta a decisão.

Este esquema é mais fácil de implementar, não exigindo grande número de pessoas com especialização em Pesquisa e Estatística.

Suas limitações são relativas ao tipo de informação que oferece. Enquanto o primeiro esquema possibilita uma resposta a ambos os tipos de questões - "os objetivos de projeto estão sendo atingidos" e - "existe efeito de tais procedimentos instrucionais na aprendizagem dos alunos", o segundo esquema responde apenas à questão do primeiro tipo, embora, neste caso, forneça informações mais detalhadas e significativas a respeito.

De qualquer modo, uma vez que o papel do avaliador é for necer aos decisores informações que lhes serão úteis, na escolha do seu

esquema de Avaliação, ele deverá levar em conta o tipo de in formação requerida pelo decisor.

Este esquema é caracteristicamente útil quando do desen- volvimento de uma inovação. Tal desenvolvimento pode ser visto como con sistindo de três processos básicos: planejamento do protótipo do produ to, produção do produto, e testagem do produto. O terceiro processo, tes tagem do produto, caracteriza-se por ensaios e revisões sucessivas do elemento produzido (Avaliação Formativa) e finalmente pelo ensaio final, para determinar a efetividade deste elemento (Avaliação Somativa). É nesta última etapa de Avaliação do produto acabado, que este esquema en contra sua plena utilidade.

6.0 - UNIDADE 6 - A ESTRUTURA DE UM PLANO DE AVALIAÇÃO

A metodologia é fundamental em toda a atividade educa tiva. É básico que se estabeleça uma metodologia clara e racional para coleta, organização e análise da informação exigida para a Avaliação de situações educacionais. Assim, uma vez que o avaliador tenha optado por uma abordagem avaliativa (no caso presente se optou, basicamente, pela de Stufflebeam et al 1971) ele precisa, em seguida, selecionar ou desenvolver um Plano de Avaliação.

Mas, como organizar um Plano de Avaliação?

Esta é uma tarefa difícil, desde que existem poucos es quemas de Planos de Avaliação que sejam adequados para alcançar as nece sidades emergentes da Avaliação. Assim, o avaliador precisa, geralmente, desenvolver o seu próprio esquema.

Muitos guias gerais, úteis, têm sido elaborados pelos es tudiosos de Avaliação. Aqui será apresentado um que tenta sintetizar os procedimentos a serem seguidos.

Estes procedimentos gerais são somente exemplos daquilo que deve ser considerado no planejamento de um estudo de Avaliação.

A Tabela F.35 é recomendada como um guia para planos de Avaliação. É essencial a consideração pelo avaliador de cada uma das suas etapas, que consistem, por sua vez, na execução dos três passos básicos do processo de Avaliação: delinear, obter e fornecer a informação.

É importante notar que a estrutura lógica do Plano de Avaliação é a mesma para qualquer tipo de Avaliação, seja Contexto, Input, Processo ou Produto, ou outros tipos de acordo com o autor da abordagem adotada. É bom notar, também, que para cada tipo de Avaliação que atenderá a um tipo especificado de decisão, é necessário um Plano de Avaliação.

A Tabela F.35 visualiza um esquema de um plano de Avaliação, cujos passos são descritos a seguir.

1) Descrição do Sistema que será Avaliado

O Sistema deve ser definido em termos de seus limites, seus elementos e as características destes elementos. Os limites se referem à localização e abrangência do sistema. Por exemplo, os limites em uma situação particular podem ser um Mestrado em Educação de uma determinada Universidade Federal. Os elementos, segundo Hammond (1973) se referem a:

- Variáveis Instrucionais:

- . Conteúdo: estrutura ou corpo de conhecimentos que constituem a instrução;
- . Metodologia: atividades de ensino, tipo de interação entre professores e alunos e meios, princípios de aprendizagem para motivação, prática e reforço;
- . Organização: organização vertical, que serve para classificar os estudantes e promovê-los desde a admissão até a saída; organização horizontal, que distribui estudantes entre professores; tempo, que envolve tanto a duração como a sequência;

TABELA F.35
UM ESQUEMA DE UM PLANO DE AVALIAÇÃO

PASSOS DE UM PLANO DE AVALIAÇÃO	
OS PASSOS BÁSICOS DO PROCESSO AVALIATIVO	DELINEAR
	OBTER
	FORNECER

1) Descrição do Sistema que será avaliado (caso necessite). a) Definição do Sistema (limites, elementos e características destes elementos). b) Identificação do tipo de decisão e níveis a serem servidos. c) Identificação dos objetivos do Sistema. d) Definição de critérios para cada momento de decisão.
2) Objetivos do Estudo de Avaliação. . Qual(is) será(ão) o(s) produto(s) do estudo de Avaliação?
3) Restrições do estudo de Avaliação que será realizado (tempo, dinheiro).
4) Questões a serem respondidas pela Avaliação e informação exigida para respondê-la.
5) Coleta da Informação a) Identificação das fontes da Informação a ser coletada. b) Identificação dos instrumentos e técnicas par. coletar a informação necessária. c) Identificação dos procedimentos de amostragem a ser empregados. d) Identificação das condições e o esquema para a coleta da informação.
6) Análise da Informação . Especificação dos procedimentos analíticos a serem empregados.
7) Relatório da Informação a) Definição das audiências para os relatórios da informação. b) Especificação dos meios para fornecer as informações para as audiências. c) Especificação do formato para os relatórios da Avaliação e/ou partes dos mesmos. d) Esquema Mestre para relatar a informação.

- . Facilidades: são o espaço, equipamento especial, e o consumo necessário para o Sistema Educacional;
- . Custo: dinheiro exigido para facilidades, manutenção e pessoal.

- Variáveis Populacionais:

- . São relativas à população direta ou indiretamente envolvida no Sistema. Uma classificação inclui estudantes, professores, administradores, especialistas educacionais, família e comunidade.

- Variáveis Comportamentais:

- . São relativas às reações das pessoas que estão na situação instrucional - domínio cognitivo, afetivo ou psicomotor.

Estes elementos ou variáveis, depois de identificados, devem ser caracterizados. Assim, tomando variáveis populacionais como exemplo, suas características podem ser definidas incluindo: dados de identificações gerais de estudantes, tais como: idade, nível de escolaridade, sexos, dados de identificação de professores, administradores e especialistas educacionais, tais como: background educacional, experiência profissional, salário, características gerais das famílias dos alunos, tais como: tamanho, residência, educação; características da comunidade, tais como: situação geográfica, população, estrutura sócio-econômica.

A identificação do tipo de decisão se refere ao momento de decisão que o Sistema enfrenta: Planejamento, Estruturação, Implementação ou Reciclagem.

A identificação dos níveis a serem servidos se refere a identificação dos decisores envolvidos no Sistema, a quem deve ser fornecida a informação.

A identificação dos objetivos do sistema deve ser feita de uma maneira operacionalizada, ou seja, eles devem ser precisos (todas as pessoas envolvidas no Sistema devem entendê-los do mesmo modo), e na medda do possível mensuráveis.

A definição de critérios inclui a especificação de variáveis para medida, e padrões para uso no julgamento do que está sendo medido. Assim, por exemplo, um critério para avaliar materiais instrucionais é a efetividade (a consecução dos objetivos) que já indica a variável a ser medida - aprendizagem -, e o padrão, que representará a base de comparação para o julgamento, neste caso, é estabelecido em termos do índice x/y , onde x representa a percentagem que alunos que atingiu uma percentagem especificada dos objetivos, e y a percentagem especificada de objetivos atingidos.

2) Objetivos do Estudo de Avaliação

Qual(is) o(s) produto(s) do estudo de Avaliação?

Os objetivos da Avaliação devem ser estabelecidos meticolosa e cientificamente. O requisito básico deste passo metodológico é o de deixar registrado o que se vai avaliar e com que finalidade. Estes objetivos são estabelecidos em função do tipo de decisão que será servido, que determinará o tipo de Avaliação a ser realizado. Utilizando a abordagem de Stufflebeam et al. (1971), se determinará objetivos de acordo com Contexto, Input, Processo ou Produto.

3) Restrições do Estudo de Avaliação que será realizado

É importante a identificação de restrições como tempo e dinheiro, que comprometem o planejamento da Avaliação.

4) Questões a serem respondidas pela Avaliação e Informação Exigida para respondê-las

As questões a serem respondidas pela Avaliação e a informação exigida para respondê-las, são gerados do tipo de decisão a ser servido e do tipo de Avaliação a ser realizado.

5) Coleta de Informação

A identificação das fontes de informação é o primeiro passo da coleta. Estas fontes de informação devem ser definidas em dois aspectos: primeiro, as origens para a informação, isto é, estudantes, professores, diretores ou pais etc., e segundo o estado presente da informação, isto é, se ela já existe registrada ou não.

Em seguida, deve-se especificar instrumentos e técnicas para coletar a informação necessária. Duas são as opções; ou se seleciona instrumentos existentes, ou se produz instrumentos novos. Exemplos incluem testes de aprendizagem, esquemas de entrevistas, pesquisas na literatura. Na elaboração de novos instrumentos deve-se cuidar dos seus aspectos de validade e confiabilidade, conforme Unidade 3.0, páginas F.179 e F.190.

Para cada informação a ser coletada deve-se, em seguida, especificar o procedimento de amostragem a ser empregado para obtê-la.

Finalmente, deve-se desenvolver um esquema para a coleta de informação, especificando:

- . Quem deve coletar?
- . Como?
- . Quando?

Esta especificação constitui um ponto crítico, metodológico, pois tem implicação direta na confiabilidade das informações. Por exemplo, um como, ou seja, uma maneira inadequada de coletar a informação, um quando inoportuno, podem resultar numa informação inútil ou tendenciosa, prejudicando a tomada de decisão.

6) Análise da Informação

O propósito desta parte é escolher o tipo de análise descritiva e/ou inferencial, para transformação dos dados em informação que

serã relatada aos decisores. A escolha do tipo de análise deve ser feita de acordo com cada caso (considerando o tipo de decisã a ser servido), para fornecer a informação mais útil a quem tomarã a decisã.

7) Relatõrio da Informaçã

O propõsito desta parte ã assegurar que os decisores te rãõ acesso imediato ã informaçã que eles precisam, e que irãõ recebẽ-la de uma maneira e forma que facilite o uso da mesma. Audiẽncias para os Relatõrios de Avaliaçã deverã ser identificadas e definidas. Entãõ, meios deverã ser definidos para fornecer informaçã a cada audiẽncia. Subsequentemente, o formato para os relatõrios de Avaliaçã e para as partes (seções) do relatõrio deverã ser identificados. E, finalmente, um esquema mestre deve ser providenciado, definindo as interrelações en tre audiẽncias, relatõrios e datas para fornecimento da informaçã.

Como jã foi dito anteriormente, a estrutura lõgica de um plano de Avaliaçã ã a mesma para qualquer tipo de Avaliaçã, seja Con texto, Input, Processo ou Produto, ou outros tipos de acordo com o autor da abordagem adotada. E, ã necessãrio um plano de Avaliaçã para cada tipo de Avaliaçã a ser realizado. Variam, no entanto, em funçã do tipo de Avaliaçã, as informações a serem delineadas, obtidas e fornecidas. As Tabelas F.4, F.5, F.6, F.7 (pãginas detalhes de informaçã pertinentes a cada um dos tipos de Avaliaçã.

Com as especificações do conteũdo, contido neste apẽndi ce, espera-se que, ao final do curso, o aluno seja capaz de planejar uma Avaliaçã, segundo o esquema da unidade 6 e utilizando os conceitos de todas as unidades.