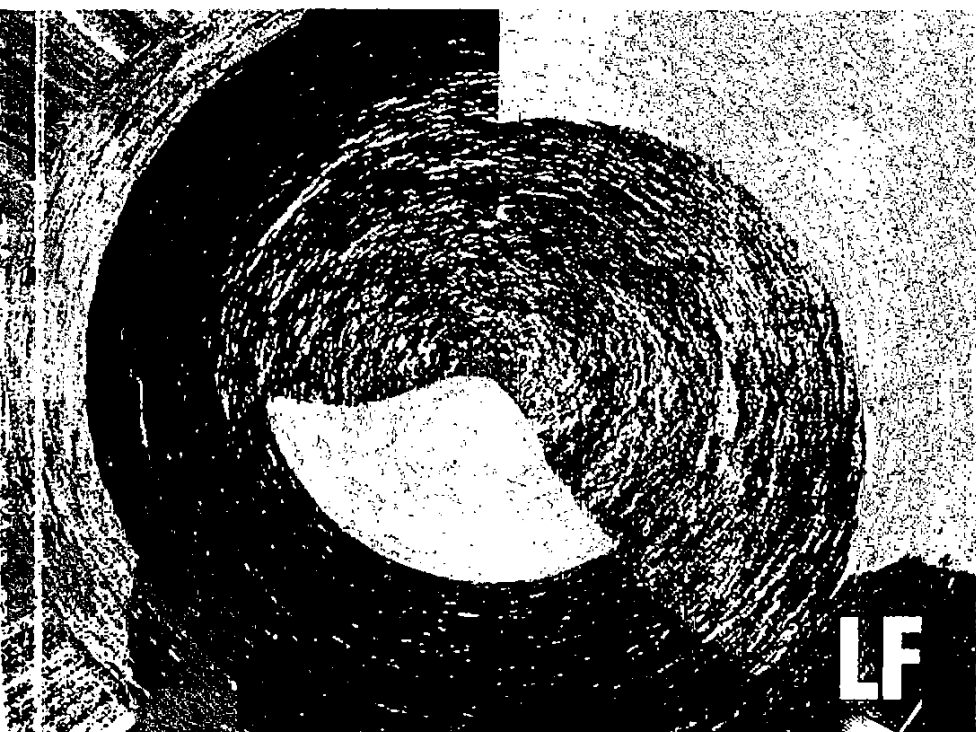


Metodologias de Pesquisa em Ensino

Marco Antonio Moreira



Metodologias de Pesquisa em Ensino

Marco Antonio Moreira



Metodologias de Pesquisa em Ensino

Marco Antonio Moreira

Copyright © 2011 Editora Livraria da Física

1ª Edição

Direção Editorial

José Roberto Marinho

Revisão ortográfica

Fabricia Carpinelli

Capa

Antonio Manuel Alves MORAIS

Diagramação

Quantumwebs

Dados Internacionais de Catalogação na Publicação (CIP)
(Câmara Brasileira do Livro, SP, Brasil)

Moreira, Marco Antonio
Metodologias de pesquisa em ensino / Marco
Antônio Moreira. -- São Paulo: Editora
Livraria da Física, 2011.

Bibliografia.

ISBN 978-85-7861-110-1

1. Ciências - Estudo e ensino 2. Pesquisa educacional 3. Pesquisa
qualitativa 4. Pesquisa quantitativa I. Título.

11-11093

CDD-507.2

Índices para catálogo sistemático:

1. Ciências : Metodologias de pesquisa em ensino : Educação 507.2

LF

EDITORA

Sumário

Apresentação.....	9
Capítulo I	11
Pesquisa em Ensino: Aspectos Metodológicos.....	11
Objetivo.....	11
Estrutura do processo de pesquisa	12
O(s) fenômeno(s) de interesse da pesquisa em ensino	15
Eventos/objetos	15
A abordagem quantitativa à pesquisa educacional	18
Fidedignidade e validade	20
Delineamentos	24
Validade interna e externa.....	28
O papel da Estatística.....	30
Algumas críticas ao enfoque quantitativo	36
A pesquisa qualitativa.....	42
Algumas características de uma abordagem qualitativa à pesquisa em ensino.....	46
O domínio metodológico da pesquisa interpretativa	50
Pesquisa quantitativa versus qualitativa: uma visão geral ...	53
Domínio Conceitual.....	54
Domínio Metodológico	54
Eventos/objetos	54
Pesquisa quantitativa versus qualitativa: uma visão detalhada	57
A questão da compatibilidade	58
Fidedignidade e validade na pesquisa qualitativa	64
O domínio conceitual da pesquisa qualitativa.....	67
Referências.....	70
Capítulo II.....	73
Pesquisa em Ensino: Métodos Qualitativos	73
Introdução	73
Os dois paradigmas clássicos.....	75

A pesquisa qualitativa (interpretativa)	76
A etnografia	80
O estudo de casos	86
A pesquisa-ação	90
Outros tipos de pesquisa qualitativa	95
A fenomenografia	96
A hermenêutica	96
A pesquisa participativa	97
A história oral	98
Fidedignidade, generalização e validade	99
Triangulação	105
Os paradigmas de pesquisa em educação: rumo à acomodação	107
Conclusão	110
Referências	113
Capítulo III	117
Pesquisa em Ensino: Métodos Quantitativos	117
Introdução	117
Dados e tratamento estatístico	119
Significância estatística e significância prática	120
A estatística apropriada	121
O papel do computador	121
Variáveis	122
Controle de variáveis	124
Delineamentos de pesquisa	126
Delineamentos não experimentais ou pré-experimentais	129
Delineamentos experimentais	131
Delineamentos quase-experimentais	134
Um pouco de estatística	137
Amostragem aleatória	137
Medidas de tendência central	140
Medidas de variabilidade	142
Curva normal	144
Intervalos de confiança	145
Distribuição de médias amostrais	146
Inferências a respeito do valor da média da população	147
Comparações entre múltiplas amostras	148

Tipos de hipóteses	148
Distribuição de médias de pequenas amostras.....	150
Fidedignidade e validade	154
O que é a correlação entre duas variáveis.....	154
Cálculo do coeficiente de correlação	156
Validade e fidedignidade de testes.....	157
Fidedignidade	158
Um algoritmo para o cálculo do coeficiente de fidedignidade de testes - Análise de Consistência Interna	163
Validade.....	169
Análise da Variância	171
O teste F - análise da variância (anova)	172
O método da análise da variância	175
Análise fatorial da variância.....	178
Interação estatística	179
O método da análise fatorial da variância.....	182
Testes estatísticos paramétricos e não paramétricos	183
Vantagens de testes estatísticos não-paramétricos	186
Testes não paramétricos - o caso de uma amostra.....	187
Testes não paramétricos - o caso de duas amostras relacionadas.....	189
Teste da probabilidade exata de Fisher	191
O teste χ^2 para duas amostras independentes.....	193
O teste da Mediana.....	194
O teste U de Mann-Whitney	194
O teste de duas amostras de Kolmogorov-Smirnov.....	194
O teste de Wald-Wolfowitz.....	194
O teste de Moses de reações extremas	195
O teste da aleatoriedade para duas amostras independentes	195
Discussão	195
Bibliografia.....	198
Capítulo IV	201
Sobre Monografias, Dissertações, Teses, Artigos e Projetos de Pesquisa: Significados e Recomendações para Iniciantes da Área de Educação Científica	201
Introdução.....	202

Parte I - Significados	204
A monografia	204
A dissertação e a tese	204
O artigo de pesquisa	207
Referências	208
Anexos e apêndices	208
Detalhes técnicos	209
O projeto de pesquisa	209
Parte II - Recomendações	210
Como fazer uma boa pesquisa ?	210
Como identificar a estrutura de uma pesquisa?	214
Como escolher a metodologia de pesquisa?	218
Contexto	219
Ensino	219
Conclusão	222
Referências bibliográficas	223
Apêndice I	227
Apresentação de um Projeto de Pesquisa em Educação	227
Roteiro	229
Conclusão	232
Bibliografia	233
Apêndice II	235
Tipos de Revisão da Literatura e/ou Teórica a Serem Evitados	235
Apêndice III	239
Sugestões sobre redação de trabalhos acadêmicos	239
Dicas de estilo	239
Palavras do inglês parecidas com palavras do português, mas com outro sentido:	240
Algumas expressões úteis em inglês e espanhol	241
Palavras do espanhol parecidas com palavras do português, mas com outro sentido:	241
Críticas a projetos	241
Dicas para projetos	242

Apresentação

Este livro é uma recopilação de textos de apoio sobre metodologias de pesquisa em ensino, particularmente de ciências. Foram escritos para serem usados independentemente uns dos outros. Por consequência, apresentam certas superposições.

O primeiro capítulo é mais abrangente, enfocando tanto aspectos quantitativos como qualitativos da metodologia de pesquisa em ensino. O segundo trata somente de métodos qualitativos enquanto o terceiro se ocupa apenas de métodos quantitativos. O quarto, por sua vez, aborda a redação de trabalhos científicos, já que a pesquisa deve sempre gerar publicações científicas.

Ainda que alguns tópicos possam estar repetidos nos diferentes capítulos, os textos são complementares e podem auxiliar na diferenciação progressiva de alguns conceitos.

É preciso também chamar a atenção para o fato de que os textos que compõem os capítulos são apenas textos de apoio, ou textos introdutórios, que buscam fornecer subsídios metodológicos para aqueles que desejam pesquisar em ensino.

Naturalmente, os métodos quantitativos e qualitativos neste livro abordados são aqueles usados na área da educação de um modo geral, mas a preocupação aqui é o ensino, ainda que seja interdependente com aprendizagem, currículo e contexto, os chamados lugares comuns do fenômeno educativo.

Porto Alegre, Dezembro 2011

Marco Antonio Moreira

Capítulo I

Pesquisa em Ensino: Aspectos Metodológicos¹

M.A. Moreira

Resumo

Texto introdutório que focaliza a pesquisa em ensino à luz de diferenças entre os enfoques quantitativo e qualitativo à pesquisa educacional, com o objetivo de prover uma introdução ao assunto a possíveis pesquisadores nessa área. Inicialmente, cada enfoque é examinado em separado; após, faz-se várias comparações entre ambos e, ao final, discute-se a questão da compatibilidade.

Objetivo

Este texto tem o objetivo de abordar a pesquisa em ensino principalmente à luz de diferenças paradigmáticas e metodológicas entre os enfoques quantitativo e qualitativo à pesquisa educacional. A intenção não é a de defender ou sugerir uma ou outra dessas abordagens, mas, sim, a de prover subsídios

¹ Adaptado do Capítulo 2 do livro *Pesquisa em ensino: o Vê epistemológico* de Gowin, de M. A. Moreira. São Paulo, Editora Pedagógica e Universitária Ltda., 1990. Porto Alegre, Instituto de Física da UFRGS, Monografias do Grupo de Ensino, Série Enfoques Metodológicos, N° 01. Revisado em 1999 e usado, em espanhol, como Texto de Apoio N° 1 do Programa Internacional de Doutorado em Ensino de Ciências da Universidade de Burgos, Espanha, em convênio com a UFRGS. Em português, é o Texto de Apoio N° 19 desse Programa, 2003. Publicado em *Actas del PIDEC*, 1:05-38, 1999.

sobre o tema, procurando ser útil como uma espécie de introdução à pesquisa em ensino.

Como meio de organizar o desenvolvimento do assunto, introduz-se primeiramente um dispositivo heurístico que procura esquematizar o processo de produção de conhecimento, seja ele quantitativo ou qualitativo. Logo após, para clarificar o que se está entendendo por pesquisa em ensino, discute-se, de forma sucinta, o fenômeno de interesse da pesquisa em ensino. Feito isso, concentra-se a atenção em aspectos metodológicos e paradigmáticos da pesquisa em ensino, considerando-a como um subconjunto da pesquisa educacional e, como tal, influenciada diretamente pelo debate da pesquisa quantitativa versus pesquisa qualitativa em educação. A fim de evitar que o texto se torne excessivamente longo, dá-se pouca ênfase a aspectos teórico-conceituais, sem querer sugerir com isso que devam ser preteridos pelos pesquisadores educacionais em função da metodologia.

Estrutura do processo de pesquisa

A estrutura do processo de pesquisa pode ser esquematizada sob a forma de um "V" (Gowin, 1981), tal como ilustrado na Figura 1. O lado esquerdo dessa figura refere-se ao domínio conceitual do processo de investigação: ali estão os conceitos-chave e os sistemas conceituais usados na pesquisa, os quais geram princípios que, por sua vez, dão origem a teorias que têm, subjacentes, determinados sistemas de valores ou filosofias.

Na ponta do "V" estão os eventos que acontecem naturalmente, ou que o pesquisador faz acontecer, a fim de fazer registros através dos quais os fenômenos de interesse possam ser estudados e/ou os objetos selecionados para análise.

O lado direito do "V" tem a ver com a parte metodológica da pesquisa. Sob o rótulo de registros e transformações, estão incluídas observações, anotações, medidas, dados, categorias,

tabelas, gráficos e estatísticas utilizadas em uma investigação. As asserções se referem aos resultados, podendo ser de conhecimento (i.e., relativas ao conhecimento produzido) ou de valor (i.e., referentes ao valor desse conhecimento).

No centro do "V" está(ão) a(s) questão(ões) básica(s), a(s) qual(is), na verdade, pertence(m) tanto ao domínio conceitual como ao metodológico. A questão básica, questão-chave, ou questão-foco, de uma pesquisa não só pergunta alguma coisa, mas também diz algo. É a questão que identifica o fenômeno de interesse da pesquisa de tal forma que é provável que alguma coisa seja descoberta, medida ou determinada ao respondê-la; é a pergunta que informa sobre o ponto central da pesquisa, dizendo, em essência, o que foi investigado.

Em termos simples, pode-se dizer que o lado esquerdo do "V" corresponde ao pensar, enquanto que o direito é relativo ao fazer. Todavia, tudo o que é feito é guiado por conceitos, teorias e filosofias, ou seja, pelo pensar. Por outro lado, novas asserções, i.e., respostas às questões básicas, podem levar a novos conceitos, à reformulação de conceitos já existentes, ou, ocasionalmente, a novas teorias e filosofias. Isso significa que existe uma constante interação entre os dois lados do "V". Essa interação, que na Figura 1 está simbolizada pela linha curva, é necessária para que se chegue nas respostas às questões básicas formuladas sobre os eventos que acontecem ou se faz acontecer para estudar o fenômeno de interesse.

Gowin propôs esse "V" como um instrumento heurístico para a análise da estrutura do processo de produção de conhecimento (entendida como as partes desse processo e a maneira como elas se relacionam) ou para desvelar conhecimentos documentados sob a forma de artigos de pesquisa, livros, ensaios, a fim de tornar esses conhecimentos adequados para propósitos instrucionais (Moreira, 1985).

Esse dispositivo ficou conhecido como "diagrama V", "V

epistemológico de Gowin" ou, simplesmente, "V de Gowin". Usar um "V" para esquematizar a estrutura da produção de conhecimento é útil porque permite convergir este processo em eventos e/ou objetos uma vez que, em última análise, toda investigação procura responder questões sobre eventos e/ou objetos. Todavia, o formato de "V" em si não é importante, poderia ser um "U", um "V" invertido ou outra figura. O importante é a ênfase dada à interação entre os domínios conceitual e metodológico para responder questões-foco sobre fenômenos de interesse (eventos e/ou objetos).

O "V de Gowin" não equaciona pesquisa com pesquisa científica nem metodologia com método científico.

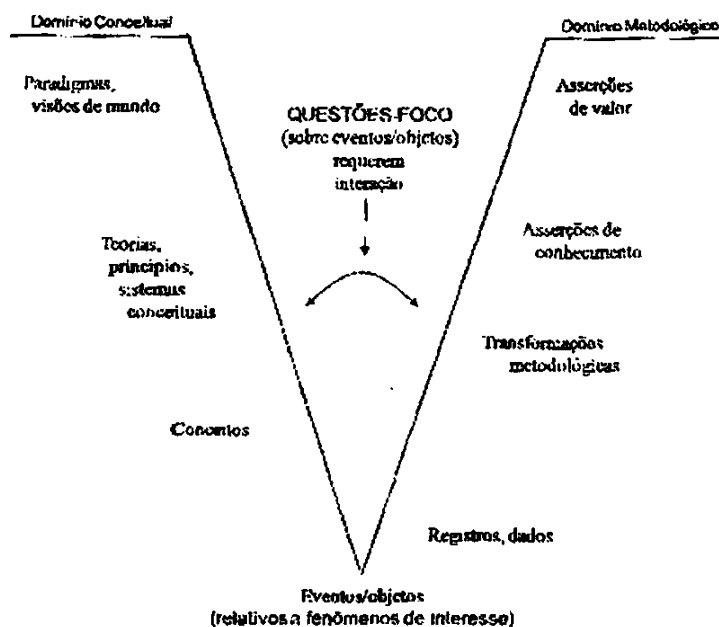


Figura 1 - O "V" epistemológico de Gowin mostrando os elementos conceituais e metodológicos que interagem na pesquisa (adaptado de Gowin, 1981).

O que ele sugere é que qualquer investigação é conduzida a partir de determinados paradigmas e bagagens teórico-conceituais, fazendo uso de certas metodologias para estudar certos fenômenos de interesse, os quais, obviamente, são selecionados em função de tais paradigmas e referenciais teórico-conceituais. A hipótese deste texto, no entanto, é a de que o mesmo fenômeno de interesse pode ser investigado sob diferentes paradigmas, referenciais teórico-conceituais e metodologias.

O(s) fenômeno(s) de interesse da pesquisa em ensino

Como sugere o próprio nome, a pesquisa em ensino tem como foco o ensino. Todavia, embora não haja, necessariamente, uma relação de causa e efeito entre ensino e aprendizagem, não faz muito sentido falar em ensino sem relacionar essa atividade à de aprender. Ou seja, o ensino tem sempre como objetivo a aprendizagem e, como tal, perde significado se for tratado isoladamente. Entretanto, aprendizagem é uma atividade idiossincrática que pode não ser consequência necessária do ensino recebido.

Por outro lado, para saber se houve aprendizagem é preciso avaliá-la. A avaliação da aprendizagem pode, em princípio, prover evidências não só sobre o que foi aprendido, mas também sobre até que ponto o ensino foi responsável por isso. Naturalmente, é possível também avaliar o ensino de outras maneiras como, por exemplo, através da opinião do aluno.

Ensino, aprendizagem e avaliação não serão aqui definidos face à existência de diversas maneiras de fazê-lo, refletindo diferentes posicionamentos teóricos. O que se pretende é destacar a inter-relação entre estes conceitos de tal modo que não se pode considerar apenas o ensino, propriamente dito, como o fenômeno de interesse da pesquisa em ensino.

Nessa questão, contudo, há também que se levar em conta

outro elemento: o que ensinar? Isto é, aquilo que se espera que seja aprendido e que será objeto da avaliação da aprendizagem. Tomando esse elemento como sendo aquele que define o currículo, poder-se-ia dizer que o(s) fenômeno(s) de interesse da pesquisa em ensino envolve(m) não só o ensino, em si, mas também aprendizagem, avaliação e currículo. Este, por sua vez, pode também ser avaliado; portanto, a avaliação, nessa perspectiva, pode se referir à aprendizagem, ao ensino e ao currículo.

Finalmente, torna-se ainda necessário considerar que o ensino formal se dá em um ambiente – a sala de aula – integrante de um ambiente mais abrangente – a escola. Esta, por sua vez, se insere no sistema escolar e na sociedade de um modo geral. A sala de aula é um micromundo, uma microcultura, com certos vínculos e determinada organização social. A escola tem suas regras formais e informais; a sociedade tem seus valores e suas expectativas. O que acontece na sala de aula é influenciado pelo que acontece em outros níveis de organização social e cultural. Tudo isso indica que o ensino se dá em um certo contexto que a pesquisa em ensino não pode ignorar e que, a rigor, é parte inseparável do fenômeno de interesse dessa pesquisa. Em um determinado momento, por exemplo, o foco de uma pesquisa em ensino pode estar no estudo do efeito de alguns aspectos contextuais do ensino ou, usando a mesma terminologia anterior, na avaliação do contexto.

Assim, pode-se então dizer, de uma maneira bem abrangente, que o fenômeno de interesse da pesquisa em ensino tem a ver com ensino, aprendizagem, avaliação, currículo e contexto. Isso é o que se tenta esquematizar no mapa conceitual da Figura 2.

Nessa perspectiva, os eventos focalizados pela pesquisa em ensino são episódios, acontecimentos, situações, relativos ao ensino, aprendizagem, currículo, contexto e avaliação ou à combinação

deles. Uma aula, um procedimento de avaliação, um novo currículo, a influência de uma certa variável sobre a aprendizagem, um experimento de laboratório, a percepção mútua de alunos e professores, são exemplos de eventos que interessam à pesquisa em ensino.

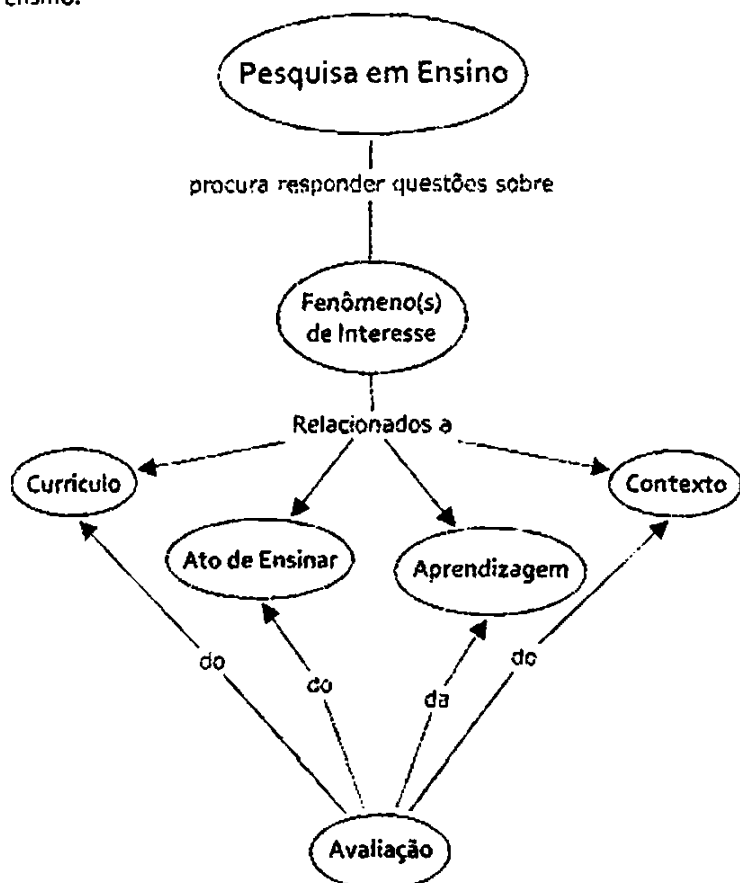


Figura 2— Um mapa conceitual para os elementos envolvidos no(s) fenômeno(s) de interesse da pesquisa em ensino.

Esses eventos ocorrem naturalmente ou são feitos acontecer pelo pesquisador que faz, então, registros dos eventos. Um

videoteipe de uma aula, ou de parte dela, é uma maneira de registrar esse evento. Anotações em uma ficha de observação ou de controle, gravações de entrevistas, mapas conceituais, respostas a testes, são exemplos de registros de eventos. Tais registros são transformados e analisados – quantitativa e/ou qualitativamente –, de modo a conduzir a explicações e/ou descrições que procuram responder questões-foco sobre o fenômeno de interesse. Essas questões, na perspectiva usada neste texto, se referem ao fenômeno de interesse da pesquisa em ensino, ou seja, ao ensino propriamente dito, à aprendizagem, à avaliação, ao currículo e ao contexto.

Naturalmente, cada um desses conceitos tem significados variados em diferentes contextos. Fugiria completamente à intenção do texto entrar nessa discussão. Ao invés disso, supõe-se que tais conceitos tenham, além dos contextuais, significados que são compartilhados por usuários com diferentes posicionamentos teórico-metodológicos.

Supõe-se, também, que o fenômeno de interesse da pesquisa em ensino se refere a esses conceitos independente do enfoque paradigmático, teórico-conceitual e metodológico. Quer dizer, na pesquisa em ensino investiga-se sempre fenômenos de interesse relacionados com ensino, aprendizagem, currículo, avaliação e contexto, porém de diferentes maneiras.

Dois enfoques têm predominado na pesquisa em educação e, por extensão, na pesquisa em ensino: quantitativo e qualitativo. O restante deste texto procura destacar diferenças entre tais enfoques.

A abordagem quantitativa à pesquisa educacional

A pesquisa quantitativa procura estudar os fenômenos de interesse da pesquisa em educação geralmente através de estudos experimentais ou correlacionais, caracterizados, primordialmente, por medições objetivas e análises quantitativas.

De um modo geral, esse tipo de enfoque se baseia no modelo de pesquisa, dito científico, das ciências exatas, do qual a Física é provavelmente o melhor exemplo. A definição de experimento dada por Kerlinger (1980, pp. 94 e 125) ilustra bem essa característica da pesquisa quantitativa experimental:

"Um experimento é uma pesquisa onde se manipulam uma ou mais variáveis independentes e os sujeitos são designados aleatoriamente a grupos experimentais [...] é um estudo no qual uma ou mais variáveis independentes são manipuladas e no qual a influência de todas ou quase todas as variáveis relevantes possíveis não pertinentes ao problema da investigação é reduzida a um mínimo."

Não é exatamente esse tipo de manipulação e controle de variáveis que um cientista procura fazer em seu laboratório?

É verdade que em estudos correlacionais, o pesquisador educacional quantitativo não necessariamente manipula variáveis: às vezes ele pode simplesmente procurar saber se há correlação entre variáveis que não decorrem de manipulações experimentais. Mas, de um modo geral, a ideia básica do enfoque quantitativo é a manipulação e controle objetivo de variáveis.

Pode-se classificar e distinguir variáveis de diversas maneiras, aqui, todavia, por uma questão de simplicidade, far-se-á apenas a distinção entre independentes e dependentes (Best, 1970, p. 143):

"Variáveis independentes são condições ou características que o experimentador manipula em sua tentativa de determinar sua relação com os fenômenos observados. Variáveis dependentes são condições ou características que aparecem, desaparecem ou mudam quando o experimentador introduz, remove ou muda variáveis independentes."

Na pesquisa educacional, uma variável independente pode ser um certo método de ensino, um tipo de material instrucional, uma recompensa, um período de exposição a uma certa condição. A variável dependente pode ser o escore de um teste, o número de erros ou o tempo gasto para executar uma tarefa. Portanto, as variáveis dependentes são mudanças medidas no desempenho dos alunos atribuíveis à influência das variáveis independentes."

Como bem sugere o exemplo dado por Best nesta citação, muitas vezes lida-se com aspectos quantitativos das variáveis, os quais geralmente deseja-se medir a fim de obter dados. A exemplo das ciências exatas, a questão de como medir as variáveis traz consigo outra característica importante da pesquisa quantitativa em ensino: os instrumentos de medida. Testes de conhecimento sob os mais diversos formatos, escalas de atitude, fichas de observação, questionários, são exemplos de instrumentos de medida usados na pesquisa em ensino.

Questionários, a propósito, são instrumentos altamente utilizados, mas sua elaboração requer cuidados especiais sob pena de não serem respondidos ou de fornecerem informações distorcidas (Moreira e Koff, 1985).

Obviamente, ao se utilizar tais instrumentos, espera-se ter o maior grau de confiança possível nas medidas feitas. De fato, há sempre por parte do pesquisador uma grande preocupação com pelo menos duas características básicas de um instrumento de medida, relacionados com a confiabilidade dos valores medidos: a *fidedignidade* e a *validade*.

Fidedignidade e validade

A fidedignidade de um instrumento refere-se à estabilidade, à reproduzibilidade, à precisão das medidas com ele obtidas, i.e., ao grau de consistência dos valores medidos. A validade, por sua

vez, tem a ver com até que ponto o instrumento está, de fato, medindo o que se supõe que esteja medindo.

Examinemos primeiramente a questão da fidedignidade. No dizer de Fox (1969, p. 353):

"Por fidedignidade entende-se a exatidão dos dados no sentido de sua estabilidade, repetitividade ou precisão. Um instrumento de coleta de dados perfeitamente fidedigno é aquele que se administrado duas vezes nas mesmas circunstâncias forneceria os mesmos dados."

Vianna (1978, p. 145) dá uma ideia clara da importância da fidedignidade de um instrumento:

"Se um teste é aplicado ao mesmo grupo um grande número de vezes, espera-se que os resultados sejam os mesmos, desde que o grupo não se modifique. Se cada vez que o teste for aplicado, satisfeitas determinadas condições, os escores forem diferentes para o mesmo grupo, não se poderá ter confiança no instrumento, porque não haverá consistência nas medidas."

Isso sugere que se poderia investigar a fidedignidade de um instrumento aplicando-o várias vezes ao mesmo grupo e observando a estabilidade dos dados obtidos. Entretanto, na prática, o grupo geralmente é testado uma só vez. (A rigor, é impossível retestar o mesmo grupo nas mesmas condições.) Por isso, recorre-se a procedimentos estatísticos que permitem estimar a fidedignidade do instrumento a partir dos dados de um certo número de indivíduos.

A estatística básica usada para estimar a fidedignidade é a correlação. Correlações (ou coeficientes de fidedignidade) de + 1,00 indicam perfeita fidedignidade, enquanto correlações perto

de zero indicam ausência de fidedignidade. Correlações entre zero e 1,00 significam níveis intermediários de fidedignidade. Na prática, os valores aceitáveis do coeficiente de fidedignidade dependem do que se está medindo. Na área de atitudes e interesses, por exemplo, em que os dados são mais flexíveis e mutáveis, correlações da ordem de 0,70 são aceitáveis. Em outras áreas, são esperados valores acima de 0,85.

Existem várias maneiras e fórmulas para estimar a fidedignidade, as quais não serão aqui abordadas por fugir ao escopo do texto. Descrições dessas técnicas e fórmulas podem ser encontradas na maioria dos livros de pesquisa ou de testes em educação (e.g., Vianna, 1978).

A fidedignidade de um instrumento é, no entanto, condição necessária, porém não suficiente para que este possa ser usado em um procedimento de pesquisa. Para isso, é preciso que tenha também validade, pois de nada adianta um instrumento que meça fidedignamente algo diferente daquilo que se pretende medir. Por exemplo, um teste que mede com muita fidedignidade o conhecimento de dados factuais em História não será válido na medida em que se pretende avaliar compreensão de fatos históricos com esse instrumento. Por outro lado, esse teste poderia ser válido para avaliar conhecimento factual; isso significa que um instrumento não é válido ou "não válido" de um modo geral, mas, sim, em relação a determinado objetivo ou situação.

A validade depende, pois, da finalidade com que é usado o instrumento. Esta é uma das razões que tornam o problema da validade muito mais complexo do que o da fidedignidade que, a rigor, é um problema resolvido (não se justificando, portanto, o uso de instrumentos de baixa fidedignidade). Dentre os vários tipos de validade existentes, destacam-se para a pesquisa em ensino os seguintes:

Validade de conteúdo – um instrumento tem validade de conteúdo na medida em que se constitui em uma amostra

representativa do conteúdo (conhecimentos e comportamentos) que está sendo medido. É também chamada de validade curricular, amostral ou lógica. Não é determinada estatisticamente, mas, sim, resulta do julgamento de diferentes examinadores que analisam a representatividade dos itens em relação às áreas de conteúdo e à relevância dos objetivos a medir. (Vianna, 1985, pp.172-173)

Validade concorrente – um instrumento apresenta esse tipo de validade quando os resultados de sua aplicação se correlacionam com os de outro instrumento já validado e que mede a mesma coisa. É o mesmo que validade congruente.

Validade preditiva – se o instrumento tem esse tipo de validade, isso significa que o pesquisador usou resultados obtidos com esse instrumento para fazer previsões sobre o comportamento futuro dos respondentes e tais previsões foram confirmadas. Ou seja, há uma alta correlação entre os escores do teste cuja validade se está argumentando e os escores no desempenho futuro, segundo algum critério obtido independentemente. É também conhecida como validade empírica ou relativa ao critério.

Em um estudo experimental, as variáveis independentes, i.e., condições ou características que o experimentador manipula (um novo método de ensino, para dar um exemplo simples), são muitas vezes chamadas de *tratamentos*. Suponhamos que um pesquisador queira estudar o efeito de um certo tratamento (variável independente). Ele faz então algumas hipóteses (i.e., o tratamento poderia ter tais e tais efeitos) e planeja um experimento para testar essas hipóteses. Nesse plano, conhecido como delineamento de pesquisa, ele leva em consideração questões tais como: quem expor ao tratamento, i.e., quem serão os sujeitos da pesquisa? Como observar (medir) os efeitos previstos (hipóteses)? Quando e quantas vezes medi-los? Como ter certeza que os efeitos medidos devem-se, de fato, à variável

independente? Se for necessário trabalhar com amostras, como proceder à amostragem? Delineamento de pesquisa é o plano e a estrutura da pesquisa.

Delineamentos

Pode-se distinguir entre delineamentos pré-experimentais, experimentais e quase-experimentais.

Provavelmente, o melhor trabalho sobre delineamento que se pode encontrar na literatura é o artigo *Experimental and quasi-experimental designs for research on teaching*, de Campbell e Stanley (1963), traduzido para o português sob o título de *Delineamentos experimentais e quase-experimentais de pesquisa* e publicado sob a forma de um pequeno livro (Campbell e Stanley, 1979). A leitura desse livro é fortemente recomendada para quem pretende fazer pesquisa experimental em ensino. Aqui, serão apenas dados alguns exemplos.

Um delineamento muito usado na pesquisa em ensino, mas que, na verdade, é um delineamento pré-experimental, porque não satisfaz o critério anterior, é o seguinte (Campbell e Stanley, 1979, p.24):

$$O_1 \times O_2$$

Neste delineamento, aplica-se um pré-teste O_1 a um grupo, submete-se esse grupo a um tratamento X e aplica-se, então, um pós-teste O_2 . O_1 e O_2 significam que o mesmo grupo é observado antes e depois do tratamento que pode ser, por exemplo, um novo método de ensino ou um recurso didático alternativo. Diferenças entre O_2 e O_1 (que podem ser simples testes de conhecimento) evidenciarão a eficácia (ou ineficácia) do tratamento X. O problema com este delineamento é que não controla outras variáveis, além de X, que poderiam explicar diferenças entre O_2 e O_1 . Por exemplo, os alunos poderiam ter melhores resultados no pós-teste porque amadureceram durante

o curso, e não porque o tratamento X tivesse sido eficiente.

Um delineamento experimental muito usado é o seguinte (Campbell e Stanley 1979, p. 26):

$$\begin{array}{c} A \quad O_1 \quad X \quad O_2 \\ A \quad O_3 \quad \quad O_4 \end{array}$$

Neste delineamento, trabalha-se com dois grupos e os sujeitos da pesquisa são designados aleatoriamente a um deles (este é o significado do A). Aplica-se um pré-teste a ambos os grupos ($O_1 = O_3$), i.e., "observa-se" os grupos antes de se manipular a variável independente X. Um dos grupos (grupo experimental) é então submetido ao tratamento X e o outro (grupo controle) não. Após, aplica-se um pós-teste ($O_2 = O_4$) a ambos os grupos. Na prática, o pré e o pós-teste podem ser iguais; diferenças entre os resultados do pré e pós-testes em ambos os grupos ($O_2 - O_1$ e $O_4 - O_3$) podem fornecer evidências sobre o efeito do tratamento X.

Este delineamento controla variáveis, exceto X, na medida em que elas influenciarão igualmente ambos os grupos e, portanto, seu efeito não pesará na comparação das diferenças $O_2 - O_1$ e $O_4 - O_3$.

Além disso, a aleatoriedade da designação dos sujeitos a um dos grupos, embora não garanta equivalência entre os grupos, reduz ao mínimo a probabilidade de que sejam diferentes. Segundo Kerlinger (1980, p.102):

"Casualização é a designação de objetos (sujeitos, tratamentos, grupos) de um universo a subconjuntos do universo de tal maneira que, para qualquer designação dada a um subconjunto, todo membro do universo tem igual probabilidade de ser escolhido para a designação. Não há total garantia de que a casualização 'igualará' os grupos, mas a

probabilidade de igualar é relativamente alta. Há outra forma de expressar esta ideia: [...] já que em procedimentos aleatórios todo membro de uma população tem igual probabilidade de ser escolhido, membros com certas características distintas – homem ou mulher, alto ou baixo grau de inteligência, dogmático ou não dogmático, e assim por diante – se selecionados, provavelmente serão contrabalançados em longo prazo pela seleção de outros membros da população com a quantidade ou a qualidade 'oposta' n da característica."

A aleatoriedade na designação de objetos é, portanto, a mais adequada segurança de que não há diferenças ou vieses iniciais entre os grupos. Nesse caso, o pré-teste não é condição essencial para que um delineamento seja verdadeiramente experimental. Assim, o delineamento anteriormente apresentado poderia ser simplesmente (Campbell e Stanley, 1979, p.46):

$$A \quad X \quad o_1$$
$$A \quad o_2$$

A rigor, esse delineamento não só pode ser usado ao invés do anterior como também é mais adequado, pois elimina qualquer influência do pré-teste no experimento. Entretanto, talvez por razões psicológicas, muitos pesquisadores não abrem mão de saber "com certeza" se os grupos experimental e controle são "iguais" no início do experimento, de modo que o segundo exemplo aqui apresentado é, provavelmente, mais usado que o terceiro, embora menos apropriado logicamente.

Um terceiro tipo de delineamento identificado por Campbell e Stanley é o quase-experimental, ou seja, aquele em que falta ao pesquisador o pleno controle da aplicação dos tratamentos experimentais – quando e a quem expor e a capacidade de

casualizar exposições (op. cit. p. 61). O delineamento "série temporal" exemplifica essa situação (op. cit., p.67):

$$O_1 O_2 O_3 O_4 \quad X \quad O_5 O_6 O_7 O_8$$

Nesse delineamento, observa-se os sujeitos várias vezes antes de se aplicar o tratamento X e várias vezes após a aplicação. Supondo que antes do tratamento as observações fossem quase que homogêneas, sem variações, que houvesse um salto na observação feita logo após o tratamento e que a partir daí houvesse nova estabilização nas observações, esse salto na série temporal poderia ser tomado como evidência do efeito de X.

Observe-se que esse delineamento é semelhante ao primeiro apresentado como exemplo, porém implica em muito mais observações, o que minimiza, embora não exclua, as deficiências do primeiro. Note-se também que implica na existência de um só grupo, o que, na prática, é uma vantagem, pois, muitas vezes, é difícil obter-se dois grupos de sujeitos.

Um exemplo simples da aplicação desse delineamento seria aquele em que o professor observasse cuidadosamente seus alunos durante algumas semanas do curso, fazendo várias medições (que podem ser, por exemplo, testes de aproveitamento ou de atitude) antes de fazer uso de uma nova estratégia de ensino. Da mesma forma, voltasse a observar seus alunos, durante algum tempo, fazendo novos registros após o uso da estratégia. Diferenças qualitativas ou quantitativas no desempenho dos alunos após o uso da estratégia que se mantivessem ao longo do tempo poderiam ser tomadas como evidência do efeito da estratégia sobre a aprendizagem cognitiva ou afetiva dos alunos.

Até aqui foi dada relativa ênfase à questão básica e ao delineamento porque são cruciais na realização de uma pesquisa experimental em ensino. Assim como o pesquisador deve formular uma questão de pesquisa clara, orientadora e relevante,

ele deve também investigá-la usando um delineamento adequado. Um mau delineamento pode invalidar as asserções de conhecimento (resultados) e de valor de uma pesquisa. Isso leva a uma ideia de *validade* do experimento.

Validade interna e externa

Dois aspectos básicos devem ser considerados ao se tirar conclusões de estudos experimentais: até que ponto se pode concluir com segurança que os efeitos observados podem ser atribuídos ao tratamento X e, satisfeita essa condição, até que ponto os resultados obtidos podem ser generalizados além dos dados específicos do estudo feito. O primeiro desses aspectos tem a ver com a validade interna do experimento e o segundo com a validade externa.

Segundo Best (1970, p. 143), o pesquisador tem dois grandes objetivos ao realizar um experimento:

"1. Deve tentar determinar se os fatores que foram modificados têm realmente um efeito sistemático no contexto experimental e se as ocorrências observadas não foram influenciadas por fatores estranhos ou não controlados. A medida em que este objetivo é atingido, é uma medida da validade interna do experimento. Mas o pesquisador terá atingido pouca coisa de valor prático se essas relações forem válidas apenas na restrita situação experimental e apenas para os indivíduos dela participantes.

2. Ele deve também determinar se as relações sistemáticas que foram identificadas, isoladas e medidas podem ser generalizadas – usadas para prever relações fora do contexto do experimento. A medida em que este objetivo é atingido, é uma medida da validade externa do experimento."

Assim, no caso de um experimento realizado pelo professor em sala de aula, a primeira preocupação ao analisar os resultados obtidos é com a possibilidade de que possam ser explicados por outros fatores que não o tratamento X. Uma vez eliminadas as explicações alternativas, os resultados terão validade interna e a preocupação passará a ser com a possibilidade de que resultados semelhantes sejam obtidos com outros alunos, em outras escolas, outros professores, outros níveis de escolaridade. Na medida em que isso for verdadeiro, os resultados terão validade externa.

Obviamente, a validade interna e a externa de um experimento dependem do delineamento utilizado, o que vem a reforçar o que já foi dito acerca da importância do delineamento de pesquisa.

Usando a linguagem de Gowin, o delineamento tem a ver com os eventos, com os registros dos eventos e com as transformações desses registros. O quando e quem expor ao tratamento X, por exemplo, estão na essência dos eventos que o pesquisador fará acontecer para estudar o fenômeno de interesse e responder a questão-foco. As observações, por sua vez, têm a ver com os registros desses eventos para que ele possa ser estudado.

Tanto o evento como o registro dos eventos envolvem condições ou características que o experimentador manipula, controla ou observa, as quais são chamadas de variáveis.

A questão da fidedignidade e da validade foi aqui abordada com relativa extensão porque, do ponto de vista da pesquisa educacional quantitativa, esse é um aspecto crucial da pesquisa em ensino. Se os dados coletados não forem válidos e fidedignos, é inútil prosseguir com a pesquisa: a Estatística não fará o milagre de transformar maus dados em bons resultados. Por outro lado, se os dados forem bons, a Estatística poderá facilitar grandemente sua organização e interpretação.

O papel da Estatística

Voltando a usar a linguagem de Gowin, cabe lembrar que se está ainda analisando o domínio metodológico da pesquisa experimental quantitativa em ensino. Nesse domínio já foi examinada a questão do delineamento de pesquisa – quando houve referência a diversos tipos de delineamento bem como à validade interna e externa de um experimento – e da mensuração das variáveis quando se abordou a fidedignidade e a validade dos instrumentos de medida.

Tudo isso se refere ao planejamento, à ocorrência e ao registro dos eventos (vide Figura 1, p.14). Passa-se agora às transformações desses registros, i.e., à reconstrução de registros (tabelas, histogramas, gráficos), à representação de registros (cálculo de médias, desvios padrão e outras grandezas que representem conjuntos de registros), à comparação de registros reduzidos e/ou reconstruídos (procura de diferenças significativas), à inferência a partir dessa comparação e, finalmente, aos resultados e sua interpretação.

Para essas transformações, a Estatística tem, naturalmente, um papel fundamental. Aliás, antes de prosseguir com alguns comentários relativos ao uso da Estatística na pesquisa em ensino, é interessante destacar a definição de Estatística dada por Kerlinger (1980, p. 353):

"Estatística é a teoria e o método de analisar dados obtidos de amostras de observações com o fim de descrever populações, estudar e comparar fontes de variância, para ajudar a tomar decisões sobre aceitar ou rejeitar relações entre fenômenos e para ajudar a fazer inferências fidedignas de observações empíricas."

Pode-se distinguir entre dois tipos de Estatística, a *descritiva* e a *inferencial*. A descritiva, como sugere o próprio nome, tem por

finalidade descrever o conjunto de dados que se dispõe e o faz através de tabulações e representações numéricas ou gráficas. Procura sumariar, sintetizar, reduzir, de modo a tornar manipuláveis as propriedades de uma massa de dados.

Por exemplo, um professor aplica determinado teste a seus alunos e o corrige atribuindo um escore a cada aluno. Vê-se, então, de posse de um conjunto de escores. Se o conjunto for pequeno, ele poderá ter uma ideia do desempenho do grupo simplesmente olhando os escores. Porém, se o grupo for grande, não terá outra saída senão organizar esses dados de alguma forma a fim de ter uma visão do desempenho do grupo como um todo. Recorrerá, então, à Estatística Descritiva.

A primeira coisa a fazer é tabular os dados e organizar uma distribuição de frequências. Feito isso, poderá representar graficamente a distribuição através de um histograma ou de uma curva de frequências. Mas poderá querer uma maneira mais conveniente, mais sintética, de descrever a distribuição, um "escore típico" talvez. Em Estatística há vários tipos de "escores típicos" que são chamados de medidas de tendência central, dentre as quais destacam-se (Elzey, 1967):

- A moda* - escore que ocorre mais vezes em uma distribuição de frequências.
- A mediana* - é o ponto em uma distribuição de frequências abaixo do qual está a metade dos escores.
- A média* - é a média aritmética dos escores em uma distribuição de frequências.

Entretanto, para melhor descrever uma distribuição de frequências precisa-se não só de uma medida da tendência central, mas também de uma medida da variabilidade dos escores, ou seja, de como eles estão espalhados entre os indivíduos. Uma das medidas mais úteis para isso é o desvio padrão: em uma distribuição normal (i.e., cuja representação

gráfica é uma curva normal), aproximadamente 68% dos escores situam-se entre um desvio padrão abaixo da média e um acima. Uma outra medida da dispersão dos escores também muito usada é a variância, a qual, em termos simples, é o quadrado do desvio padrão.

Enquanto que a Estatística Descritiva é uma ferramenta que procura organizar e facilitar a manipulação dos dados, a Estatística Inferencial busca resolver outro grande problema com que se defronta quem está analisando dados quantitativos: inferir propriedades de uma população a partir de uma amostra da mesma. Por exemplo, o professor que aplica um teste a uma turma de alunos de Física 1 obtém dados de uma amostra da população de alunos de Física 1. O termo população refere-se, pois, a todos os indivíduos com uma certa característica, enquanto que amostra refere-se a uma parte da população. O problema é, então, o seguinte: até que ponto a amostra representa a população? Como é virtualmente impossível para um grupo menor ser exatamente representativo de um muito maior, há sempre um erro de amostragem. Com que margem de segurança poderia, então, o professor supor que os resultados obtidos pelo grupo que ele testou em Física 1 são válidos para a população de alunos de Física 1?

Esse tipo de questão identifica a finalidade da Estatística Inferencial, i.e., fazer inferências sobre uma população a partir de uma amostra da mesma, e também evidencia sua importância para a pesquisa quantitativa em ensino, pois o pesquisador nessa área seguramente defrontar-se-á com problemas dessa natureza. Talvez o primeiro desses problemas seja o da amostra: se for possível escolher uma amostra da população alvo, como fazê-lo de modo a minimizar o erro de amostragem, a incluir representantes de subpopulações, a dar igual oportunidade a todos de entrar na amostra? E se não for possível escolher uma amostra, o que fazer para diminuir o efeito do fato de que o grupo

disponível não foi escolhido por um processo de amostragem? Todas essas perguntas sugerem que a amostragem é uma questão que o pesquisador deve considerar seriamente sob pena de sacrificar a validade de seu experimento. Cabe aqui lembrar que os delineamentos verdadeiramente experimentais (p. 26) são aqueles nos quais os sujeitos da pesquisa são designados aleatoriamente.

Há várias técnicas de amostragem, assim como há maneiras de minimizar o problema da não aleatoriedade na escolha dos sujeitos, as quais não serão aqui discutidas por fugir à intenção do texto e porque podem ser facilmente encontradas em livros de Estatística e de pesquisa em educação. Um problema relacionado com o da amostragem, e muito comum na pesquisa em ensino, é o seguinte: muitas vezes o pesquisador em ensino (às vezes o próprio professor) dispõe de dois grupos, um experimental e um controle, aos quais aplica um determinado teste. Obtém, então, dois conjuntos de escores, cujas médias e desvios padrões calcula obtendo valores diferentes. Considerando as médias, fica então a questionar se são realmente diferentes ou se fazem parte de uma mesma distribuição de médias de uma mesma população.

Imagine-se uma população normal de indivíduos com uma média x . Cada amostra dessa população terá uma média provavelmente diferente de x , algumas até muito diferentes. Mas se for feita uma tabulação de um grande número de médias de amostras dessa população, a distribuição de frequências será normal e a média das médias será igual à média da população (x). Isso significa que duas médias bastante diferentes podem fazer parte da mesma distribuição de frequências e, portanto, as amostras correspondentes serem da mesma população, ou seja, não há diferença estatisticamente significativa entre os grupos, é tudo uma questão de amostragem.

No caso da diferença entre as médias dos grupos experimental e controle, o que o pesquisador quer então saber é

se a diferença entre elas é ou não estatisticamente significativa. De maneira análoga, ele vai querer saber se a diferença entre a dispersão dos escores (medida pela variância) nos dois grupos é ou não estatisticamente significativa.

Para resolver esse tipo de problema, a Estatística fornece ao pesquisador os chamados *testes de significância estatística*, como, por exemplo, o "teste t" e o "teste F", também facilmente encontrados em textos introdutórios de Estatística.

Frequentemente, no entanto, pesquisas em ensino envolvem mais de duas amostras. O problema do pesquisador é basicamente o mesmo: a variação entre os escores médios dos diferentes grupos representa uma real diferença entre eles ou é apenas erro de amostragem (ou seja, as médias são de amostras de uma mesma população)?

Nesse caso, a técnica estatística apropriada é a chamada *Análise da Variância*, a qual envolve comparar a variância entre as médias dos grupos com a variância dentro dos grupos. Trata-se aqui de um tratamento estatístico dos dados já um pouco mais sofisticado, mas ainda muito comum na pesquisa educacional quantitativa.

Finalmente, cabe ainda destacar outra forma de análise estatística também muito usada na pesquisa quantitativa em ensino: a correlação. Em várias oportunidades, o pesquisador dispõe de escores para duas variáveis relativas a cada indivíduo e quer saber se há uma relação entre esses variáveis. Por exemplo, em um curso de Física, o professor pode estar interessado em saber se há uma relação entre o conhecimento prévio dos alunos em Matemática e o seu desempenho nas provas de Física. A técnica estatística apropriada para investigar essa questão é a correlação e o resultado de sua utilização é um coeficiente de correlação que varia entre -1,00 e +1,00. O coeficiente igual a +1,00 indica perfeita correlação positiva entre as variáveis consideradas, no sentido de que um aumento nos

escores de uma variável corresponde sempre a um aumento proporcional nos escores da outra. Coeficiente de correlação nulo indica ausência de relação entre as variáveis, enquanto que um coeficiente igual a $-1,00$ indica uma perfeita correlação negativa entre essas variáveis, no sentido de que um aumento nos valores de uma variável é acompanhado sempre de um decréscimo proporcional nos valores da outra.

Naturalmente, ao usar a técnica da correlação, o pesquisador procura descobrir ou confirmar relações causais entre as variáveis, entretanto, o coeficiente de correlação apenas indica quantitativamente a magnitude da relação entre as variáveis. Ou seja, não há necessariamente entre elas uma relação causal. O fato de duas variáveis estarem relacionadas não implica obrigatoriamente que uma esteja causando a outra. Para chegar a uma relação de causa e efeito entre duas variáveis, o pesquisador tem de ser capaz de descartar explicações alternativas para a relação encontrada.

Resumindo, procedimentos estatísticos são praticamente indispensáveis à pesquisa quantitativa em ensino, na medida em que auxiliam o pesquisador a descrever dados, fazer inferências e investigar relações causais. Em outras palavras, tais procedimentos são recursos dos quais lança mão para transformar (i.e., traduzir, representar, comparar, inferir) registros de eventos. A partir dessas transformações o pesquisador chega, então, às asserções de conhecimento e de valor. As primeiras são respostas à(s) questão(ões) básica(s) que ele se propôs a investigar, enquanto as últimas têm a ver com o valor dessas respostas ou com o valor do experimento em si.

Tais respostas, contudo, não são frutos exclusivos do domínio metodológico da pesquisa. Como bem indica o "V" epistemológico de Gowin apresentado na Figura 1, as asserções de conhecimento requerem interação entre os dois lados do "V",

i.e., entre o domínio metodológico e o conceitual, o qual não será aqui discutido. Neste ponto, a título de recapitulação do que foi abordado em relação ao domínio metodológico da pesquisa educacional quantitativa, apresenta-se na Figura 3 um mapa conceitual para esse domínio, i.e., um diagrama destacando hierarquicamente seus principais conceitos e relações significativas entre eles.

Algumas críticas ao enfoque quantitativo

Do que foi dito nas últimas seções relativamente a aspectos importantes de uma abordagem experimental quantitativa à pesquisa em ensino – tais como delineamento, instrumentos de medida, validade e fidedignidade, procedimentos estatísticos – depreende-se facilmente que tal abordagem procura estudar os fenômenos de interesse da pesquisa em ensino primordialmente através de medições objetivas e análises quantitativas. Infere-se também que nessa perspectiva procura-se isolar variáveis e eventos a fim de estudá-los objetivamente. O pesquisador, por sua vez, busca se “desprender” da pesquisa para não introduzir viés. Tais posturas, no entanto, são passíveis de crítica. Segundo Gohn (1984, pp. 4-6), por exemplo:

“Os fatos isolados são abstrações, momentos artificialmente separados do todo, os quais, só quando inseridos no todo correspondente, adquirem verdade e concreticidade [...] estudo das partes e dos processos isolados não é suficiente; ao contrário, o problema essencial consiste em relações organizadas que resultam da interação dinâmica, e fazem com que o comportamento da parte seja diverso, se porventura for examinado isoladamente no interior de um todo [...] Toda pesquisa deve revelar não apenas o movimento presente dos fenômenos, mas como eles se

reproduzem e se transformam. A medida que a pesquisa avança, que se inicia o processo de conhecimento, revela-se o que há de universal no particular."

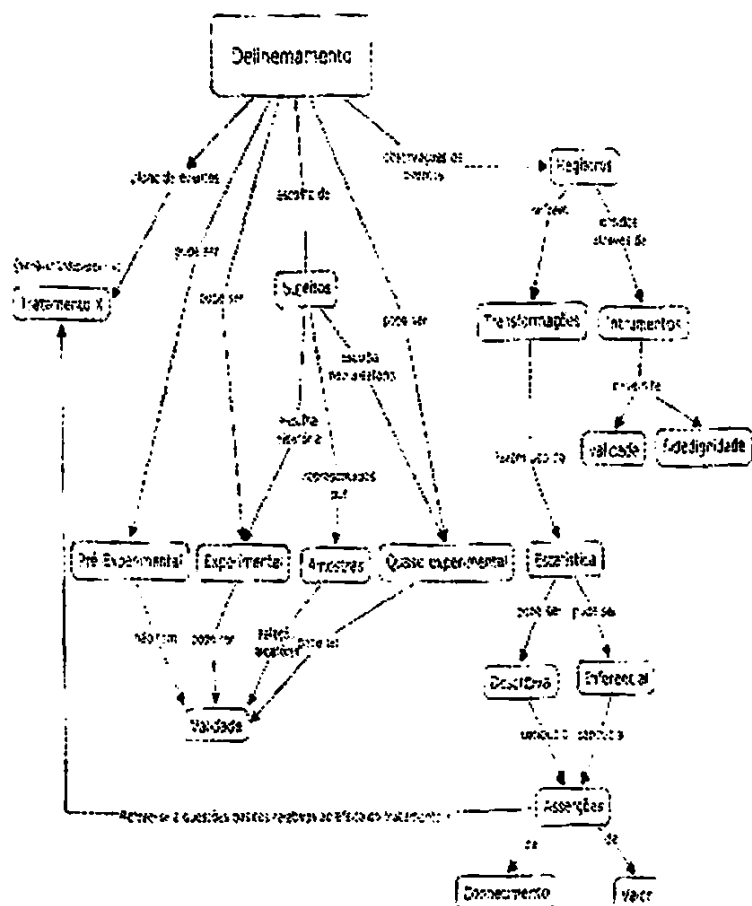


Figura 3 - Um mapa conceitual relativo ao domínio metodológico do "V" de Gowin aplicado à pesquisa educacional quantitativa (Moreira, M.A., 1999)

Da mesma forma, há também que registrar outro ponto destacado pela mesma autora (ibid., p.9):

"Fundamentalmente, uma pesquisa nunca é neutra, ou seja, ela é sempre influenciada, marcada pelos pressupostos teórico-metodológicos de seu autor, ainda que este o negue. Assim, a chamada fase de coleta, registro e sistematização dos dados, na qual é necessária a utilização de alguns instrumentos metodológicos, como mediadores, também não é neutra [...] Não existe coleta de dados sem pressupostos teóricos, e estes estão imbuídos de ideologias e juízos valorativos [...] Portanto, as técnicas de pesquisa não devem ser explicadas como meros instrumentos neutros."

Esta questão é também salientada por Kerlinger (1980) quando fala nos ataques à objetividade implícita na pesquisa empírica lógico-formal (pp. 297-298):

"A objetividade, dizem, leva ao distanciamento, à frieza, à inumanidade. O distanciamento e a frieza da ciência destroem os valores humanos e desumanizam o homem [...] O segundo argumento é mais sofisticado e exerce maior influência, principalmente na Europa, onde faz parte dos ataques marxistas à objetividade. Diz-se – com bastante verdade, por sinal – que ninguém pode ser verdadeiramente objetivo. Quando os cientistas colocam a objetividade como um ideal, estão enganando-se a si próprios e aos outros. Todos nós, inclusive os cientistas, somos conduzidos por nossos valores e motivos. Não podemos ser objetivos. Marxistas e ideólogos semelhantes vão mais além. Dizem, por exemplo, que os valores da sociedade influenciam as hipóteses e as pesquisas de cientistas burgueses e, se esses valores forem corruptos, como o são na sociedade capitalista, então a pesquisa e os resultados são inevitavelmente corruptos. A

objetividade, então, é um mito burguês; é uma arma de opressão. Dizem ainda que é mais importante conhecer a história de uma hipótese do que testá-la."

O próprio Kerlinger, todavia, contra-argumenta (op.cit., p. 298) dizendo que os cientistas não reivindicam nenhuma objetividade pessoal; eles insistem na objetividade como um procedimento metodológico que pode, e deve, ser colocado à parte dos cientistas e de suas predileções. Em relação ao primeiro argumento, de que a objetividade é distanciada, fria e desumana, ele concorda que é correto e que deve ser desta forma. Segundo ele (ibid.):

"É precisamente esta separação da pesquisa científica das preferências humanas, aliada à insistência em testes objetivos empíricos das hipóteses – que, uma vez enunciados publicamente, ficam eles próprios fora dos seres humanos – que aumentou de maneira tão notável nosso conhecimento."

Quanto ao segundo argumento, Kerlinger acredita que também tem pouco peso, a não ser com aqueles que querem acreditar nele (op.cit., p. 299):

"Naturalmente todos nós somos influenciados por nossas preferências. O fato de podermos ou não ser pessoalmente objetivos é discutível. Mas a questão não é essa. A questão, como já observamos, é que os procedimentos da ciência são objetivos – e não os cientistas. Os cientistas, como todos os homens e mulheres são opináticos, dogmáticos, ideológicos – influenciados pelas forças que influenciam todos nós. Esta é a verdadeira razão para insistir em objetividade de procedimento: levar a questão para fora de nós mesmos, sujeitá-la à investigação crítica pública. Não há verdades absolutas, cientificamente falando. Não

podemos saber nada completamente. Há apenas graus relativos de conhecimento válido e fidedigno. Os procedimentos objetivos aumentam a probabilidade de obter conhecimento mais fidedigno e mais válido através da pesquisa."

Como se vê, Kerlinger defende com veemência a abordagem quantitativa à pesquisa educacional, mas seus argumentos, embora carreguem o peso de sua autoridade como pesquisador educacional, não parecem ser suficientes para rebater críticas cada vez mais insistentes (e contundentes) a essa abordagem.

Uma maneira talvez mais direta de criticar a pesquisa quantitativa em ensino é dizer que simplesmente tomou emprestado o modelo das ciências naturais aplicadas, no caso a agricultura. Quer dizer, a pesquisa em ensino nesse enfoque é, essencialmente, similar a uma pesquisa em agricultura na qual estuda-se quantitativamente o efeito de um determinado tratamento (adubo, quantidade de luz, por exemplo) ou de certas condições (tais como tipo de solo e/ou de clima) sobre a qualidade de algum produto agrícola. Nessa pesquisa, manipula-se algumas variáveis, controla-se outras, faz-se medições, compara-se tratamentos, utiliza-se técnicas estatísticas, procura-se correlações. Tudo muito semelhante aos procedimentos usados na pesquisa em ensino. Entretanto, o argumento é de que isso é possível na área de agricultura porque, apesar das variações climáticas de um ano para outro, das diferenças de solo de uma região para outra, as variáveis fundamentais – tais como os componentes químicos, as estruturas genéticas das plantas, a bioquímica do crescimento e metabolismo das plantas – são suficientemente constantes em forma e limitadas em escopo de modo a permitir pesquisa e desenvolvimento através de repetidas medições, previsões e intervenção experimental controlada (Erickson, 1986). Em ensino, contudo, não há nos fenômenos estudados esse grau de uniformidade.

A suposição básica do "modelo agrícola" aplicado à pesquisa em ensino é a de que o que existe de invariante entre diferentes situações de ensino deve emergir de estudos experimentais. Ao investigar-se tais situações, as variações entre elas serão, então, explicadas como erro de variância. O problema é que muitos anos de pesquisas sob essa suposição ainda não levaram ao surgimento de tais invariantes. Essa questão é examinada por Erickson (op.cit. p.131), em um artigo sobre pesquisa qualitativa incluído no terceiro *Handbook for Research on Teaching* (Wittrock, 1986):

"No primeiro Handbook for Research on Teaching só esta teoria e delineamento de pesquisa eram estimulados no capítulo introdutório – o modelo positivista de ciência emprestado das ciências naturais [...] O primeiro 'Handbook' continha o que desde então se tornou um artigo clássico sobre delineamento experimental (Campbell e Stanley, 1963), segundo o qual uma espécie de pesquisa e desenvolvimento agrícolas poderiam ser conduzidos [...] Vinte anos depois parece que existe tanta variabilidade através das salas de aula e tanta variabilidade na implementação de tratamentos que a avaliação de programas em larga escala através de métodos quase-experimentais é muito problemática. A medida que isso se tornou aparente, estudo após estudo, o próprio Campbell (1975) sugeriu o uso de métodos observacionais mais naturalísticos – estudos de caso feitos por observadores participantes, ou estudos de 'documentação', os quais dariam uma visão detalhada da real estrutura e processo de implementação de programas."

De um modo geral, as críticas à pesquisa quantitativa em ensino, ou em educação para ser mais abrangente, partem de defensores de uma outra orientação, um outro enfoque: a pesquisa qualitativa.

A pesquisa qualitativa

Embora raízes recentes do que é comumente chamado de pesquisa educacional qualitativa possam ser encontradas em métodos antropológicos desenvolvidos no início do século passado, o uso acentuado dessa abordagem na pesquisa em ensino é um movimento atual, iniciado há pouco mais de três décadas.

Para quem tem familiaridade com a pesquisa quantitativa, mesmo que seja apenas através de leitura de artigos de pesquisa ou sobre pesquisa, uma boa maneira de entender os significados atribuídos à pesquisa qualitativa é contrastar as duas abordagens. De uma perspectiva bem abrangente, pode-se iniciar dizendo que as visões de mundo, as filosofias, as metodologias são diferentes.

"A pesquisa quantitativa está baseada em uma filosofia positivista que supõe a existência de fatos sociais com uma realidade objetiva independente das crenças dos indivíduos, enquanto que a qualitativa tem raízes em um paradigma segundo o qual a realidade é socialmente construída [...] A pesquisa quantitativa procura explicar as causas de mudanças em fatos sociais, primordialmente através de medição objetiva e análise quantitativa, enquanto a qualitativa se preocupa mais com a compreensão do fenômeno social, segundo a perspectiva dos atores, através de participação na vida desses atores [...] A pesquisa quantitativa tipicamente emprega delineamentos experimentais ou correlacionais para reduzir erros, vieses e outros ruídos que impedem a clara percepção dos fatos sociais, enquanto o protótipo do estudo qualitativo é a etnografia [...] O pesquisador quantitativo ideal é desprendido para evitar vieses, enquanto o pesquisador qualitativo fica 'imerso' no fenômeno de interesse." (Firestone, 1987, pp.16-17)

Embora seja comum usar-se o termo pesquisa qualitativa e distingui-la da pesquisa quantitativa, muitos autores não gostam

dessa terminologia. Eisner (1981), por exemplo, argumenta que qualquer forma de pesquisa quantitativa deve necessariamente preocupar-se com qualidades, de modo que a distinção que se deve procurar fazer não deve ser entre formatos qualitativos e não qualitativos de pesquisa, mas, sim, entre o que é estudado de uma maneira científica e o que é estudado artisticamente. Eisner (op. cit., pp. 5-9) identifica várias dimensões, através das quais abordagens científicas e artísticas diferem:

"A pesquisa científica está sempre preocupada com questões de validade [...] Em abordagens artísticas, os cânones de fidedignidade de testes e de amostras não se aplicam [...] Sua validade, se é que o termo é apropriado, é determinada por seu grau de credibilidade [...] Validade nas artes é o produto da capacidade de persuasão de uma visão pessoal [...] Abordagens científicas tendem a focalizar comportamentos manifestos de grupos ou indivíduos [...] Abordagens artísticas enfocam mais as experiências dos indivíduos e o significado que suas ações têm para outros [...] O foco das abordagens artísticas à pesquisa está nos significados e experiências das pessoas que atuam na subcultura em estudo [...] A pesquisa científica em ciências sociais procura ir do particular para o geral e está interessada em particulares somente na medida em que representam o geral. Seleção aleatória é a base do processo [...] O pesquisador artisticamente orientado procura tornar o particular vivido de modo que suas qualidades possam ser experienciadas e porque ele acredita que o particular tem uma contribuição a dar para a compreensão do geral [...] Examinando o estilo dos trabalhos publicados em revistas de pesquisa, fica claro que a padronização de estilo é considerada uma virtude. Espera-se que o pesquisador identifique o

problema, resuma a literatura, descreva os instrumentos e as amostras, explique o tratamento, apresente, discuta os resultados e, finalmente, aponte possíveis implicações. Neste formato, qualquer traço de personalidade do investigador deve ser neutralizado [...] Em abordagens artísticas à pesquisa, a padronização da forma é contraprodutiva [...] o que essas abordagens procuram é explorar a forma para informar [...] Portanto, ser hábil em formas artísticas de expressão – ser capaz de escrever – é criticamente importante para o pesquisador educacional artisticamente orientado [...] A pesquisa científica visa a produção de ideias que permitirão antecipar eventos futuros, se não controlá-los [...] A pesquisa artisticamente orientada não procura controlar ou produzir asserções formais preditivas. Ela busca é a explicação. Está mais próxima de uma atividade hermenêutica do que tecnológica [...] Não está atrás de um algoritmo, e, sim, de uma heurística [...] Abordagens artísticas buscam um pluralismo metodológico [...] estão mais preocupadas com a criação de significados do que com a descoberta de verdades [...] procuram a criação de imagens que terão significado para as pessoas [...] Verdade implica singularidade e monopólio. Significado implica relativismo e diversidade.

Ao distinguir entre abordagens artísticas e científicas à pesquisa, Eisner assume uma posição similar a de Erickson – ao falar do “modelo positivista de ciência emprestado das ciências naturais” (1986, p.131) usado na pesquisa em ensino – no sentido de que a pesquisa quantitativa tem origem na metodologia das ciências físicas. Quer dizer, cientistas sociais, e pesquisadores educacionais em particular, tomaram “emprestada” a

metodologia das ciências físicas, especialmente da Física, para investigar o mundo social e humano. É claro que adaptações foram feitas, mas certas premissas implícitas na pesquisa educacional quantitativa revelam ainda hoje a influência da pesquisa nas ciências físicas. Por exemplo, segundo Smith (1983), o pesquisador nesta orientação tende a adotar o papel de observador de uma realidade com existência própria, tenta eliminar quaisquer vieses e ideias pré-concebidas, procura não se envolver emocionalmente e colocar-se "acima" de crenças e valores contextuais, busca ficar "isento", não participar, limitar-se (objetivamente) ao "que é" e não (subjetivamente) ao que "deveria ser". Historicamente, a distinção entre pesquisa quantitativa e qualitativa remonta ao debate entre o positivismo e o idealismo ocorrido no final do século XIX. Sem querer equacionar realismo com positivismo, nem idealismo com uma perspectiva filosófica idealista mais profunda, Smith (1983, pp. 8 e 9) argumenta que debates contemporâneos sobre abordagens quantitativas e qualitativas podem ser conduzidos à luz de diferenças entre realismo e idealismo:

"Realismo está baseado na ideia de que a realidade existe independente de nós [...] conhecimento e verdade é uma questão de correspondência, o que é verdade é o que corresponde à realidade [...] A investigação da realidade através do método particular que chamamos científico (daí o realismo científico) pode ser conduzida independente dessa realidade; a atividade de investigar não afeta o que está sendo investigado [...] Idealismo, em contraste, argumenta que o que existe depende da mente humana. O sujeito e o objeto, percebidos como dois elementos pelos realistas, tornam-se um para os idealistas que não percebem nenhuma realidade independente dos esforços mentais de moldar e criar [...] o que é investigado não é independente do processo de investigação [...] o que conta como conhecimento ou

que deve ser considerado verdade é uma questão de concordância em um contexto histórico e socialmente vinculado [...] Para idealistas, instrumentos não têm lugar independente daquilo a que se destinam medir, São extensões dos pesquisadores e operam como um elemento em sua tentativa de construir ou de dar força à realidade. Para realistas, instrumentos são uma maneira de atingir uma medição acurada de um objeto com existência própria. Nesse contexto, instrumentos válidos são os que produzem representações acuradas [...] Na perspectiva de uma realidade independente, se o investigador parar de estudar algo, esse algo continuará a existir e permanecerá ligado a outras coisas da mesma maneira [...] baseados na ideia de que a realidade é feita ou pelo menos moldada, idealistas acreditam que a realidade não pode ter existência prévia à investigação e deixará de existir se a pesquisa for abandonada."

A questão da pesquisa qualitativa versus pesquisa quantitativa pode também ser discutida à luz de outras perspectivas um tanto quanto dicotômicas como naturalismo versus racionalismo ou relativismo versus objetivismo (Smith e Heshusius, 1986). Entretanto, a menos que se queira aprofundar muito no assunto, pouco haveria a acrescentar ao que já foi dito sobre essa questão. Deixando um pouco de lado a distinção entre pesquisa quantitativa e qualitativa, seria interessante focalizar agora apenas a abordagem qualitativa a exemplo do que já foi feito com a quantitativa

Algumas características de uma abordagem qualitativa à pesquisa em ensino

Pesquisa qualitativa é um termo que tem sido usado alternativamente para designar várias abordagens à pesquisa em ensino, tais como pesquisa etnográfica, participativa observacional, estudo de caso, fenomenológica construtivista,

interpretativa, antropológica cognitiva. Cada uma dessas abordagens forma um todo coerente, englobando suposições internamente consistentes sobre natureza humana, sociedade, objeto de estudo e metodologia (Jacob, 1987, p.1), porém compartilham muitas semelhanças e por questão de simplicidade são comumente chamadas de pesquisa qualitativa. Erickson (1986), no entanto, prefere o termo *interpretativa* porque é mais inclusivo, não dá à pesquisa a conotação de ser essencialmente não quantitativa e, principalmente, porque sugere a característica básica comum de todas essas abordagens – o interesse central da pesquisa na questão dos significados que as pessoas atribuem a eventos e objetos, em suas ações e interações dentro de um contexto social e na elucidação e exposição desses significados pelo pesquisador.

Erickson (1986, p.126) vê a pesquisa interpretativa (qualitativa) em ensino fundamentalmente sob a ótica do significado:

"Seres humanos, supõe a perspectiva interpretativa, criam interpretações significativas do ambiente físico e comportamental que os rodeia [...] Através da cultura, seres humanos compartilham significados aprendidos e em determinadas situações frequentemente parecem ter criado interpretações significativas similares. Mas estas similaridades superficiais mascaram uma diversidade subjacente; em uma dada ação, não se pode supor que os comportamentos de dois indivíduos, atos físicos de forma similar, tenham o mesmo significado para os dois indivíduos [...] Portanto, uma distinção analítica crucial em pesquisa interpretativa é entre comportamento, o ato físico, e ação, que é o comportamento mais as interpretações de significados atribuídas por quem atua e por aqueles com os quais o ator interage [...] O objeto da pesquisa interpretativa social é ação, não comportamento [...]"

Pode-se distinguir entre significados denotativos e conotativos. Denotativos são aqueles compartilhados culturalmente, aqueles que permitem a comunicação entre os indivíduos de uma determinada cultura; conotativos são os idiossincráticos, pessoais, diferentes para cada indivíduo. Quer dizer, um certo objeto ou evento dentro de uma certa cultura tem significados comuns aos membros dessa cultura, mas, ao mesmo tempo, cada indivíduo pode interpretá-lo de maneira diferente. Naturalmente, nessa diversidade de interpretações significativas está presente na microcultura da sala de aula:

"A pesquisa interpretativa presume que as microculturas das salas de aula diferem de uma para outra, independente do grau de similaridade em características gerais entre elas [...] Dessa perspectiva, as similaridades superficiais entre grupos e salas de aula é que são triviais e ilusórias na pesquisa em ensino [...] Isso não quer dizer que a pesquisa em ensino não está interessada na descoberta de universais, mas sim que ela segue uma rota diferente para descobri-los [...] A busca não é de universais abstratos, aos quais se chega através de generalizações estatísticas de amostras para populações, mas, sim, de universais concretos atingidos através do estudo detalhado de um caso específico e da comparação desse caso com outros estudados igualmente com grande detalhe." (op.cit., p.130)

Assim como indivíduos compartilham certos significados, alguns aspectos do que ocorre em qualquer situação de ensino são generalizáveis a outras situações, outros são específicos de uma dada situação ou específicos de indivíduos em particular:

"A tarefa do pesquisador é então desvelar os diferentes níveis de universalidade e particularidade confrontados em um caso específico – o que é largamente universal, o

que é generalizável a outras situações, o que é peculiar a esse caso [...] Portanto, a preocupação primordial da pesquisa interpretativa é particularização ao invés de generalização. Descobre-se universais manifestados concreta e especificamente, não em abstração e generalidade." (ibid., p.130)

Retomando o "V" epistemológico de Gowin (Figura 1, p. 14) como instrumento orientador da discussão, percebe-se que muito do que foi dito até aqui tem a ver com a "filosofia" ou, poder-se-ia dizer, o paradigma da pesquisa qualitativa. Mas não só isso: pelo menos implicitamente falou-se de questões básicas e de eventos de interesse da pesquisa em ensino sob a ótica interpretativa. Os fenômenos de interesse da pesquisa qualitativa em ensino têm também a ver com ensino propriamente dito, aprendizagem, currículo, avaliação e contexto, mas são analisados sob outros pontos de vista. A sala de aula, por exemplo, é vista como um ambiente organizado social e culturalmente, no qual ações mudam constantemente, significados são adquiridos, trocados, compartilhados. Naturalmente, o contexto assume então um papel de destaque, pois os significados e as ações são contextuais. A pesquisa interpretativa procura analisar criticamente cada significado em cada contexto. O pesquisador, nessa perspectiva, pergunta-se continuamente que significados têm as ações e os eventos de ensino, aprendizagem, avaliação e currículo para os indivíduos que deles participam. Indaga-se permanentemente sobre o que está acontecendo e como isso se compara com o que está acontecendo em outros contextos.

As diferenças entre os enfoques quantitativo e interpretativo não decorrem então do fenômeno de interesse estudado, mas da maneira como ele é estudado. Em um estudo qualitativo observacional, o pesquisador não procura testar hipóteses, e, sim, desenvolvê-las. Ao invés de começar o estudo com hipóteses, ele

parte de suposições tentativas sobre o fenômeno a ser investigado. Tais suposições servem de guia ao pesquisador (Lutz e Ramsey, 1974, p.5).

O domínio metodológico da pesquisa interpretativa

E o domínio metodológico, propriamente dito, da pesquisa interpretativa? Aparentemente, a metodologia da pesquisa qualitativa guarda semelhanças com a da quantitativa, principalmente examinando o lado direito do "V" epistemológico de uma perspectiva bem abrangente: o pesquisador interpretativo registra eventos, obtém dados, transforma-os e faz asserções. Mas a natureza de seus procedimentos é diferente. O pesquisador quantitativo faz uso de instrumentos de medida (e.g., testes, questionários), seleciona amostras, aplica tratamentos, procura correlações, faz inferências, usa testes estatísticos, busca validade interna e externa. O investigador interpretativo observa participativamente, de dentro do ambiente estudado, imerso no fenômeno de interesse, anotando cuidadosamente tudo o que acontece nesse ambiente, registrando eventos – talvez através de audiotapes ou de videotapes –, coletando documentos tais como trabalhos de alunos, materiais distribuídos pelo professor, ocupa-se não de uma amostra no sentido quantitativo, mas de grupos ou indivíduos em particular, de casos específicos, procurando escrutinar exhaustivamente determinada instância tentando descobrir o que há de único nela e o que pode ser generalizado a situações similares.

O pesquisador qualitativo também transforma dados e eventualmente faz uso de sumários, classificações e tabelas, mas a estatística que usa é predominantemente descritiva. Ele não está preocupado em fazer inferências estatísticas, seu enfoque é descritivo e interpretativo ao invés de explanatório ou preditivo. Interpretação dos dados é o aspecto crucial do domínio metodológico da pesquisa qualitativa. Interpretação do ponto de

vista de significados. Significados do pesquisador e significados dos sujeitos.

Em ensino de ciências, por exemplo, grande parte da pesquisa conduzida nas últimas décadas situou-se na área de concepções alternativas, inicialmente chamadas de intuitivas ou espontâneas.

Naturalmente, a análise interpretativa dos dados gera asserções de conhecimento, as quais são publicadas pelo pesquisador sob a forma de um relatório ou artigo de pesquisa. Nessa etapa assume grande importância outra faceta da pesquisa qualitativa: a narrativa. Ao invés de usar gráficos, coeficientes, tabelas estatísticas para apresentar resultados e asserções de conhecimento, o pesquisador interpretativo narra o que fez e sua narrativa concentra-se não nos procedimentos, mas nos resultados. Suas asserções dependem de sua interpretação e só terão validade para o leitor (que pode ser um colega pesquisador, um professor, um administrador, o próprio sujeito da pesquisa) na medida em que este concordar com essa interpretação.

Para isso, o pesquisador enriquece sua narrativa com trechos de entrevistas, excertos de suas anotações, vinhetas, exemplos de trabalhos de alunos, entremeados de comentários interpretativos procurando persuadir o leitor, buscando apresentar evidências que suportem sua interpretação e, ao mesmo tempo, permitam ao leitor fazer julgamentos de modo a concordar ou não com as asserções interpretativas do pesquisador. Nesse sentido é que, como disse Eisner (1981, p. 6), a validade de uma pesquisa qualitativa é determinada por sua credibilidade.

Em muitos casos, a coleta de dados era feita através de entrevistas clínicas gravadas em audiotapes. No relato da investigação, o pesquisador geralmente apresentava ao leitor vários trechos de transcrições das entrevistas através dos quais procurava persuadi-lo que determinado aluno possuía uma certa

concepção alternativa. Uma outra maneira de investigar concepções alternativas dos alunos é através de mapas conceituais (Moreira e Buchweitz, 1987; Moreira, 1987, 2010) – diagramas indicando relações entre conceitos.

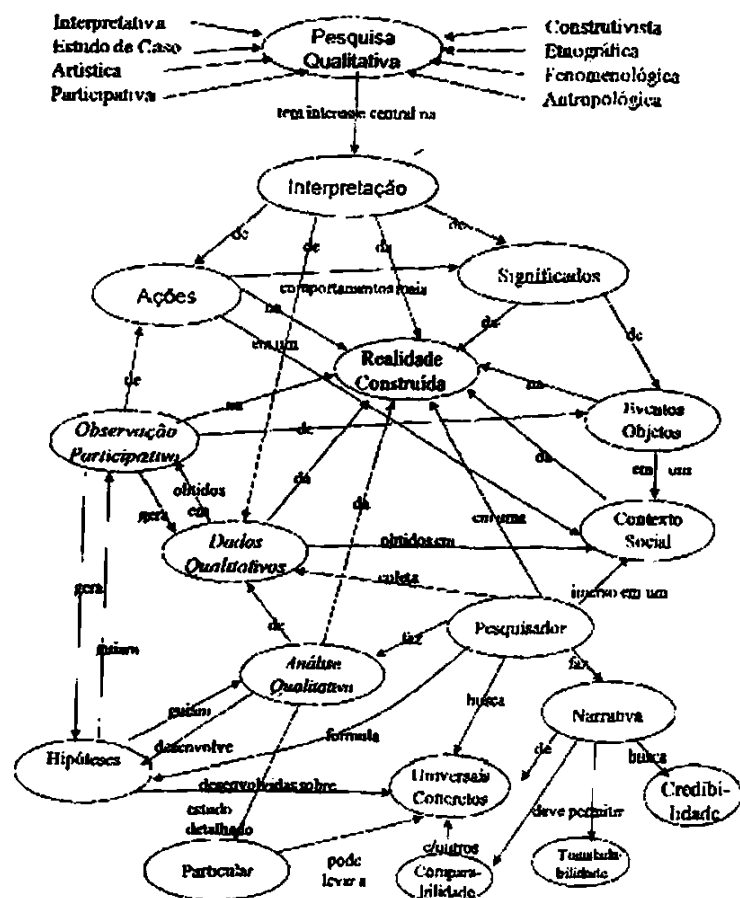


Figura 4 – Um mapa conceitual para o domínio metodológico do "V" de Gowin aplicado à pesquisa educacional qualitativa (M. A. Moreira, 1999)

De um modo geral, o professor (pesquisador) solicita ao estudante que trace um mapa conceitual para determinado

conjunto de conceitos, ou para um determinado conteúdo. Feito isso, o aluno explica o mapa, oralmente ou por escrito, enquanto o pesquisador observa, registra e interage com ele, procurando captar os significados atribuídos (pelo aluno) às relações explicitadas no mapa. Ou seja, procura colocar-se na perspectiva de quem faz o mapa para tentar interpretá-lo segundo a ótica do aluno.

O importante é descobrir os significados que o aluno atribui ao que está no mapa; a partir daí é possível, por exemplo, inferir concepções alternativas. No relato desse tipo de estudo, o pesquisador apresenta exemplos de mapas conceituais traçados pelos alunos juntamente com trechos de suas explicações, aos quais adiciona seus comentários interpretativos.

Na Figura 4 (acima), apresenta-se, a título de síntese, um mapa conceitual para o domínio metodológico da pesquisa educacional interpretativa.

Pesquisa quantitativa versus qualitativa: uma visão geral

Tal como esquematizado na Figura 5, as abordagens qualitativa e quantitativa à pesquisa em educação subscrevem diferentes paradigmas, diferentes visões de mundo, que levam a diferentes maneiras de ver os fenômenos educacionais, de selecionar objetos e eventos para estudar esses fenômenos, de formular questões, de fazer registros, de transformar registros em dados e de analisar dados. Naturalmente, isso leva a diferentes asserções de conhecimento sobre as quais são feitos diferentes tipos de asserções de valor.

Antes de prosseguir com esse paralelo entre os enfoques qualitativo e quantitativo à pesquisa educacional, deve ser destacado que a palavra "diferente" é uma palavra-chave no contexto da comparação que está sendo feita, isto é, tais abordagens estão sendo consideradas diferentes não alternativas.

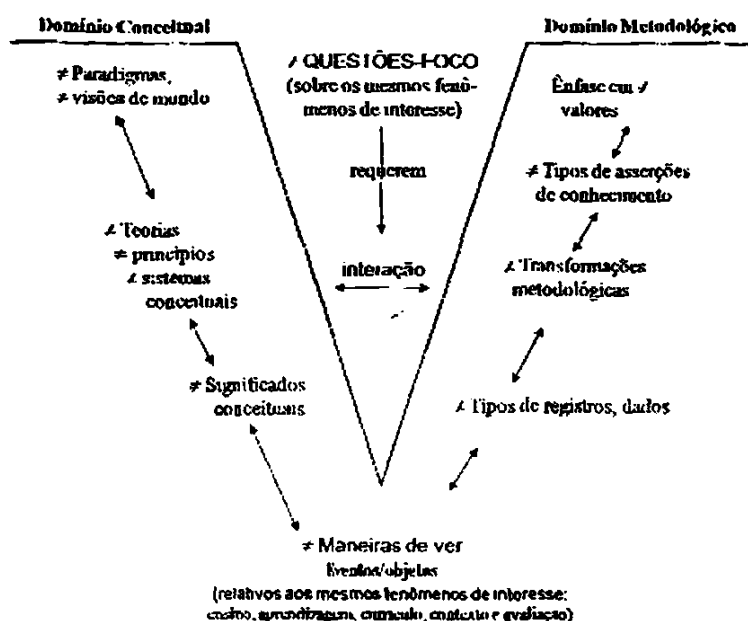


Figura 5 - Uma visão geral de onde estão as diferenças entre as perspectivas qualitativa e quantitativa à pesquisa educacional à luz do "V" epistemológico de Gowin.

Por outro lado, tal como sugerido na Figura 6, elas são diferentes perspectivas para examinar os mesmos fenômenos de interesse. A suposição aqui é aquela feita no início, i.e., a de que existem cinco lugares comuns em educação: professor (ensino), aluno (aprendizagem), currículo, avaliação e contexto (Schwab, 1983, apud Gowin, 1981; Novak e Gowin, 1984). Ou seja, direta ou indiretamente, os fenômenos educacionais sempre envolvem professores, alunos e um currículo em um certo contexto, bem como a avaliação. Isso, no entanto, não implica uma relação de causa e efeito entre ensino e aprendizagem, mas, sim, que esses cinco lugares comuns estão presentes nas experiências educacionais e devem ser considerados no estudo de eventos educativos. É nesse sentido que os fenômenos de interesse das

perspectivas qualitativa e quantitativa à pesquisa em educação estão sendo considerados os mesmos.

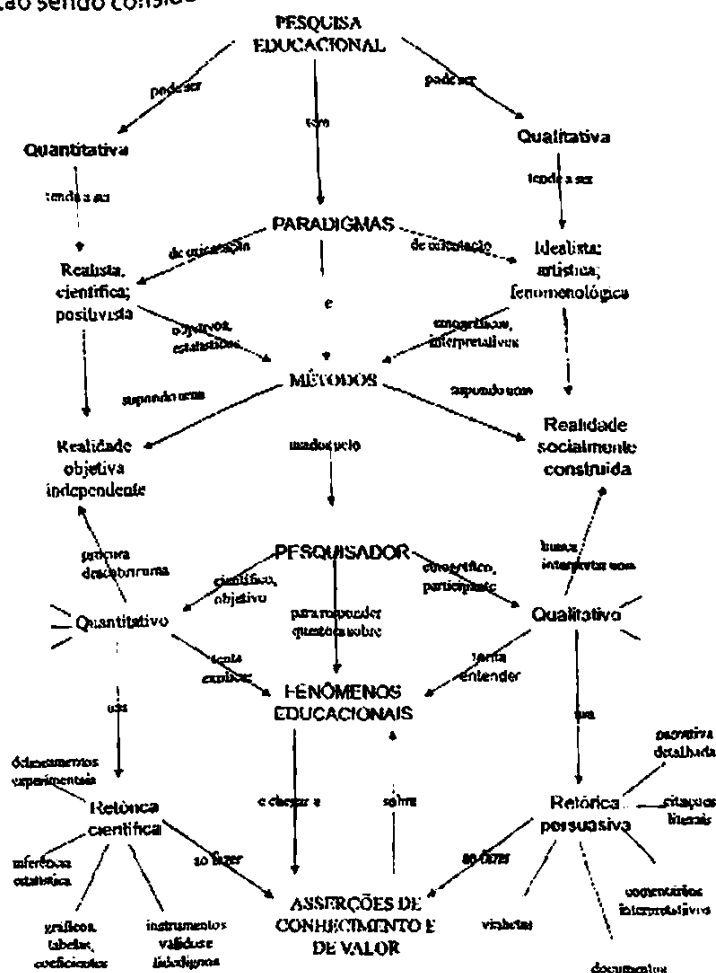


Figura 6 - Um esquema conceitual mostrando diferenças entre as abordagens quantitativa e qualitativa à pesquisa educacional (M.A. Moreira, 1999).

Naturalmente, cada perspectiva seleciona diferentes eventos

ou objetos para estudar esses fenômenos e essa seleção é influenciada pela bagagem conceitual do pesquisador. Alguns pesquisadores qualitativos pretendem estarem livres de teorias e hipóteses quando fazem suas observações, mas esse compromisso com uma pureza virginal é, como colocam Lutz e Ramsey (1974, p. 6), apenas uma nobre aspiração, ou uma romântica noção de trabalho de campo, como diz Erickson (1986, p. 240).

A posição deste autor é também a de que toda pesquisa, qualitativa ou quantitativa, é conduzida sob referenciais de teorias, princípios e sistemas conceituais, embora muitas vezes não explicitos. Obviamente, tais referenciais diferem. A pesquisa quantitativa, por exemplo, pode estar mais interessada em teorias explanatórias e preditivas do tipo das teorias da Física, enquanto a qualitativa pode ter mais interesse em sistemas de referência mais interpretativos, descritivos, mas em ambos os casos os referenciais existem. De maneira análoga, as duas abordagens são igualmente influenciadas por conceitos; entretanto, nesse caso é provável que a maioria dos rótulos conceituais sejam os mesmos e as diferenças residam nos significados atribuídos a esses rótulos. Isto é, pesquisadores qualitativos e quantitativos constituem diferentes comunidades de usuários de conceitos e, em consequência, os mesmos rótulos conceituais como ensino, aprendizagem, currículo e contexto, podem ter significados diferentes em cada comunidade. Todavia, essas comunidades provavelmente compartilham alguns significados para esses conceitos de tal modo que possam comunicar-se e que a asserção de que os cinco lugares comuns da educação, que são o ensino, a aprendizagem, o currículo, o contexto e a avaliação, seja válida para ambas.

Resumindo, o argumento que está sendo elaborado nesta seção é o de que as diferenças entre os modos qualitativo e quantitativo de pesquisa educacional estão nos seus domínios

conceituais e metodológicos. Em ambos os casos, esses domínios interagem e conduzem a diferentes questões-foco e diferentes asserções de conhecimento sobre os mesmos fenômenos de interesse, envolvendo os cinco lugares comuns da educação: professor (ensino), aluno (aprendizagem), currículo, contexto e avaliação. Além disso, levam também a enfatizar diferentes tipos de asserções de valor. Por exemplo, a pesquisa qualitativa tende a destacar valores sociais das asserções de conhecimento, enquanto que a quantitativa provavelmente está mais preocupada com valores instrumentais dos resultados.

Pesquisa quantitativa versus qualitativa: uma visão detalhada

A Figura 6 procura fornecer subsídios para diferenciar ainda mais as perspectivas qualitativa e quantitativa à pesquisa educacional, deixando de lado outras possíveis perspectivas justamente para enfatizar as mais utilizadas.

No topo dessa figura está o conceito de pesquisa educacional, a qual pode ser conduzida sob diferentes abordagens com diferentes paradigmas subjacentes: o qualitativo relacionado com perspectivas idealistas, artísticas e fenomenológicas e o quantitativo vinculado a posturas realistas, científicas e positivistas.

A parte desses rótulos aos quais cada perspectiva está usualmente associada, as principais diferenças paradigmáticas entre essas perspectivas estão em suas visões de mundo, suas maneiras de ver a realidade: do ponto de vista quantitativo, existe uma realidade objetiva, independente, a ser descoberta através da pesquisa; verdade é uma questão de correspondência com a realidade; verdadeiro é o que corresponde à realidade. Na ótica da pesquisa qualitativa, não existe realidade independente, ela é socialmente construída, depende da mente humana, verdade é uma questão de concordância, não de correspondência.

Estes paradigmas subscrevem diferentes metodologias:

métodos qualitativos são etnográficos, interpretativos, descritivos, enquanto procedimentos quantitativos são ditos objetivos, científicos, acurados. Consistentemente com esses métodos e paradigmas, o pesquisador qualitativo procura um entendimento interpretativo de uma realidade socialmente construída na qual ele está imerso, enquanto o pesquisador quantitativo busca descobrir uma realidade com existência própria, da qual ele ou ela deve estar o mais desvinculado possível a fim de evitar qualquer viés. As metas desses procedimentos metodológicos são chegar às respostas às questões sobre fenômenos educacionais, i.e., às asserções de conhecimento que são apresentadas através de distintas retóricas, usando diferentes meios de persuadir a audiência: a retórica quantitativa é fria, procurando ser isenta de valores, neutra, padronizada; faz uso de números, gráficos, tabelas, coeficientes e outros indicadores objetivos, buscando persuadir o leitor da acuidade, objetividade, cientificidade da metodologia e dos resultados da pesquisa. A retórica qualitativa não evita a linguagem cotidiana carregada de valores; é detalhada, provendo elementos em quantidade suficiente – citações, vinhetas, documentos, comentários interpretativos – para persuadir o leitor que as asserções de conhecimento são interpretações válidas daquilo que os eventos significam do ponto de vista dos sujeitos da pesquisa.

Todas essas diferenças estão resumidas na Tabela 1, a qual complementa o esquema conceitual apresentado na Figura 6. Por outro lado, tal como foi sugerido antes, muitos pesquisadores parecem estar combinando as metodologias, ou simplesmente misturando-as, e ignorando os paradigmas subjacentes a cada uma delas. Este ponto será discutido a seguir.

A questão da compatibilidade

Parece não haver consenso em relação às respostas referentes às questões sobre a possível compatibilidade entre as abordagens qualitativa e quantitativa à pesquisa em ensino. Há respostas afirmativas e negativas. Alguns pesquisadores acreditam que as duas abordagens são plenamente compatíveis,

outros acham que há uma incompatibilidade fundamental. Naturalmente, há também posições intermediárias. Firestone (1987, pp. 16 e 20) coloca o assunto da seguinte maneira:

"Com a crescente aceitação dos métodos qualitativos em educação, o debate passou a ser sobre qual deve ser sua relação com os métodos quantitativos. Nos extremos há dois grupos: os puristas e os pragmatistas. Os puristas acreditam que os dois tipos de métodos são incompatíveis porque estão baseados em paradigmas que fazem suposições diferentes sobre o mundo e sobre o que é uma pesquisa válida. [...] Os pragmatistas veem uma relação mais instrumental entre paradigma e métodos [...] métodos são conjuntos de técnicas [...] ambos os tipos de métodos podem ser associados aos atributos do paradigma qualitativo ou do quantitativo. [...] Estudos quantitativos geralmente estão baseados em um paradigma positivista enquanto a pesquisa qualitativa está frequentemente baseada em um paradigma fenomenológico. [...] Em decorrência, cada tipo de método usa diferentes técnicas de apresentação para projetar suposições divergentes sobre o mundo e diferentes meios de persuadir o leitor sobre suas conclusões. Ainda assim não são diametralmente opostos. [...] Usados separadamente, proveem diferentes tipos de informação. Quando enfocam a mesma questão, estudos qualitativos e quantitativos podem triangular – isto é, usar diferentes métodos para avaliar a robustez ou estabilidade dos resultados. Quando estudos usando diferentes métodos têm resultados similares, pode-se ter mais certeza que os resultados não são influenciados pela metodologia. Quando os resultados divergem, é preciso mais pesquisa, mas uma comparação de estudos pode frequentemente sugerir importantes linhas de pesquisa a seguir."

	Pesquisa Quantitativa	Pesquisa Qualitativa
Paradigmas	Realidade objetiva, independente de crenças, com existência própria. Investigar não afeta o que está sendo investigado. Instrumentos são uma maneira de atingir medições precisas de objetos e eventos com existência própria; Instrumentos válidos são os que produzem representações acuradas da realidade. Se o investigador parar de estudar algo, esse algo continuará a existir e permanecerá ligado a outras coisas da mesma maneira. Verdade é uma questão de correspondência com a realidade (Smith, 83).	Realidade socialmente construída; não há realidade independente dos esforços mentais de criar e moldar o que existe depende da mente humana. O que é investigado não é independente do processo de investigação. Instrumentos não têm lugar independente daquilo a que se destinam medir; são extensões dos pesquisadores em sua tentativa de construir ou de dar forma à realidade. A realidade não tem existência prévia à investigação e deixará de existir se a pesquisa for abandonada. Verdade é questão de concordância em um contexto (Smith, 83).
Questões	Procuram explicar causas de mudanças em fatos sociais, principalmente através de medição objetiva e análise quantitativa (Firestone, 87). Focalizam comportamentos de grupos ou indivíduos (Eisner, 82). Buscam a predição e controle de eventos, algoritmos, verdades, universais abstratos aos quais se chega através de generalizações estatísticas de amostras para populações (Erickson, 86).	Buscam a compreensão do fenômeno social segundo a perspectiva dos atores através de participação em suas vidas (Firestone, 87). Focalizam significados e experiências; ação ao invés de comportamento (Eisner, 82). Procuram a explicação interpretativa; heurísticas ao invés de algoritmos; universais concretos atingidos através do estudo detalhado de um caso e da comparação com outros estudados com igual detalhe (Erickson, 86).
Métodos	Tomam emprestado o modelo das ciências físicas para investigar o mundo social e humano. Ocupam-se de delineamentos experimentais, quase-experimentais e correlacionais; testes de hipóteses; Instrumentos válidos e fidedignos; testes de significância; amostragem, inferência estatística; generalização.	Usam técnicas etnográficas, estudos de caso, antropologia educacional. Ocupam-se de observação participativa; significados individuais e contextuais; interpretação; desenvolvimento de hipóteses; indicadores de baixa inferência; casos, grupos ou indivíduos específicos, particularização. Eventualmente fazem uso de estatística descritiva.

Pesquisado	Distante para evitar viés (Firestone, 87); objetivo. Limita-se ao que é. Quantifica registros de eventos. Usa meios científicos. Busca fidedignidade e validade.	Imerso no fenômeno de interesse (Firestone, 87), participante. Anota, ouve, observa, registra, documenta, busca significados, interpreta. Procura credibilidade.
	Padronizada, estatística, objetiva. Extenso uso de tabelas, gráficos, coeficientes. Procura neutralizar a personalidade do pesquisador. Fria, científica, buscando convencer o leitor que a análise feita é neutra, impessoal: (Firestone, 87).	Persuasiva, descritiva, detalhada. Extenso uso de transcrições, vinhetas, documentos, exemplos, comentários interpretativos. Usa a linguagem cotidiana com suficiente detalhe para evidenciar que são válidas as interpretações dos significados tidos pelos atores (Erickson, 86)

Tabela 1. Um paralelo entre os enfoques quantitativo e qualitativo à pesquisa educacional.

A posição de Eisner é também conciliadora (1981, p.9):

"Cada abordagem ao estudo de situações educacionais provê de maneira única a sua própria perspectiva. Cada uma ilumina a seu modo as situações que os seres humanos procuram compreender. O campo da educação, em particular, precisa evitar o monismo metodológico. Nossos problemas devem ser atacados de todas as maneiras que forem frutíferas [...] A questão não é contrastar qualitativo e não qualitativo, mas como abordar o mundo educacional. É para o artístico que devemos nos voltar não como uma rejeição ao científico, mas porque com ambos podemos atingir visão binocular. Olhar através de um só olho nunca proporcionou muita profundidade de campo."

Shulman (1981, p.12) parece assumir uma posição pragmatista:

"Temos que primeiramente entender nosso problema, decidir que questões queremos formular

e então escolher o modo de investigação disciplinada mais apropriado a essas questões. Se os métodos adequados são altamente quantitativos e objetivos, ótimo. Se forem mais subjetivos ou qualitativos, podemos também usá-los responsabilmente."

Smith (1983, p. 12) tende a considerar incompatíveis os enfoques qualitativo e quantitativo:

"Cada abordagem adota diferentes procedimentos e tem diferentes implicações epistemológicas. Uma abordagem assume uma posição sujeito-objeto em relação à matéria de ensino, a outra uma posição sujeito-sujeito. Uma separa fatos e valores, a outra os percebe inseparavelmente misturados. Uma busca leis, a outra procura compreensão. Considerando nosso estado atual de pensamento, tais posições não parecem ser compatíveis. Isso não quer dizer que os dois enfoques nunca poderão ser conciliados, somente que no momento atual as divisões são mais notáveis do que as possibilidades de unificação."

Em outro trabalho (Smith e Heshusius, 1986), Smith manifesta-se explicitamente contra a suposta compatibilidade entre as perspectivas qualitativa e quantitativa. No resumo deste trabalho (op.cit., p.4) lê-se:

"O argumento neste artigo é que a alegada compatibilidade e apelo à cooperação entre a pesquisa quantitativa e a qualitativa são insustentáveis. Além disso, essas posições têm o desafortunado efeito de encerrar um importante debate. A fim de elaborar esse argumento, este

trabalho revisa sucintamente a transição do conflito para a cooperação entre as duas perspectivas e então mostra como a compatibilidade está baseada em uma confusão sobre duas diferentes definições de método. Finalmente, a discussão centraliza-se em porque esse debate – por envolver questões cruciais para a compreensão de quem somos e o que fazemos como pesquisadores – deve ser revigorado.”

Ao final do mesmo trabalho (ibid., p.11), os autores reiteram sua posição:

“Como todas essas questões são cruciais para nós [...] não devem ser deixadas para os filósofos na esperança de que resolvam nossos problemas. Em termos simples, encerrar o debate com base em injustificadas suposições de compatibilidade e cooperação é um passo errado no momento errado.”

Alguns autores, como Miles e Huberman (1984, pp. 20 e 21), veem a questão da compatibilidade de um ponto de vista prático, funcional, mecanicista:

“Olhando cuidadosamente para a pesquisa que está sendo feita em nome de uma ou outra epistemologia, parece que poucos pesquisadores não estão combinando as duas perspectivas. Uma consequência dessa combinação é que mais e mais estudos incluem não somente dados quantitativos, mas também qualitativos [...] Tanto o neopositivismo como o neoidealismo constituem um continuum epistemológico, não uma dicotomia [...] muito da pesquisa atual representa combinações paradigmáticas [...] no mundo da pesquisa real vê-se cada vez mais estudos ligando dados qualitativos e quantitativos [...] Apesar do crescente interesse em estudos qualitativos,

nos ressentimos de métodos claramente definidos para tirar significados válidos de dados qualitativos. O que precisamos são métodos práticos, comunicáveis e não autoenganosos: científicos no sentido positivista da palavra e dirigidos para um entendimento interpretativo no melhor sentido do termo."

A posição mecanicista de Miles e Huberman é duramente criticada por Marshal (1986, pp. 26-28):

"Tentativas de aproximar critérios e métodos positivistas para assegurar validade podem minar os pontos fortes da pesquisa qualitativa [...] O valor especial da pesquisa qualitativa – a exploração de significados em situações do mundo real – deve ser preservado [...] Este é o momento errado para esforços de padronização e de busca de técnicas mais refinadas para provar como sabemos o que sabemos, como fizeram Miles e Huberman. É o momento de captar ideias dos desafios e divisar planos de pesquisa que permitam perspectivas alternativas [...] A pesquisa qualitativa deve preservar suas qualidades únicas e sua habilidade de explorar em busca de significados. [...] não deve ser submetida a aproximações dos métodos e critérios estabelecidos por positivistas, não agora. Deve preservar a oportunidade de explorar e capitalizar em análises divergentes."

Considerando tais pontos de vista, talvez a melhor postura em relação ao debate pesquisa qualitativa x quantitativa seja mesmo a de continuá-lo, sem assumir apressada e simplistamente a posição de que são apenas métodos alternativos, nem radicalizar, considerando-os totalmente incompatíveis. A pesquisa em ensino só pode crescer com a continuação desse tipo de debate.

Fidedignidade e validade na pesquisa qualitativa

Alguns pesquisadores da linha etnográfica como, por exemplo, Marshal (1984, p. 28), veem a tentativa de aplicar

conceitos da pesquisa quantitativa à investigação qualitativa como uma ameaça às suas potencialidades e como um esforço para submetê-la aos padrões da pesquisa quantitativa. Sem descuidar desse tipo de preocupação, examinemos, a título de ilustração, a aplicabilidade de dois conceitos fundamentais para a pesquisa quantitativa – fidedignidade e validade – à pesquisa interpretativa. Para isso, tomemos como base o trabalho de Lecompte e Goetz (1982, pp. 31-60) sobre problemas de fidedignidade e validade na pesquisa etnográfica, lembrando que a fidedignidade na perspectiva quantitativa se refere ao grau de reproducibilidade das medidas (ou estudos), enquanto a validade tem a ver com a acuidade dos resultados, com o grau em que as conclusões efetivamente representam a realidade empírica, com o grau em que os instrumentos realmente estão medindo o que se pretende medir.

Para ser justo, uma vez que anteriormente foram apontadas várias críticas à abordagem quantitativa, cabe frisar aqui que talvez as maiores críticas ao enfoque qualitativo (pelos adeptos da perspectiva quantitativa) é que os estudos etnográficos geralmente não têm fidedignidade e validade. Naturalmente, os puristas ignoram essa crítica sob o argumento de tais conceitos não terem significados, ou possuírem outros significados, na perspectiva etnográfica. Outros, como Lecompte e Goetz, preocupam-se com a questão da credibilidade da pesquisa etnográfica e apresentam propostas para aumentar a fidedignidade e a validade dos estudos interpretativos.

O conceito de fidedignidade aplicado à pesquisa etnográfica refere-se à medida em que estudos etnográficos podem ser replicados. Naturalmente, essa é uma tarefa gigantesca para pesquisadores na área, pois a pesquisa etnográfica ocorre em ambientes naturais, às vezes é planejada para registrar processos de mudança, lida com comportamentos humanos, de modo que, a rigor, nenhum estudo pode ser replicado exatamente. Além

disso, o processo etnográfico é também personalista; nenhum etnógrafo trabalha como outro.

Em função de tudo isso, falhas do pesquisador em especificar precisamente o que foi feito podem criar sérios problemas de fidedignidade. Precisamente aí é que pesquisadores etnográficos podem procurar aumentar a fidedignidade de seus estudos. Ou seja, reportando clara e precisamente aspectos como o status do pesquisador em relação ao grupo estudado, seu grau de participação, fontes de informação, contexto físico e social em que foram coletados os dados, métodos de coleta e de análise dos dados, pressupostos teóricos. Enfim, descrevendo de maneira precisa tudo o que possa facilitar a realização, com boa aproximação, de um estudo replicativo. Isso aumentaria, em princípio, a fidedignidade externa do estudo. A fidedignidade interna se refere à questão, dentro de um único estudo, de se vários observadores concordarão entre si. Isto é, se os significados tidos por vários observadores serão suficientemente congruentes de modo que venham a descrever fenômenos da mesma maneira e chegar às mesmas conclusões sobre eles. Pesquisadores etnográficos geralmente usam várias estratégias para reduzir ameaças à fidedignidade interna (op.cit., p. 41): indicadores e narrativas com baixo nível de inferência participação de vários pesquisadores, registro mecânico de dados (áudio e videoteipes, por exemplo).

Em relação à validade, a situação é bastante favorável para a pesquisa etnográfica, pois a validade talvez seja seu ponto forte. A maneira como os dados são coletados e analisados, o escrutínio de situações particulares diminuiu os riscos de invalidade interna, i.e., de que não se esteja "medindo" o que se pensa estar "medindo".

As ameaças à validade externa de um estudo etnográfico têm a ver com efeitos que obstruem ou reduzem a comparabilidade e a transladabilidade do estudo. Enquanto comparabilidade e transladabilidade são fatores que podem contribuir para a generalização de estudos experimentais, eles praticamente

definem o grau de validade externa de um estudo etnográfico; são fatores cruciais para a replicação da pesquisa etnográfica (op.cit., p.34)

"Comparabilidade requer que o etnógrafo delineie as características do grupo estudado, ou de construtos gerados, tão claramente que possam servir de base para comparação com outros grupos semelhantes e não semelhantes. Translabilidade supõe que métodos de pesquisa, categorias analíticas e características de fenômenos e grupos sejam identificadas tão explicitamente que comparações possam ser conduzidas com confiança. Assegurar comparabilidade e translabilidade provê a base para fazer comparações. Para etnógrafos, ambas funcionam como análogas das metas de pesquisa mais estreitamente controladas: generalização dos resultados de pesquisa e produção de asserções causais."

Portanto, aumentar a validade externa de um estudo qualitativo implica em aumentar seu grau de comparabilidade e translabilidade. Todas essas considerações sobre fidedignidade e validade convergem em um ponto: a necessidade de se descrever com precisão e detalhe tudo o que foi feito. Na prática, isso se reflete em relatos relativamente longos, com uma retórica persuasiva, rica em descrições que, ao mesmo tempo em que procura convencer o leitor, permitem, com maior ou menor grau de aproximação, replicar o estudo. Nessa perspectiva, os relatos de uma pesquisa interpretativa procuram dar evidências de validade e fidedignidade dos estudos feitos.

O domínio conceitual da pesquisa qualitativa

Para encerrar essa discussão sobre pesquisa qualitativa e retomando mais uma vez o fio da meada através do "V" de

Gowin, é necessário examinar ainda uma questão pendente: o aspecto teórico-conceitual da pesquisa qualitativa em ensino.

Novamente aí há posições divergentes, mas parece haver mais consenso no sentido de que a pesquisa interpretativa é também conduzida a partir de concepções teóricas.

Por exemplo, Lutz e Ramsey (1974, p.17), falando do uso de métodos de campo antropológicos na pesquisa em educação, fazem os seguintes comentários:

"De um lado da questão estão aqueles que não querem ser afetados por teorias ou hipóteses quando começam um estudo de campo. Esse compromisso com uma pureza virginal é uma nobre aspiração. Entretanto, a maioria dos pesquisadores já está enviesada por uma bagagem conceitual da qual não podem se livrar. Recomenda-se a estes que estão tão sobrecarregados que usem essa bagagem a seu favor e conscientizem seus leitores dos perigos dos vieses trazidos com ela para o campo. [...] Não é provável que alguém esteja totalmente livre de conceitos que o predisponham a ver certas coisas. [...] É difícil acreditar que pesquisadores possam ir ao campo sem vieses conceituais. [...] Mas sem hipóteses e teorias de trabalho, a coleta de dados é aleatória e sem objetivos. Um foco é tão importante como uma mente aberta, e alguma delimitação é tão importante quanto abrangência. [...] É suficiente dizer que uma base teórica para um estudo etnográfico é extremamente importante."

Erickson (1986, p. 140) classifica como romântica a possibilidade de conduzir um trabalho observacional de campo totalmente livre de concepções:

"Antropólogos, em especial, criaram a mística concepção de que o trabalho de campo não é ensinável. [...] Esta é uma noção extremamente romântica de trabalho de campo. Entra-se nele sem concepções e aprende-se os métodos usando-os (da mesma maneira que uma pessoa pode aprender a nadar sendo jogada dentro de uma piscina). [...] Pode-se argumentar que não há induções puras. Sempre trazemos padrões interpretativos para a experiência. Desse ponto de vista, a tarefa do trabalho de campo é tornar-nos cada vez mais reflexivamente conscientes dos padrões de interpretação daqueles que observamos e dos nossos próprios padrões interpretativos culturalmente aprendidos que trazemos para o campo."

Em outro ponto do mesmo trabalho (op.cit., p.152), Erickson refere-se outra vez a essa questão:

"Há significados que não podem ser completamente antecipados teoricamente antes de se entrar no ambiente a ser estudado. Devido a esses significados locais desconhecidos e as dimensões não identificadas do problema de pesquisa é que o trabalho de campo é necessário. Mas como destacamos antes, o pesquisador de campo é sempre guiado por um conjunto geral de interesses de pesquisa e muitas vezes por um conjunto bastante específico de questões de pesquisa."

De fato, é difícil imaginar a possibilidade de conduzir uma pesquisa em ensino – ou em qualquer outra área – sem a ótica de uma bagagem teórico-conceitual. Observe-se que o domínio conceitual é inerente à essência do processo de pesquisa tal como

ilustrado pelo "V" de Gowin, instrumento heurístico que está sendo usado como referencial na presente discussão sobre pesquisa em ensino (quantitativa ou qualitativa). Seria talvez incoerente tentar analisar a pesquisa qualitativa à luz do "V" epistemológico se não se acreditasse que o domínio conceitual tem também, nesse tipo de pesquisa, um papel fundamental.

Referências

- Best, J.W. (1970). *Research in education*. Englewood Cliffs, NJ: Prentice Hall.
- Campbell, D.T. (1978). Qualitative knowing in action research. In: Brenner, M. et al. (Eds.), *The social context of method*. New York: St. Martin's.
- Campbell, D.T. and Stanley, J.C. (1963). Experimental and quasi-experimental design for research on teaching. In Gage, N.L. (Ed.), *Handbook of research on teaching*. (2nd. ed.). Chicago: Rand McNally.
- Campbell, D.T. e Stanley, J.C. (1979). *Delineamentos experimentais e quase-experimentais de pesquisa*. Tradução de R.A.T. Di Dio. São Paulo: EPU-EDUSP.
- Eisner, E.W. (1981). On the differences between scientific and artistic approaches to qualitative research. *Educational Researcher*, 10 (4): 5-9.
- Elzey, F.F. (1967). *A first reader in statistics*. Belmont, CA: Brooks Cole Publishing Co..
- Erickson, F. (1986). Qualitative methods in research on teaching. In: Wittrock, M.C. (Ed.), *Handbook of research on teaching*. (3rd. ed.) New York : Macmillan Publishing Co.
- Firestone, W.A. (1957). Meaning in method: the rethoric of quantitative and qualitative research. *Educational Researcher*, 16(7): 16-21.
- Fox, D.J. (1969). *The research process in education*. New York: Holt, Rinehart and Winston.

- Gohn, M.G.M. (1984). A pesquisa nas ciências sociais: considerações metodológicas. *Cadernos CEDES*, 12: 3-14.
- Gowin, D.B. (1970). The structure of knowledge. *Educational Theory*, 20(4): 319-28.
- Gowin, D.B. (1981). *Educating*. Ithaca, NY: Cornell University Press.
- Gowin, D.B. and Millman, J. (1969). Research methodology - a point of view. *Review of Educational Research*, 39(5): 553-59.
- Jacob, E. (1987). Qualitative research traditions: a review. *Review of Educational Research*. 57 (1): 1-50.
- Kerlinger, F.N. (1980). *Metodologia da pesquisa em ciências sociais*. Tradução de H.M. Rotundo. São Paulo e Brasília: EPU-EDUSP e INEP.
- Lecompte, M.D. e Goetz, J.P. (1962). Problems of reliability and validity in ethnographic research. *Review of Educational Research*, 52(1): 31-60.
- Lutz, F.W. and Ramsey, M.A. (1974). The use of anthropological field methods in education. *Educational Researcher*, 3(10): 5-9.
- Marshall, C. (1984). The wrong time for mechanistics in qualitative research. *Educational Researcher*, 13(9): 26-28.
- Miles, M.B. and Huberman, A.M. (1984). Drawing valid meaning from qualitative data: toward a shared craft. *Educational Researcher*, 13(5): 20-30.
- Moreira, M.A. (1985). O "V" de Gowin na análise de experimentos: uma alternativa instrucional. In: Moreira, M.A. *Atividade docente na universidade: alternativas instrucionais*. Porto Alegre e Rio Grande: D.C. Luzzato Editores e Editora da FURG.
- Moreira, M.A. (1987). Concept mapping as a possible strategy to detect and to deal with misconceptions in physics. In : Novak, J.D. (Ed.), *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics Education*. Ithaca, NY, Cornell University, Department of Education.

Moreira, M.A. (1990). *Pesquisa em ensino: o Vê epistemológico de Gowin*. São Paulo: Editora Pedagógica e Universitária Ltda.

Moreira, M.A. (2010). *Mapas conceituais e aprendizagem significativa*. São Paulo: Centauro Editora.

Moreira, M.A. e Buchweitz, B. (1987). *Mapas conceituais*. São Paulo: Editora Moraes.

Moreira, M.A. e Koff, E.D. (1985). O questionário como instrumento de coleta de informações sobre o ensino. In : Moreira, M.A. (Org.). *Ensino na universidade: sugestões para o professor*. Porto Alegre: Editora da Universidade.

Runkel, P.J. and McGrath, J.E. (1972). *Research on human behavior: a systematic guide to method*. New York : Holt, Rinehart and Winston.

Shulman, L.S. (1981) Disciplines of inquiry in education: an overview. *Educational Researcher*, 10 (6): 5-12.

Smith, J.K. (1983). Quantitative versus qualitative research: an attempt to clarify the issue. *Educational Researcher*, 12 (3): 6-13.

Smith, J.K. and Heshusius, L. (1986). Closing down the conversation: the end of the quantitative-qualitative debate among educational inquirers. *Educational Researcher*, 15(1): 4-13.

Vianna, H.M. (1978). *Testes em educação*. São Paulo: IBRASA.

Wittrock, M.C. (Ed.). (1986). *Handbook of research on teaching*. (3 ed.). New York: Mcmillan Publishing Co.

Capítulo II

Pesquisa em Ensino: Métodos Qualitativos¹

M. A. Moreira

Resumo

Após uma breve comparação entre os paradigmas quantitativo e qualitativo, o texto é dedicado apenas ao enfoque qualitativo descrevendo, com mapas conceituais, a etnografia, o estudo de casos e a investigação-ação. Ao final, aborda-se a questão da triangulação metodológica.

Introdução

Neste texto, a pesquisa em educação em ciências é entendida como a produção de conhecimentos resultante da busca de respostas a perguntas sobre ensino, aprendizagem, currículo e contexto educativo em ciências, assim como sobre o professorado de ciências e sua formação permanente dentro de um quadro epistemológico, teórico e metodológico consistente e coerente. Porém, o mesmo se ocupará só do domínio metodológico dessa pesquisa e, nesse domínio, o foco será particularmente a *metodologia qualitativa*.

A metodologia da pesquisa em educação em ciências é a mesma da pesquisa em educação e esta foi dominada, ao longo do século XX, por dois paradigmas clássicos: um inspirado na metodologia das ciências naturais, enfatizando observações

¹ Programa Internacional de Doctorado en Enseñanza de las Ciencias. Universidad de Burgos, Espanha, Universidade Federal do Rio Grande do Sul, Brasil. Texto de Apoio n.º 14. Publicado em *Actos del PIDEC*, 4:25-55, 2002.

empíricas quantificáveis e adequadas para tratamentos estatísticos; o outro derivado da área humanística com ênfase em informações holísticas e qualitativas e em abordagens interpretativas.

O filósofo alemão Wilhelm Dilthey argumentava (apud Husén, 1988), já em 1890, que as humanidades tinham sua própria lógica de pesquisa e que a diferença entre as ciências naturais e as humanidades era que estas buscavam compreender enquanto que as primeiras procuravam explicar (op. cit., p. 17). Essa distinção parece hoje muito simplificada, mas serve para mostrar que o debate é antigo.

A pesquisa em educação começa (segundo Landsheere, 1988), por volta de 1900, sob o nome de "pedagogia experimental", com pesquisadores como Meumann na Alemanha, Binet na França, Thorndike nos Estados Unidos e Claparède na Suíça, pouco tempo depois da "psicologia experimental" iniciada por Wundt em Leipzig, por volta de 1880, e fortemente influenciada por ela. De acordo com esse mesmo autor (op. cit., p. 11), nas três primeiras décadas do século passado, a pesquisa educativa teve uma acentuada ênfase quantitativa, dirigida ao estudo da eficácia no ensino, particularmente nos Estados Unidos.

Posteriormente, nos anos 30 a 50, a crise econômica e a guerra levaram a uma grande redução na atividade de pesquisa em educação, em especial na Europa. Porém, nessa mesma época aparecem como campo de interesse dos pesquisadores os estudos de natureza sociológica questionando a escola como mecanismo de reprodução de distinções sociais e práticas discriminatórias (ibid., p. 13). Nas décadas de 60 e 70, novamente, houve um período de muito apoio financeiro à pesquisa educativa, em particular aquela voltada ao desenvolvimento curricular em ciências e matemática. Foi também uma época de domínio da abordagem quantitativa, porém a reação a essa "tradição positivista" começava a ser cada vez mais forte no contexto da pesquisa educativa em nível internacional. A tal ponto que nos anos 80 e 90 houve um claro domínio da

abordagem qualitativa na pesquisa em educação em geral e em ciências em particular.

Esta pequena e pouco rigorosa resenha histórica foi feita somente para reforçar a asserção de que os dois paradigmas clássicos – o quantitativo e o qualitativo – dominaram a pesquisa educativa no século XX com uma certa alternância.

Desde o ponto de vista epistemológico, em todo esse tempo surgiram teses de incompatibilidade paradigmática kuhneana (Smith, 1983; Smith e Heshusius, 1986; Marshal, 1986), de compatibilidade prática, funcional, pragmática (e.g., Shulman, 1981; Miles e Huberman, 1984), de conciliação e triangulação metodológicas (e.g., Eisner, 1981; Firestone, 1987) ou integradoras, como a de Keeves (1988) e a de Bericat (1998). Essas colocações são discutidas em um texto anterior e complementar a este (Moreira, 2000).

Neste texto ficaremos na perspectiva integradora, mas este tema será deixado para o final. Por enquanto, continuaremos na distinção entre os dois paradigmas. Mesmo que esta distinção e inclusive a ideia de paradigma possam ser objeto de críticas por parte de pesquisadores (e.g., Walker e Evers, 1988), continuaremos nela, por algum tempo, por razões didáticas. Por essas razões, no item seguinte distinguiremos os dois paradigmas clássicos como se constituíssem uma dicotomia. Uma vez estabelecida essa distinção, o texto abordará sucessivamente metodologias qualitativas como a etnografia, o estudo de caso, a pesquisa ação e outras. Na conclusão, abandonaremos a visão dicotômica e defenderemos uma postura integradora.

Os dois paradigmas clássicos

Na Tabela 1² se estabelece uma comparação dicotômica entre os paradigmas quantitativo e qualitativo em termos de pressupostos, objetivos, métodos, o papel do pesquisador e a retórica de apresentação do conhecimento produzido. Por ser

² Esta tabela é a mesma Tabela 1 do capítulo anterior, no qual foi usada para fazer uma espécie de fechamento do texto. Neste, ao contrário, ela é usada como um tipo de organizador prévio.

autoexplicativa, tal tabela não será comentada.

A pesquisa qualitativa (interpretativa)

Na Figura 1 se apresenta um mapa conceitual para a abordagem qualitativa na pesquisa educativa. Já que os mapas conceituais não são autoexplicativos, faremos um breve comentário. No topo aparece o conceito de *pesquisa qualitativa* ao qual estão associados atributos como *interpretativa, holística, naturalista, participativa, interacionista simbólica, construtivista, etnográfica, fenomenológica e antropológica*.

O interesse central dessa pesquisa está em uma *interpretação dos significados* atribuídos pelos sujeitos à suas ações em uma *realidade socialmente construída*, através de *observação participativa*, isto é, o pesquisador fica *imerso* no fenômeno de interesse. Os *dados* obtidos por meio dessa participação ativa são de *natureza qualitativa* e analisados de forma correspondente. As *hipóteses* são *geradas* durante o processo investigativo. O pesquisador busca *universais concretos* alcançados através do estudo profundo de *casos particulares* e da comparação desse caso com outros estudados também com grande profundidade. Através de uma *narrativa detalhada*, o pesquisador busca *credibilidade* para seus modelos interpretativos.

A pesquisa qualitativa também é chamada de *naturalista*, porque não envolve manipulação de variáveis, nem tratamento experimental (é o estudo do fenômeno em seu acontecer natural); *fenomenológica*, pois enfatiza os aspectos subjetivos do comportamento humano, o mundo do sujeito, suas experiências cotidianas, suas interações sociais e os significados que dá a essas experiências e interações; *interacionista simbólica*, uma vez que toma como pressuposto que a experiência humana é mediada pela interpretação, a qual não se dá de forma autônoma, mas na medida em que o indivíduo interage com outro, é por meio de interações sociais como vão sendo construídas as *interpretações*, os significados, a visão de realidade do sujeito (André, 1998, pp. 17-18).

Erickson (1986, p. 119), um pesquisador muito conhecido em

educação, prefere o termo *pesquisa interpretativa* para se referir a toda uma família de abordagens de pesquisa participativa observacional, em lugar de pesquisa qualitativa, por ser mais abrangente, por evitar a ideia de que seja essencialmente não quantitativa e por apontar ao interesse central dessa pesquisa, que é o significado humano em um contexto social e sua elucidação e exposição pelo pesquisador.

	Paradigma quantitativo realista/racionalista	Paradigma qualitativo idealista/naturalista
Pressupostos	Realidade objetiva, independente de crenças, com existência própria. O pesquisador não afeta ao que se está pesquisando. Os instrumentos são uma maneira de alcançar medições precisas de objetos e eventos com existência própria; instrumentos válidos são os que produzem representações exatas da realidade. Se o pesquisador deixa de estudar algo, esse algo continuará existindo e permanecerá ligado a outras coisas da mesma maneira. Dualismo sujeito-objeto. Verdade é uma questão de correspondência com a realidade (Smith, 83).	Realidade socialmente construída; não há realidade independente dos esforços mentais de criar e moldar; o que existe depende da mente humana. O que se pesquisa não é independente do processo de pesquisa. Os instrumentos não têm lugar independentemente de aquilo que têm que medir, são extensões dos pesquisadores na sua tentativa de construir ou de dar forma à realidade. A realidade não tem existência prévia à pesquisa e deixará de existir se a pesquisa for abandonada. Não há dualismo sujeito-objeto. Verdade é questão de concordância em um contexto (Smith, 83).
Objetivos	Procuram explicar causas de mudanças em fatos sociais, principalmente através de medição objetiva e análise quantitativa (Firestone, 87). Focam comportamentos de grupos ou indivíduos (Eisner, 81). Buscam a predição e controle de eventos, algoritmos, verdades universais abstratos aos que se chega através de generalizações estatísticas de amostras para populações (Erickson, 86).	Procuram a compreensão do fenômeno social segundo a perspectiva dos atores através de participação em suas vidas (Firestone, 87). Focam significados e experiências, ações em vez de comportamentos (Eisner, 81). Procuram a explicação interpretativa; heurísticas em vez de algoritmos; universais concretos alcançados através do estudo detalhado de um caso e da comparação com outros estudados com igual detalhe (Erickson, 86).

Métodos	Tomam emprestado o modelo das ciências físicas para pesquisar o mundo social e humano. Ocupam-se de desenhos experimentais, quase-experimentais e correlacionais; testes de hipóteses; instrumentos válidos e fidedignos; testes de significância; amostragem; inferência estatística; generalização. Seguem um modelo hipotético-dedutivo.	Usam técnicas etnográficas, estudos de caso, antropologia educativa. Ocupam-se de observação participativa; significados individuais e contextuais; interpretação; desenvolvimento de hipóteses; indicadores de baixa inferência; casos, grupos ou indivíduos específicos; particularização. Podem fazer uso de estatística descritiva. São bem mais indutivos.
Papel do pesquisador	Distante para evitar vies (Firestone, 87); objetivo. Limita-se ao que é. Quantifica registros de eventos. Faz uso de meios científicos. Busca fidedignidade e validade.	Imerso no fenômeno de interesse (Firestone, 87); participante. Faz anotações, ouve, observa, registra, documenta, busca significados, interpreta. Procura credibilidade.
Retórica	Padronizada, estatística, objetiva. Extenso uso de tabelas, gráficos, coeficientes. Procura neutralizar a personalidade do pesquisador. Fria, científica, buscando convencer o leitor de que a análise feita é neutra impessoal (Firestone, 87).	Persuasiva, descritiva, detalhada. Extenso uso de transcrições, vinhetas, documentos, exemplos, comentários interpretativos. Usa a linguagem cotidiana com suficiente detalhe para evidenciar que são válidas as interpretações dos significados tidos pelos atores (Erickson, 86).

Tabela 1. Um paralelo entre os paradigmas quantitativo e qualitativo na pesquisa educativa (M.A. Moreira, 2000)

Para ele (op. cit., p. 121), a pesquisa interpretativa envolve: a) intensa e ampla participação no contexto pesquisado, b) cuidadosos registros do que ocorre nesse contexto juntamente com outras fontes de evidência (e.g., anotações, documentos, exemplos de coisas feitas pelos sujeitos, gravações em áudio ou em vídeo) e c) análise reflexiva de todos esses registros e evidências assim como descrição detalhada (i.e., utilizando a narrativa e transcrições literais de verbalizações dos sujeitos).

Para Erickson (op. cit., p. 129), a tarefa da pesquisa interpretativa é descobrir maneiras específicas através das quais formas locais e não locais de organização social e cultural se relacionam com atividades de pessoas específicas em suas

sistemas como a ideologia política, os pressupostos das subculturas étnicas e sociais a respeito do papel de mulheres e homens, das relações adequadas entre adultos e crianças, etc., isto é, por aculturação.

Uma vez apresentadas algumas características gerais da pesquisa qualitativa, passaremos a focar, sucessivamente, três metodologias principais dentro dessa abordagem: a *etnografia*, o *estudo de casos* e a *pesquisa-ação*.

A etnografia

A etnografia é uma metodologia, uma ferramenta para estudar e compreender uma cultura, a maneira de vida de um grupo de pessoas, isto é, suas ideias, crenças, valores e pressupostos, seus comportamentos e as coisas que fazem (Ogbu et al., 1988, p. 48). Em outras palavras, *a etnografia é uma tentativa de descrever uma cultura* (André, 1998, p. 19).

A pesquisa etnográfica consta essencialmente de uma descrição de eventos que ocorrem no cotidiano da vida de um grupo, com especial atenção às estruturas sociais e condutas de indivíduos a respeito do seu status de pertença ou membresia ao grupo, e uma interpretação do que significa todo isso para a cultura do grupo (Taft, 1988, p. 71).

Na etnografia, o pesquisador participa, o quanto é possível, da vida normal do grupo pesquisado, da cultura pesquisada. A pesquisa é conduzida no cenário natural dos eventos, no contexto em que ocorrem os acontecimentos, através de observação participativa. Para chegar a uma *compreensão descritiva contextualizada* da cultura, o pesquisador tem de se meter em tal cultura, aprender a "*linguagem nativa*", como disse o célebre antropólogo Malinowski, *interagir* com os membros dessa cultura, desenvolver uma *compreensão empática* da vida das pessoas tal como elas a percebem, assim como uma *perspectiva holística* do grupo. Tudo isso, é lógico, implica um longo "tempo de residência" nessa cultura. Isto é, o pesquisador deve permanecer

"imerso" na cultura pesquisada durante um período de tempo "suficientemente grande" para, de acordo com Malinowski, contextualizar os dados em um "account" holístico e coerente e descrever "a vida tal como é vivida" (Ogbu et al., 1988, p. 50).

O pesquisador etnográfico tem, por conseguinte, um duplo papel: participante e observador. Por um lado, ele tem de se envolver com o grupo, "aculturar-se" nele. Por outro, deve ser capaz de observar, interpretar, discernir, desenvolver uma perspectiva holística. Sendo ao mesmo tempo observadores e participantes, os pesquisadores etnográficos não são desprendidos do fenômeno de interesse; eles influenciam seus dados e são influenciados por estes em todas as etapas de observação, interpretação e descrição (Taft, 1988, p. 72). A grande vantagem de ser observador participante parece ser, ao mesmo tempo, a principal dificuldade que deve enfrentar o pesquisador participante. Ao mesmo tempo em que tenta "pertencer" à cultura pesquisada, ele deve também ser capaz de "mirá-la desde fora", interpretá-la, descrevê-la.

A observação participativa é a principal técnica de pesquisa etnográfica. Porém, as entrevistas são muito utilizadas. Os dados gerados por essas duas técnicas são frequentemente complementados por outros como documentos, narrativas, histórias de vida, artefatos e diagramas produzidos no grupo pesquisado. No geral, o pesquisador etnográfico busca recolher toda a informação possível, não só através de observação participativa e entrevistas, para interpretá-la indutivamente e construir uma realidade social que é sua compreensão descritiva contextualizada da cultura pesquisada.

A metodologia etnográfica é qualitativa e holística, faz uso da intuição, empatia e outras habilidades do pesquisador para interpretar descritivamente uma cultura. Seu interesse está em descobrir (no sentido de construir uma descrição compreensiva contextualizada), e não em verificar. Porém, isso não implica não

ter nenhuma hipótese ou teoria inicial. O pesquisador etnográfico não começa um trabalho de campo "sem ter nada na cabeça". Isso não existe. Ele sempre terá conhecimentos teóricos prévios que, de alguma maneira, vão orientar seus passos iniciais, mas não deve ter hipóteses e teorias que serão verificadas ou rejeitadas no estudo. Isto é, o pesquisador etnográfico não deve ter ideias pré-concebidas, tal como recomendou Malinowski (apud Taft, 1988, p. 74).

As hipóteses são formuladas recursivamente durante o processo e o desenvolvimento da pesquisa. De forma gradativa, pode emergir uma base teórica para a compreensão dos processos grupais. Essa base teórica é conhecida como *teoria fundamentada*, isto é, fundamentada no próprio processo de pesquisa (ibid.), ou fundamentada nos dados. Tal teoria, que foi desenvolvida de maneira indutiva, provavelmente gerará hipóteses úteis para guiar, inicialmente, novas observações participativas. Porém, não se está falando aqui do indutivismo científico ingênuo tão criticado epistemologicamente, nem de hipóteses que serão "comprovadas" em estudos "mais rigorosos". (O tema da credibilidade dos "resultados" da pesquisa qualitativa será discutido mais adiante em outro item.)

A etnografia é uma metodologia de pesquisa em antropologia que chegou à pesquisa em educação não faz muito, nos anos 60 do século passado. Uma grande diversidade de etnografias educativas têm sido desenvolvidas desde essa época, porém o conceito de cultura permaneceu como constructo unificador. Três orientações principais podem ser identificadas (Ogbu et al., 1988, pp. 50-51) a partir de distintos níveis de análise e diferentes ênfases em suas definições de cultura: *etnografia holística* (também conhecida como etnografia tradicional, velha etnografia ou macroetnografia), a *etnociência* (também chamada de nova etnografia ou antropologia cognitiva) e a *microetnografia* (etnografia da comunicação).

A *etnografia holística*, da qual falamos até agora, tenta descrever a cultura, ou o grupo como um todo, enquanto a *etnociência* e a *microetnografia* focam unidades muito menores como palavras, indivíduos ou cenas (ibid.).

A *microetnografia* é uma *etnografia* focada, isto é, uma *etnografia* que se ocupa de olhar repetidas vezes e de analisar de maneira detalhada os registros audiovisuais de interações humanas em cenas-chave, em situações-chave de interação social, acompanhadas de observação participativa do contexto mais amplo no qual ocorrem tais cenas (op. cit., p. 51). É uma *etnografia* da comunicação, focando sujeitos individuais e seu discurso em certos cenários.

A *etnociência* se afasta da *etnografia* holística tradicional ao definir cultura primariamente em termos de cognições das pessoas. Seus pressupostos básicos são que o conteúdo dos dados culturais consta de regras, códigos e um ordenamento ideativo da sociedade que está organizado em diversos domínios culturais de conhecimento.

As experiências são codificadas em "lexemas" ou palavras; portanto, a linguagem é a principal fonte de dados culturais e as técnicas de estudo da linguagem podem ser aplicadas ao estudo da cultura ideativa ou cognição. Por conseguinte, há menos ênfase na observação participativa e mais ênfase na coleta de vocabulários sobre eventos particulares, assim como nos esquemas classificatórios (op. cit., p. 52).

Relacionada com a *etnociência* está a *etnometodologia* que, segundo André (1998, p. 18), não é exatamente uma metodologia, mas um campo de pesquisa: é o estudo de como os indivíduos compreendem e estruturam seu cotidiano, isto é, é a tentativa de descobrir "os métodos" que as pessoas utilizam em seu dia a dia para entender e construir a realidade que as envolve. Por consequência, seus principais focos de interesse são os conhecimentos tácitos, as formas de compreensão do sentido

comum, as práticas cotidianas e as atividades rotineiras que moldam as condutas dos atores sociais (ibid.).

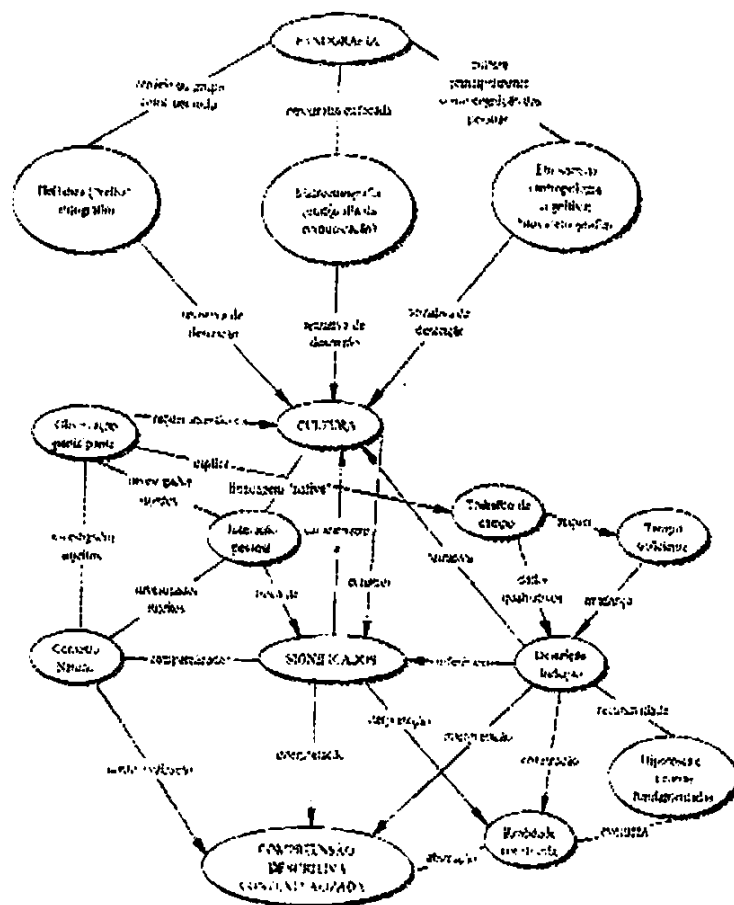


Figura 2 – Um mapa conceitual para a etnografia (Moreira, 2002).

Independente desses aparentemente diferentes tipos de etnografia podemos caracterizá-la de maneira geral como a tentativa de descrição de uma cultura. A principal preocupação na etnografia se refere ao significado que têm as ações e eventos para as pessoas ou grupos estudados (op. cit., p. 19). A etnografia

é um esquema de pesquisa desenvolvido pelos antropólogos para estudar uma cultura e uma sociedade. Etimologicamente, etnografia significa "descrição cultural" (ibid., p. 27).

Em educação, rigorosamente falando, o que se faz são estudos etnográficos, isto é, uma adaptação da etnografia à educação, uma vez que o fenômeno de interesse da pesquisa educativa é, em última análise, o processo educativo, não uma cultura ou um grupo social em si mesmos. Tais estudos etnográficos incluem, por exemplo, uma sala de aulas em particular, um pequeno grupo em uma sala de aula ou em uma escola, cenas ou diálogos na sala de aulas, relações escola-comunidade, etc.

A modo de conclusão deste item a respeito da etnografia, se apresenta na Figura 2 um mapa conceitual para a etnografia. No topo aparece como conceito mais abrangente o próprio conceito de *etnografia*, que pode ser *holística* (a etnografia tradicional ou "velha" etnografia), *microetnografia* (a etnografia da comunicação) ou *etnociência* (a antropologia cognitiva ou a "nova" etnografia).

Porém, a etnografia é sempre uma tentativa de descrever uma cultura (ou uma microcultura), que é caracterizada principalmente por *significados*, construídos e partilhados pelo grupo social, isto é, pelo "ponto de vista nativo" (ideias, crenças, valores, pressupostos), segundo Ogbu et al. (1988, p. 50).

A descrição de uma cultura requer *observação participante*, *trabalho de campo* (durante um tempo suficiente) e *interação pessoal* (intersubjetividade, empatia) em um *contexto natural*. A etnografia busca *descrição* e utiliza a *indução* para chegar a uma *realidade construída*; as *hipóteses* são desenvolvidas ao longo do processo e as teorias emergem dos dados, isto é, são *teorias fundamentadas* (nesse sentido, os métodos qualitativos são indutivos). O resultado de todo o processo é uma *compreensão descritiva contextualizada* de um grupo social, de umas cenas, de um discurso, de umas cognições ou, em termos mais abrangentes e originais, de uma cultura.

O estudo de casos

De acordo com Sturman (1988, p. 61), estudo de caso é um termo genérico para a pesquisa de um indivíduo, um grupo ou um fenômeno. Enquanto as técnicas usadas nessa pesquisa podem variar e incluir tanto enfoques qualitativos como quantitativos, a característica que mais distingue o estudo de caso é a crença de que os sistemas humanos desenvolvem uma completude e integração, isto é, não são simplesmente um conjunto de partes ou de traços. Por conseguinte, o estudo de caso se encaixa em uma tradição holística de pesquisa segundo a qual as características de uma parte são determinadas grandemente pelo todo ao qual pertence. A compreensão das partes requer a compreensão de suas inter-relações no todo. É uma visão sistêmica que pressupõe que os elementos de um evento educativo, por exemplo, são interdependentes e inseparáveis e uma mudança em um elemento implica uma mudança no resto.

Portanto, fazer uma pesquisa do tipo estudo de caso, isto é, para entender um caso, para compreender e descobrir como as coisas ocorrem e por que ocorrem, para talvez predizer algo a partir de um único exemplo ou para obter indicadores que possam ser usados em outros estudos (talvez quantitativos), é necessária uma profunda análise das interdependências das partes e dos padrões que emergem. O que se requer é um estudo de padrões, não de variáveis isoladas (ibid.). Para tudo isso, as técnicas de pesquisa qualitativa são frequentemente as mais adequadas.

O estudo de caso é utilizado, há muito tempo, em campos como o direito, a medicina, a psicologia e a administração. Porém, na pesquisa educativa, o uso dessa metodologia é mais recente e dentro de uma concepção bastante restringida, ou seja, o estudo descritivo de uma unidade que pode ser uma escola, um professor, um aluno ou uma aula (André, 1998, p. 21). Pode também, por exemplo, ser o estudo de um currículo ou de um sistema escolar.

Qual seria a diferença entre estudo de caso e etnografia? Segundo André (ibid.), pode-se fazer um estudo de caso etnográfico, isto é, aplicar a abordagem etnográfica ao estudo de um caso. Porém, nem todos os tipos de estudos de caso se encaixam na perspectiva etnográfica, nem todos os estudos etnográficos são estudos de caso.

Quais seriam, então, os distintos tipos de estudos de caso?

Em primeiro lugar é necessário distinguir entre estudo de caso e conceitos afins. De acordo com Serrano (1998, p. 85), o *trabalho com casos* denota os procedimentos corretivos, remediais, de desenvolvimento ou ajuste que seguem ao diagnóstico das causas de desajuste; o *método de casos* é uma estratégia didática na qual os elementos principais do estudo de casos são apresentados aos estudantes com propósitos ilustrativos, sem necessidade de oferecer uma visão completa dos fatos (o propósito é mais o de estabelecer um marco de discussão e debate); a *história de casos* é a busca do passado de uma pessoa, grupo ou instituição; o *estudo de casos*, por sua vez, pode ser definido como uma descrição intensiva, holística e uma análise profunda de uma entidade singular, um fenômeno ou unidade social.

Antes de falar de tipos de estudos de casos, é interessante também destacar seus traços essenciais. Segundo essa mesma autora (op. cit., p. 91), as *propriedades essenciais* de um estudo de casos qualitativo são a *particularização* (se centram em uma situação, evento, programa ou fenômeno particular), a *descrição* (o produto final é uma descrição rica e densa do objeto de estudo), a *heurística* (iluminam a compreensão do leitor respeito ao objeto de estudo) e a *indução* (baseiam-se no raciocínio indutivo; as teorias, os conceitos ou as hipóteses surgem de um exame dos dados fundados no contexto mesmo).

Quanto aos tipos de estudos de caso, Serrano (ibid., p. 97) argumenta que podem classificar-se pela natureza do informe

final, independentemente de sua orientação disciplinar ou área de interesse, em *descritivos*, *interpretativos* e *avaliativos*.

Estudos de caso descritivos (ibid.) se caracterizam por um informe detalhado de um fenômeno, objeto de estudo sem fundamentação teórica prévia; são inteiramente descritivos, não se guiam por generalizações estabelecidas ou hipotéticas, nem desejam formular hipóteses ou teorias.

Estudos de caso interpretativos (op. cit., p. 98) contêm descrições ricas e densas; porém, os dados descritos são utilizados para desenvolver categorias conceituais ou para ilustrar, defender ou desafiar pressupostos teóricos difundidos antes do estudo. O pesquisador deve reunir tanta informação sobre o objeto de estudo quanto seja possível, com a pretensão de interpretar ou teorizar sobre o fenômeno.

Estudos de caso avaliativos (ibid.) implicam descrição, explicação e juízo; sobretudo, este tipo de estudo de casos examina a informação para emitir um juízo; a emissão de juízos é o ato final e essencial da avaliação.

Não obstante, a mesma autora coloca que mesmo que se possa estabelecer esta classificação e mesmo que alguns estudos de casos possam ser puramente descritivos, em educação, a maioria dos estudos de casos é uma combinação de descrição e avaliação ou de descrição e interpretação.

O estudo de caso interpretativo nos remete outra vez ao tema da *teoria fundamentada* referido na etnografia. Este tipo de estudo de caso parece ser uma metodologia ideal para fundamentar uma teoria, isto é, para induzir uma teoria a partir de dados descritivos muito ricos. Porém, não se trata de uma teoria formal no sentido usado nas ciências naturais, também não do "método indutivista". São mais bem categorias, hipóteses compreensivas.

Outra classificação de tipos de estudos de caso é fornecida por Stenhouse (1985, apud Sturman, 1988, p. 63):

Estudo de caso etnográfico, do qual já falamos, que envolve o estudo profundo de uma entidade singular geralmente através de observação participante e entrevistas.

Estudo de caso pesquisa-ação, no qual o foco está em gerar uma mudança no caso em estudo.

Estudo de caso avaliativo que envolve avaliação de programas e no qual, muitas vezes, um trabalho de campo mais condensado substitui a abordagem etnográfica mais demorada.

Estudo de caso educativo que está desenhado para melhorar a compreensão da ação educativa.

Como se pode perceber nas classificações de Serrano e Stenhouse, é difícil separar o estudo de caso de outros tipos de pesquisa qualitativa, como a etnografia e a pesquisa-ação. Podemos caracterizar bem o estudo de caso argumentando que sua preocupação central é a compreensão de uma instância singular, o que significa que o objeto estudado é caracterizado como único, como uma representação singular da realidade que é multidimensional e historicamente localizada (André, 1998, p. 21). Porém, como vimos, um estudo de caso pode ser feito através de uma etnografia ou de uma pesquisa-ação, por exemplo.

As características essenciais dos estudos de casos, assim como seus distintos tipos, estão mapeadas conceitualmente na Figura 3. Os dois conceitos-chaves são *estudo de casos* e *instância singular*. Na parte superior do mapa, acima do conceito de estudo de casos, estão suas propriedades essenciais (*indução, particularização, heurística e descrição*); na parte inferior, embaixo do conceito de instância singular, aparecem as instâncias de tal conceito. No eixo central do mapa estão os diferentes tipos de estudo de caso identificados por Serrano (1998) e Stenhouse (1985). Os conectores tentam explicitar as relações entre os conceitos e as flechas sugerem certas convergências.

A pesquisa-ação

O objetivo fundamental da pesquisa-ação consiste em melhorar a prática em vez de gerar conhecimentos. A produção e utilização do conhecimento se subordinam a este objetivo e estão condicionadas por ele (Elliott, 1993, p. 67). A melhora na prática consiste em implantar aqueles valores que constituem seus fins, por exemplo, a educação no ensino (ibid.). Porém, o conceito de educação como fim do ensino transcende a conhecida distinção entre processo e produto. A melhora da prática supõe levar em conta ao mesmo tempo os resultados e os processos.

Segundo Kemmis e McTaggart (1988; apud Kemmis, 1988, p. 174), a pesquisa-ação é definida como uma forma de pesquisa *coletiva, autorreflexiva, empreendida por participantes de situações sociais para melhorar a produtividade, racionalidade e justiça de suas próprias práticas sociais ou educativas, assim como a sua compreensão em relação a tais práticas e as situações em que ocorrem. Os participantes podem ser professores, alunos, diretores, pais e outros membros da comunidade, isto é, qualquer grupo que partilha uma preocupação, um objetivo. É uma pesquisa colaborativa, porém, é importante enfatizar que esta ação colaborativa depende de que cada indivíduo examine criticamente suas próprias ações (ibid.).*

Na educação, quando se pretende melhorar a prática, é preciso considerar conjuntamente os processos e os produtos. Este tipo de reflexão simultânea sobre a relação entre processos e produtos constitui, segundo Elliott (op. cit., p. 68), uma característica fundamental da pesquisa-ação. É uma *prática reflexiva* que aspira a melhorar a concreção dos valores do processo, muito diferente do raciocínio técnico que versa sobre os meios para conseguir um fim. Para Elliot (op. cit., p. 69), é ao mesmo tempo ético e filosófico.

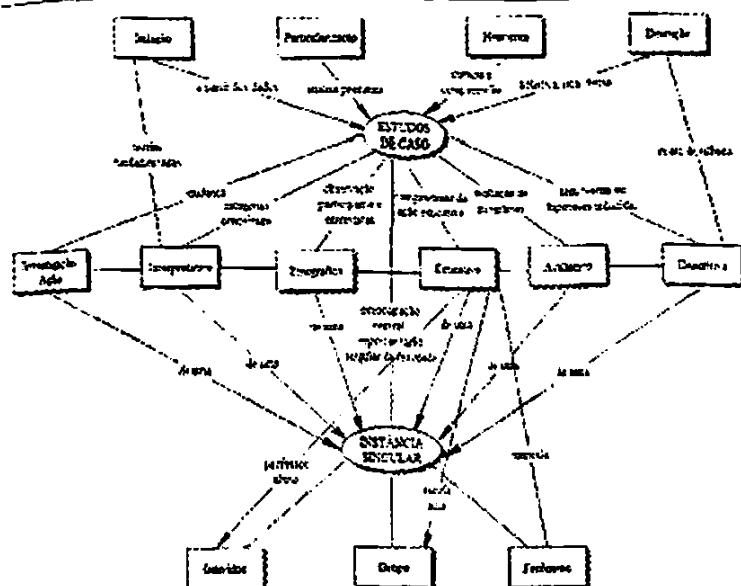


Figura 3 – Um mapa conceitual para estudo de casos (Moreira, 2002).

Na medida em que a reflexão trata da escolha de um curso de ação em um determinado conjunto de circunstâncias para levar à prática os próprios valores, reveste-se de caráter ético. Mas como a eleição ética supõe a interpretação dos valores que não se traduzir à prática – a reflexão sobre os meios não se pode separar da reflexão sobre os fins –, a reflexão ética tem uma dimensão filosófica (ibid.).

Na pesquisa-ação, os professores são incentivados a questionar suas próprias ideias e teorias educativas, suas próprias práticas e seus próprios contextos como objetos de análise e crítica (Kemmis, 1988, p. 174). A partir de uma reflexão cuidadosa, os professores podem desvelar ideias ou suposições teóricas que resultam injustificadas e os deixam perdidos em sua tarefa docente; por exemplo, tem-se suposições muito rígidas a respeito da natureza de habilidades inatas dos estudantes (ibid.).

Analogamente, os docentes, através da reflexão crítica, podem concluir que práticas antigas moldadas por hábito e tradição são inúteis ou irrelevantes nos tempos atuais; por exemplo, práticas disciplinares que funcionavam antes, hoje já não são aceitáveis ou são contraproducentes (ibid.). Quanto ao contexto, eles podem chegar à conclusão de que sua estrutura é inadequada e obstaculiza o alcance de metas educativas; por exemplo, a estrutura física da aula pode dificultar o trabalho em grupos, a interação pessoal, o ensino centrado no aluno.

Não obstante, não se pode esquecer que a pesquisa-ação é uma pesquisa coletiva, colaborativa. A reflexão pessoal é importante, mas a verdadeira mudança vem da autorreflexão coletiva. Que os participantes sintam a necessidade de iniciar mudanças, de inovar, isso é condição necessária antecedente da pesquisa-ação, mas não suficiente.

O processo de pesquisa-ação, segundo Kemmis e McTaggart (1988) e Elliot (1993, p. 88), se caracteriza por uma espiral de ciclos de reconhecimento (descoberta de fatos): reconhecimento de uma situação que se quer mudar; *planejamento* geral da ação objetivando a mudança; *desenvolvimento, implementação e avaliação* dessa ação; reflexão à luz da evidência recolhida na implementação; revisão do plano geral; *planejamento* de nova ação; implementação, avaliação, reflexão, revisão do plano; *planejamento e implementação* de uma terceira ação...

Naturalmente, este caráter cíclico não significa um processo linear, automático, mecânico. Tal como se disse no começo deste item, a pesquisa-ação, através dessa espiral de ciclos, tem por objetivo a *melhoria* das práticas e da compreensão de situações, e o *envolvimento* tanto quanto possível de todos os afetados intimamente pelas ações em todas as fases do processo investigativo. A pesquisa-ação é um processo colaborativo, autorreflexivo, no qual o envolvimento direto dos professores e outros implicados, na coleta de dados, análise crítica, reflexão,

cria imediatamente um sentido de responsabilidade quanto à melhora da prática (Kemmis, 1988, p. 174).

A pesquisa-ação unifica processos considerados com frequência independentes, por exemplo: o ensino, o desenvolvimento do currículo, a avaliação, a pesquisa-educativa e o desenvolvimento profissional (Elliott, 1993, p. 72). O ensino, por exemplo, no contexto da pesquisa-ação, se concebe como uma forma de pesquisa dirigida a compreender como traduzir os valores educativos à formas concretas de prática.

O desenvolvimento do currículo não é um processo que antecede o ensino; o desenvolvimento de programas curriculares se produz através da prática reflexiva do ensino (ibid.). A pesquisa-ação não reforça a postura dos professores enquanto conjunto de indivíduos que operam de forma independente e autônoma, que não partilham suas reflexões com os demais.

De maneira geral, pode-se dizer que a pesquisa-ação sempre implica um plano de ação baseado em objetivos de mudança (melhora), a implementação e controle desse plano através de fases de ação, assim como a descrição concomitante do processo cíclico resultante. Porém, Kemmis e McTaggart (1988, apud Kemmis, 1988) identificam várias características básicas da pesquisa-ação que ajudam a distingui-la de outros tipos de pesquisa qualitativa. Segundo eles, a pesquisa-ação:

- é uma abordagem para melhorar a educação através de mudanças e para aprender desde as consequências das mudanças;
- se desenvolve através de uma espiral autorreflexiva de ciclos de planificação, ação, observação sistemática, reflexão, replanificação, nova ação, observação e reflexão;
- é participativa, as pessoas trabalham para melhorar suas próprias práticas;
- é colaborativa, cria grupos autocríticos que participam e colaboram em todas as fases do processo investigativo;

- envolve os participantes em um processo de teorização sobre suas práticas, questionando circunstâncias, ações e consequências dessas práticas;
- requer que as pessoas ponham em xeque suas ideias e suposições com relação às instituições;
- é aberta com relação ao que conta como evidência, ou dados, mas sempre implica manter e analisar registros das consequências das ações implementadas;
- permite que os participantes, ao mesmo tempo, mantenham registros de suas próprias mudanças pessoais e analisem criticamente as consequências dessas mudanças;
- começa pequena, normalmente com pequenas mudanças que um pequeno grupo, ou talvez uma só pessoa, possa tentar, mas se desloca, gradativamente, rumo a mudanças mais extensivas;
- requer que os participantes analisem criticamente as situações (salas de aulas, escolas, sistemas educativos) nos quais trabalham;
- é um processo político porque envolve mudanças nas ações e interações que constituem e estruturam práticas sociais. Tais mudanças tipicamente afetam as expectativas e interesses de outros, além dos participantes imediatos nessas ações e interações.

Tal como se fez nas seções anteriores, esta finaliza com um mapa conceitual com relação ao tema abordado. A Figura 4 apresenta um mapa conceitual para pesquisa-ação. O conceito central é *mudança*: a pesquisa-ação tem como meta melhorar a *prática* através da mudança.

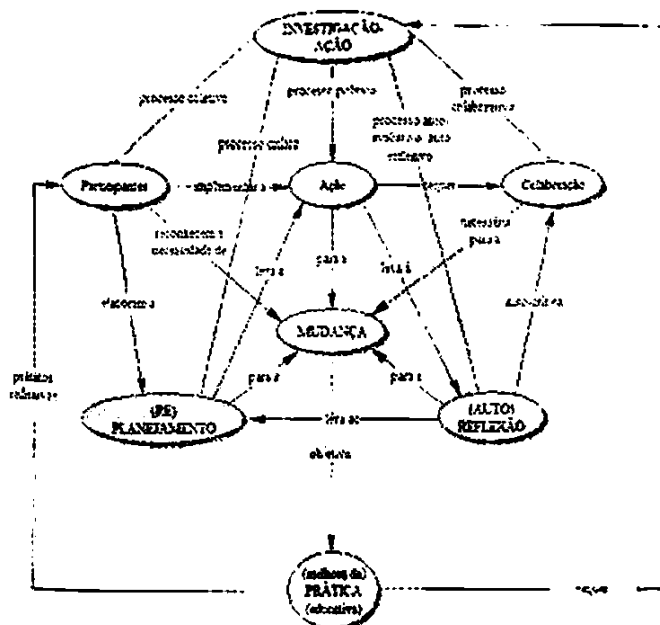


Figura 4 – Um mapa conceitual para pesquisa-ação (Moreira, 2002).

É também central a triade (re) *planificação ↔ ação ↔ (auto) reflexão* que caracteriza o *processo cíclico* da pesquisa-ação. Por outro lado, a pesquisa-ação é um processo *participativo, coletivo, colaborativo, político, autorreflexivo, autocrítico, autoavaliativo* que requer o envolvimento dos *participantes* em todas as fases e em todos os aspectos característicos desse processo.

Outros tipos de pesquisa qualitativa

Este texto começou comparando as abordagens qualitativa e quantitativa em termos bastante dicotômicos; depois, foram focadas características gerais da pesquisa qualitativa e passou-se logo a descrever a *etnografia*, o *estudo de casos* e a *pesquisa-ação*, consideradas as *três metodologias principais* dessa abordagem. Porém, há outras, das quais serão apresentadas algumas, sucintamente, a seguir.

Uma vez concluídas essas breves apresentações voltar-se-á a

temas gerais, no sentido de que se aplicam a várias metodologias qualitativas, focando tópicos como a apresentação e a análise de dados, a fiabilidade e validade dos estudos qualitativos, possibilidades de generalização, triangulação e outros. No final, será retomada a questão dos paradigmas a partir de um olhar integrador.

A fenomenografia

A fenomenografia é o estudo empírico dos diferentes modos através dos quais as pessoas vivenciam, percebem, apreendem, compreendem ou conceituam vários fenômenos no, e aspectos do, mundo em seu entorno. As palavras vivência, percepção, compreensão ou conceituação são usadas de maneira intercambiável. Porém, isso não significa que não há diferenças em seus significados, mas, sim, que o número limitado de maneiras, através das quais certo fenômeno é interpretado pelas pessoas pode ser identificado, por exemplo, independente de se estão embebidas na experiência imediata do fenômeno ou em uma reflexão sobre o mesmo fenômeno (Marton, 1988, p. 95-97).

As diferentes experiências, compreensões, percepções, etc., são caracterizadas em termos de "categorias de descrição" logicamente relacionadas e hierarquizadas. Tais categorias representam diferentes capacidades de lidar com (ou entender) um fenômeno. Como algumas maneiras de experienciar um fenômeno são mais eficientes que outras com relação a algum critério, é possível estabelecer uma hierarquia de categorias de descrição (ibid.). A fenomenografia se aplica, por exemplo, ao estudo da resolução de problemas ou em pesquisas que procuram diferenças críticas nos significados atribuídos a certos fenômenos, conceitos ou princípios chaves em certo campo de conhecimentos.

A hermenêutica

A hermenêutica foi originalmente definida como a arte, ou a ciência, da interpretação, em particular da Bíblia. Porém,

contemporaneamente, é definida como a teoria e a prática da interpretação e compreensão em diferentes tipos de contextos humanos (religiosos, seculares e cotidianos). Isto é, a hermenêutica não se refere só à exegese e interpretação de textos, mas considera a compreensão e a interpretação como uma marca definitiva da existência humana e da vida social (Ödman e Kerdeman, 1988, p. 185).

Assim como a hermenêutica, a fenomenologia também se ocupa da estrutura da compreensão; porém, a fenomenologia constrói a compreensão primeiramente em termos de construtos e funções cognitivos, enquanto para a hermenêutica a compreensão não é somente uma função cognitiva, é também a condição ontológica da existência humana (op. cit., p. 186).

Assim como a teoria crítica, a hermenêutica assume que compreensão e significado são constituintes da vida social; não obstante, a hermenêutica mantém que uma vez que o sujeito está sempre envolvido em algum processo de compreensão, é impossível captar em qualquer forma final ou definitiva os significados embebidos em uma tradição. A hermenêutica, portanto, evita envolver-se na tentativa de fundamentar a compreensão em um marco teórico ou em um método e se concentra em interpretar culturas desde dentro de certas situações e contextos (ibid.).

Na pesquisa educativa, a hermenêutica pode aprofundar a compreensão do fenômeno educativo focando os significados que subjazem a estratégias e práticas educativas específicas. Por exemplo, como devem ser interpretadas certas práticas administrativas? Quais são os significados subjacentes? Analogamente, a hermenêutica pode aprofundar a compreensão dos significados e propósitos subjacentes a um currículo.

A pesquisa participativa

A pesquisa participativa é descrita por Hall (1988, p. 198), de uma maneira geral, como um processo que combina três

atividades: pesquisa, educação e ação. É uma ação social em favor dos dominados, explorados, pobres, excluídos. A preocupação por poder e democracia e suas interações é central na pesquisa participativa. É também crítica a atenção a gênero, raça, etnia, orientação sexual, habilidades físicas e mentais, e outros fatores sociais (ibid.). A pesquisa participativa está desenhada para contribuir nos processos de mudança de poder ou democratização em uma variedade de contextos. Na pesquisa participativa não há receitas nem ortodoxias metodológicas: as questões e os métodos de trabalho devem fluir dos sujeitos envolvidos e de seu contexto.

(Não se deve confundir pesquisa participativa com observação participante, que é uma mistura de várias técnicas, um estilo ou estratégia de pesquisa, na qual, como já vimos, o principal instrumento é o próprio pesquisador que deve ficar imerso no cenário para ouvir, ver e começar a experienciar a vida como os sujeitos a vivem; Ball, 1988, p. 310).

A história oral

Como uma técnica de pesquisa, a história oral vai além do que o entrevistado responde às perguntas do entrevistador. Todos os matizes de seu testemunho são significativos: a não resposta, o silêncio, a vacilação, tudo conta como evidência. As fontes de história oral são, portanto, mais que uma questão de evidência falada e mais do que simplesmente registrar fatos (Hyams, 1988, p. 91).

Como uma abordagem para a pesquisa, a história oral permite dar voz aos "sem voz", aos anônimos. É como construir a história de baixo para cima. Mesmo que sofra restrições entre os historiadores (por exemplo, só os sobreviventes são entrevistados; não se pode generalizar), esta abordagem é usada em outros campos, como o da educação. Por exemplo, as atitudes dos professores com relação ao sistema educativo ou às

reformas educativas podem ser obtidas de maneira mais expansiva, mais aberta. As reminiscências dos professores podem permitir uma melhor compreensão da implementação das políticas educativas. A história oral tem potencial para explicar interpretações não oficiais de problemas educativos.

O caráter único do testemunho individual é considerado importante para construir uma história. Porém, como foi dito antes, é preciso levar em conta muito mais que a informação factual provida. Trata-se de procurar informação que não está nos registros escritos, nas biografias.

Fidedignidade, generalização e validade

A *fidedignidade* se refere ao grau em que se podem replicar as medidas e os estudos. Em uma abordagem quantitativa, a *fidedignidade* das medições e dos instrumentos é um requisito básico, porém, em uma abordagem qualitativa, tal conceito não tem sentido ou deve ter outro significado, pois, em grande medida, o pesquisador é o principal instrumento ou, em outras palavras, o instrumento é uma extensão do pesquisador.

Em relação aos estudos, a *fidedignidade* exige que um pesquisador que utilize os mesmos métodos que outro chegue a idênticos resultados (Goetz e Lecompte, 1988, p. 214). Isso coloca um enorme problema nas pesquisas sobre o comportamento natural ou dos fenômenos únicos. O estabelecimento da *fidedignidade* de um delineamento interpretativo se complica ainda mais pela natureza dos dados e do processo de pesquisa, pelos usos da apresentação dos resultados e pela visão de mundo dos pesquisadores nesse campo (ibid.).

De acordo com esses autores (op. cit., p. 215), se comparados com os delineamentos experimentais de laboratórios, estritamente controlados, ou com os experimentos de campo, os delineamentos da pesquisa naturalística parecem resistir-se a toda tentativa de réplica. Por exemplo, os problemas de unicidade e idiossincrasia podem levar a afirmar que é impossível

replicar um estudo etnográfico ou um estudo de caso. Além disso, sendo que o comportamento humano nunca é estático, nenhum estudo, independentemente de seus métodos e desenhos, pode ser replicado com exatidão (ibid.).

Porém, se pode contra-argumentar que tais unicidades e idiossincrasias não são tão extremas que não tenham nenhum grau de "representatividade", alguns traços semelhantes a outras. Isto é, os grupos ou fenômenos sociais ou indivíduos pesquisados podem não ser tão únicos e idiossincráticos que não tenham nada a ver com outros grupos, fenômenos ou indivíduos. Ou se, de fato, o são, não são de interesse da pesquisa educativa. Para que serve uma etnografia de uma aula de ciências que não tem absolutamente nada a ver com outras aulas de ciências ou que não gere uma compreensão contextualizada com algum valor para elas? Para que serve estudar um professor de ciências tão único, que não tem nada a ver com outros professores de ciências?

Isso nos remete ao tema da *generalização*: devem-se procurar generalizações em estudos qualitativos, através de, por exemplo, estudos de "casos representativos"?

Segundo André (1998, p. 58), a generalização no sentido de leis que se aplicam universalmente não é um objetivo da pesquisa qualitativa. Porém, a ideia de generalização é bastante aceita nessa abordagem, no sentido de que os dados de um estudo podem ser úteis para compreender dados de outros estudos (ibid.). Por isso, a descrição densa é considerada vital quando se pretende fazer comparações ou transferências de uma situação a outra; a análise de similitudes e diferenças torna possível julgar em que medida as compreensões construídas em um estudo podem ser consideradas hipóteses sobre o que pode ou não ocorrer em outras situações. Essa posição é compartilhada por Ogbu et al. (1988, p. 53), que diz que a generalização se torna possível porque o conhecimento construído através da abordagem qualitativa é profundo e contextualizado; as

descrições detalhadas compreensivas permitem aos leitores fazer comparações e tomar decisões bem fundamentadas a respeito da generalização.

Posição semelhante é tomada por Taft (1988, p. 74), que argumenta que para generalizar de um caso individual a outros é necessário alcançar uma compreensão suficientemente detalhada sobre a significatividade dos eventos a respeito do contexto em que ocorrem para poder estender interpretações a outros contextos e grupos. Quando um pesquisador tenta compreender um grupo, ele é ajudado por conhecer outros grupos; as generalizações são feitas através da capacidade do pesquisador de mediar entre um grupo e outros. Portanto, a descrição etnográfica de uma escola, por exemplo, deriva seu valor em grande parte do fato de que o pesquisador – assim como os leitores – tem familiaridade com outras escolas, e com escolas em geral (ibid.).

A postura de Erickson (1986, p. 130), como vimos no começo deste texto, é que na pesquisa interpretativa o que se busca não é “universais abstratos”, alcançados através de inferências estatísticas de amostras para populações, mas “universais concretos” aos quais se chega estudando um caso com muito detalhe e comparando-o com outros casos estudados com igual detalhe.

Não obstante, nenhuma generalização dessa natureza deve ser considerada final, mas apenas como hipótese de trabalho para outros estudos interpretativos ou como subsídio para questionários, entrevistas ou testes (Taft, 1988, p. 74).

O tema da generalização tem a ver com o da *validade externa* dos estudos, isto é, em que medida os construtos e universais concretos criados pelos pesquisadores são aplicáveis a mais de um grupo? Por outro lado, é necessário considerar também a questão da *validade interna*, ou seja, os pesquisadores estão interpretando o que acreditam interpretar? (Do ponto de vista quantitativo, a pergunta seria se os pesquisadores estão medindo o que acreditam medir.)

A validade pode ser pensada como uma qualidade das

conclusões e dos processos através dos quais são alcançadas, mas seu significado exato depende do critério de verdade que se está utilizando. Considerando que, na abordagem qualitativa, verdade é uma questão de concordância em contexto, o melhor significado de validade nessa abordagem parece ser o de *credibilidade*, como sugere Taft (op. cit., p. 73). A credibilidade depende do convencimento da comunidade de pesquisadores e leitores com relação às evidências apresentadas e aos processos utilizados. Sturman (1988, p. 65) propõe as seguintes estratégias para alcançar credibilidade:

- os procedimentos de coleta de dados devem ser explicados;
- os dados recolhidos devem ser apresentados e estarem prontos para reanálise;
- instâncias negativas devem ser relatadas;
- vieses devem ser reconhecidos;
- análises de trabalhos de campo devem ser documentadas;
- a relação entre asserção e evidência deve ser esclarecida;
- evidências primárias devem ser distinguidas das secundárias, assim como as descrições das interpretações;
- diários ou *logs* devem dar conta do que foi feito durante as distintas fases do estudo;
- técnicas devem ser desenhadas para "checar" a qualidade dos dados.

Erickson (1985, p. 140) se reporta ao mesmo tema, a partir de outra perspectiva, ao indicar cinco tipos de inadequações das evidências apresentadas:

1. *quantidade inadequada de evidências*; o pesquisador tem pouca evidência para garantir certas asserções-chave;
2. *diversidade inadequada de tipos de evidências*; o pesquisador não tem diferentes fontes de dados (por

exemplo, observações, entrevistas, documentos); não busca triangulação de dados;

3. *interpretações incorretas das evidências*; o pesquisador não compreende bem certos aspectos-chave da complexidade da ação ou dos significados atribuídos pelos atores no contexto;
4. *inadequadas evidências desconfirmadoras*; o pesquisador não tem dados que possam desconfirmar uma asserção-chave, ou, mais importante, não apresenta evidência de que tenha sido conduzida uma busca deliberada de dados potencialmente desconfirmadores;
5. *análise inadequada de casos discrepantes*; o pesquisador não indaga as instâncias desconfirmadoras nem as compara com as confirmadoras para determinar quais aspectos dos casos desconfirmadores eram iguais ou diferentes dos aspectos análogos dos casos confirmadores.

O mesmo autor (op. cit., p. 145) recomenda que o informe de estudos etnográficos contenha nove elementos principais:

1. *asserções empíricas* (uma tarefa básica da análise de dados é gerar tais afirmações em grande medida através de indução);
2. *vinhetas narrativas analíticas* (uma vinheta narrativa é uma representação vívida de um evento no cotidiano do caso ou grupo pesquisado. A vinheta tenta persuadir o leitor de que as coisas no contexto eram como o autor diz que eram);
3. *citações das notas de campo* (pode-se citar diretamente as notas no informe, indicando a data em que foram tomadas. Uma série de extratos de notas de campo pode servir como evidência de que o modo particular em que ocorreu um certo evento foi típico);

4. *citações de entrevistas* (as palavras dos entrevistados são um meio de transmitir aos leitores os pontos de vista dos sujeitos do estudo);
5. *informes sinópticos dos dados* (mapas, tabelas de frequências, diagramas);
6. *comentários interpretativos enquadrando uma certa descrição*;
7. *comentários interpretativos enquadrando a descrição geral* (a descrição geral tem como principal objetivo estabelecer a possibilidade de generalizar os padrões que foram ilustrados nas descrições particulares);
8. *discussão teórica* (comentários interpretativos com relação ao significado mais amplo dos padrões que emergiram dos dados);
9. *informe sobre a história natural da indagação no estudo* (isto é, uma discussão/descrição com relação a como certos conceitos-chave na análise evoluíram ou como padrões não esperados foram encontrados durante o trabalho de campo e na reflexão subsequente).

Os comentários interpretativos que enquadram as descrições particular e geral podem ser de três tipos: os que precedem e seguem uma descrição particular no texto, a discussão teórica que aponta à significatividade mais ampla dos padrões identificados nos eventos mencionados, e uma resenha das mudanças que ocorreram desde o ponto de vista do pesquisador durante o transcurso da indagação (op. cit., p. 152).

Para Erickson (ibid, p. 145), cada um desses nove elementos, separadamente e em conjunto, permitem ao leitor três coisas: em primeiro lugar, possibilitam experimentar de forma vicária o cenário descrito e confrontar instâncias de asserções-chave e construtos analíticos; em segundo, permitem examinar todo o espectro de evidências no qual está baseada a interpretação do pesquisador; e em terceiro, deixam que o leitor considere os

fundamentos teóricos e pessoais da perspectiva do autor tal como mudou ao longo do estudo.

Esta seção do texto foi dedicada ao tema da *fidedignidade*, *generalização* e *validade* dos estudos qualitativos. Mesmo que se pudesse argumentar que são conceitos típicos de estudos quantitativos, toda a seção foi desenvolvida com o objetivo de mostrar que tais conceitos têm sentido no contexto de uma pesquisa interpretativa, desde que se dê a eles o significado apropriado. Tal significado, como vimos, parece ser o de credibilidade, o qual nos remete a outro conceito importante na metodologia da pesquisa interpretativa: a *triangulação*, uma estratégia central para alcançar credibilidade (Sturman, 1988, p. 65).

Triangulação

A triangulação pode envolver o uso de diferentes fontes de dados, diferentes perspectivas ou teorias, diferentes pesquisadores ou diferentes métodos; é uma resposta holística à questão da fidedignidade e da validade dos estudos interpretativos (ibid.). Para Denzin (1988, p. 318), a triangulação é o emprego e combinação de várias metodologias de pesquisa no estudo de um mesmo fenômeno.

Não é uma estratégia típica da pesquisa qualitativa, também não é uma estratégia nova: o uso de múltiplas medições e métodos, de modo que se superem as debilidades inerentes ao uso de um único método ou um único instrumento tem uma longa história nas ciências naturais e sociais. Na pesquisa quantitativa, a triangulação é usada, tradicionalmente, como uma estratégia de validação de observações (ibid.).

Não obstante, segundo Denzin (op. cit.), a apropriação desse conceito pelos pesquisadores interpretativos e sua aplicação a problemas típicos da pesquisa qualitativa é mais recente e representa um compromisso com um sofisticado rigor

metodológico por parte dos pesquisadores, no sentido de que estão comprometidos a tornar seus esquemas empíricos e interpretativos o mais públicos possível. De acordo com esse autor (p. 319), há cinco tipos básicos de triangulação:

1. *triangulação de dados*, envolvendo tempo, espaço e pessoas;
2. *triangulação de pesquisadores*, que consiste no uso de múltiplos observadores, em lugar de um só;
3. *triangulação de teorias*, que consiste em utilizar mais de um esquema teórico na interpretação do fenômeno pesquisado;
4. *triangulação metodológica*, que envolve o uso de mais de um método e pode consistir em estratégias intramétodos ou entre métodos;
5. *triangulação de verificação por sujeitos*, na qual os pesquisados examinam e confirmam, ou desconfirmam, o que se escreveu sobre eles.

Existe também a *triangulação múltipla*, na qual o pesquisador combina em uma pesquisa múltiplos observadores, perspectivas teóricas, fontes de dados e metodologias.

Por outro lado, mesmo que Denzin associe a triangulação com um compromisso com o rigor metodológico, o uso dessa estratégia na pesquisa qualitativa não está livre de críticas: os argumentos são, por exemplo, que a triangulação de dados tem um viés positivista, que dois pesquisadores nunca observam o mesmo fenômeno da mesma maneira, que diferentes métodos geram diferentes imagens e recortes da realidade e que a triangulação de teorias não tem sentido epistemologicamente.

Porém, o próprio Denzin (op. cit., p. 321) contra-argumenta dizendo que a triangulação nos estudos qualitativos não deve ser comparada com a análise de correlação nas pesquisas quantitativas e que nunca deve ser uma estratégia eclética. Na triangulação de pesquisadores não se espera que estes observem

exatamente da mesma maneira e que um corrobore o que o outro observa, mas que suas diferentes observações expandam a base interpretativa do estudo e que revelem aspectos do fenômeno pesquisado que não seriam necessariamente observados por um único pesquisador.

Com relação à *triangulação metodológica*, o importante é justamente que possam emergir diferentes imagens. Com relação à *triangulação de teorias*, Denzin (ibid.) diz que ela, em vez de requerer que as interpretações sejam consistentes com dois ou mais teóricos, simplesmente requer que o pesquisador seja consciente das diferentes maneiras através das quais pode ser interpretado o fenômeno.

O tema da triangulação, em particular as triangulações teóricas e metodológicas, está muito vinculado à questão dos paradigmas, focada desde uma perspectiva dicotômica no começo deste texto. Nessa oportunidade foi dito que o olhar dicotômico estava sendo usado por razões didáticas e que no final seria retomado o assunto dos paradigmas sob uma visão integradora. Então, chegou o momento de voltar aos paradigmas.

Os paradigmas de pesquisa em educação: rumo à acomodação

A dicotomia estabelecida anteriormente entre os "dois paradigmas clássicos" é simplificadora. A metodologia da pesquisa nas ciências sociais não pode ser pensada simplesmente como "não positivista" em contraposição a uma suposta tradição positivista da pesquisa nas ciências naturais. Se assim fosse, não teria sentido falar de acomodação de paradigmas. Porém, considerando paradigma como "um conjunto básico de crenças que orienta a ação" (Guba, 1990, apud Alves-Mazzotti, 1996, p. 17), quer dizer, uma concepção de mundo que guia o pesquisador, não só na escolha de método, mas também em suas posições ontológicas e epistemológicas, pode-se distinguir pelo menos

três paradigmas como sucessores do positivismo: o *pós-positivismo*, a *teoria crítica* e o *naturalismo/construtivismo* (ibid).

O primeiro seria uma versão modificada do positivismo, revisando pontos insustentáveis (e.g., agora se supõe que a realidade existe, mas que nunca será totalmente apreendida pela pesquisa); a teoria crítica é ideologicamente orientada, uma vez que rejeita a neutralidade: o processo de pesquisa é mediado pelo pesquisador, e o termo crítica se refere tanto à crítica interna que resulta do questionamento analítico da argumentação e do método como à análise das condições de regulação social, desigualdade e poder; ao naturalismo/construtivismo subjaz à ideia de que os resultados de qualquer pesquisa são sempre influenciados pela interação pesquisador/pesquisado, de maneira que o conhecimento é sempre produto da atividade humana e, portanto, nunca pode ser visto como algo definitivo, mas como algo que está sempre se modificando (Alves-Mazzotti, 1996, pp. 17-20).

O paradigma naturalista/construtivista é o que foi enfatizado neste texto: a realidade é socialmente construída (o que implica que sempre há múltiplas realidades; as ideias, os valores e a interação pesquisador/pesquisado influenciam na configuração dos "fatos", o que implica que a teoria é subdeterminada). Estas características, por sua vez, implicam um relativismo que é problemático para os outros dois paradigmas: *se alguém se propõe compreender os significados atribuídos pelos atores às situações e eventos dos quais participam, se tenta entender a "cultura" de um grupo ou organização no qual coexistem diferentes visões correspondentes aos subgrupos que os compõem, então o relativismo não constitui um problema; mas se nos propomos à construção de teorias (pós-positivismo) ou à transformação social (teoria crítica), o que exige acordo ao redor de decisões ou princípios que possibilitem a ação conjunta, o relativismo passa a ser um problema* (op. cit., p. 21)

Com a identificação destes três paradigmas "pós-clássicos", e

possivelmente outros, já se percebe que a questão dos paradigmas em ciências sociais não é dicotômica, que as ciências sociais são multiparadigmáticas e que a acomodação ou integração de paradigmas não deve ser considerada impossível ou, pelo menos, é uma questão em aberto, como disse Alves-Mazzotti (op.cit., p. 22).

Um bom exemplo de argumento em favor da integração paradigmática é dado por Bericat (1998). Este autor considera que existem três razões fundamentais que podem motivar o desenho multimétodo em uma pesquisa social: *complementação*, *combinação* e *triangulação* (p. 37).

A *complementação* existe quando, no marco de um mesmo estudo, se obtém duas imagens, uma procedente de métodos de orientação qualitativa e outra de métodos de orientação quantitativa (ibid.), resultando assim um duplo e diferenciado conjunto de asserções de conhecimento sobre o fenômeno de interesse. O que se obtém são duas perspectivas diferentes sem pretensão alguma de solapamento ou convergência. As asserções de conhecimento são apresentadas com duas partes bem diferenciadas, cada uma das quais expõe os resultados alcançados pela aplicação do respectivo método. Segundo Bericat (ibid.), na *complementação*, o grau de integração metodológica é mínimo, e sua legitimidade se suporta sobre a crença de que cada orientação é capaz de revelar diferentes zonas da realidade social, assim como que é necessário contar com essa dupla visão para um melhor entendimento do fenômeno (ibid.).

Na *combinação*, a estratégia é integrar subsidiariamente uma metodologia, seja a qualitativa ou a quantitativa, na outra, com o objetivo de fortalecer a validade desta compensando suas debilidades mediante a incorporação de informações que procedem da aplicação da outra metodologia. O que se busca não é a convergência de resultados, que finalmente procederão de uma única metodologia, mas bem mais uma adequada combinação metodológica (op. cit., p. 39).

Finalmente, na *triangulação*, o que se pretende é um

solapamento ou convergência de resultados. Não se trata de complementar a visão de realidade com dois olhares, mas de utilizar duas metodologias para o estudo de um mesmo e idêntico aspecto de uma realidade social. As metodologias, tal como na complementação, são implementadas de forma independente, mas se focam em um mesmo objeto de estudo buscando resultados convergentes. *A legitimidade desta estratégia depende de se acreditamos que ambas as metodologias realmente podem captar idêntico aspecto da realidade, isto é, se o solapamento é possível. Na medida em que pensemos que conduzem a visões incomensuráveis da realidade, então estaremos no caso da complementação* (op. cit., p. 38).

Na Figura 5, estas três estratégias de integração estão esquematizadas em um mapa conceitual. Trata-se sempre de captar aspectos de uma realidade social. Porém, isso se pode tentar com uma única metodologia ou integrando-as de modo complementar, triangular ou combinatório. A triangulação e a combinação só são possíveis na medida em que se aceita pelo menos certo grau de comensurabilidade paradigmática. Caso contrário, a única alternativa é a complementação.

Conclusão

Na Figura 6 se apresenta, a modo de conclusão deste texto, um diagrama V. Este tipo de diagrama, também conhecido como "V" epistemológico, foi desenhado por D.B. Gowin (1981) para esquematizar a estrutura do processo de produção de conhecimento. Aqui é usado para refletir tal produção no contexto do paradigma qualitativo, particularmente na educação.

Este diagrama pretende ser uma espécie de resumo de todo o texto. Como tal, é preciso levar em conta que nenhum dos itens que aparecem no diagrama V está completo. São dados só exemplos do que poderia integrar cada item. É esta a razão dos três pontos que aparecem no final deles.

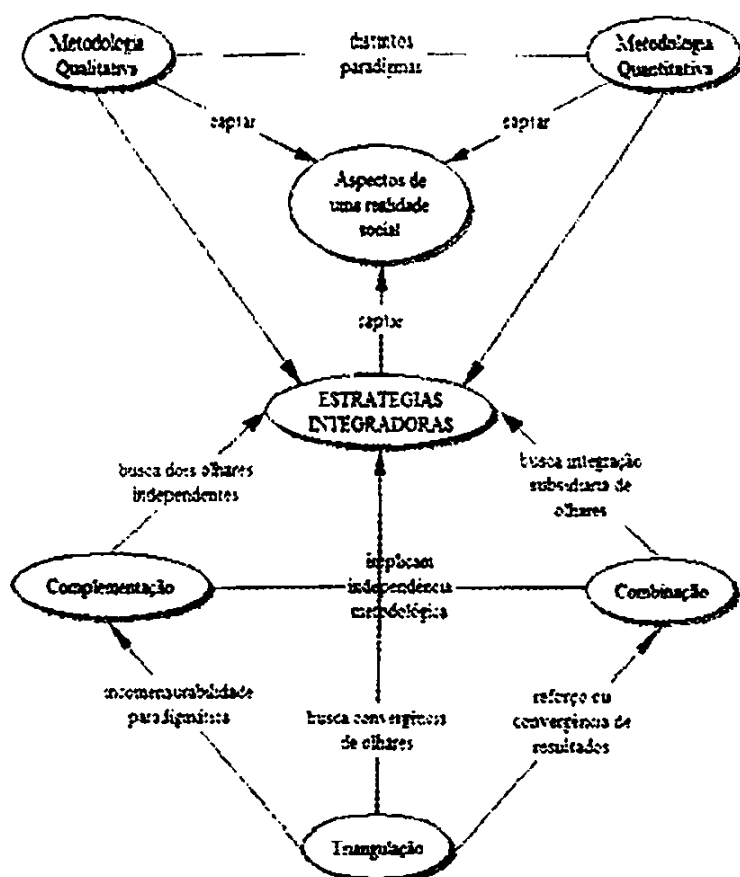


Figura 5 – Um mapa conceitual para estratégias integradoras entre a metodologia qualitativa e quantitativa (Moreira, 2002).

Dominio Conceitual

Filosofias:
Os humanos criam
interpretações significativas
dos eventos
e objetos do seu entorno
e tais interpretações podem
ser estruturadas através de
metodologias metodológico-
construtivistas. A realidade
é socialmente construída.
Verdade é questão de
consistência em contexto
...

Teorias: estão fundamentadas
no próprio processo de
pesquisa, emergem dos dados;
são geradas a partir da análise
indutiva dos dados, são
hipóteses compreensivas

Princípios: A interação
pesquisador/pesquisado influencia na
configuração dos conhecimentos
produzidos e nas teorias conceituais.
O pesquisador é o principal
instrumento de pesquisa. A realidade
socialmente construída: aplica
múltiplas realidades.
...

Conceitos: diversos conceitos, realidade:
construída, ação, significado,
compreensão contextualizada,
participação, interpretação subjetiva,
dados qualitativos, análise qualitativa,
tempo

Tipos de perguntas, temas de interesse

Significados atribuídos pelas
pessoas às situações e eventos nos
quais participam. Interpretação
em contexto; compreensão
contextualizada de significados;
quais são os contextos de
significados que criam juntos as
pessoas (e.g., professores e
alunos)? Há diferenças nas
perspectivas de significados das
pessoas (e.g., alunos e professores
da sala)? Como são criados e
usados os sistemas de
significados nas interações
cotidianas? Como é a vida
vivida? compreensão de
idéias, crenças, valores,
e sentimentos

As pessoas: busca de
padrões de explicação
de significados de
ações

(construção de
interpretações
significativas)

Dominio Metodológico

Aserções de valor:
dado que o fenômeno
educativo é
socialmente social, a
pesquisa qualitativa
é potencialmente útil para
estudá-lo.

Aserções de conhecimento:
compreensão
contextualizada; descrições
de significados de realidades
socialmente construídas.

Procedimentos analíticos:
indução, comparação,
contrastar, busca de
categorias (tipologias),
construção, escolha de unidades
de análise e análise, interpretação
fundamental, interpretação.

Tipos de pesquisa: etnografia,
estudo de caso, perspectiva-ação.

Metodologias técnicas: observação
participativa, entrevistas, análise de
discurso, vídeo-gravação, conversas.

Registros e dados: notas de campo,
transcrições de entrevistas, vídeos,
documentos, produções dos sujeitos.

Eventos/objetos de estudo:
Indivíduos, grupos ou fenômenos em seu acontecer natural

Figura 6 – Um diagrama V para a pesquisa qualitativa, particularmente em educação (M.A. Moreira, 2002)

Referências

- Alves-Mazzotti, A.J. (1996). O debate atual sobre os paradigmas de pesquisa em educação. *Cadernos de Pesquisa*, São Paulo, n. 96: 15-23.
- André, M.E.D.A. (1998). *Etnografia da prática escolar*. 2ª ed. São Paulo: Papirus Editora.
- Ball, S.J. (1988). Participant observation. In Keeves, J.P. (Ed). *Educational research, methodology, and measurement. An international handbook*. Oxford: Pergamon Press. pp. 310-314.
- Bericat, E. (1988). *La integración de los métodos cuantitativo y cualitativo en la investigación social*. Barcelona: Editor al Ariel.
- Denzin, N.K. (1988). Triangulation in educational research. In Keeves, J.P. (Ed). *Educational research, methodology, and measurement. An international handbook*. Oxford: Pergamon Press. pp. 318-322.
- Eisner, E.W. (1981). On the differences between scientific and artistic approaches to qualitative research. *Educational Researcher*, 10(4): 5-9.
- Elliott, J. (1993). *El cambio educativo desde la investigación-acción*. Madrid: Ediciones Morata.
- Erickson, F. (1986). Qualitative methods in research on teaching. In Wittrock, M.C. (Ed.). *Handbook of research on teaching*. New York: Macmillan Publishing Co. pp. 119-161. Traducción al español: Erickson, F. (1989) Métodos cualitativos de investigación sobre la enseñanza. In Wittrock, M.C. (Comp.). *La investigación en la enseñanza*, II. Barcelona, Paidós. pp. 195-301.
- Firestone, W.A. (1987). Meaning in method: the rethoric of quantitative and qualitative research. *Educational Researcher*, 16(7): 16-21.
- Goetz, J.P. y Lecompte, M.D. (1988). *Etnografía y diseño cualitativo en investigación educativa*. Madrid: Ediciones Morata.

- Guba, E.G. (1990). The alternative paradigm dialog. In: Guba, E.G. (Ed.) *The paradigm dialog*. London, Sage. Apud: Alves-Mazzotti (1996). O debate atual sobre os paradigmas de pesquisa em educação. *Cadernos de Pesquisa*, São Paulo, n. 96: 15-23.
- Hall, B.L. (1988). Participatory research. In Keeves, J.P. (Ed). *Educational research, methodology, and measurement. An international handbook*. Oxford: Pergamon Press. pp. 198-204.
- Husén, T. (1988). Research paradigms in education. In Keeves, J.P. (Ed). *Educational research, methodology, and measurement. An international handbook*. Oxford: Pergamon Press. pp. 16-21.
- Hyams, B.K. (1988). *Oral history*. In Keeves, J.P. (Ed). *Educational research, methodology, and measurement. An international handbook*. Oxford: Pergamon Press.
- Keeves, J.P. (1988). Towards a unified view of educational research. In Keeves, J.P. (Ed). *Educational research, methodology, and measurement. An international handbook*. Oxford, Pergamon Press. pp. 1-7.
- Kemmis, S. (1988). Action research. In Keeves, J.P. (Ed). *Educational research, methodology, and measurement. An international handbook*. Oxford: Pergamon Press. pp. 173-179.
- Kemmis, S. and McTaggart, R. (Eds.). (1988). *The action research reader*. 3rd ed. Geelong: Deakin University Press.
- Landsheere, G. de (1988). History of educational research. In Keeves, J.P. (Ed). *Educational research, methodology, and measurement. An international handbook*. Oxford: Pergamon Press. pp. 8-16.
- Marshall, C. (1984). The wrong time for mechanistics in qualitative research. *Educational Researcher*, 13(9): 26-28.
- Marton, F. (1988). Phenomenography. In Keeves, J.P. (Ed). *Educational research, methodology, and measurement. An international handbook*. Oxford: Pergamon Press. pp. 95-101.
- Miles, M.B. and Huberman, A.M. (1984) Drawing valid meaning from qualitative data. toward a shared craft. *Educational Researcher*, 13(5): 20-30.

- Moreira, M.A. (2000). Investigación en enseñanza: aspectos metodológicos. In *Actas de la I Escuela de Verano sobre Investigación en Enseñanza de las Ciencias*. Burgos: Servicio de Publicaciones de la Universidad de Burgos. pp. 13-51.
- Moreira, M.A. (2002). Investigación en educación en ciencias: métodos cualitativos. *Actas del PIDECE*, 4:25-54.
- Ödman, P.J. & Kederman, D. (1988). Hermeneutics. In Keeves, J.P. (Ed). *Educational research, methodology, and measurement. An international handbook*. Oxford: Pergamon Press. pp. 185-192.
- Ogbu, J.U., Sato, N.E. and Kim, E.Y. (1988). Anthropological inquiry. In Keeves, J.P. (Ed). *Educational research, methodology, and measurement. An international handbook*. Oxford: Pergamon Press. pp. 48-54.
- Serrano, G.P. (1998). Investigación cualitativa. Retos e interrogantes. I. Métodos. Madrid: La Muralla S.A.
- Shulman, L.S. (1981). Disciplines of inquiry in education: an overview. *Educational Researcher*, 10(6): 5-12.
- Smith, J.K. (1983). Quantitative versus qualitative research: na attempt to clarify the issue. *Educational Researcher*, 12(3): 6-13.
- Smith, J.K. and Heshusius, L. (1986). Closing down the conversation: the end of the quantitative-qualitative debate among educational inquirers. *Educational Researcher*, 15(1): 4-13.
- Stenhouse, L. (1985). Case study methods. In Husén, T. & Postlethwaite, T.N. (Eds.). *International Encyclopedia of Education*. Oxford, Pergamon Press.
- Sturman, A. (1988) Case study methods. In Keeves, J.P. (Ed). *Educational research, methodology, and measurement. An international handbook*. Oxford: Pergamon Press. pp. 61-66.
- Taft, R. (1988). Ethnographic research methods. In Keeves, J.P. (Ed). *Educational research, methodology, and measurement. An international handbook*. Oxford: Pergamon Press. pp. 71-75.

- Walker, J.C. & Evers, C.W. (1988). Research in education: epistemological issues. In Keeves, J.P. (Ed). *Educational research, methodology, and measurement. An international handbook*. Oxford: Pergamon Press. pp. 22-31.

Capítulo III

Pesquisa em Ensino: Métodos Quantitativos¹

M.A. Moreira

P.R.S. Rosa

Resumo

A finalidade deste texto é a de resgatar a metodologia quantitativa aplicada à pesquisa em ensino. Não há nada de novo nele. Todos os tópicos enfocados encontram-se muito mais elaborados em livros de metodologia da pesquisa ou de Estatística. Contudo, chamar a atenção para questões como tratamento de variáveis, delineamentos de pesquisa, amostragem, significância estatística, fidedignidade e validade, análise de variância, testes não paramétricos, pode ser uma contribuição para a atual pesquisa em ensino que, há muito tempo, usa praticamente apenas métodos qualitativos e, assim fazendo, acaba enviesada.

Introdução

D. B. Gowin (1981, 2005) propôs um dispositivo heurístico – que hoje é conhecido como diagrama V (Moreira, 2006) – para desvelar a estrutura do processo de produção do conhecimento. A Figura 1 apresenta esse dispositivo aplicado à pesquisa quantitativa em educação. O centro desse diagrama é o *domínio de interesse* da pesquisa (por exemplo, o ensino e a aprendizagem) e as perguntas feitas dentro desse domínio. Para respondê-las, confirmá-las ou refutá-las, é preciso fazer registros

¹ Texto de Apoio Nº 29, Programa Internacional de Doctorado en Enseñanza de las Ciencias da Universidade de Burgos, Espanha, desenvolvido no marco de um convênio com a Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil.

de um evento que se faz acontecer ou que acontece naturalmente. O lado esquerdo do diagrama é o *domínio conceitual* (ou teórico) da pesquisa; nele estão conceitos, princípios, modelos, teorias e filosofias que fundamentam teórica e epistemologicamente a pesquisa e que interagem com os registros, transformações e asserções que constituem o *domínio metodológico* que aparece no lado direito do diagrama.

Neste texto o foco estará no domínio metodológico. Como se vê na Figura 1, o lado metodológico começa com os registros. Sem registros não se faz pesquisa empírica. A partir daí, um passo fundamental desse tipo de pesquisa é a conversão desses registros em índices numéricos.

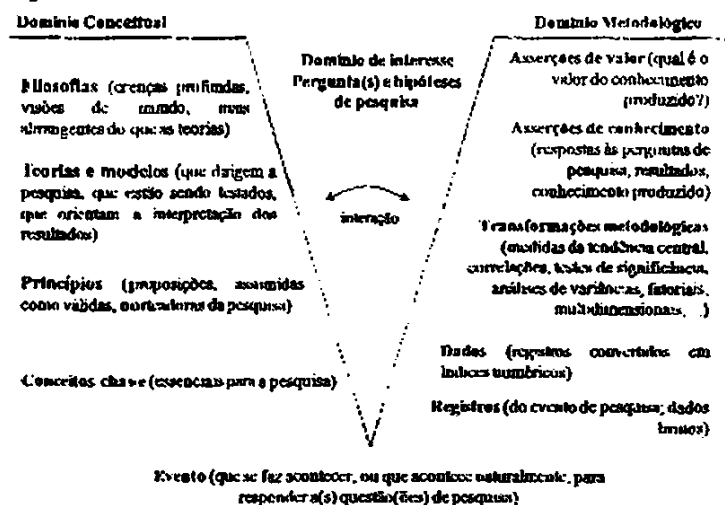


Figura 1 – O diagrama V aplicado à pesquisa quantitativa em educação.

Por exemplo, no caso em que os registros sejam mapas conceituais, é preciso definir critérios como, digamos, tantos pontos para a hierarquia, tantos para conectivos, etc., para chegar a um escore para cada mapa. Ou, então, definir categorias como, suponhamos, muito bom, bom, regular, falho, e atribuir pontos a cada categoria.

É certo que também se pode trabalhar quantitativamente

com índices não numéricos como, por exemplo, variáveis dicotômicas do tipo sim ou não, feminino ou masculino, mas na pesquisa empírica predomina o uso de índices numéricos.

É igualmente correto que o mais importante na pesquisa, seja ela de qualquer natureza, é a pergunta de pesquisa, aquela cuja busca de respostas gera conhecimentos. O conhecimento humano é construído, reconstruído, refutado, modificado, sempre pela busca, muitas vezes obstinada, de respostas a perguntas sobre determinados fenômenos de interesse.

A identificação de uma questão de pesquisa que valha a pena investigar, que possa gerar conhecimentos, é a parte mais difícil da pesquisa. No entanto, é também muito importante um delineamento de pesquisa que permita fazer registros relevantes que, por sua vez, originem dados (tipicamente índices numéricos, na pesquisa quantitativa) de máxima relevância à pergunta de pesquisa.

Dados e tratamento estatístico

Tipicamente, na pesquisa educacional empírica, o pesquisador se questiona sobre qual a evidência que o apoiará em relação a certas hipóteses de pesquisa. Faz, então, registros, converte-os em números e trata-os estatisticamente para ver se servem como evidência.

Costuma-se dizer que bons dados falam por si mesmos. Isso significa que se os dados forem de boa qualidade, os procedimentos estatísticos são imediatos. Por outro lado, se os dados forem ruins, não há tratamento estatístico que os transforme em bons.

O importante é a qualidade dos dados, não as manipulações estatísticas. A relevância das conclusões estatísticas nunca será maior do que a adequação dos dados numéricos trabalhados estatisticamente. A interpretação dos resultados estatísticos depende do que está por detrás dos dados. Os números a serem

analisados não são entidades sagradas representando puras abstrações. Ao contrário, sua utilidade na pesquisa empírica reside no fato que têm referentes no contexto da pesquisa, que significam algo no mundo real (Millman, 1970).

Uma boa estratégia para melhor abordar a questão da análise e qualidade dos dados é considerá-la antes, ou seja, pensar na análise dos dados antes de coletá-los, evitando-se, assim, que os números obtidos não tenham muito a ver com as questões investigadas.

Outro aspecto a ser considerado nessa questão é o da fidedignidade e validade dos instrumentos. Sem instrumentos fidedignos e válidos, os números que deles resultarem não serão confiáveis e de nada servirá tratá-los estatisticamente.

No que se refere a instrumentos, é conveniente testá-los previamente, perguntando a alguns sujeitos como interpretam determinados itens, ou fazer uma análise do tipo "tem sentido atribuir o mesmo número a uma não resposta e a uma resposta neutra?", quer dizer, "um item deixado em branco pode ser considerado como resposta neutra?", "sem opinião é o mesmo que opinião neutra?"(op.cit.).

Significância estatística e significância prática

Na análise dos dados é importante levar em conta que o nível de significância estatística ($p < 0,05$, por exemplo) não é uma medida da importância ou da significância prática de um resultado, pois esse nível depende do número de casos e da eficiência do delineamento de pesquisa.

Quando um resultado é estatisticamente significativo, o pesquisador deve analisar a magnitude dos efeitos, ou seja, qual é a importância, no contexto da investigação, de uma diferença de médias, estatisticamente significativa, por exemplo, entre 7,5 e 7,2, ou qual é a relevância de uma correlação, estatisticamente significativa ao nível 0,05, de um coeficiente de correlação de 0,23 entre duas variáveis?

Outros aspectos a serem considerados na questão da significância estatística e significância prática são o tamanho e a variabilidade da amostra. Mesmo que se obtenham diferenças, correlações ou fatores estatisticamente significativos para amostras de 12 ou 13 sujeitos, é preciso perguntar-se qual é a significância prática desses resultados. Por outro lado, amostras pequenas e, inclusive, amostras maiores de 30 ou 40 sujeitos, por exemplo, podem apresentar grande variabilidade. Em uma amostra de 10 sujeitos, se dois tiverem o escore máximo de dez pontos, dois obtiverem zero pontos e os demais alcançarem cinco ou seis pontos, a média será cinco ou mais pontos. Mas o que significa essa média na prática? É claro que nesse caso é melhor trabalhar com a moda, mas, mesmo assim, é preciso perguntar qual é a significância prática desse resultado. Analogamente, esse questionamento deve ser feito quando é grande a variabilidade dos escores em amostras maiores.

A estatística apropriada

Na escolha da técnica estatística adequada para o tratamento dos índices numéricos é conveniente considerar primeiramente se o que se quer é *descrever* características de um conjunto de números ou se o que se pretende é *estimar* valores da população. No primeiro caso, a estatística a ser usada é a *descritiva*, no segundo a *inferencial*.

É também conveniente voltar à questão de pesquisa e, tendo em conta sua natureza, considerar se medidas da tendência central, correlações, testes de significância para diferenças de médias, análises de variância, análises fatoriais ou multidimensionais são, de fato, necessárias para respondê-la.

O papel do computador

Há algoritmos de computador para realizar quaisquer análises estatísticas de dados numéricos. Basta injetar esses

dados nesses algoritmos para que rapidamente saiam médias, desvios padrão, coeficientes de correlação, variâncias, fatores, etc., etc.. O computador atende aos comandos e seus procedimentos atuam sobre os dados que lhe são fornecidos. E aí voltamos ao começo desse assunto de análise de dados: se os dados forem ruins, não há computador que gere bons resultados. Não adianta ficar dando comandos para que ele gere mais e mais coeficientes, tabelas, fatores. Tudo será lixo!

Por outro lado, se os dados forem bons, é preciso saber interpretar os produtos dos procedimentos estatísticos aos quais eles são submetidos. Não é preciso usar lápis e papel, ou calculadora, para, tediosamente, calcular estatísticas (valores da amostra), parâmetros (valores da população), coeficientes, correlações, fatores. O computador faz isso muito mais rapidamente e sem erros. Mas é fundamental saber interpretar os resultados (op.cit.).

A análise dos dados é, como foi dito no início, uma etapa fundamental da pesquisa empírica. Sem dúvida. Mas mais importante do que ela é a questão de pesquisa e os registros que se faz dos eventos usados para responder essa questão. São esses registros que serão convertidos em índices numéricos que, por sua vez, serão analisados estatisticamente. Os procedimentos estatísticos estão disponíveis em profusão e o computador os executa rapidamente. Mas a análise, em si, tem de ser feita pelo pesquisador.

Passemos agora a descrever e tratar de aspectos e técnicas dos métodos quantitativos de pesquisa em educação.

Variáveis

Condições que podem ser variadas ou selecionadas pelo investigador são chamadas de variáveis independentes. As medidas das respostas feitas durante o experimento constituem as variáveis dependentes.

Os níveis de uma variável referem-se às categorias da variável. Sexo, por exemplo, tem dois níveis. O número de níveis de uma variável como idade pode ser arbitrariamente determinado e pode variar de dois (e.g., acima de 35 ou abaixo ou igual a 35 anos) ao infinito (quando idade é tratada como uma variável contínua).

A palavra *fator* é frequentemente usada como sinônimo da expressão *variável independente*. Uma *variável independente manipulada* é uma condição que está sob controle direto do experimentador. *Variáveis de tratamento* são, normalmente, variáveis independentes manipuladas pelo pesquisador, cujos efeitos ele está querendo observar. Uma *variável independente normativa* é aquela na qual o pesquisador não está livre para produzir a condição em si mesma, embora ele seja livre para decidir quais níveis da variável serão incluídos na investigação². Sexo e idade são exemplos de variáveis deste tipo.

Medidas iniciais, ou seja, obtidas antes do início da investigação, que são usadas para formar grupos homogêneos (relativamente à variável dependente), são chamadas de *variáveis de agrupamento*. Observações suplementares às observações antes do tratamento relativamente a possíveis diferenças são chamadas de *observações concomitantes* ou *covariáveis*. Uma observação concomitante pode ser usada como uma alternativa ao agrupamento em delineamentos experimentais³ ou, em alguns delineamentos quase-experimentais, ela pode ser usada como um esforço para superar as deficiências do delineamento.

Outro tipo de caracterização de variáveis diz respeito à natureza do processo de contagem da variável. Podemos definir quatro grupos de variáveis, com respeito à escala usada para medir as variáveis:

² Este tipo de variável também é chamada de parâmetro.

³ Aqueles nos quais a amostragem é aleatória; ver p. 26.

1. **Variáveis nominais** – estamos lidando com esse tipo de variável quando apenas apontamos se a grandeza medida pela variável está presente ou não. Assim, por exemplo, uma pessoa pode ser professor ou não. Normalmente, atribui-se valor 1 se o atributo medido pela variável está presente e atribui-se o valor 0 se o atributo não está presente. Ou seja, neste caso, professor (1) – não professor (0).
2. **Variáveis ordinais** – são aquelas em que os dados apresentam algum tipo de ordenação. É atribuído um grau, conforme alguma escala, a cada sujeito na medida em que ele apresenta o atributo que está sendo medido. Por exemplo, um teste classificatório de proficiência em matemática, no qual os sujeitos são ordenados de acordo com as notas obtidas no teste.
3. **Variáveis intervalares** – este tipo de variável se caracteriza por valores que estão distribuídos em uma escala com uma diferença constante entre dois valores consecutivos. Idade pode ser um exemplo deste tipo de variável, se contarmos apenas os anos efetivamente completados.
4. **Variáveis racionais** – são aquelas em que a escala de medida é composta por números racionais e, além disso, existe um zero que define a ausência da propriedade medida pela variável. Por exemplo, notas em um teste são variáveis desse tipo se atribuirmos valores fracionários às respostas.

Controle de variáveis

Variáveis que não são de interesse direto do pesquisador podem ser removidas ou terem a sua influência minimizada por vários métodos:

1. **Remoção de variáveis** – algumas variáveis podem ser eliminadas selecionando-se casos com características uniformes (usando-se, por exemplo, apenas mulheres, ou apenas homens, para eliminar-se a influência da variável sexo).
2. **Pareamento de casos** – selecionando-se pares ou conjuntos de indivíduos com características idênticas (ou quase idênticas) e distribuindo-os aos grupos experimental e de controle. Entretanto, o pareamento não é considerado satisfatório a menos que os sujeitos dos pares ou conjuntos sejam distribuídos aleatoriamente aos grupos experimental ou de controle. Uma limitação desse método ocorre pela dificuldade em parear-se sujeitos usando-se duas ou mais variáveis.
3. **Balanceamento de casos** – distribui-se os sujeitos aos grupos experimental e de controle de tal modo que médias e variâncias dos grupos sejam semelhantes tanto quanto possível. Este método também apresenta uma dificuldade similar àquela observada no pareamento de casos: a dificuldade de equacionar grupos com base em mais do que uma variável.
4. **Análise de covariância** – este método permite ao experimentador eliminar diferenças iniciais em várias variáveis entre os grupos experimental e de controle por métodos estatísticos. Usando escores de pré-testes como covariáveis, esse método é considerado preferível ao convencional pareamento de grupos.
5. **Aleatoriedade** – a aleatoriedade pode ser obtida através da seleção ao acaso dos sujeitos, dentre aqueles da população que se quer estudar, que vão participar dos

grupos de controle e experimental. A aleatoriedade nos dá um método efetivo de eliminar erros sistemáticos e de minimizar o efeito de variáveis externas. O princípio da aleatoriedade é baseado na hipótese de que através da seleção a eatória, quaisquer diferenças entre os grupos sejam simplesmente devido à amostragem ou ao acaso. Essas diferenças são conhecidas como erros de amostragem ou erros de variância e suas intensidades podem ser estimadas pelo pesquisador. Em um experimento, diferenças na variável dependente que podem ser atribuídas ao efeito da variável independente são conhecidas como variância experimental. A significância de um experimento pode ser testada comparando-se a variância experimental com o erro de variância. Se ao término do experimento as diferenças entre os grupos experimental e de controle forem muito grandes para serem atribuídas ao erro de variância, pode-se presumir que estas diferenças são atribuíveis à variância experimental. A aleatoriedade é o método mais efetivo de se formar grupos e controlar variáveis externas, devendo ser usado sempre que as circunstâncias o permitirem. (Best, 1970).

Delineamentos de pesquisa

Entende-se por delineamento de uma pesquisa ao conjunto composto pelo plano de trabalho do pesquisador, a maneira como este seleciona as suas amostras e analisa os seus dados. Pode-se dizer que de nada valem a observação cuidadosa e a exaustiva e detalhada análise estatística se isto for feito para um plano de pesquisa inadequado à situação em estudo. Convém lembrar o que já foi ressaltado neste texto: não é uma boa estatística que torna boa uma pesquisa.

O assunto do delineamento experimental encontra-se muito

bem desenvolvido e apresentado na obra de Campbell e Stanley (1963, 1991), livro, aliás, de leitura obrigatória para o pesquisador interessado em métodos quantitativos. Aqui será feito apenas um resumo da classificação de Campbell e Stanley.

Seguir-se-á aqui a notação clássica de Campbell e Stanley no que se refere à designação de observações e tratamentos. Designaremos pela letra *O* uma observação. Um subíndice na letra *O* indica uma observação particular de uma série, *não necessariamente em ordem cronológica*. O índice funciona apenas como um rótulo para uma dada observação. Designaremos pela letra *X* a um tratamento. Então, como exemplo, a sequência

$$O_1 \ X \ O_2$$

indica que foi feita uma observação (denotada por O_1). Em seguida, foi aplicado um tratamento *X* e então feita uma segunda observação (denotada por O_2). Quando na frente de uma sequência de observações e tratamentos estiver a letra *A*, isso significa que as amostras foram selecionadas aleatoriamente. Cada sequência que se encontra em uma linha diz respeito a um mesmo grupo de sujeitos (a amostra). Assim, a sequência

$$A \ O_1 \ X \ O_2$$

$$A \ O_3 \ O_4$$

é lida como: observa-se um grupo de sujeitos uma vez (O_1), submete-se então o grupo de sujeitos a um determinado tratamento *X* e então observa-se o mesmo grupo de sujeitos uma segunda vez (O_2). Este grupo de sujeitos é chamado de *grupo experimental*. A segunda linha do delineamento significa que se observa um segundo grupo uma vez (O_3), não se aplica o tratamento* e então o grupo de sujeitos é observado uma segunda vez (O_4). A esse segundo grupo é dado o nome de *grupo*

* De fato, essa denominação é arbitrária uma vez que um não tratamento também é um tratamento.

de controle. A letra A indica que os dois grupos, de controle e experimental, foram escolhidos aleatoriamente. Outra convenção adotada é a de que quando duas letras se encontram na mesma coluna, significa que os eventos aconteceram simultaneamente no tempo. Assim, no exemplo, as observações O_1 e O_3 aconteceram no *mesmo instante de tempo*⁵.

Seguindo a classificação de Campbell e Stanley, os delineamentos de pesquisa podem ser divididos segundo três classes:

- Delineamentos não experimentais.
- Delineamentos experimentais.
- Delineamentos quase-experimentais.

Por *experimental*, que é o adjetivo comum a todas as classes colocadas anteriormente, entendemos condições *controladas* de pesquisa. Ou seja, o pesquisador deve ser capaz de controlar, ou de levar em conta de forma apropriada, todas as variáveis pertinentes a um determinado estudo. Os fatores de validade de cada experimento podem ser classificados como fatores de validade interna, os quais dizem respeito às variáveis que se não controladas, tornam sem significância quaisquer asserções de conhecimento sobre os resultados do experimento, ou fatores de validade externa, os quais, se não controlados invalidam, a generalização dos resultados do experimento para uma determinada população. Um determinado experimento somente é verdadeiramente experimental caso se controle a totalidade das variáveis que influenciam os fatores de validade, interna e/ou externa. Obviamente, quando se trata de pesquisa em Ciências Sociais, nem sempre é possível controlar apropriadamente todas

⁵ Claro que esta afirmação deve ser entendida como aproximadamente no mesmo instante de tempo, podendo haver um certo intervalo de tempo entre as observações.

as variáveis envolvidas em determinada situação. Portanto, o estudo será tão mais próximo de um experimento verdadeiro quanto mais controle o experimentador tiver sobre estas variáveis. A linha básica de raciocínio para classificar determinado delineamento numa ou noutra categoria é o quanto mais controle o delineamento oferece aos fatores de validade interna e externa.

Delineamentos que oferecem pouco ou nenhum controle das variáveis pertinentes são chamados de *delineamentos não experimentais* ou *pré-experimentais*. Por outro lado, delineamentos que oferecem alto grau de controle são chamados de *delineamentos experimentais*. Por fim, delineamentos que oferecem grau de controle em nível médio, porém sem oferecer controle nos níveis da categoria anterior, são chamados de *delineamentos quase-experimentais*.

Na óptica quantitativa, o pesquisador deve sempre procurar um delineamento experimental para o seu trabalho. Na impossibilidade de um tratamento deste tipo, é aceitável um delineamento quase-experimental. Um tratamento não experimental nunca é aceitável.

Delineamentos não experimentais ou pré-experimentais

Dentro da classificação de Campbell e Stanley, os delineamentos não experimentais ou pré-experimentais são de três tipos.

Delineamento de tipo 1 – Neste delineamento é observado apenas um grupo sob a ação do tratamento X. Esse delineamento é esquematizado como:

$$X \quad O_1$$

ou seja, nesse delineamento o grupo que experimentou o tratamento X, e apenas ele, é observado uma única vez.

As desvantagens desse delineamento são evidentes. Talvez a maior delas seja o fato de que não se tem controle algum sobre as

variáveis externas que atuam concomitantemente com X . Variáveis tais como *história, maturação, interação do experimentador com o tratamento, etc.*, não são de forma alguma controladas. Não há razão plausível para a utilização desse delineamento, devendo-se evitá-lo a todo o custo. Um exemplo desse tipo de delineamento é aquela situação em que o professor aplica um novo método de ensino e então uma prova. Nenhuma consequência que se tire do resultado da prova é válida devido às deficiências desse delineamento.

Delineamento de tipo 2 – Um delineamento, já referido na seção anterior, muito usado na pesquisa em ensino, mas que na verdade é um delineamento pré-experimental, é o seguinte:

$$O_1 \quad X \quad O_2$$

Nesse delineamento, aplica-se um pré-teste O_1 a um grupo, submete-se o grupo a um tratamento X e aplica-se, então, um pós-teste O_2 . O_1 e O_2 significam que o mesmo grupo é observado antes e depois do tratamento que pode ser, por exemplo, um novo método de ensino ou um recurso didático alternativo. Diferenças entre O_1 e O_2 (que podem ser simples testes de conhecimento) evidenciarão a eficácia ou ineficácia do tratamento X . O problema com esse delineamento é que não controla outras variáveis, além de X , que poderiam explicar as diferenças entre O_1 e O_2 . Por exemplo, os alunos poderiam ter melhores resultados no pós-teste porque algum evento ocorreu entre a aplicação do pré-teste e do pós-teste (variável história), e não porque o tratamento X tenha sido eficiente.

Delineamento de tipo 3 – Deve-se ter cuidado em não confundir este delineamento com um dos delineamentos experimentais que vão ser descritos mais adiante. Nesse delineamento, temos dois grupos, experimental e de controle, mas a seleção dos sujeitos que pertencem aos dois grupos não é aleatória. Dessa

forma, esse delineamento não controla a variável *seleção*. Esse delineamento tem a forma

$$\begin{array}{cc} X & O_1 \\ & O_2 \end{array}$$

em que a ausência da letra A à esquerda de X significa que não houve aleatoriedade no processo de seleção. Nesse tipo de delineamento não há evidência alguma da equivalência entre os dois grupos antes do início do experimento. Um exemplo desse tipo de delineamento ocorre quando se seleciona dois grupos de sujeitos de determinada escola para constituírem os grupos experimental e de controle, pelo simples fato de pertencerem à mesma turma. Se as turmas forem formadas por alunos que no ano anterior foram bons alunos em Matemática, ou maus alunos nesta disciplina, então a variável seleção com certeza influenciará qualquer tratamento alternativo sobre o ensino de Matemática que for aplicado a um ou outro grupo.

Delineamentos experimentais

Os delineamentos experimentais são aqueles em que se consegue controlar a maior parte, senão todas, as fontes de invalidade interna e externa. Seguindo a tradição de Campbell e Stanley, estes delineamentos podem ser classificados em três categorias.

Delineamento de tipo 4 - Um delineamento experimental, já mencionado na seção anterior, muito usado é o seguinte:

$$\begin{array}{cccc} A & O_1 & X & O_2 \\ A & O_3 & & O_4 \end{array}$$

Neste delineamento trabalha-se com dois grupos e os sujeitos da pesquisa são designados *aleatoriamente* a um deles (este é o significado de A). Observa-se ambos os grupos antes da aplicação do tratamento X, por exemplo, aplicando-se um pré-teste aos dois grupos ($O_1 = O_3$). Um dos grupos (grupo experimental) é então submetido ao tratamento X, enquanto o

outro (grupo de controle) não recebe o tratamento. Após, observa-se os grupos, aplicando-se, por exemplo, um pós-teste ($O_2 = O_4$) a ambos os grupos. Na prática, os pré e pós-testes podem ser iguais.

Um erro comum no uso de delineamentos desse tipo é analisar o resultado para determinação da eficácia do tratamento do seguinte modo: tomam-se as diferenças entre os resultados do pré e pós-teste em ambos os grupos ($O_1 - O_2$ e $O_3 - O_4$), aplicando-se a seguir um teste estatístico. Se a diferença entre as médias do grupo experimental antes e depois da aplicação do tratamento X , for estatisticamente significativa e a diferença entre as médias do grupo de controle não for estatisticamente significativa, toma-se então o tratamento como eficaz. Esta é uma forma errada de se analisar a eficácia do tratamento e não é fornecida evidência alguma sobre o efeito do tratamento X . A forma correta de se proceder é comparar o resultado *final* (as médias finais em um teste de conhecimento por exemplo) entre os grupos experimental e de controle entre si.

Esse delineamento controla variáveis na medida em que elas influenciarão igualmente ambos os grupos, exceto X , obviamente, e, portanto, os efeitos dessas variáveis não pesarão na comparação das diferenças $O_1 - O_2$ e $O_3 - O_4$.

Além disso, a aleatoriedade da designação dos sujeitos a um dos grupos, embora não garanta equivalência entre os grupos em 100 %, reduz ao mínimo a probabilidade de que sejam diferentes. Segundo Kerlinger (1980, p. 102):

Casualização é a designação de objetos (sujeitos, tratamentos, grupos) de um universo a subconjuntos do universo de tal maneira que, para qualquer designação dada a um subconjunto, todo membro do universo tem igual probabilidade de ser escolhido para o

designação. Não há total garantia de que a casualização "igualará" os grupos, mas a probabilidade de igualar é relativamente alta. Há outra forma de expressar essa ideia: [...] já que em procedimentos aleatórios todo membro de uma população tem igual probabilidade de ser escolhido, membros com certas características distintas – homem ou mulher, alto ou baixo grau de inteligência, dogmático ou não dogmático, e assim por diante –, se selecionados, provavelmente serão contrabalançados em longo prazo pela seleção de outros membros da população com a quantidade ou qualidade 'opostas' da característica.

Delineamento tipo 5^e – A aleatoriedade da designação de sujeitos aos grupos de controle e experimental é, portanto, a mais adequada segurança de que não existam diferenças ou vieses iniciais entre os grupos. Nesse caso, o pré-teste não é condição essencial para que um delineamento seja verdadeiramente experimental. Assim, o delineamento anteriormente apresentado poderia ser simplesmente:

A	X	O ₁
A		O ₂

De fato, esse delineamento não só pode ser usado ao invés do anterior como também é mais adequado, pois elimina qualquer influência do pré-teste no experimento. Entretanto,

^e Aqui se está fazendo uma Inversão entre a denominação dada por Campbell e Stanley aos delineamentos 5 e 6. Para Campbell e Stanley, o que está sendo chamado de delineamento 5 é o delineamento 6 e vice-versa.

talvez por razões psicológicas, muitos pesquisadores não abrem mão de saber "com certeza" se os grupos experimental e de controle eram iguais no início do experimento, de modo que o quarto exemplo de delineamento aqui apresentado é provavelmente mais usado que o quinto, embora menos apropriado logicamente.

Delineamento tipo 6 (Delineamento de quatro grupos de Solomon) – Este delineamento é a soma das vantagens dos delineamentos quatro e cinco. Seu esquema é:

A	O ₁	X	O ₂
A	O ₃		O ₄
A		X	O ₅
A			O ₆

Este tipo de delineamento controla variáveis como *interação do pré-teste com o tratamento, maturação e história*. A desvantagem desse tipo de delineamento é a dificuldade em se obter tantos grupos para participar da pesquisa.

Delineamentos quase-experimentais

Um terceiro grupo de delineamentos identificado por Campbell e Stanley é o dos delineamentos quase-experimentais, ou seja, aqueles em que falta ao pesquisador "*o pleno controle da aplicação dos estímulos experimentais – quando e quem expor e a capacidade de casualizar exposições*" (op. cit., p. 61). Todos os delineamentos pertencentes a esse grupo carecem do rigor e controle existentes nos delineamentos pertencentes ao grupo dos delineamentos experimentais, mas podem ser usados quando a situação não permitir o uso de delineamentos verdadeiramente experimentais.

Delineamento tipo 7 (Série temporal) – O delineamento "série temporal" exemplifica essa situação:

$$O_1 O_2 O_3 O_4 \quad X \quad O_5 O_6 O_7 O_8$$

Neste delineamento, os sujeitos são observados várias vezes antes de se aplicar o tratamento X e várias vezes após a aplicação. Supondo-se que antes do tratamento as observações fossem quase que homogêneas, sem variações, apresentando um padrão bem definido e que houvesse um salto nos escores das observações feitas após o tratamento e que, a partir daí, houvesse nova estabilização nos escores das observações, com a apresentação de outro padrão, esse salto quantitativo na série temporal seria tomado como evidência do efeito X .

Observe-se que esse delineamento é semelhante ao primeiro apresentado como exemplo, porém implica em muito mais observações, o que minimiza, embora não exclua, as deficiências do primeiro. Note-se também que implica a existência de um só grupo, o que, na prática, é uma vantagem, pois muitas vezes é difícil obter-se dois grupos de sujeitos.

Um exemplo simples de aplicação desse delineamento seria aquele em que o professor observa cuidadosamente seus alunos durante algumas semanas do curso, fazendo várias medições (que podem ser testes de aproveitamento ou de atitude) antes de fazer uso de uma nova estratégia de ensino. Da mesma forma, após o uso da estratégia, volta a observar seus alunos, durante algum tempo, fazendo novos registros. Diferenças, qualitativas ou quantitativas, no desempenho dos alunos após o uso da estratégia, e que se mantêm ao longo do tempo, podem ser tomadas como evidência do efeito da estratégia sobre a aprendizagem cognitiva ou afetiva dos alunos.

Delineamento de tipo 8 (Amostras temporais equivalentes) –
Este delineamento é, de fato, uma variação do delineamento anterior. Nele, introduz-se a variável experimental (o tratamento X) alternadamente e observa-se o grupo. O seu esquema é o seguinte:

$$O_1 \times O_2 \quad X_0 \quad O_3 \times O_4 \quad X_0 \quad O_5 \times O_6 \quad X_0 \quad O_7 \times O_8$$

Como se pode ver neste delineamento, o mesmo grupo de sujeitos é observado alternadamente na presença do tratamento e sem a presença do tratamento (aqui simbolizada por X_0). A análise é feita a partir da comparação dos valores médios do grupo com e sem tratamento experimental. Neste ponto, esse delineamento se assemelha a um delineamento com dois grupos.

Delineamento de tipo 9 (Grupo de controle não equivalente) –
Este delineamento tem a seguinte estrutura:

$$\begin{array}{ccc} O_1 & \times & O_2 \\ O_3 & & O_4 \end{array}$$

Neste caso, o grupo de controle e o grupo experimental não possuem equivalência amostral, pois não foi usada a aleatoriedade na escolha das amostras. Neste tipo de delineamento, os grupos constituem coletivos reunidos naturalmente, tais como classes escolares já compostas previamente à ação do pesquisador. O controle do pesquisador reside unicamente na decisão sobre qual dos grupos vai receber o tratamento e quando.

Seguramente, nesse tipo de delineamento haverá problemas sérios derivados do fator *seleção* e de sua interação com outros fatores importantes tais como *história*, *maturação*, etc.

Campbell e Stanley propõem ainda cinco outros delineamentos quase-experimentais que não serão aqui apresentados.

Nesta seção foi dada ênfase bastante grande ao delineamento, uma vez que esta é uma questão crucial na realização de uma pesquisa quantitativa em ensino. Assim como o pesquisador deve formular uma questão de pesquisa clara, orientadora e relevante, ele deve também investigá-la usando um delineamento adequado.

Um mau delineamento pode invalidar as asserções de conhecimento (resultados) e de valor de uma pesquisa, jogando por terra todo o trabalho realizado, seja por não controlar as fontes de invalidade interna seja por não controlar as de invalidade externa.

Sempre que possível, um dos delineamentos experimentais deve ser utilizado. Quando isso não é possível, a alternativa é o uso de um dos delineamentos quase-experimentais aqui expostos (e discutidos em muito maior extensão em Campbell e Stanley, 1963, 1991), tendo em mente as deficiências que estes delineamentos oferecem. A pesquisa não deve deixar de ser feita se a situação não permitir o uso de um delineamento puramente experimental, mas o pesquisador deve deixar claro para si e para os demais pesquisadores a limitação de escopo, no que tange à validade, de seus resultados bem como quais são os pontos em que novos trabalhos devem ser realizados de modo a estudar a interferência de fatores os quais não puderam ser controlados naquele experimento específico. Cabe lembrar que não é uma única pesquisa que constrói o corpo de conhecimentos de uma área mas, sim, um conjunto delas.

Um pouco de estatística

O objetivo desta seção e das próximas é o de discutir os principais tópicos relacionados com a análise de experimentos quantitativos em ensino. Como foi dito anteriormente, a principal ferramenta para esse tipo de análise é a Estatística. Justifica-se, portanto, dedicar algum espaço a ela.

Amostragem aleatória

O termo técnico aleatório indica que a amostra é selecionada de tal modo que cada elemento da população tem uma chance igual de entrar na amostra. O pesquisador deve ter

uma lista completa de todos os elementos da população e então selecionar sua amostra de modo tal que nenhum elemento da população seja privilegiado pelo procedimento de escolha.

O propósito da aleatoriedade não é garantir que os dois grupos se comportarão igualmente bem na ausência do tratamento. A aleatoriedade não garante igualdade. A aleatoriedade permite evitar aquele tipo de resultado que poderia ser atribuído à variabilidade da amostra. Aleatoriedade é um procedimento para selecionar amostras, e não uma característica da amostra. Ela também não assegura representatividade, nem dá indicativo de como a amostra se comportará.

Define-se por erro de amostragem aquele tipo de erro que se comete ao selecionar amostras aleatórias para representar a população. Em virtude deste tipo de erro, é virtualmente impossível para um grupo pequeno ser exatamente representativo de outro muito maior. Esse erro de amostragem está presente toda vez que se selecionar amostras, não importa o quão cuidadosa seja feita a seleção aleatória. A seguir, são definidos alguns tipos de procedimentos de seleção de amostras:

- Amostragem aleatória simples – é o processo de selecionar observações de um grupo maior de tal modo que cada sujeito na população de onde se está selecionando a amostra tenha uma probabilidade igual e independente de ser selecionada.
- Amostragem estratificada – é, algumas vezes, um modo recomendado de proceder ao escolher amostras. Divide-se a população em grupos menores homogêneos de modo a obter-se uma melhor representação. Com cada subgrupo algum processo de seleção aleatória pode ser usado. Esse processo dá ao pesquisador uma amostragem mais significativa da que seria obtida diretamente da comunidade inteira. De modo a ter-se

uma representação mais fiel da população como um todo, pode-se, ainda, fornecer pesos ao número de sujeitos pertencentes aos diferentes grupos de modo a ter-se uma representação proporcional à distribuição na população.

- **Amostragem estratificada proporcional** – este tipo de processo de amostragem ocorre quando se toma uma percentagem de cada grupo no processo de composição da amostra.
- **Amostragem sistemática** – tem-se este tipo de amostragem caso uma população tenha sido listada e então algum tipo de seleção seja feita segundo algum critério tal como, por exemplo, tomar-se o *enésimo* elemento da lista.
- **Amostragem por agrupamento** – é uma variação da amostragem aleatória simples, particularmente apropriada quando a população é grande ou quando a distribuição geográfica da população é espalhada. A cada agrupamento é dado um número e são selecionados grupos aleatoriamente. O uso da amostragem por agrupamento é geralmente escolhido por razões econômicas e aspectos administrativos.

Em qualquer tipo de amostragem, a característica da amostra inevitavelmente diferirá em algum grau, ainda que pequeno, da característica da população. Mas quando a amostragem aleatória é usada, as chances de que a ocorrência do erro de amostragem influencie a variável dependente em uma direção particular são as mesmas que existem de influenciar a mesma variável em outra direção qualquer. *Amostragem aleatória é a única forma de amostragem pela qual alguma quantidade de erro específica pode ser estimada. Pode-se dizer que é a amostragem aleatória o que diferencia as pesquisas experimentais das pesquisas não-experimentais.*

Medidas de tendência central

Entende-se por distribuição de frequências a uma tabulação (listagem) dos escores obtidos em uma certa amostra com o número de vezes em que esses escores aparecem na amostra. Assim, uma listagem com as notas obtidas pelos alunos em determinada avaliação com o número de vezes com que cada nota aparece é uma distribuição de frequências.

Quando o número de sujeitos pertencentes à(s) amostra(s) é pequeno, pode-se ter uma ideia de como se comportam os escores. No entanto, se tivermos amostras muito grandes, fica difícil ter uma ideia do comportamento da amostra e temos, então, de recorrer a transformações sobre os escores. Uma das formas de fazer isto é usando as chamadas medidas de tendência central.

Define-se uma medida de tendência central como sendo um número em torno do qual os valores da distribuição de frequências se distribuem. Em Estatística, são utilizados basicamente três tipos de medidas de tendência central: a média, a moda e a mediana.

- A Moda de uma distribuição de frequências é definida como o escore que tem a mais alta frequência.
- A Mediana é definida como aquele escore que marca o ponto médio do conjunto de dados, ou seja, aquele escore para o qual tem-se 50% dos escores com valores superiores e 50 % dos escores com valores inferiores.
- A Média aritmética simples de uma distribuição de frequências é definida como aquele número obtido a partir da divisão da soma de todos os escores obtidos pelo número de elementos na distribuição de frequências. Matematicamente:

$$\bar{X} = \frac{\sum X_i n_i}{N}$$

em que os X_i são os escores obtidos, n_i é o número de vezes em que cada escore aparece e N é o número de elementos na distribuição de frequências. O símbolo $\sum$ significa que se está somando.

Exemplo: seja a distribuição de frequências constante Tabela 2.

Tabela 2 - Distribuição de frequências para uma prova de uma disciplina hipotética.

Turma 1	9,8
Turma 2	7,4
Turma 3	6,2
Turma 4	6,0
Turma 5	5,9
Turma 6	4,5
Turma 7	3,4
Turma 8	3,4
Turma 9	3,4
Turma 10	1,0
Turma 11	0,5

A moda desta distribuição é dada pelo valor 3,4, pois é o escore que aparece o maior número de vezes na distribuição (3 vezes). A mediana da distribuição é o valor 4,5, pois, para esse valor, 50 % (5 escores) são maiores que ele e 50 % são menores. Por fim, a média aritmética simples da distribuição é dada por:

$$\bar{X} = \frac{9,8 + 7,4 + 6,2 + 6,0 + 5,9 + 4,5 + 3 \times 3,4 + 1,0 + 0,5}{11} = \frac{51,5}{11}$$

$$\bar{X} = 4.7$$

Outros tipos de média podem ser definidos como, por exemplo, a média aritmética ponderada e a média geométrica.

A média aritmética ponderada é usada quando se deseja ter uma medida de tendência central de certa distribuição de frequências em que os escores contribuem com pesos diferentes. Matematicamente, ela é definida por:

$$\bar{X}_p = \frac{\sum p_i n_i X_i}{\sum p_i}$$

em que $\bar{X}_p$ é o valor da média ponderada, p_i é o peso atribuído a cada escore X_i e n_i é o número de vezes que o escore aparece na distribuição de frequências.

A média geométrica é definida por:

$$\bar{X}_g = \sqrt[n]{\prod n_i X_i}$$

em que $\bar{X}_g$ é a média geométrica, o símbolo Π significa que devemos multiplicar os elementos que vêm a seguir e os X_i e n_i foram definidos anteriormente. Esse tipo de média é utilizado quando os escores da nossa distribuição são obtidos uns dos outros a partir de um fator multiplicativo.

Medidas de variabilidade

A média, seja ela de qual tipo for, não diz tudo a respeito de uma distribuição de frequências. Aqui cabe um comentário a respeito dos processos estatísticos e a perda de informação que necessariamente ocorre quando se usa números, tais como a média, para a representação de distribuições de frequências. No processo de mediação ocorre uma perda de informação, uma vez que se substitui a informação total, ou seja, a distribuição, por algo que pretende ser representativo dessa distribuição. Ao fazê-lo, perde-se a estrutura fina de informação provida pelo conjunto completo de escores. Assim, por exemplo, consideremos as hipotéticas Tabelas 2 e 3:

Tabela 2 - Notas na Turma 1.

aluno 1	5,1
aluno 2	5,9
aluno 3	7,2
aluno 4	5,9
aluno 5	5,9

Tabela 3 - Notas na
Turma 2.

aluno 1	9,0
aluno 2	5,0
aluno 3	6,0
aluno 4	8,0
aluno 5	2,0

Se for calculada a média aritmética simples para essas duas distribuições de frequência, obter-se-á para ambas a média 6,0. Entretanto, se essas distribuições forem representativas de notas de alunos em duas turmas diferentes, em uma disciplina cuja média de aprovação seja 6,0, observa-se, então, que na Turma 1 há um sujeito aprovado enquanto que na Turma 2 há três. Olhando somente para a média das duas turmas, essa informação seria perdida.

Uma forma de minimizar essa perda de informação, decorrente do uso de uma medida de tendência central, são as medidas de variabilidade. Uma medida de variabilidade indica o quão espalhados estão os escores na distribuição. Ou seja, uma medida de variabilidade é uma forma de ter uma ideia do quanto os escores se afastam da medida de tendência central que se está utilizando.

A mais simples das medidas de variabilidade é o intervalo (I). O intervalo diz entre quais valores se distribuem os escores da distribuição que está sendo analisada. Assim, no exemplo, o intervalo para a distribuição 1 é dado por $[7,2; 5,1]$, enquanto na distribuição 2 o intervalo é dado por $[9,0; 2,0]$.

Outra medida de variabilidade é o desvio da média (dm). Essa quantidade diz o quanto os escores se desviam da média. No exemplo dado, o escore 9,0 da Tabela 3 possui um desvio da média de 3,0 ($9-6=3$).

Uma medida de variabilidade das mais utilizadas em análises

de distribuições de frequência é o desvio padrão (dp). O desvio padrão indica o quão *espalhada* é uma distribuição. O desvio padrão tem uma interpretação muito simples, originada da equação que define a *distribuição normal* (que será analisada na próxima seção). Tomando um intervalo definido por $[\bar{X} - dp; \bar{X} + dp]$, tem-se, dentro desse intervalo, em torno de 68 % dos valores da distribuição. Fazendo uma nova medida, então, há uma *probabilidade de 68 %* de que essa nova medida pertença a esse intervalo. Considerando intervalos definidos por múltiplos do desvio padrão se estará englobando um número cada vez maior de escores dentro do intervalo em questão.

Matematicamente, o desvio padrão é dado por:

$$dp = \sqrt{\frac{\sum_{j=1}^N (X_j - \bar{X})^2}{N}}$$

em que N é o número de escores, X_j significa o j -ésimo elemento da distribuição de escores e $\bar{X}$ é o valor médio da distribuição.

Curva normal

Define-se como *curva da distribuição* a representação gráfica de uma frequência de distribuição de escores em que os valores dos mesmos são indicados no eixo horizontal e os valores das frequências dos escores particulares no eixo vertical. Curvas de distribuição podem vir em diferentes formas e tamanhos. No entanto, muitas frequências de distribuição tendem a seguir um certo padrão chamado de *distribuição normal*, especialmente quando existem muitos escores envolvidos. A forma da curva que pode ser traçada destas distribuições é chamada *curva normal* (Figura 1). Uma das características da curva normal é a *simetria*; outra característica importante é que a média, a mediana e a moda são *idênticas*.

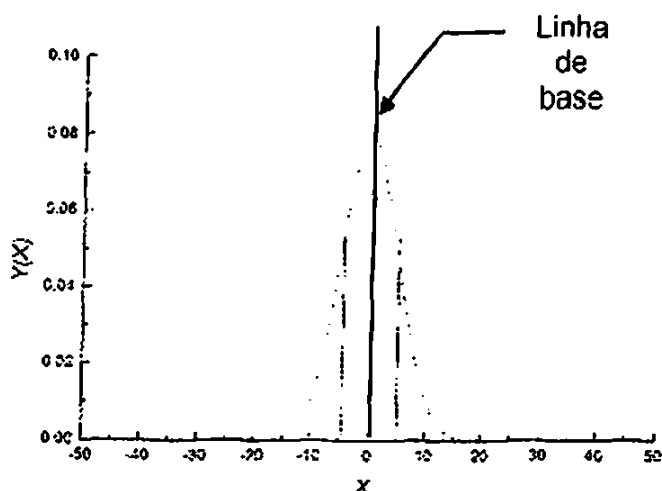


Figura 1 – A curva normal

Na curva normal, a distância da linha de base até a primeira linha pontilhada é o desvio padrão da distribuição, por definição⁷. A equação que define a curva normal é dada por:

$$Y(X) = \frac{1}{dp\sqrt{2\pi}} \exp \left[-\frac{1}{2} \frac{(X - \bar{X})^2}{dp^2} \right]$$

em que dp é o desvio padrão, $\bar{X}$ é o valor médio e X é o valor para o qual se quer calcular Y .

Intervalos de confiança

Quando se diz que um escore aleatoriamente selecionado cairá dentro de um intervalo específico dos valores dos escores obtidos, é preciso fazê-lo com algum grau de confiança, ou seja, sabendo quão provável é que se esteja correto. O intervalo de 95% é chamado de intervalo de confiança porque, se a distribuição for normal, pode-se estar certo de que 95% das vezes

⁷ No presente exemplo, a curva foi gerada com um desvio padrão assumido de 5.

em que um valor da variável em questão for escolhido aleatoriamente ele estará dentro desse intervalo. Em termos de probabilidades, o intervalo de confiança de 95% designa os dois valores entre os quais existe a probabilidade $p=0,95$ de que um escore selecionado aleatoriamente pertença a esse intervalo. O intervalo de confiança de 95% é dado por:

$$I_{95\%} = [\bar{X} - 1,96dp; \bar{X} + 1,96dp]$$

em que dp é o desvio padrão e $\bar{X}$ a média calculados para a distribuição.

O intervalo de confiança de 99% é outro intervalo normalmente utilizado. Os limites desse intervalo estão entre:

$$I_{99\%} = [\bar{X} - 2,58dp; \bar{X} + 2,58dp]$$

Distribuição de médias amostrais

Suponha-se que são selecionados de uma população hipotética um grande número de amostras, cada uma das quais, digamos, com 50 sujeitos, e seja calculado o escore médio para cada grupo em alguma variável de interesse. Pode-se então traçar uma curva representando a distribuição desses escores. Nessa distribuição de médias amostrais, a média de todas as médias é a média da população e as médias amostrais são distribuídas em torno da média da população seguindo a distribuição normal.

Quando se lida com distribuições de médias, o desvio padrão é chamado de desvio padrão da média (dpm). A interpretação dessa quantidade é semelhante àquela do desvio padrão de medida: tomando-se outra amostra da população a probabilidade de que a média obtida por essa nova amostra esteja dentro do intervalo de confiança de 95% é dada por:

$$I_{95\%} = [\bar{X} - 1,96dpm; \bar{X} + 1,96dpm]$$

Inferências a respeito do valor da média da população

Quando se sabe a média de uma amostra e toma-se uma estimativa do desvio padrão da média, não é possível inferir o valor da média da população a partir da média que se tem para aquela amostra, mas pode-se fazer hipóteses sobre o valor para a média da população e, usando a estimativa do desvio padrão da média, determinar a probabilidade de obter-se uma média amostral que difira da média hipotetizada da população tanto quanto se queira.

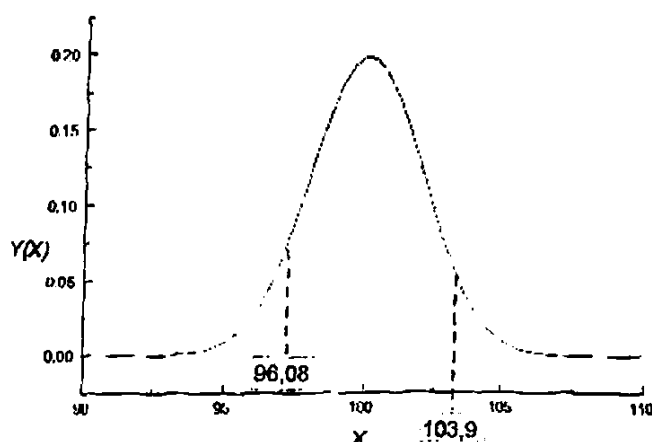


Figura 2 – Curva normal indicando intervalo de confiança de 95 %.

Suponhamos que uma média amostral seja 97 e que o desvio padrão da média estimado seja de $dpm=2$. Suponhamos também que a hipótese feita foi de que a média da população é 100. Pode-se então determinar a probabilidade de obter nossa média da amostra de 97.

A Figura 2 indica que o intervalo de confiança de 95% varia de 96,08 ($100 - 1,96 \text{ dpm}$) a 103,92 ($100 + 1,96 \text{ dpm}$). Isto diz que se a média da população for 100, então, a probabilidade de

selecionar-se uma amostra cuja média esteja dentro do intervalo de confiança considerado é $p=0,95$. Outra forma de dizer isso é que a probabilidade de obter-se uma média amostral menor que 96,08 ou maior que 103,92 é $p=0,05$. Portanto, pode-se aceitar a hipótese de que a amostra, cuja média é 97, vem de uma distribuição de médias amostrais retiradas de uma população cuja média é 100.

Comparações entre múltiplas amostras

Em situações de pesquisa, usualmente deseja-se comparar duas ou mais amostras. Por exemplo, pode-se querer determinar se existe uma diferença em aquisição de conhecimentos entre estudantes que são ensinados por um método A e aqueles que são ensinados por um método B. A questão é: *qual é a probabilidade de que diferenças entre as duas médias das amostras seja devido simplesmente a erro de amostragem?* Em outras palavras, *pode a diferença de médias entre as duas amostras ser atribuída ao erro aleatório nas amostras, ou os alunos ensinados por um método realmente aprendem mais que aqueles ensinados pelo outro método?* Temos, de fato, duas amostras pertencentes a populações diferentes, representadas por duas distribuições normais de escores?

Ao testar a hipótese nula se está assumindo que a diferença entre as duas amostras é devida simplesmente ao erro de amostragem.

Tipos de hipóteses

Quando se quer decidir se determinado procedimento é melhor que outro, formula-se a hipótese de não existe diferença entre os procedimentos (i.e., qualquer diferença observada é meramente devida a flutuações ao tomar duas amostras da mesma população). Este tipo de hipótese é a chamada Hipótese Nula, denotada por H_0 . A hipótese alternativa à H_0 , ou seja, de que a diferença observada não é meramente devido à amostragem, é denotada por H_1 .

Ao decidir se rejeita, ou não, a hipótese nula, dois tipos de erros podem acontecer. O primeiro tipo de erro acontece quando se rejeita a hipótese nula com base em dados de amostras que de fato vêm da mesma população, então diz-se que se comete um erro *Tipo I*. Se por outro lado, aceita-se a hipótese nula quando, de fato, as amostras vêm de populações diferentes, diz-se que foi cometido um erro do *Tipo II*⁸.

Na prática, conforme o caso, um tipo de erro pode ser mais sério do que o outro e, assim, uma solução de compromisso deve ser obtida em favor de uma limitação do erro que, naquela situação, seja considerado mais sério. A única forma de diminuir a influência desses dois tipos de erro é aumentar o tamanho da amostra, o que nem sempre é possível.

Ao testar a hipótese nula, a probabilidade máxima com a qual o pesquisador aceita correr o risco de cometer um erro de *tipo I* é chamada de nível de significância estatística da pesquisa. Quando o pesquisador decide o nível de probabilidade que usará ao rejeitar a hipótese nula, ele estará dando a probabilidade com a qual arriscará estar errado em sua decisão. Se ele seleciona o nível de significância como 0,05, está dizendo que há uma probabilidade de 0,05 de que esteja errado. Se não quiser correr um risco de erro tão significativo, ele pode então escolher para nível de significância $p=0,01$. Neste nível é menos provável que esteja cometendo um erro do *tipo I*; entretanto, estará incrementando a probabilidade de cometer um erro do *tipo II*.

Na prática da pesquisa em educação, os níveis de significância de 0,05 e 0,01 são usuais, embora outros níveis de significância sejam também utilizados.

Uma hipótese que não indica a direção da diferença esperada, mas meramente estabelece que existe uma diferença, é chamada de hipótese bilateral (*two-tailed*). Esse tipo de

⁸ Tecnicamente, o pesquisador não deveria aceitar a hipótese nula, mas, sim, falhar em rejeitar a hipótese nula.

hipótese é assim designada porque está preocupada com ambas as caudas da distribuição normal das diferenças entre médias amostrais.

Uma hipótese que afirma qual tratamento é melhor que o outro é chamada de hipótese *unilateral (one-tailed)*, porque está somente preocupada com um dos lados da distribuição de diferenças entre médias amostrais.

Distribuição de médias de pequenas amostras

Até aqui foram consideradas amostras que contêm um grande número de sujeitos (trinta ou mais). As propriedades da distribuição normal são válidas para grandes amostras, mas não quando há um número pequeno de sujeitos em cada amostra. A distribuição tende a ficar achatada quando, em cada amostra, o número de sujeitos é pequeno.

Para fins estatísticos isso significa que, para dados em que as amostras são pequenas, não é possível usar-se as propriedades da curva normal para decidir a favor ou contra a aceitação da hipótese nula. Ao invés disso, é necessário usar valores que reflitam esse achatamento da curva normal. Esses valores são chamados de valores *t*, para os quais também foram calculados valores para os níveis de significância $p=0,05$ e $p=0,01$ para amostras de qualquer tamanho. Existem tabelas estatísticas preparadas (uma das quais é apresentada na Tabela 4) para valores *t* para todos os tamanhos de amostras sendo comparadas, de modo que, sabendo-se quantos sujeitos estão em cada amostra que se está comparando, poderemos facilmente determinar o valor de *t* necessário ao nível de significância escolhido (normalmente $0,05$ ou $0,01$).

Se um valor *t* indica diferenças dentro do intervalo de confiança de 95%, seu valor normalmente não é comunicado. Ao invés, o pesquisador afirma que o valor *t* não é significativo. Neste caso, ele aceita a hipótese nula e atribui a diferença

observada entre suas amostras ao simples erro de amostragem.

Este tipo de teste estatístico é chamado de teste *t* e é utilizado para comparação entre médias de amostras pequenas quando, por hipótese, as amostras foram escolhidas aleatoriamente e os escores vieram de populações distribuídas segundo a distribuição normal. Outros testes estatísticos estão disponíveis se a hipótese de normalidade não puder ser feita.

Amostras com número de sujeitos menor que 30 são chamadas de pequenas amostras. Um estudo estatístico de distribuições amostrais, no qual as amostras são pequenas, é chamado de *Teoria de Pequenas Amostras*. Entretanto, um nome mais apropriado seria *Teoria Exata da Amostragem*, uma vez que os resultados obtidos mantêm-se tanto para pequenas como para grandes amostras. Uma distribuição importante é a distribuição *t* de Student⁹. Esta distribuição é dada, matematicamente, por:

$$Y = \frac{Y_0}{\left(1 + \frac{t^2}{N-1}\right)^{N/2}}$$

$$Y = \frac{Y_0}{\left(1 + \frac{t^2}{gl}\right)^{(gl+1)/2}}$$

em que Y_0 é uma constante que depende em N de tal modo que a área total sobre a curva é 1¹⁰. A quantidade $N-1$ é chamada de *grau de liberdade (gl)*. Para grandes valores de gl ou N ($N > 30$), as curvas da figura se aproximam da curva normal padrão. A Figura 3 mostra várias curvas dessa distribuição para vários valores de graus de liberdade gl .

Para fins de cálculo entre duas amostras, com N_1 e N_2 ,

⁹ Mantém-se aqui o nome em inglês por ser consagrado na literatura e pelo uso na área.

¹⁰ A isto chama-se de condição de normalização.

sujeitos, de médias e desvios padrão dados por $\bar{X}_1$, dp_1 , $\bar{X}_2$ e dp_2 respectivamente, o valor de t é dado por:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sigma \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}}$$

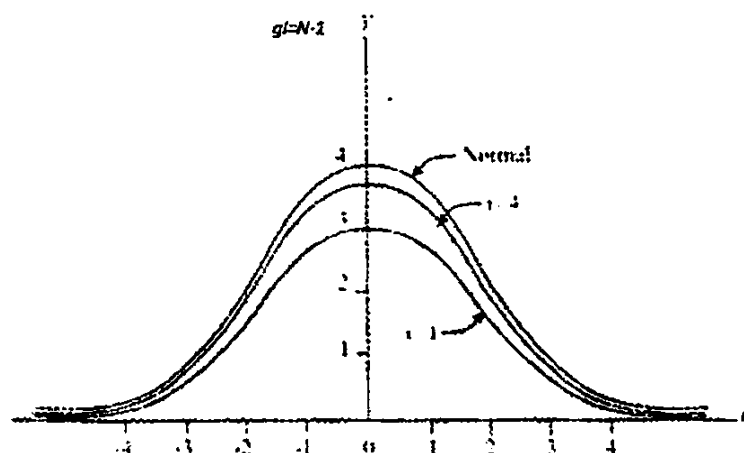


Figura 3 – A distribuição t de Student para distintos graus de liberdade.

em que a quantidade σ é definida por:

$$\sigma = \sqrt{\frac{N_1 dp_1^2 + N_2 dp_2^2}{N_1 + N_2 - 2}}$$

Valores de t foram calculados, para níveis de significância de 0,05 e 0,01, para qualquer tamanho de amostra. Os estatísticos prepararam tabelas estatísticas destes valores t para todos os tamanhos das amostras sendo comparadas, de modo que, se soubermos quantos sujeitos existem em cada amostra, podemos compará-las facilmente e determinar o valor t necessário para o nível de significância desejado (0,05 ou 0,01). Uma tabela desse tipo é exemplificada na Tabela 4.

Tabela 4 - Tabela t para hipóteses unilaterais e bilaterais.

gl	Nível de significância para hipótese unilateral									
	.40	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
	Nível de significância para hipótese bilateral									
	.80	.50	.20	.10	.05	.02	.01	.005	.002	.001
1	0.32	1.000	3.078	6.314	12.706	31.821	63.657	127.32	318.31	536.62
2	.589	0.816	1.885	2.920	4.303	6.965	9.925	14.083	22.326	31.598
3	.777	.765	1.638	2.353	3.182	4.941	5.841	7.453	10.214	12.924
4	.727	.741	1.533	2.132	2.775	3.747	4.604	5.556	7.173	8.610
5	.675	.727	1.475	2.015	2.571	3.455	4.232	4.773	5.893	6.869
6	.659	.718	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	.650	.711	1.415	1.895	2.365	2.998	3.459	4.028	4.785	5.408
8	.642	.706	1.397	1.860	2.306	2.895	3.352	3.859	4.501	5.141
9	.636	.703	1.381	1.843	2.282	2.821	3.250	3.650	4.297	4.781
10	.630	.700	1.372	1.832	2.268	2.784	3.169	3.581	4.194	4.587
11	.625	.697	1.363	1.796	2.201	2.718	3.106	3.497	4.075	4.437
12	.620	.695	1.356	1.782	2.179	2.681	3.055	3.426	3.930	4.318
13	.615	.694	1.350	1.771	2.160	2.650	3.012	3.372	3.851	4.221
14	.610	.692	1.345	1.761	2.145	2.626	2.977	3.326	3.787	4.140
15	.605	.691	1.341	1.751	2.131	2.602	2.947	3.286	3.733	4.071
16	.600	.690	1.337	1.746	2.120	2.583	2.921	3.252	3.686	4.015
17	.597	.689	1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18	.595	.688	1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19	.593	.688	1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20	.591	.687	1.325	1.725	2.086	2.528	2.845	3.151	3.552	3.850
21	.589	.686	1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22	.588	.685	1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23	.586	.685	1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.767
24	.585	.685	1.318	1.711	2.064	2.492	2.797	3.091	3.467	3.745
25	.584	.684	1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26	.583	.684	1.315	1.706	2.056	2.478	2.779	3.067	3.435	3.707
27	.582	.684	1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690
28	.581	.683	1.313	1.701	2.048	2.467	2.763	3.047	3.408	3.674
29	.580	.683	1.311	1.699	2.045	2.462	2.755	3.038	3.395	3.659
30	.579	.682	1.310	1.697	2.042	2.457	2.750	3.030	3.385	3.646
40	.565	.681	1.303	1.684	2.021	2.423	2.704	2.971	3.307	3.551
60	.554	.679	1.296	1.671	2.000	2.390	2.660	2.915	3.232	3.460
120	.544	.677	1.289	1.658	1.982	2.358	2.617	2.860	3.160	3.373
∞	.535	.675	1.282	1.645	1.960	2.325	2.575	2.807	3.090	3.291

Fidedignidade e validade

Antes que qualquer análise dos dados possa ser feita, é necessário que o pesquisador se pergunte: *a forma de obtenção dos dados me dá segurança sobre a sua validade?*

Em seções anteriores foi discutida a estrutura de um experimento analisando os vários tipos de delineamentos possíveis (destacando o que diferencia um delineamento do tipo experimental dos que não possuem esta propriedade), bem como os aspectos relacionados com as características de um bom teste. No entanto, de nada adianta ser cuidadoso na escolha do delineamento da pesquisa se os instrumentos de coleta de dados forem inadequados. Inadequação significa dizer que o instrumento não mede corretamente o que se propõe a medir (questão relacionada com a fidedignidade do instrumento) ou mede outra coisa diferente daquela que se propõe a medir (questão relacionada com a validade do instrumento). Para usar uma analogia comum quando se fala em fidedignidade e validade, pode-se imaginar a seguinte situação: um atirador de dardos acerta repetidamente o mesmo lugar do alvo. Neste caso, diz-se que há fidedignidade, pois em várias repetições o atirador acerta sempre no mesmo lugar ou, em outras palavras, consegue reproduzir o mesmo resultado (posição) com o mesmo instrumento. Porém se a posição acertada não for o centro do alvo, dizemos que não há validade, pois o objetivo do jogo é acertar o centro. Se a posição acertada for o centro, sempre, então dizemos que há fidedignidade e validade.

A ferramenta básica para a análise de fidedignidade é a correlação estatística entre variáveis. Vejamos como se calcula esta quantidade e qual é o seu significado.

O que é a correlação entre duas variáveis

Conceitualmente, a *correlação* ou o *coeficiente de correlação* (r) indica como se comportam duas ou mais variáveis umas em

relação às outras. Quando temos uma correlação alta, isto indica que o crescimento de uma variável é acompanhado pelo crescimento (no caso de um coeficiente de correlação perto de $+1$) ou pelo decréscimo (no caso de um coeficiente de correlação perto de -1) da outra variável. Convém lembrar aqui que o fato de duas variáveis serem correlacionadas (tanto positiva quanto negativamente), não implica uma relação causal entre as duas variáveis. Para o estabelecimento de uma relação causal entre elas deve-se recorrer a outras ferramentas de análise, ou seja, devemos procurar na teoria as razões dessa dependência e os fatores de comprovação dessa dependência.

O grau de correlação é indicado pelo valor do coeficiente de correlação. Para uma correlação perfeitamente positiva, tal como mostrado na *Figura 4.a*, em que o coeficiente de correlação r tem valor $+1$. O coeficiente para uma correlação perfeitamente negativa é mostrado na *Figura 4.b*, na qual o coeficiente de correlação r tem valor de -1 . Estes dois valores são os valores máximos para r . O coeficiente $r=0$ indica a inexistência de correlação. Neste caso, o comportamento de uma variável não é relacionado de qualquer modo ao comportamento da outra variável. Isso é mostrado na *Figura 4.c*.

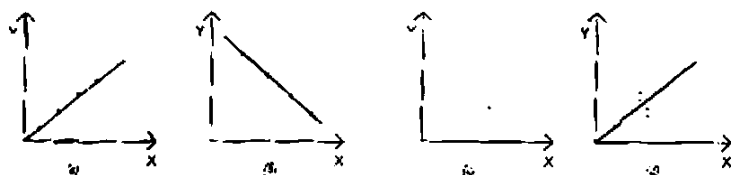


Figura 4 – Representação gráfica dos vários tipos de correlação possíveis entre variáveis.

A *Figura 4.d* nos mostra um caso em que a correlação existe, mas é imperfeita. Pela simples visão dessa figura, podemos ver que os pontos tendem a se alinhar ao longo de uma direção específica, embora não estejam todos sobre uma mesma reta. Isto nos indica que a correlação é positiva, embora não seja perfeita.

Tal como ocorre com os escores médios de amostras, os coeficientes de correlação, calculados a partir de dados amostrais são afetados por erros de amostragem. Assim, a exemplo do que foi feito com as médias amostrais, uma pergunta se impõe: *quão provável é que o coeficiente de correlação obtido a partir dos dados amostrais não seja fruto do erro de amostragem e reflita um verdadeiro relacionamento existente na população?* É razoável se tomar por hipótese de que, como em toda inferência feita a partir de um processo de amostragem, exista um erro devido ao próprio processo de amostragem. Entretanto, a exemplo do que acontece para outros tipos de testes estatísticos, existem tabelas para vários valores de tamanho de amostra, em qualquer nível de significância estatística desejado²².

Cálculo do coeficiente de correlação

A forma de cálculo do coeficiente de correlação é função do tipo de variável com a qual estamos lidando. Como já foi discutido, estas podem ser divididas em quatro grupos: nominais, ordinais, intervalares ou racionais. Para cada pareamento de variáveis, a forma de cálculo do coeficiente de correlação é diferente, devendo-se levar em conta os tipos de variáveis envolvidos. A fórmula de cálculo do coeficiente de correlação, definida a seguir, é válida somente quando as variáveis envolvidas são ambas dos tipos *intervalares* ou *racionais*. Para outros tipos de variáveis sugere-se a consulta a Glass e Stanley (1970).

O termo *correlação*, tal como está sendo usado aqui, significa *correlação linear*. Nesse caso, o coeficiente de correlação entre duas variáveis X e Y é dado por:

$$r_{xy} = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$$

em que $x = X - \bar{X}$ e $y = Y - \bar{Y}$ ²³.

²² Usualmente os níveis 0,01 e 0,05.

²³ Os valores com barra, seguindo a convenção anterior, são as médias em X e Y .

Esta equação é chamada de *fórmula produto-momento*¹³. Deve-se observar a simetria entre as variáveis que compõe essa equação: se trocarmos x com y , o resultado é o mesmo. Com isso, $r_{xy} = r_{yx}$, ou seja, tanto faz calcularmos o coeficiente de correlação chamando uma das variáveis de X e a outra de Y como calcularmos denominando inversamente as variáveis.

Validade e fidedignidade de testes

Testes de conhecimento tentam medir o que um indivíduo apreendeu em uma certa área, seu nível atual de domínio de um certo conteúdo. Muitos testes usados em escolas são testes de conhecimento. Com frequência, escores em testes de conhecimento são usados na avaliação de cursos, professores, métodos de ensino e outros fatores considerados significativos na prática educacional. São usados na classificação, promoção ou retenção de estudantes em determinados níveis de ensino. Eles são usados como ferramentas para diagnosticar pontos fracos e fortes dos estudantes e como uma base para prêmios, recompensas, etc., dentro do ambiente escolar.

Testes de aptidão tentam prever o grau de conhecimento que pode ser esperado de indivíduos em uma atividade particular. Estes testes procuram prever a capacidade de um indivíduo para adquirir um melhor desempenho a partir da instrução. Testes de aptidão podem ser usados para dividir os estudantes em grupos relativamente homogêneos com fins instrucionais.

Na pesquisa, testes são instrumentos de medida usados pelos pesquisadores para, de um modo geral, coletar dados. Deste modo, se a pesquisa for de qualidade, é de se supor que os procedimentos usados para a coleta de dados (e sua análise) devem ser bons. A qualidade da pesquisa não pode ser melhor do que a qualidade dos procedimentos e instrumentos usados para coletar e analisar os dados.

¹³ Observe-se que essa equação fornece automaticamente o sinal correto para o coeficiente de correlação.

Como já foi dito, esses instrumentos devem ter duas características indispensáveis:

1. **Fidedignidade:** um teste é fidedigno se suas medidas são acuradas e consistentes quando aplicado, em tempos diferentes, nas mesmas condições.
2. **Validade:** em geral, um teste possui validade se mede o que se propõe a medir.

Um teste pode ser fidedigno, mesmo se não for válido. Um teste válido sempre é fidedigno. Não existe uma única forma de validade ou fidedignidade de um instrumento. Existem muitos tipos de validade e fidedignidade. Em geral, a fidedignidade está relacionada com a consistência e a validade com a interpretação do teste. Um teste pode ser válido para um objetivo e não ser válido para outro.

Fidedignidade¹⁴

Reiterando, por fidedignidade entendemos o quão acurados são os dados no sentido de sua estabilidade, reproducibilidade ou precisão. Um instrumento de coleta de dados fidedignamente perfeito é um que se fosse aplicado duas vezes sob as mesmas circunstâncias nos forneceria os mesmos resultados. Como foi dito, a *correlação* é a ferramenta estatística básica na análise de fidedignidade de testes. Uma correlação 1,00 indicaria perfeita fidedignidade, enquanto a correlação 0,00 indicaria nenhuma fidedignidade. Correlações intermediárias indicariam diferentes níveis de fidedignidade.

São várias as formas de medir-se a fidedignidade de um teste:

1. **Teste-reteste** – é exatamente o que o nome implica; produz-se o primeiro conjunto de dados através da aplicação do teste em

¹⁴ Parte dessa seção é baseada em *The Research Process in Education*, de D.J. Fox, 1969.

um dado instante de tempo e, após decorrido algum intervalo de tempo, longo o suficiente para esquecer o teste, mas curto o suficiente para que não sejam produzidas alterações significativas nos respondentes, um segundo conjunto de dados é obtido pela aplicação do mesmo teste ao mesmo conjunto de respondentes (o *reteste*). Após a segunda aplicação, os dois conjuntos de dados são correlacionados e a correlação obtida estima a fidedignidade do teste.

O maior problema nesta estratégia envolve o intervalo de tempo. Como definir o que se entende por intervalo de tempo grande o suficiente para esquecer do teste, mas curto o suficiente de modo que não se produzam alterações nos respondentes? A resposta a esta questão depende do tipo de pesquisa que está sendo feita.

O procedimento de teste-reteste é mais aplicável àquelas situações em que a aptidão, habilidade ou conhecimento que está sendo avaliado possui estabilidade, não mudando no período entre o teste e o *reteste*. Esta, obviamente, é uma limitação de ordem prática, sendo, sob o ponto de vista técnico, muito difícil garantir a integridade dos respondentes entre o teste e o *reteste*. Outro aspecto perigoso da estratégia do teste-reteste é que estamos calculando a correlação entre dados provenientes de duas exposições dos sujeitos ao mesmo conteúdo.

2. *Forma alternada*: nesta forma de medição de fidedignidade o pesquisador deve desenvolver duas formas paralelas ou equivalentes de seu instrumento, digamos formas A e B, administrar ambas às mesmas pessoas e correlacionar os dois conjuntos de dados obtidos. A administração das duas formas pode ser com um intervalo de tempo entre as duas aplicações ou de uma única vez, com as duas formas sendo aplicadas consecutivamente. Em ambos os casos, é aconselhável alternar-se a ordem de aplicação entre as duas formas do teste, isto é, metade dos respondentes respondem à forma A do teste e depois

à forma B e a outra metade responde primeiro à forma B e depois à forma A. Esta maneira de aplicar o teste elimina efeitos que poderiam mascarar os resultados obtidos, tais como cansaço ou aborrecimento durante a aplicação da segunda forma do teste.

Este tipo de procedimento possui as seguintes dificuldades:

- em muitos casos, é difícil desenvolver uma forma do instrumento, o que dirá duas;
 - há situações em que é difícil, senão impossível, perguntar a mesma coisa duas vezes de forma equivalente ou similar;
 - se o procedimento envolver certo intervalo de tempo entre a aplicação do teste e de sua forma alternada, da mesma forma que ocorreu para a forma teste-reteste, pode surgir o problema de reagrupar o mesmo grupo de sujeitos para responder o teste pela segunda vez;
 - administrar consecutivamente duas formas do mesmo teste exige duas vezes mais tempo e requer a manutenção de um interesse suficiente por parte dos respondentes, uma vez que eles trabalharão, basicamente, sobre os mesmos conteúdos.
3. Método da metade ou par-ímpar: um terceiro procedimento, chamado *da metade* ou *par-ímpar*, soluciona cada um dos problemas apontados anteriormente, quando da discussão das outras formas de cálculo de fidedignidade. Ele envolve somente a administração de um instrumento em apenas uma única forma. O instrumento é aplicado uma única vez a um grupo de respondentes; entretanto, ele é analisado de forma a separar os escores de cada respondente em duas metades. Estes dois conjuntos de escores são então correlacionados. Desse modo, se um instrumento possui 100 itens, para obter-se a estimativa de fidedignidade na forma *par-ímpar*, obter-se-ia os escores de cada respondente nas questões ímpares e então um escore separado para as outras 50 questões pares. É possível, mas não o mais aconselhável, usar o procedimento de separação em *metades*, isto é, obter um escore para cada respondente baseado na

primeira metade do teste e outro baseado na segunda metade do teste.

A forma *par-ímpar* é a preferível devido às seguintes vantagens:

- normalmente, um instrumento de medida cobre várias áreas do conhecimento em diferentes seções as quais geralmente são estanques e bem diferenciadas;
- fatores tais como fadiga ou perda de interesse poderiam causar omissão por parte do respondente nas questões finais do teste.

No entanto, não importa qual seja o procedimento usado, estimativas de fidedignidade obtidas a partir do uso de metades de um instrumento também apresentam seus problemas, principalmente relacionados com o fato de que a fidedignidade é relacionada com o número de itens de um instrumento.

O seguinte procedimento foi desenvolvido de modo a calcular a fidedignidade de um instrumento como um todo a partir do cálculo da fidedignidade usando metades do teste. É a chamada fórmula de Spearman-Brown, que recebe esse nome em homenagem aos pesquisadores que, de forma independente, a desenvolveram:

$$\alpha_{sb} = \frac{2 \times \alpha}{1 + \alpha}$$

em que α_{sb} é a chamada *estimativa de fidedignidade de Spearman-Brown*, α é a correlação entre as duas metades do teste. O que esta fórmula nos dá é somente uma predição ou estimativa da fidedignidade que o pesquisador poderia esperar para o instrumento como um todo a partir dos valores de fidedignidade obtidos para cada metade do teste.

As maiores vantagens de usar-se correlação entre metades como uma estimativa de fidedignidade do teste total são de natureza prática: necessita-se de apenas um teste e somente uma sessão de testagem por parte dos respondentes. Suas desvantagens são as mesmas.

Expectativas para a fidedignidade de um instrumento diferirão dependendo da natureza da informação que está sendo procurada. Se o que se estiver procurando é uma informação tipo demográfica, como, por exemplo, lugar de nascimento, escolaridade e experiência profissional, poderíamos esperar uma alta fidedignidade de um instrumento. Em termos de correlações, isso implica em coeficientes de correlação da ordem ou superiores a 0,90. Por outro lado, se a procura for sobre conhecimento e habilidades, as quais são informações de caráter não tão fixo quanto o exemplo anterior, a expectativa de fidedignidade será menor e um coeficiente de fidedignidade de 0,85 é aceitável. Por fim, se o tipo de informação procurada for mais movediça, como, por exemplo, atitudes e interesses, um coeficiente da ordem de 0,70 será aceitável.

Uma pergunta que surge naturalmente nesse ponto é a seguinte: *como melhorar a fidedignidade de um teste?* Como comentado anteriormente, a fidedignidade de um teste está diretamente relacionada com a extensão do teste. Portanto, um procedimento óbvio de modo a melhorar a fidedignidade de um teste é alterar a sua extensão. *Mas em quanto a fidedignidade poderia ser melhorada com o acréscimo de uma ou mais questões ao teste?* A resposta a essa pergunta encontra-se embutida na fórmula de cálculo do coeficiente de Spearman-Brown, a qual fornece a fidedignidade máxima que poderia ser obtida pelo aumento no tamanho do teste. Essa estimativa é dada pela equação na sequência:

$$r_k = \frac{kr}{1 + (k-1)r}$$

em que k é o número de vezes que o novo teste é maior que o anterior, r é a fidedignidade apresentada pela forma atual do teste e r_k é a estimativa de fidedignidade da nova forma do teste.

Exemplo: suponhamos que a fidedignidade de um teste é estimada em 0,50. Este teste tem a sua extensão duplicada pelo acréscimo de itens. Qual será a estimativa da nova fidedignidade?

$$r_k = \frac{kr}{1 - (k-1)r}$$
$$r_1 = \frac{2 \times 0,5}{1 + (2-1) \times 0,5}$$
$$r_k = 0,67$$

Um algoritmo para o cálculo do coeficiente de fidedignidade de testes - Análise de Consistência Interna

Uma das aplicações do coeficiente de correlação é na Análise da Consistência Interna de testes. É muito comum na prática docente que o professor some escores de itens isolados de testes, compondo, assim, um escore bruto, o qual é usado para análises e inferências. No entanto, esse procedimento somente é aceitável quando todos os itens do teste se referem a um mesmo conjunto de conceitos e/ou habilidades. A análise de consistência interna de um teste tem por objetivo verificar o quanto de verdade existe nessa hipótese a respeito de determinado teste. A ideia geral é comparar o desempenho em cada item dos respondentes ao teste com o desempenho dos respondentes ao teste como um todo. Se o desempenho dos respondentes a um item não se correlaciona com seu desempenho como um todo, isto significa que aquele item em particular não está avaliando as mesmas características das demais questões que compõem o teste e, portanto, deve ser descartado (ou modificado). A Análise de Consistência Interna é parte indispensável do processo de pesquisa. O pesquisador não pode usar um teste sem verificar a sua consistência interna. Sem essa etapa, a soma de escores atribuídos a itens particulares não pode ser feita, e toda inferência obtida a partir desse escore total será sem significado.

A seguir, é apresentado um roteiro¹⁵ para a execução de

¹⁵ Este roteiro foi elaborado pelo professor Fernando Lang da Siveira, do Instituto de Física da UFRGS.

análise de consistência interna. Os dados utilizados no exemplo são retirados da Tabela 5:

Tabela 5 - Dados para o exemplo de análise de consistência interna.

ITEM → INDIVÍDUO ↓	1	2	3	4	5	6	7	8	Total
1.	5	5	5	5	5	5	5	2	37
2.	5	4	4	5	3	5	5	3	34
3.	5	5	4	5	3	4	5	3	34
4.	4	4	5	4	4	5	5	2	33
5.	5	4	5	5	2	4	4	3	32
6.	4	5	4	4	3	5	5	2	32
7.	4	4	5	5	5	5	3	1	32
8.	4	5	5	4	4	4	5	1	32
9.	3	4	5	4	3	5	5	2	32
10.	4	3	4	5	2	5	5	3	32
11.	5	5	5	4	1	4	3	4	32
12.	4	4	3	4	5	4	4	3	32
13.	4	4	4	3	3	3	4	4	29
14.	4	4	4	4	2	4	3	3	28
15.	3	4	4	4	3	4	4	2	28
16.	3	4	5	3	2	4	5	2	28
17.	4	3	3	5	4	3	2	4	28
18.	4	4	4	3	3	5	4	1	28
19.	3	3	4	4	4	4	3	2	27
20.	4	4	3	3	2	3	4	4	27
21.	4	3	3	3	4	3	4	3	27
22.	3	4	3	3	4	4	4	2	27
23.	3	4	3	3	1	3	4	5	26
24.	3	3	3	4	3	4	3	3	26
25.	2	3	3	3	5	4	3	1	24
26.	3	2	3	3	4	3	4	2	24
27.	2	2	3	2	5	3	3	5	24
28.	4	3	3	3	4	1	1	4	23
29.	2	2	2	1	4	3	3	4	21

30.	3	3	1	2	3	2	2	2	18
31.	1	1	2	1	3	1	2	4	25
F(1)	2	1	1	2	2	2	1	4	
F(2)	2	3	2	2	5	1	3	10	
F(3)	9	8	11	10	10	8	8	8	
F(5)	13	14	9	10	9	12	10	7	
F(6)	5	5	8	7	5	8	9	2	

1. Transformar a resposta de cada indivíduo a cada item em um escore.
2. Calcular o escore total de cada indivíduo somando os escores obtidos em cada item por aquele indivíduo.
3. Ordenar os indivíduos, em ordem decrescente, pelo escore total.
4. Construir a matriz dos resultados (ver a Tabela 5).
5. Determinar a frequência F (número de vezes em que o escore aparece) de cada escore em cada item (ver as últimas cinco linhas da Tabela 5).
6. Calcular a média do escore total. Isto é feito somando-se os escores totais de cada um dos sujeitos e dividindo-se pelo número de sujeitos (N). Assim, no exemplo, toma-se a soma da última coluna da tabela 5 e divide-se por 31 (número de sujeitos):

$$\bar{T} = \frac{37 + 34 + \dots + 15}{31} = \frac{868}{31} = 28,00$$

7. Calcular a variância V_t do escore total. No exemplo:

$$V_t = \frac{\sum T^2}{N} - (\bar{T})^2$$

$$V_t = \frac{24986}{31} - (28,00)^2$$

$$V_t = 22,00$$

8. Calcular a média e a variância de cada item. Por exemplo, para o item 1 tem-se:

$$I = (1 \times 2) + (2 \times 2) + (3 \times 9) + (4 \times 13) + (5 \times 5) = 110 \text{ (soma dos escores)}$$

$$I^2 = (1^2 \times 2) + (2^2 \times 2) + (3^2 \times 9) + (4^2 \times 13) + (5^2 \times 5) = 424 \text{ (soma dos quadrados dos escores)}$$

$$\bar{I} = \sum I / N = 110 / 31 = 3,548 \text{ (média dos escores atribuídos ao item)}$$

$$V_i = 424 / 31 - (3,548)^2 = 1,089 \text{ (variância no item).}$$

A Tabela 6 traria os valores da média e da variância para cada item do nosso exemplo:

Tabela 6 - Cálculos para o primeiro sujeito da Tabela 5.

ITEM	1	2	3	4	5	6	7	8
MÉDIA	3,548							
VARIÂNCIA	1,089	0,946						1,273
		6						3

9. Calcular o somatório da variância dos itens. Para os dados do exemplo (ver a Tabela 5):

$$V_i = 1,089 + 0,946 + \dots + 1,273 = 9,144$$

10. Calcular o coeficiente de fidedignidade (coeficiente α de Cronbach):

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum V_i}{V_r} \right)$$

$$\alpha = \frac{8}{7} \left(1 - \frac{9,144}{22,00} \right)$$

$$\alpha = 0,668$$

em que k é o número de itens do teste.

Como foi dito, o coeficiente de fidedignidade é um número entre -1 e 1. Quanto mais próximo de 1 (em módulo), melhor é o teste. No exemplo, o coeficiente de fidedignidade obtido está bastante longe de um valor aceitável. A fim de melhorá-lo, é preciso descobrir quais os itens do teste estão levando a esse

valor baixo. Para isso, deve-se calcular o índice de discriminação de cada item, descartando-se aqueles itens que possuírem baixo índice de discriminação. A lógica por trás desse procedimento é a seguinte: se uma questão do teste tem, aproximadamente, a mesma frequência de escores em dois grupos com escores totais diferentes (por exemplo, os 25% com escores superiores e os 25% com escores inferiores), então esse não é um bom item do teste. Essa é a próxima fase.

11. Dividir os sujeitos em três grupos (superior, médio e inferior) através do escore total. Os grupos superior e inferior devem perfazer, aproximadamente, 25 % dos indivíduos.

12. Calcular a média de cada item no grupo superior e no grupo inferior.

No exemplo, média do item 1 no grupo superior:

$$\bar{I}_s = (4 \times 5 + 4 \times 4) / 8 = 4,5$$

Média do item 1 no grupo inferior:

$$\bar{I}_i = (2 \times 3 + 4 + 2 \times 2 + 2 \times 1) / 7 = 2,286$$

A Tabela 7 traria os valores calculados para cada um dos itens do exemplo.

Tabela 7 - Valores médios para cada item relativos aos grupos superior G_s e inferior G_i para os dados do exemplo.

1. ITEM	1	2	3	4	5	6	7	8
2. MÉDIA G_s	4,5							
3. MÉDIA G_i	2,286							
	6							

13. Calcular o índice de discriminação de cada item. O índice de discriminação é definido como a diferença entre os valores médios dos escores do item nos grupos superior e inferior:

$$D_i = \bar{I}_s - \bar{I}_i$$

$$D_i = 4,5 - 2,286 = 2,214$$

no exemplo, para o item 1.

A Tabela 8 apresentaria os valores do índice de discriminação para todos os itens de nosso exemplo.

Tabela 8 - Valores do índice de discriminação para cada item do exemplo.

ITEM	1	2	3	4	5	6	7	8
D_i	2,21							
	4							

14. Calcular a seguir o coeficiente de correlação item-total. Esse coeficiente indicará quais são os itens que estão prejudicando a fidedignidade do teste e que devem ser eliminados. O coeficiente de correlação item-total é definido por:

$$r_{iT} = \frac{D_i}{K \times \sqrt{V_i}}$$

A constante K que aparece nesta equação depende da extensão dos grupos extremos; para os dados do exemplo, seu valor é de 2,57. Valores dessa constante para várias combinações de grupos encontram-se tabelados em livros de Estatística.

Para o item 1 do exemplo, o coeficiente de correlação item-total é dado por:

$$r_{iT} = 2,214 / [2,57 \times \sqrt{(1,089)}] = 0,826$$

Para os demais itens, organizar-se-ia uma tabela como a Tabela 9:

Tabela 9 - Valores do coeficiente de correlação item-total para os dados do exemplo.

ITEM	1	2	3	4	5	6	7	8
r_{iT}	0,82							
	6							

15. Verificar se existem itens com coeficiente de correlação item-total próximo de zero ou negativo. Se existirem, esses devem ser eliminados. No exemplo, os itens 5 e 8 devem ser eliminados.

16. Recalcular o escore total de cada indivíduo eliminando os

itens deficientes. Note-se que agora os itens a serem levados em conta são seis e não mais oito. Assim, por exemplo, o escore total do primeiro sujeito passará a ser 30, do segundo 28 e assim por diante.

17. Repetir as etapas de 6 a 9 com o objetivo de encontrar o novo coeficiente de fidedignidade. Para os dados do exemplo, após a eliminação dos itens 5 e 8 e da realização do novo cálculo, o coeficiente de fidedignidade passará a ser 0,904.

Na prática, esse procedimento é feito, em segundos, por um programa de computador apropriado. O exemplo detalhado, passo a passo, foi dado com a intenção de facilitar a compreensão do que significa "calcular o coeficiente de fidedignidade".

Validade¹⁶

Enquanto a fidedignidade é pré-requisito básico para qualquer instrumento de pesquisa, a validade é a característica mais importante que ele deve possuir. Isto porque a validade lida com o relacionamento entre os dados obtidos e o propósito para o qual eles foram coletados. Deste modo, validade é definida como o grau com que o procedimento realmente mede o que se propõe a medir.

A fidedignidade é uma etapa que precede a da validade, sendo essencial à mesma, e coloca um limite superior à validade do instrumento. Assim, um instrumento com fidedignidade 0,00 não pode ter qualquer validade; no outro extremo, um instrumento com validade 1,00 pode *possivelmente* ser perfeitamente válido. Para os graus intermediários de fidedignidade, a validade máxima pode ser estimada a partir da raiz quadrada do coeficiente de fidedignidade. Entretanto, enquanto a fidedignidade coloca limites na validade, ela não é garantia desta última. Por exemplo, pode acontecer que um instrumento com uma fidedignidade de 0,60 tenha uma validade

¹⁶ Adaptado de Fox, 1969

muito menor que $0.77 \sqrt{0.60} = 0.77$ e, de fato, possa ter nenhuma validade.

Portanto, fidedignidade é pré-requisito para que um instrumento seja válido, garantindo que ele mede de forma correta, mas isto não é garantia de que ele meça o que está se propondo a medir. A validade deve ser estimada separadamente, uma vez que a fidedignidade do instrumento já tenha sido estabelecida e que os valores encontrados sejam satisfatórios.

São vários os tipos de análise que podem ser feitas de modo a estabelecer a validade do instrumento. Contudo, não há como, na fidedignidade, uma forma matemática de estabelecer a validade de um instrumento. A seguir são apresentados os vários tipos de análises possíveis que levam ao estabelecimento da validade de um instrumento:

1. Validade de face: este tipo de validade é estabelecido a partir da análise superficial da natureza do instrumento, isto é, pela apresentação do instrumento. Obviamente, esse é o tipo de análise de validade mais fraco.
2. Validade de conteúdo: mais apropriada que a anterior, esta técnica de análise de validade é, muitas vezes, a forma mais adequada disponível ao pesquisador para analisar certos tipos de instrumentos, tais como questionários e entrevistas. Entretanto, é uma técnica que depende do julgamento de quem faz a análise de validade o que, sem dúvida, é seu ponto fraco. Esta técnica verifica se o instrumento está medindo o que se propõe a medir através da análise, feita por especialistas no conteúdo do instrumento, da existência de razões racionais para a escolha dos itens do instrumento ou de uma base, lógica ou empírica, para esta escolha.
3. Validade de construção: esta técnica é definida como a habilidade do instrumento de distinguir grupos para os quais se sabe previamente que se comportam de forma diferente na variável ou construto sob estudo. Em nível de procedimento,

determinar a validade de construção tem dois estágios. O primeiro consiste na definição de um *critério* para identificar os grupos que diferem no construto que o novo instrumento se propõe a medir. O segundo estágio consiste em administrar o instrumento a estes grupos e determinar se eles diferem significativamente no novo instrumento também. Se assim for, então tem-se argumento para defender a validade do novo instrumento.

4. **Validade congruente ou concorrente:** o pesquisador que oferece validade concorrente ou congruente apresenta dados de natureza correlacional, mostrando que o desempenho no seu novo instrumento correlaciona-se com o desempenho de algum instrumento de medir a variável já existente e aceito como válido. Se o critério que está sendo empregado pelo novo instrumento é da mesma família que o instrumento antigo, então se fala em validade congruente. Se, por outro lado, o critério empregado é de natureza diferente entre os dois instrumentos, fala-se em validade concorrente.

5. **Validade preditiva:** existe validade preditiva quando o pesquisador pode prever comportamentos dos respondentes, na área de interesse da pesquisa, a partir dos dados obtidos pelo instrumento. Este tipo de processo de validação implica que o pesquisador deva esperar algum tempo para saber se as predições feitas se concretizaram ou não e em que extensão. Esses dados podem ser apresentados de várias formas. Dentre elas, a correlação entre os escores preditos e o escore real, porcentagens de predições corretas, etc.

Análise da Variância

Como foi visto nas seções sobre Estatística, a tendência central (a média, por exemplo) não é suficiente para distinguir distribuições de frequência. A média aritmética de duas amostras, por exemplo, pode ser praticamente a mesma, mas a variabilidade dos escores pode ser muito diferente.

O teste F - análise da variância (anova)

Quando se quer determinar se os escores em uma amostra são mais variáveis que os escores em outra amostra, pode-se usar a técnica chamada de teste *F*. Usando o teste *F* podemos determinar se a variabilidade em um conjunto de dados é significativamente maior que a variabilidade em outro conjunto de dados. Ao conduzir um teste *F*, usa-se uma medida da variabilidade chamada de *variância*, ao invés do desvio padrão. A *variância*, em termos simples, é o quadrado do desvio padrão.

Para executar um teste *F* entre duas variâncias, simplesmente divide-se a maior variância pela menor. Isto dá o que é chamado de *razão F* entre as duas variâncias. A questão a responder é se a variância obtida de uma amostra difere significativamente da variância obtida em outra amostra. Neste caso, a hipótese nula é que não existe diferença entre a variabilidade dos escores em uma amostra em relação a variabilidade dos escores da outra amostra. Então é utilizada uma tabela de valores *F*, existente em livros de Estatística, para determinar se rejeita-se ou não a hipótese nula, no nível de significância escolhido.

O teste *F* também pode ser usado para analisar a variabilidade entre médias de escores de três ou mais amostras quando se pode assumir que as amostras foram obtidas através de seleção aleatória e a partir de uma população distribuída normalmente. O teste *F*, usado para comparação entre várias médias de escores, é chamado de **Análise da Variância (ANOVA)** e envolve a comparação de duas variâncias estimadas.

O que se pretende é comparar uma estimativa da variância da população obtida a partir dos escores dentro de cada amostra com uma estimativa da variância obtida dos escores médios das várias amostras. Uma das variâncias estimadas é obtida pela comparação da variância estimada para cada uma das amostras separadamente e então as combinando de modo a obter uma estimativa única chamada de estimativa de variância dentro dos grupos.

A outra variância estimada é calculada a partir dos escores médios para cada uma das amostras e calculando-se a variância estimada, usando-se estes escores médios e o tamanho da amostra no cálculo. Esta variância é chamada de *estimativa de variância entre os grupos*.

O objetivo é determinar se a estimativa de variância entre grupos é significativamente maior que a estimativa de variância dentro dos grupos. Se a estimativa de variância entre os grupos for significativamente maior que a estimativa de variância dentro dos grupos, poderemos rejeitar a hipótese nula e dizer que as amostras não provêm da mesma população.

De modo a aplicar a análise da variância, uma *razão F* é calculada entre as duas variâncias estimadas: usando a variância entre grupos estimada como numerador e a variância dentro dos grupos como denominador. A partir daí, usando-se uma tabela de valores *F*, podemos determinar, para qualquer tamanho das amostras, a *razão F* necessária para rejeitar a hipótese nula, no nível de significância especificado.

Exemplo: suponhamos que queiramos determinar se os níveis de iluminação afetam a produtividade no trabalho em uma empresa de produtos eletrônicos. Para estudar isto, seleciona-se, aleatoriamente, quatro amostras de 40 empregados cada e os distribui-se em diferentes níveis de iluminação. Mede-se então a produtividade de cada grupo e obtém-se os dados da *Tabela 10*.

É evidente que a produtividade média das quatro amostras é diferente. Mas é preciso saber se a variabilidade entre as médias amostrais (i.e., a diferença entre as médias) ocorreu como resultado de erro de amostragem ou se essa variabilidade pode ser atribuída à quantidade de iluminação.

Tabela 10 - Dados para o exemplo da produtividade.

Nível	Valor médio
I	40
II	38
III	27
IV	26

Neste exemplo, a hipótese nula a ser testada é a de que *não há diferença na produtividade dos empregados como resultado dos diferentes níveis de iluminação*. Para determinar isto, pode-se ou não desprezar a hipótese nula, examina-se esse conjunto de dados usando a análise da variância. A interpretação dos resultados da análise da variância é feita do mesmo modo que aquela do teste *t*. Suponhamos que a razão *F* no exemplo dado seja significativa ao nível 0,01. Poder-se-ia então rejeitar a hipótese nula e concluir que o nível de iluminação é relacionado à produtividade.

A inspeção das médias das amostras diz que a produtividade entre os níveis I e II difere somente por dois pontos. Da mesma forma, a diferença entre os níveis III e IV é somente de um ponto. A diferença maior aparece entre os níveis II e III. A análise da variância somente nos diz que existe uma diferença geral entre as quatro médias, porém não informa qual dos grupos é o responsável pela diferença significativa. Como será visto mais adiante, a análise de qual é a causa da diferença observada é feita através da Análise Fatorial da Variância. A técnica de análise de variâncias pode ser usada para a análise de diferenças entre qualquer número de amostras e é também aplicável para a análise de diferenças entre grupos dentro de amostras, tais como masculino-feminino ou agrupamentos por idade. No exemplo, as amostras poderiam ter sido divididas em grupos de empregadas mulheres e grupos de empregados homens e, além disso, seria possível ainda tê-los agrupado por idade. Naturalmente, neste caso seriam necessários muito mais empregados na amostra. Usando a análise de variâncias, poderiam ter sido examinadas diferenças em produtividade entre sexos, de acordo com a idade dos empregados e entre diferentes níveis de iluminação. Isto poderia ser chamado de análise de variâncias tri-lateral, uma vez que poderia ter sido analisada a produtividade como função de idade, sexo e quantidade de iluminação.

O método da análise da variância

De modo a descrever o método de cálculo da razão *F*, será utilizado outro exemplo. Suponhamos que um investigador educacional esteja interessado na eficácia relativa de dois métodos de ensino, denotados por *A*₁ e *A*₂. Após selecionar dez estudantes¹⁴, o pesquisador os divide em dois grupos, aleatoriamente, um experimental e outro de controle. Após algum tempo, ele mede a aprendizagem dos sujeitos de ambos os grupos, usando algum tipo de teste. Os resultados são mostrados na Tabela 11:

Tabela 11 - Dois conjuntos de dados experimentais para um exemplo hipotético.

<i>A</i> ₁	<i>x</i>	<i>x</i> ²	<i>A</i> ₂	<i>x</i>	<i>x</i> ²
4	0	0	3	0	0
5	1	1	1	-2	4
3	-1	1	5	2	4
2	-2	4	2	-1	1
6	2	4	4	1	1
Σx	20		15		
Σx^2		10			10
<i>N</i>	4		3		
					$\Sigma x_i = 3$
					5
					$M_i = 3,5$

O trabalho a ser feito com os dados anteriores é o de localizar e computar as diferentes variâncias que compõem a variância total. Para calcular a variância total usa-se a fórmula:

$$V_t = \frac{\sum x^2}{N - 1}$$

em que $\sum x^2$ é a soma de quadrados (sq), $x = X - \bar{X}$ é o desvio dos escores médios e *N* é o número de casos na amostra como um

¹⁴ Seguramente, nenhum pesquisador faria um estudo desse tipo com apenas cinco estudantes em cada grupo. É um exemplo fictício.

todo. Deste modo, reagrupando os dados da tabela, sem a preocupação quanto ao grupo a que o sujeito pertence, obtém-se para a variância total o valor de $V_t = 2,5$ ($22,5/10-1$) (ver a Tabela 12).

Tabela 12 - Cálculos de V_t para os dados do exemplo.

X	x	x^2
4	0,5	0,25
5	1,5	2,25
3	-0,5	0,25
2	-2,5	2,25
6	2,5	6,25
3	-0,5	0,25
1	-2,5	6,25
5	1,5	2,25
2	-1,5	2,25
4	0,5	0,25
ΣX	35	
M	3,5	
Σx^2		22,5
V_t		2,5

Observe-se que existe também uma variância entre os grupos, a qual é, presumivelmente, devida às manipulações experimentais. Isto é, o experimentador fez algo para um grupo e algo diferente para o outro. Estes tratamentos diferentes poderiam ocasionar diferenças entre os grupos, expressas pelas médias diferentes. Deste modo, se os grupos forem diferentes, haverá uma variância entre os grupos:

$$V_b = \frac{\sum x_b^2}{k-1}$$

em que k é o número de grupos e x_b é a diferença entre a média do grupo e a média entre os grupos. O cálculo da variância entre os grupos, V_b , é mostrado na Tabela 13.

Tabela 23 - Cálculos de V_b para os dados do exemplo 1.

	X	x	x^2
	4	0,5	0,25
	3	0,5	0,25
ΣX	7		
M	3,5		
Σx^2			0,50

$$V_b = 0,50$$

Existe ainda uma fonte remanescente de variância devida ao erro de amostragem. Leva-se em conta esse tipo de erro quando se calcula a variância dentro de cada grupo separadamente e então toma-se a média desses valores de modo a obter V_w . Para o exemplo:

$$\frac{\sum x_{A1}}{n_{A1} - 1} = 10 / 4 = 2,5 \quad \frac{\sum x_{A2}}{n_{A2} - 1} = 10 / 4 = 2,5$$

A média desses valores dá $V_w = 2,5$.

Foi dito antes que a variância total é composta de fontes separadas de variância: a variância entre grupos (V_b) e a variância dentro dos grupos (V_w). Logicamente, essas variâncias deveriam adicionar-se de modo a fornecer a variância total (V_t). A equação teórica seria então: $V_t = V_b + V_w$. Como se pode ver nos dados anteriores, não é isso que acontece. A razão é que os graus de liberdade é que foram usados nos cálculos feitos anteriormente, como denominadores, ao invés de N , n e k . O cálculo das variâncias usando-se N , n e k é matematicamente correto, porém estatisticamente incorreto.

Outro aspecto importante da análise é a estimativa dos valores para a população. Pode ser mostrado que utilizando-se os graus de liberdade nos denominadores, a fórmula da variância dará estimativas imparciais dos valores da população, um tópico de grande preocupação estatística. Deve-se, portanto, lembrar que variâncias, como as usadas na análise de variâncias, não são aditivas. Somas de quadrados, por outro lado, são sempre

aditivas. As somas de quadrados também são, naturalmente, uma medida da variabilidade. Exceto na fase final da análise, quando as somas de quadrados (sq) são calculadas, estudadas e analisadas.

Comparando a análise de variância com o teste t , pode-se dizer que a forma de encarar o problema é conceitualmente similar, com diferenças na metodologia. O método é geral, porém enquanto diferenças de mais que dois grupos podem ser testadas em termos de significância estatística usando a análise de variância, o teste t aplica-se somente a dois grupos. A análise de variâncias para dois grupos fornece os mesmos resultados que o teste t . A razão formada pela divisão da variância entre grupos (V_b) pela variância dentro dos grupos (V_w) é chamada de razão F :

$$F = \frac{V_b}{V_w}$$

Os valores F dos dados experimentais são calculados e comparados contra uma tabela de valores F . Se os valores obtidos são maiores que os valores tabelados, no nível escolhido de significância estatística e graus de liberdade, as diferenças expressas por V_b refletem diferenças significativas. Neste caso, a hipótese nula, de que não há diferença alguma entre as médias, é rejeitada naquele nível de significância.

Na prática, podemos calcular a razão F a partir dos escores brutos. No exemplo precedente, variâncias padrão foram usadas de modo a mostrar os aspectos fundamentais do método. Naturalmente, todo esse cálculo pode ser feito rapidamente com um aplicativo apropriado.

Análise fatorial da variância

Na análise fatorial da variância, duas ou mais variáveis variam independentemente ou interagem uma com a outra para produzir variações na variável dependente. A análise fatorial da variância é o método estatístico que analisa os efeitos independentes e interativos de duas ou mais variáveis independentes em uma variável dependente.

No passado, muitos pesquisadores (e isto ainda é válido para muitos pesquisadores nos dias de hoje) acreditavam que o método de pesquisa mais efetivo era permitir que uma variável independente evoluísse enquanto as demais variáveis independentes eram controladas, tanto quanto possível. A análise fatorial da variância veio mudar esse quadro, permitindo que se possa analisar o efeito de várias variáveis independentes ao mesmo tempo.

A análise fatorial serve para vários propósitos. Por exemplo, o delineamento fatorial e a análise fatorial da variância permitem ao pesquisador manipular e controlar duas ou mais variáveis simultaneamente¹⁵. Podemos não somente estudar os efeitos de determinado método de ensino sobre a aprendizagem, mas também analisar os efeitos, digamos, de tipos de reforço sobre as respostas. Além disso, podemos controlar variáveis tais como sexo, inteligência e classe social.

Variáveis que não são manipuladas podem ser controladas. Ao invés do procedimento disseminado de parear os sujeitos para testes sobre inteligência ou atitudes, pode-se construir essas variáveis (e muitas outras) em delineamentos com um caráter fatorial. Não somente controlar-se essas variáveis como também obtém-se informação adicional de grande valor e significância.

Interação estatística

Relativamente à interação estatística, a hipótese nula (H_0) é de que não há interação entre as variáveis independentes, ou seja, de que não existe influência da combinação de variáveis. Esta hipótese poderia ser chamada de *hipótese da diferença constante* ou *hipótese dos escores relativos constantes*. O que se quer dizer com isto é que as diferenças entre células de linhas diferentes permanecerão constantes de coluna para coluna;

¹⁵ Embora seja possível o controle de mais que três variáveis, estes delineamentos se tornam pouco práticos devido à dificuldade de obter-se um número suficiente de sujeitos de modo a ocupar todas as células. A forma mais simples de uma análise fatorial de variância é $2 \times 2 \times 2$.

equivalentemente poder-se-ia dizer que as diferenças entre células de colunas diferentes permanecerão constantes de linha para linha. Como um exemplo, considere-se a Tabela 14 uma hipotética tabela de médias, de 12 células.

Tabela 14 - Dados de um exemplo onde inexistente interação.

	Colunas			
	1	2	3	4
1	3	6	5	7
2	2	5	4	6
3	5	8	7	9
a				
s				

Esse é um exemplo típico de falta de interação. A diferença das médias em células na segunda linha em relação à primeira linha é constante e igual a 1, enquanto a diferença entre os valores na primeira linha e os da terceira também é constante e igual a 2. As diferenças entre linhas sendo constantes, então as diferenças entre colunas também o serão. A Figura 5 mostra graficamente o que significa não haver interação entre as variáveis.

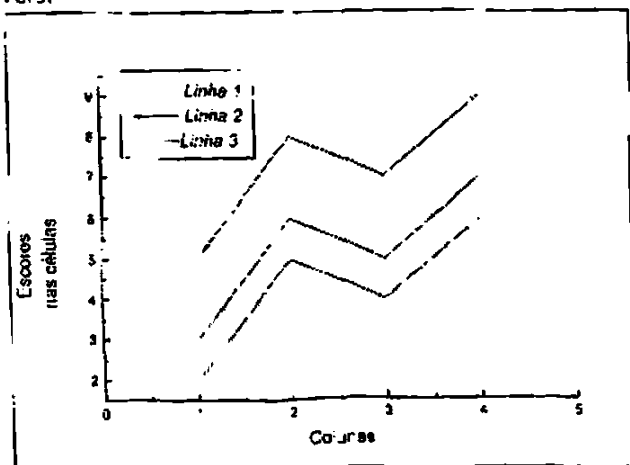


Figura 5 - Um exemplo onde não há interação.

Como já comentado, graficamente a não interação aparece como linhas paralelas em um gráfico em que se representa os valores médios das células em cada linha. No gráfico, marca-se para cada linha da tabela os valores médios, os quais são ligados por uma linha. Se não houver interação (*hipótese nula*), as linhas obtidas no gráfico para cada linha da tabela não se cruzarão, ou mais precisamente, serão paralelas. No caso de haver algum tipo de interação, então as linhas não mais serão paralelas, não havendo necessidade de que elas se cruzem. Um exemplo de interação (fictício) é mostrado na Figura 6, que é uma representação gráfica dos dados da Tabela 15:

Tabela 15 - Dados de um exemplo onde existe interação.

Linhas	Colunas				
		1	2	3	4
1		4	5	7	5
2		3	1	4	4

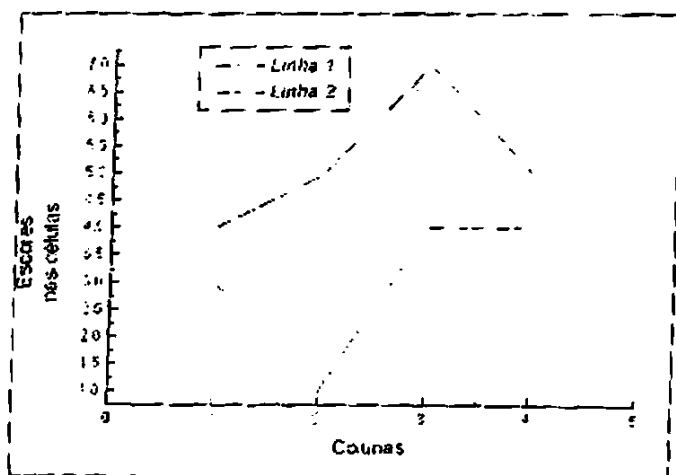


Figura 6 - Um exemplo onde existe interação.

Como um último comentário desta seção, convém discutir o que se entende por análise de variância unilateral e o que se

entende por análise de variância bilateral. Não é o número de variáveis que se está utilizando que define o tipo de análise de variância, é o número de classes de variáveis utilizadas. Assim, por exemplo, se em um determinado experimento analisa-se a influência de três métodos de ensino diferentes, faz-se uma análise de variâncias unilateral, uma vez que as variáveis analisadas (métodos de ensino) pertencem a uma única classe (métodos). Por outro lado, se além de métodos de ensino estiverem sendo analisados diferentes tipos de motivação, o caso será de uma análise de variâncias bilateral, pois, nesse caso, haverá duas classes: métodos de ensino e tipos de motivação. Esse exemplo será discutido na próxima seção de modo a discutir o método propriamente dito.

O método da análise fatorial da variância

De modo a estudar-se o método usado na análise fatorial da variância, é apresentado a seguir um exemplo hipotético¹⁹.

Um pesquisador está interessado no estudo da influência de dois fatores na aprendizagem. O primeiro desses fatores é o método de ensino, que pode ser um de dois métodos a serem estudados, e que serão denotados por A_1 e A_2 , e o segundo um de dois tipos de motivação, os quais serão denotados por B_1 e B_2 . Desse modo, os sujeitos são distribuídos em células para o estudo em que interagem um método de ensino e um tipo de motivação como, por exemplo, $A_1 B_1$ que nos indica que os sujeitos serão submetidos ao método de ensino A_1 e ao tipo de motivação B_1 . As possibilidades estão dispostas na Tabela 16. A amostra hipotética é composta por oito sujeitos, distribuídos na forma de dois sujeitos por célula.

¹⁹ Extraído de Kerlinger, 1964.

Tabela 16 - Delineamento fatorial para variáveis método de ensino e motivação.

Métodos			
Motivação		A ₁	A ₂
	B ₁	A ₁ B ₁	A ₂ B ₁
	B ₂	A ₁ B ₂	A ₂ B ₂

Testes estatísticos paramétricos e não paramétricos

Um tópico central na moderna teoria estatística é a *Estatística Inferencial*. A estatística inferencial está preocupada em resolver dois tipos de problemas: a estimativa dos parâmetros da população e testes de hipóteses. Na inferência estatística, a preocupação é como tirar conclusões sobre um grande número de eventos com base em observações de uma porção deles. A Estatística dá ferramentas com as quais se formaliza e são padronizados procedimentos para tomarmos decisões.

Um problema comum em estatística inferencial é o de determinar, em termos de probabilidades, se as diferenças observadas entre duas amostras significam que as populações a partir das quais as amostras são retiradas são realmente diferentes. As diferenças podem ocorrer apenas devido ao acaso durante o processo de amostragem.

No desenvolvimento dos métodos estatísticos, as primeiras técnicas estatísticas de inferência que apareceram foram aquelas que faziam muitas hipóteses sobre a natureza da população da qual os escores eram retirados. Uma vez que os valores da população são parâmetros, essas técnicas estatísticas foram chamadas de *paramétricas*²⁰. Por exemplo, a técnica de inferência pode ser baseada na hipótese de que os escores foram retirados de uma população cujos escores seguem a distribuição normal. Ou a técnica pode ser baseada na hipótese de que os conjuntos

²⁰ A média, o desvio padrão e a variância de uma população, ou qualquer outra medida da população, são parâmetros.

de escores foram retirados de populações tendo a mesma variância ou o mesmo espalhamento dos escores.

Mais tarde surgiu um grande número de técnicas estatísticas de inferência que não fazem hipóteses por demais numerosas ou restritivas sobre os parâmetros da população. Essas técnicas, as quais são *independentes de distribuição* ou *não paramétricas*, levam a conclusões que apresentam poucas limitações. Algumas técnicas não paramétricas são chamadas de *testes de ordenamento*. Este nome tem origem no fato de que enquanto as técnicas paramétricas têm o seu foco central nas diferenças de médias e variâncias, as técnicas não paramétricas têm seu foco no *ordenamento* dos escores e não em seus valores numéricos.

Enquanto um parâmetro é um valor de uma determinada população, uma estatística por outro lado é uma medida calculada de uma amostra. *Um teste estatístico não paramétrico é um teste cujo modelo não especifica condições sobre os parâmetros da população da qual as amostras foram retiradas.*

Quando se faz alguma afirmativa a respeito da natureza da população e sobre o processo de amostragem, se está estabelecendo um *modelo estatístico*. Associado com todo teste estatístico existe um modelo e uma prescrição de medida; o teste estatístico em questão é válido sobre certas condições, e o modelo e a prescrição de medida especificam estas condições. Algumas vezes, é possível testar se as condições de um modelo estatístico particular se encontram presentes, mas, muitas vezes, é necessário tomar por hipótese a presença dessas condições. Desse modo, as condições de validade de um teste, o modelo estatístico (ou seja, as hipóteses feitas quando da construção do teste), são muitas vezes chamadas de *hipóteses do teste*.

É óbvio que quanto menos ou mais fracas forem as hipóteses por trás do modelo subjacente a um teste estatístico, menos restrições é preciso fazer sobre as conclusões obtidas pelo teste estatístico associado ao modelo. Ou seja, quanto menor ou menos restritivas forem as hipóteses por trás do modelo, mais

gerais as conclusões e/ou resultados obtidos.

Entretanto, os testes mais poderosos são justamente aqueles que têm as hipóteses mais fortes ou condições mais restritivas. Os testes paramétricos, por exemplo, o teste t , ou o teste F , tem uma variedade de hipóteses fortes embasando o seu uso. Quando aquelas hipóteses são válidas, esses testes são os que mais provavelmente rejeitarão a hipótese H_0 quando esta hipótese for falsa.

As condições que devem ser satisfeitas para fazer o teste t , o teste mais poderoso e, de fato, necessárias para que se tenha confiança em qualquer inferência feita a partir de resultados obtidos com o teste, são:

1. As observações devem ser independentes – com isto se quer dizer que a seleção de qualquer caso da população para inclusão na amostra não deve influenciar a probabilidade de inclusão de qualquer outro caso. Da mesma forma, o escore atribuído a um caso não deve influenciar o escore atribuído a outro caso.
2. As observações devem ser retiradas de uma população normal – Como visto anteriormente, uma das hipóteses por trás do teste t é que a população segue a distribuição normal.
3. As variâncias das populações de onde as amostras são retiradas devem ser iguais.
4. As escalas devem ser intervalares ou racionais – isso é necessário para que se possa executar operações de caráter aritmético sobre os números representativos dos escores.
5. Aditividade (condição para validade do teste F) – as médias das distribuições normais devem ser combinações lineares dos efeitos devidos a colunas e/ou linhas. Isto é, os efeitos devem ser aditivos.

Deste modo, um teste estatístico paramétrico é um teste

cujo modelo especifica certas condições sobre os parâmetros da população da qual a amostra foi retirada. Uma vez que essas condições normalmente não são testadas, elas são assumidas como verdadeiras. A significância e validade de um teste paramétrico dependem do grau de acerto ao supor-se corretas estas hipóteses.

Um teste estatístico não paramétrico é um teste cujo modelo não especifica condições sobre os parâmetros da população da qual as amostras foram retiradas.

Vários critérios poderiam ser considerados na escolha de um teste estatístico para uso quando um processo de tomada de decisão sobre a hipótese de pesquisa está em curso. Estes critérios são:

1. O poder do teste.

Poder = $1 - \text{probabilidade de um erro tipo II}$.

Isto é, a probabilidade de aceitar-se H_0 quando de fato ela é falsa.

2. A aplicabilidade do modelo estatístico no qual o teste é baseado aos dados da pesquisa.

3. Poder-eficiência – é relativo à quantidade de aumento no tamanho da amostra a qual é necessária para tornar um teste B tão poderoso quanto um teste A . O poder-eficiência do teste B em relação ao teste A é definido por:

$$P_R = 100 \times \frac{N_A}{N_B}$$

em que N_A e N_B são, respectivamente, os tamanhos das amostras submetidas aos testes A e B .

4. O tipo da medida usada na pesquisa.

Vantagens de testes estatísticos não-paramétricos

1. Afirmções probabilísticas obtidas a partir de testes estatísticos não paramétricos são probabilidades exatas

2. Se os tamanhos das amostras são pequenos ($N \leq 6$), não existe alternativa ao uso de testes estatísticos não paramétricos, a menos que se conheça exatamente as características da distribuição seguida pela população.
3. Existem testes estatísticos não paramétricos apropriados para tratamento de amostras as quais provêm de várias populações diferentes. Nenhum dos testes paramétricos pode manipular dados deste tipo sem exigir que se acredite em hipóteses irrealísticas.
4. Há testes estatísticos não paramétricos para tratar dados os quais são inerentemente *ordinais*, isto é, o pesquisador é apenas capaz de dizer que um sujeito apresenta mais ou menos de determinada característica, mas não sabe dizer *o quanto* a mais ou a menos em relação a outro sujeito.
5. Os testes não paramétricos são capazes de tratar dados os quais são simplesmente classificatórios, seguindo uma escala tipo *nominal*. Nenhum dos testes paramétricos é capaz de lidar com dados com essas características.
6. Por fim, os testes não paramétricos são mais fáceis de aprender e aplicar que os testes paramétricos.

Entretanto, se todas as hipóteses do modelo estatístico paramétrico são de fato encontradas nos dados e se as medidas são do tipo exigido pelo teste, então testes estatísticos não paramétricos são desnecessários. Uma medida da necessidade do uso de um teste não-paramétrico em uma situação desse tipo é dada pelo *poder-eficiência* do teste não-paramétrico. Suponhamos que o poder-eficiência do teste não paramétrico seja de 90%. Isto significa que um teste não paramétrico na mesma situação exigiria somente 10 % da amostra para ser tão efetivo quanto o teste não paramétrico.

Testes não paramétricos - o caso de uma amostra

Esta situação acontece quando, aleatoriamente, retira-se uma amostra de determinada população e testa-se a hipótese de

que essa amostra vem de uma população com uma distribuição especificada.

Uma técnica paramétrica comum no caso de uma amostra é usar-se o teste t para a diferença entre as médias observada (a da amostra) e esperada (a da população). Entretanto, existem muitos tipos de dados para os quais o teste t é inaplicável. Esses fatores de inaplicabilidade do teste t podem ter várias fontes, dentre as quais cabe salientar:

1. as hipóteses e exigências para a aplicabilidade do teste não são satisfeitas para os dados do experimento;
2. pode ser preferível frente à determinada situação, evitar-se as hipóteses sobre as quais o teste t é construído e ganhar-se com isso em generalidade das afirmações resultantes da pesquisa;
3. os dados da pesquisa são, inerentemente, dos tipos nominal ou ordinal e, portanto, não são passíveis de análise pelo teste t ;
4. o pesquisador não está de fato interessado somente em diferenças na localização, mas, sim, em qualquer tipo de diferença não importando a sua origem;

Neste caso, o experimentador pode escolher usar um dos seguintes testes estatísticos não paramétricos:

1. O teste binomial: a distribuição binomial é a distribuição de amostragem de proporções que podemos observar em amostras retiradas de uma população composta por duas classes. Isto é, ela fornece os vários valores os quais podem ocorrer sob H_0 . Portanto, quando os escores da pesquisa estão em duas classes, a distribuição binomial pode ser utilizada para testar H_0 .
2. O teste χ^2 para uma amostra: este teste é utilizável quando queremos analisar dados os quais caem em mais que duas categorias. Este teste é útil para decidir se a diferença observada entre o número de objetos que recaem em cada categoria e o número esperado com base na hipótese nula é significativa.

3. O teste de Kolmogorov-Smirnov para uma amostra: este teste está preocupado com o grau de concordância entre a distribuição de um conjunto de escores observados e alguma previsão teórica específica. Ele determina se os escores na amostra podem, razoavelmente, serem pensados como vindo de uma população que obedece à distribuição predita teoricamente.
4. Teste do período para uma amostra: neste caso se está interessado em saber se os dados sob análise provêm de uma distribuição aleatória. Ele é baseado na sequência em que os escores aparecem originalmente, isto é, o número de períodos que a amostra apresenta.

Testes não paramétricos - o caso de duas amostras relacionadas

Testes estatísticos de duas amostras são usados quando o pesquisador deseja estabelecer se dois tratamentos são diferentes ou se um tratamento é "melhor" do que outro. Neste tipo de comparação, algumas vezes diferenças significativas são observadas, as quais não resultam do tratamento. Uma maneira de transpor-se a dificuldade imposta por diferenças estranhas entre os grupos é usar duas amostras relacionadas no experimento. Isto é, podemos parear ou relacionar de alguma outra maneira as duas amostras estudadas. Este pareamento pode ser atingido usando cada sujeito como seu próprio controle, ou pelo pareamento de sujeitos e então distribuindo os membros do par aleatoriamente aos grupos de controle e experimental. Quando o sujeito serve como seu próprio controle, ele é exposto a ambos os tratamentos em tempos diferentes. Quando o método do pareamento é utilizado, o esforço é para selecionar para cada par de sujeitos indivíduos que sejam tão iguais quanto possível na(s) variável(eis) estranha(s) a(s) qual(is) pode(m) influenciar o experimento.

Quando possível, o método de usar cada sujeito como seu próprio controle é preferível ao método do pareamento. A razão

para isto é que é difícil parear pessoas por desconhecimento de variáveis relevantes que determinam o comportamento. O delineamento de pareamento é somente uma boa opção na medida em que o pesquisador seja hábil em parear os sujeitos, sendo esta habilidade, frequentemente, bastante limitada.

A técnica paramétrica mais usada para comparações entre duas amostras correlacionadas é o teste t. Uma diferença de escores pode ser obtida de dois escores provenientes de cada um dos membros dos sujeitos pareados ou dos escores de cada sujeito sob as duas condições. *O teste t assume que estas diferenças nos escores são normalmente distribuídas em uma população da qual a população foi retirada.*

Em inúmeros casos, o teste t é inaplicável. Nestas situações o pesquisador pode escolher um dos seguintes testes estatísticos não paramétricos:

1. O teste de McNemar para a significância de variações: este teste é particularmente aplicável para aqueles delineamentos de "antes" e "depois", nos quais cada pessoa é usada como seu próprio controle e em medidas nas quais os escores são obtidos usando-se variáveis nominais ou ordinais. Deste modo, por exemplo, o teste pode ser usado para testar a efetividade de um tratamento particular (encontro, editorial de um jornal, propaganda via correio, visita pessoal, etc.), nas preferências de votos entre vários candidatos. Ou ele pode ser usado para averiguar a influência das migrações do campo para a cidade na preferência política, e assim por diante.
2. O teste do Sinal: o teste do sinal recebe o seu nome do fato que ele usa como dados os sinais de mais e menos, mais do que medidas quantitativas. Este teste é particularmente útil para a pesquisa na qual medidas quantitativas são impossíveis ou impraticáveis, mas com a possibilidade de ordenar com respeito um ao outro os membros de pares.

3. O teste de Walsh: se o experimentador puder assumir que a diferença nos escores que ele observa, em duas amostras relacionadas, provém de populações que obedecem a distribuições simétricas, ele pode usar um teste bastante poderoso desenvolvido por Walsh. Note que a hipótese não é que as diferenças nos escores são de populações normais (para as quais pode-se usar o teste paramétrico t); note-se também que as diferenças também não provêm da mesma população. O que o teste faz é assumir que as populações são simétricas, de modo que a média é uma adequada medida de tendência central e é igual à mediana.

Testes não paramétricos - o caso de duas amostras não relacionadas

Quando o uso de duas amostras relacionadas é impraticável ou inapropriado, pode-se fazer uso de duas amostras independentes. Neste delineamento, as duas amostras podem ser obtidas por um dos dois métodos:

1. as amostras podem ser retiradas aleatoriamente de duas populações;
2. as amostras podem originar-se devido à atribuição, aleatória, de dois tratamentos a membros, cujas origens são arbitrárias, de uma mesma amostra.

Nos dois casos não é necessário que as duas amostras tenham o mesmo tamanho. A técnica paramétrica normal de averiguar a diferença entre duas amostras não relacionadas é também aplicar-se o teste t nas médias das amostras. No caso da não aplicabilidade do teste paramétrico (por exemplo, não ter certeza que as distribuições são normais), o pesquisador pode escolher para analisar os seus dados um dentre vários testes não paramétricos, os quais são apresentados a seguir.

Teste da probabilidade exata de Fisher

Esta é uma técnica extremamente útil para analisar dados

discretos quando as duas amostras independentes têm tamanhos pequenos. Este teste é usado quando os escores das duas amostras independentes, escolhidas aleatoriamente, caem em uma de duas categorias mutuamente exclusivas. Em outras palavras, todo sujeito em ambos os grupos obtém um de dois escores possíveis. Os escores são representados por frequências em uma tabela de contingência 2×2 , como a Tabela 17.

Tabela 17 - Tabela de contingência para o teste de Fischer.

	-	+	
Grupo I	A	B	A+B
Grupo II	C	D	C+D
Total	A+C	B+D	N

Para os dados nessa tabela (em que A, B, C e D são frequências), poderia ser determinado se o Grupo I e o Grupo II diferem significativamente na proporção de mais ou menos atribuídos a eles. A probabilidade exata de se observar um particular conjunto de frequências em uma tabela 2×2 , quando os totais marginais são mantidos fixos, é dada pela *distribuição hipergeométrica*:

$$p = \frac{\binom{A+C}{A} \binom{B+D}{B}}{\binom{N}{A+B}} = \frac{(A+C)! (B+D)!}{A! C! B! D!} \times \frac{(A+B)! (C+D)!}{N!}$$

Deste modo:

$$p = \frac{(A+B)! (C+D)! (A+C)! (B+D)!}{N! A! B! C! D!}$$

Ou seja, a probabilidade exata da ocorrência observada é obtida tomando-se a razão entre o produto de fatoriais de quatro totais marginais e o produto dos fatoriais dos escores

encontrados em cada célula por fatorial de N , o número total de observações independentes²¹.

Exemplo: Considere-se os dados mostrados na Tabela 18:

Tabela 18 - Dados do Exemplo.			
	-	+	
Grupo I	10	0	10
Grupo II	4	5	9
Total	14	5	19

Nesta tabela, $A=10$, $B=0$, $C=4$ e $D=5$. Os totais marginais são $A+B=10$, $C+D=9$, $A+C=14$ e $B+D=5$. N , o número total de observações independentes é 19. A probabilidade exata de que estes 19 casos recaíssem nas células, do modo como o que foi observado, é dada por:

$$p = \frac{10!9!14!5!}{19!10!0!4!5!} = 0,0108$$

Portanto determinamos que a probabilidade de obter-se essa distribuição dos escores, sob H_0 , é $p=0,0108$.

O teste χ^2 para duas amostras independentes

Quando os dados consistem de categorias discretas, esse teste pode ser usado para determinar a significância estatística de diferenças entre dois grupos independentes. As medidas envolvidas podem ser inclusive as de uma escala nominal. Por exemplo, deseja-se saber se dois grupos de professores, de Física e de Química, diferem quanto a sua opinião em relação a uma certa estratégia de ensino. Pode-se medir essa opinião com um simples "a favor" ou "contra", calcular as frequências e aplicar o teste χ^2 .

²¹ O fatorial de um número a , denotado por $a!$ (lê-se a fatorial), é obtido pelo produto dos inteiros, começando em 1, até o número a . Assim, por exemplo: $3! = 1 \times 2 \times 3 = 6$. O 0! é igual a 1, por definição.

O teste da Mediana

Este é um procedimento para testar se duas amostras independentes diferem em tendência central. Ele fornecerá informação de quão provável é que as duas amostras independentes (não necessariamente de mesmo tamanho) tenham sido retiradas de populações com a mesma mediana.

O teste U de Mann-Whitney

Quando se dispõe de pelo menos uma medição ordinal, esse teste pode ser usado para verificar se duas amostras independentes foram tiradas da mesma população. É um dos mais potentes testes não paramétricos e é uma das melhores alternativas ao teste paramétrico t quando o pesquisador quer evitar suposições subjacentes ao teste t ou quando a medição feita é mais fraca do que uma escala intervalar (Siegel, 1956, p. 116).

O teste de duas amostras de Kolmogorov-Smirnov

Este é também um teste para verificar se duas amostras independentes foram retiradas da mesma população. A forma bilateral do teste é sensível somente a qualquer tipo de diferença na distribuição da qual as duas amostras foram retiradas: diferenças em localização (tendência central), em dispersão, em simetria, etc. O teste unilateral é utilizado para testar se os escores de um grupo experimental serão "melhores" que aqueles do grupo de controle.

O teste de Wald-Wolfowitz

Se quisermos testar a hipótese nula de que duas amostras independentes foram selecionadas da mesma população contra a hipótese alternativa de que os dois grupos diferem completamente, poderemos utilizar este teste. Isto é, com amostras suficientemente grandes, este teste pode rejeitar H_0 se as duas populações diferirem em qualquer aspecto: tendência central, variabilidade, simetria ou algum outro fator. Deste modo,

ele pode ser usado em uma larga classe de hipóteses alternativas. Enquanto muitos testes são endereçados a tipos específicos de diferenças entre dois grupos, o teste de Wald-Wolfowitz analisa qualquer tipo de diferença.

O teste de Moses de reações extremas

Em ciências do comportamento, algumas vezes se espera que uma condição experimental cause em alguns sujeitos o aparecimento de comportamentos extremos em uma determinada direção, enquanto que em outros sujeitos o comportamento será extremado na direção oposta. Desse modo, pode-se pensar que depressão econômica e instabilidade política provocarão em algumas pessoas reações extremamente conservadoras, enquanto outras reagirão de uma forma extremamente progressista em temas de opiniões políticas.

O teste de Moses é especificamente projetado para uso com dados coletados para testar esse tipo de hipóteses. Ele poderia ser usado quando é esperado que a condição experimental afete alguns sujeitos de um modo e outros de maneira oposta.

O teste da aleatoriedade para duas amostras independentes

Esta é uma técnica não paramétrica poderosa e útil para a testagem da significância da diferença entre as médias de duas amostras independentes quando N_1 e N_2 são pequenos. Com o teste da aleatoriedade podemos determinar a probabilidade exata, sob H_0 , associada com as nossas observações e podemos fazê-lo sem assumir a distribuição normal ou homogeneidade da variância nas populações envolvidas (as quais devem ser assumidas se o teste paramétrico equivalente, o teste t, for usado).

Discussão

Todos os testes não paramétricos para duas amostras independentes testam se é provável que as duas venham da mesma população. Entretanto, os vários testes apresentados são

mais ou menos sensíveis aos diferentes tipos de diferenças entre as duas amostras. Por exemplo, quando se quer testar se duas amostras representam populações que diferem em localização (tendência central), existem testes que são mais sensíveis a este tipo de diferença e, portanto, poderiam ser escolhidos: o teste da mediana, o teste de Fisher (para N pequeno), o teste U de Mann-Whitney, o teste de Kolmogorov-Smirnov (para duas amostras; unilateral) e o teste da aleatoriedade. Por outro lado, se fosse desejo do pesquisador determinar se suas duas amostras provêm de populações que diferem em qualquer aspecto em geral, i.e., localização ou dispersão ou simetria, etc., ele poderia escolher um dos seguintes testes: o χ^2 , o teste de Kolmogorov-Smirnov (bilateral) ou o teste de Wald-Wolfowitz. A técnica restante, o teste de Moses, serve unicamente para testar se um grupo experimental está exibindo reações extremas, sejam extremistas ou defensivas, em comparação com as reações exibidas por um grupo de controle independente.

A escolha entre os testes que são sensíveis às diferenças em localização são determinadas pelo tipo de medida obtida na pesquisa e pelo tamanho das amostras. O teste mais poderoso em termos de localização é o da aleatoriedade. Entretanto, esse teste pode ser somente usado quando os tamanhos das amostras forem pequenos e quando tivermos confiança na natureza numérica da medida obtida. Com grandes amostras ou medidas fracas (medidas ordinais), a alternativa sugerida é o teste U de Mann-Whitney o qual, quase sempre, é mais poderoso que o teste da aleatoriedade. Se as amostras são muito pequenas, o teste de Kolmogorov-Smirnov é levemente mais eficiente que o teste U . Se a medida for de um tipo tal que é significativo, somente dicotomizar as observações acima ou abaixo da mediana, então o teste da mediana é aplicável. Este teste não é tão poderoso quanto o teste U de Mann-Whitney em termos de proteção contra diferenças de localização, mas é mais apropriado

que aquele quando os dados das observações não podem ser completamente hierarquizados. Se o tamanho das amostras sob consideração for muito pequeno, quando aplicando o teste da mediana, o pesquisador poderia fazer uso do teste de Fisher.

A escolha entre os testes que são sensíveis a todas as diferenças é decidida a partir da intensidade das medidas obtidas, o tamanho das amostras e o poder relativo dos testes disponíveis. O teste χ^2 é apropriado para os dados que são medidos em escalas nominais ou mais fortes. Quando N é pequeno e os dados estão em uma tabela de contingência 2×2 , poderíamos usar o teste de Fisher ao invés do teste χ^2 . Em muitos casos, o teste χ^2 pode não fazer uso eficiente de toda a informação contida nos dados. Se os escores das populações forem continuamente distribuídos, nós podemos escolher ou o teste de Kolmogorov-Smirnov (bilateral) ou o teste de Wald-Wolfwitz ao invés do teste χ^2 . De todos os testes para qualquer tipo de diferença, o teste de Kolmogorov-Smirnov é o mais poderoso. Se ele for usado com dados os quais não assumem a hipótese de continuidade, ele ainda é aceitável mas opera mais conservadoramente, isto é, os valores de p obtidos serão levemente maiores. Se a hipótese nula for rejeitada a partir desses dados, pode-se seguramente ter confiança na decisão. O teste de Wald-Wolfwitz também protege contra todos os tipos de diferenças, mas não é tão poderoso quanto o anterior.

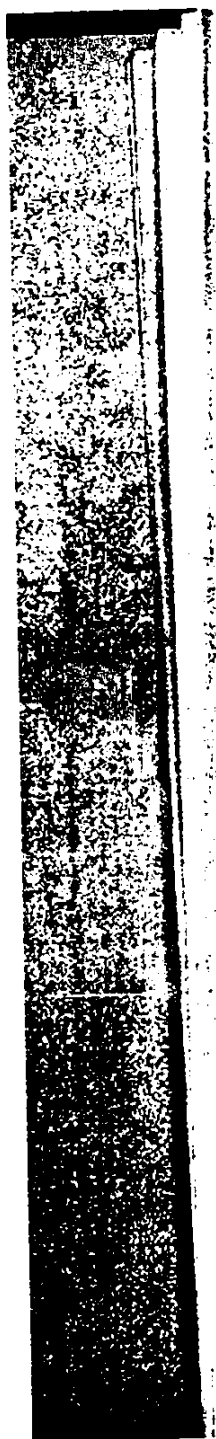
Dois pontos devem ser enfatizados sobre o uso de testes do segundo grupo. Primeiro, quando se está interessado em testar hipóteses alternativas de que os grupos diferem em tendência central, isto é, de que um grupo tem média maior que o outro, então se deve usar um teste especificamente projetado para capturar diferenças em localização, um dos testes do primeiro grupo listado anteriormente. Segundo, quando se rejeita a hipótese nula com base em um teste que capta qualquer tipo de diferença (um dos testes do segundo grupo), deve-se assegurar que os dois grupos provêm de populações diferentes, porém não se pode dizer em que elas diferem.

Bibliografia

- Best, J. W. (1970). *Research in education*. 2. ed. Englewood Cliffs: Prentice Hall.
- Campbell, D.R.; Stanley, J.C. (1963). Experimental and quasi-experimental designs for research on teaching. In Gage, N. L. *Handbook of research on teaching*, Chap. 5. Chicago: Rand Mc Nally.
- Campbell, D.R.; Stanley, J.C. (1991). *Diseños experimentales y cuasiexperimentales en la investigación social*. Buenos Aires: Amorrortu Editores.
- Darlington, R. B. (1975). *Radicals and squares*. Ithaca, N. Y.: Logan Hill Press.
- Esey, F. F. (1967). *A first reader in statistics*. Belmont, CA: Brooks/Cole Publishing Co.
- Fox, D. J. (1969). *The research process in education*. New York: Holt, Rinehart and Winston.
- Glass, G.V.; Stanley, J.C. (1970). *Statistical methods in education*. Englewood Cliffs, N. J.: Prentice Hall.
- Gowin, D.B. (1981). *Educating*. Ithaca, N.Y.: Cornell University Press.
- Gowin, D.B.; Alvarez, M. (2005). *The art of educating with V diagrams*. New York: Cambridge University Press.
- Kerlinger, F.N. (1964). *Foundations of behavioral research*. New York: Holt, Rinehart and Winston.
- Kerlinger, F.N. (1980). *Metodologia da pesquisa em ciências sociais*. São Paulo: E.P.U., EDUSP, INEP.
- Millman, J. (1970). *Data analysis*. Conferência convidada proferida no Simpósio Nacional de Professores de Pesquisa Educacional, St. Louis, USA.
- Moreira, M.A. (2006). *Mapas conceituais e diagramas V*. Porto Alegre: Editora do Autor.
- Runkel, P.J.; Mc Grath, J.E. (1972). *Research on human behavior*. New York: Holt, Rinehart and Winston.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill Book Co.

Spiegel, M.R. (1973). *Statistics*. New York: Schaum Publishing Co.



Capítulo IV

Sobre Monografias, Dissertações, Teses, Artigos e Projetos de Pesquisa: Significados e Recomendações para Iniciantes da Área de Educação Científica¹

M.A. Moreira

Consider, for example, where "knowledge" comes from. It isn't just there in a book waiting for someone to come along and "learn" it. Knowledge is produced in response to questions. And new knowledge results from the asking of new questions; quite often new questions about old questions. Here is the point: Once you have learned how to ask questions - relevant and appropriate and substantial questions - you have learned how to learn and no one can keep you from learning whatever you want or need to know. (Postman and Weingartner, 1969, p. 23)

Resumo

O objetivo deste texto é o de fornecer subsídios, principalmente a estudantes de pós-graduação da área de educação em ciências, sobre a redação de monografias, dissertações, teses, artigos e projetos de pesquisa. São discutidos os significados de cada um destes trabalhos acadêmico-científicos e feitos comentários e recomendações.

¹ Programa Internacional de Doutorado em Ens.no de Ciências, Universidade de Burgos, Espanha; Universidade Federal do Rio Grande do Sul, Brasil. Texto de Apoio N°13. Publicado em: *Actas del PIDEIC*, 4:03-24, 2002

Introdução

Na vida acadêmica, não há como fugir de monografias, dissertações, teses, artigos e projetos de pesquisa. São atividades que fazem parte dos rituais acadêmicos. Sem dúvida, a relevância do tema, o rigorismo do tratamento, a originalidade do enfoque, a importância do conhecimento produzido são variáveis determinantes do valor de cada um destes trabalhos acadêmicos. Contudo, a elaboração de tais trabalhos tem suas regras e sem elas dificilmente um tema relevante será abordado, adequadamente, de modo a gerar um produto de reconhecido valor acadêmico. Este texto pretende servir de apoio aos que se iniciam na elaboração de trabalhos acadêmicos dessa natureza, na área de educação científica.

A *monografia* é um texto acadêmico que enfoca um *tema específico*, em geral, através de pesquisa bibliográfica, revisão da literatura, acompanhada de alguma análise crítica, síntese ou categorização feita pelo autor. Quer dizer, a monografia não deve ser uma colagem desarticulada de resumos de artigos. O material bibliográfico levantado deve ser organizado de modo a dar uma visão integradora do tema focalizado. As monografias são muito usadas como trabalho de final de curso, tanto em cursos, propriamente ditos, de graduação ou de especialização, como em disciplinas de pós-graduação.

A *dissertação* é um trabalho acadêmico que se situa entre a monografia e a tese. Normalmente, a dissertação é o requisito final do mestrado e a tese do doutorado. As dissertações e teses diferem das monografias quanto ao nível em que o assunto é tratado e também porque, em geral, envolvem o relato de uma pesquisa feita para responder alguma questão sobre algum fenômeno de interesse, envolvendo muito mais do que revisão da literatura, por mais articulada que seja. Por sua vez, dissertações e teses também diferem entre si quanto ao nível de aprofundamento teórico e metodológico em que é tratada e

respondida a questão de pesquisa. O mestrado acadêmico é uma etapa intermediária entre a graduação e o doutorado. Consequentemente, a dissertação de mestrado ocupa uma posição intermediária entre a monografia e a tese de doutorado.

É como se as três estivessem ao longo de um contínuo no qual a monografia ocupasse uma extremidade e a tese a outra.

Projeto de pesquisa, como o próprio nome sugere, é uma proposta de pesquisa na qual o investigador deve dizer o que pretende investigar, por que vale a pena investigar o que se propõe a investigar e a partir de qual fundamentação teórica e metodológica conduzirá sua investigação. O desenvolvimento da pesquisa deve resultar na produção de algum conhecimento novo, alguma contribuição para a área. O relato da pesquisa pode originar uma dissertação, uma tese ou, no mínimo, um relatório de pesquisa. No entanto, as dissertações e teses, assim como as monografias, em geral, não são publicadas; são muito volumosas, muito densas, para serem divulgadas em larga escala; ficam à disposição nas bibliotecas. Relatórios de pesquisa, usualmente, se destinam às agências financiadoras do projeto de pesquisa. Qual é, então, o mecanismo, por excelência, para divulgar, amplamente, os resultados da pesquisa, i.e., os conhecimentos produzidos? É o *artigo científico*. Mesmo que a investigação tenha sido descrita em relatório de pesquisa, dissertação ou tese, é praticamente uma obrigação do pesquisador comunicar seus achados à comunidade acadêmico-científica através de artigos científicos.

A seguir, na primeira parte deste trabalho, serão feitos comentários específicos, quase definitivos, relativos ao significado de monografias, dissertações, teses, artigos e projetos de pesquisa. Relatórios de pesquisa por seu caráter mais técnico e, frequentemente, por seguirem o formato proposto pelas agências, não serão aqui abordados. Livros também não serão enfocados, porque sua publicação depende mais de critérios

editoriais do que acadêmicos. Ao final, na segunda parte, os comentários – em muitos casos, recomendações – serão gerais, aplicáveis a quaisquer dos trabalhos acadêmicos em questão.

Parte I - Significados

A monografia

Trata-se, repetindo, de um texto acadêmico que aborda um assunto específico elaborado predominantemente através de revisão da literatura (pesquisa bibliográfica). Embora possa parecer óbvio, é preciso frisar que a monografia tem começo, meio e fim. Tecnicamente, fala-se em *introdução*, *desenvolvimento* e *conclusão* (Bastos, Paixão, Fernandes e Deluiz, 1998, p. 13).

Introdução é a parte inicial na qual é definido o tema e os objetivos da monografia, levantam-se os antecedentes e controvérsias relativas ao assunto, indica-se a relevância do estudo e delimita-se o âmbito do trabalho.

Desenvolvimento é o núcleo do trabalho, o conteúdo da monografia. Comumente, está dividido em seções e subseções. É a parte em que o autor disserta sobre o assunto e desenvolve o tema a partir do estudo (geralmente bibliográfico) feito.

Conclusão é o fechamento do texto, no qual se faz um resumo, uma síntese integradora, uma análise crítica das diversas ideias trabalhadas ao longo do desenvolvimento.

A dissertação e a tese

A estrutura é similar a da monografia, porém mais complexa e profunda: *introdução*, *revisão da literatura*, *fundamentação teórica ou conceitual*, *metodologia*, *apresentação e discussão dos resultados*, *conclusão e recomendações*, cada uma destas partes correspondendo a um capítulo.

A *introdução* diz respeito, fundamentalmente, à definição do

problema, ou da questão básica investigada. Nesta parte, fala-se do assunto escolhido, de sua relevância, dos antecedentes e controvérsias, se for o caso, dos objetivos, das questões específicas a serem pesquisadas, dos limites do estudo, das hipóteses, se for o caso; define-se termos, se necessário, e, ao final, dá-se uma ideia dos capítulos seguintes. Mas o mais importante é explicitar claramente o que foi pesquisado e a relevância prática ou teórica da pesquisa.

A *revisão de literatura* tem por objetivos identificar o que já foi publicado sobre o assunto, o que já foi feito, por quem, quando e como. Em outras palavras, mapear o que já foi pesquisado na área. É um capítulo indispensável de uma dissertação ou tese e deve ser o mais completo e atual possível. Não se pode fazer uma pesquisa ignorando o que já foi realizado na área, para não correr o risco de, ingenuamente, "redescobrir a pólvora", para dar o devido crédito a pesquisadores que já abordaram o problema em pauta e para poder contribuir para a construção do conhecimento no campo de estudos da dissertação ou tese. Frequentemente, a revisão da literatura, em geral o segundo capítulo da dissertação ou da tese, é uma monografia, embora se possa argumentar que a revisão da literatura de uma dissertação ou tese deva ser mais exaustiva do que aquela feita em uma monografia (Moura, Ferreira e Paine, 1998, p. 36).

A *fundamentação teórica*, ou o marco conceitual, é também parte essencial de uma dissertação ou tese. Pesquisa é produção de conhecimento a partir da interação entre teoria e metodologia, entre pensar e fazer. Investigação sem referencial teórico não contribui para o crescimento da área, para o progresso do conhecimento em relação ao fenômeno de interesse sobre o qual se formula(m) a(s) questão(ões) de pesquisa. Estudos sem base teórica são erráticos. Na fundamentação teórica, o pesquisador deve explicitar os conceitos-chave da investigação, os construtos, os princípios

teóricos norteadores, as teorias específicas e, se for o caso, os que estão por detrás da abordagem ao problema, que subjazem às questões de pesquisa. Não é necessário afiliar-se a uma teoria em particular, mas é imprescindível esclarecer o domínio conceitual da pesquisa.

A *metodologia* se refere a como foi feita a investigação, a sequência de passos seguida para ir da pergunta-chave da pesquisa às respostas encontradas. A metodologia trata de variáveis, construtos, amostras, casos, populações, delineamentos, tratamentos, categorias de análise, instrumentos de medida, registros dos eventos, dados, transformações quantitativas ou qualitativas dos dados. Por instrumentos de medida, entendem-se testes, entrevistas, escalas de atitude, questionários ou qualquer outro meio de fazer registros dos eventos ou objetos de estudo, gerando, assim, dados a serem analisados. Por transformações, compreende-se, por exemplo, o tratamento estatístico dos dados em um estudo quantitativo ou o enfoque interpretativo típico de um estudo qualitativo.

A *apresentação e discussão dos resultados* é um capítulo geralmente estruturado em seções que correspondem a cada questão investigada. Os resultados são respostas às questões de pesquisa. *O relato dos resultados visa oferecer evidências que esclareçam cada questão levantada ou cada hipótese formulada na proposição do problema* (Bastos et al., 1998, p. 11). Na discussão e interpretação dos resultados, o autor se reporta à fundamentação teórica e à revisão da literatura, chamando atenção para a consistência ou desacordo de seus resultados com a teoria e com estudos prévios sobre o assunto.

O capítulo final é o da *conclusão e recomendações*. Nele, pode-se até chegar a uma conclusão propriamente dita, porém, mais comumente, faz-se uma síntese da pesquisa, dá-se uma visão integradora de toda a tese ou dissertação. As *recomendações* são sugestões práticas para a implementação dos

resultados ou para pesquisas adicionais (ibid.). A conclusão é também o local apropriado para as asserções de valor em relação ao conhecimento produzido, i.e., para argumentar sobre o valor deste conhecimento.

Cabe aqui reiterar que, embora a dissertação e a tese tenham a mesma estrutura, elas diferem quanto ao nível de aprofundamento teórico e metodológico em que são abordadas e respondidas as questões de pesquisa. Basta lembrar que a dissertação é o requisito final do mestrado e a tese (às vezes, o único requisito) do doutorado, titulação esta que deve significar capacidade de pesquisar com independência.

O artigo de pesquisa

Artigos de pesquisa também relatam a investigação feita e os resultados obtidos, porém de uma forma muito compacta e pública. As monografias, dissertações e teses são, normalmente, trabalhos não publicados, enquanto o artigo é, por excelência, a publicação científica. É no artigo que o pesquisador expõe seus resultados (e se expõe) à comunidade científica. Os artigos são publicados em revistas e estas têm suas normas, mas a estrutura básica de um artigo é, praticamente, sempre a mesma: *introdução* (frequentemente incluindo revisão da literatura e marco teórico), *materiais e métodos* (ou metodologia), *apresentação e discussão dos resultados, conclusão*. É, em essência, a mesma estrutura da dissertação ou tese, mas, ao invés de capítulos, o artigo contém seções compactas relativas a cada um desses itens. Sem o menor rigorismo, apenas para dar uma ideia da extensão de um artigo, diria que não vai muito além de dez páginas. No entanto, artigos sobre pesquisas qualitativas geralmente são mais extensos, podendo chegar a vinte ou trinta páginas, devido à natureza de sua retórica que é narrativa, persuasiva. (As monografias, em geral, têm algumas dezenas de páginas, enquanto as dissertações e teses facilmente chegam à casa das centenas.) Finalmente,

cabe destacar que a comunidade acadêmico-científica, usualmente, só reconhece o valor acadêmico-científico de artigos publicados em revistas arbitradas, ou seja, revistas que só publicam artigos que passaram por um processo de revisão, às cegas, por, pelo menos, dois pesquisadores experientes, preferentemente especialistas na área em questão.

Referências

Qualquer um dos trabalhos focalizados anteriormente tem, ao final, uma lista de referências bibliográficas. Há regras técnicas para isso, as quais podem ser encontradas em manuais, como os relacionados nas referências bibliográficas deste texto que, por sua vez, constituem um exemplo de como podem ser apresentadas as referências. *Contudo, mais importante do que as regras técnicas para listar referências, é, de fato, dá-las ao longo do texto, incluindo número de página, cada vez que o autor de um trabalho acadêmico usar palavras ou ideias de outro pesquisador ou pesquisadora.* Nos artigos, é mais difícil deixar de dar as referências adequadas, ou porque quem o fizer pode cair em descrédito ou porque o artigo é devolvido para que sejam incorporadas as citações pertinentes. Nas dissertações e teses, no entanto, às vezes tomam-se como próprias a redação e as ideias de outrem. Isso é imperdoável e deve ser motivo para não aprovar a dissertação ou tese. Há também regras sobre como fazer citações, as quais estão também nos manuais, porém, insisto, o mais importante é realmente citar, ao longo do texto, quem deve ser citado.

Anexos e apêndices

Os trabalhos acadêmicos, muitas vezes, incluem anexos e/ou apêndices com materiais adicionais ou complementares que quebrariam a continuidade do corpo do trabalho se fossem nele inseridos. Por exemplo, neste texto de apoio há um apêndice (I)

sobre apresentação de um projeto de pesquisa. Trata-se, basicamente, de um roteiro para elaboração de um projeto de pesquisa. Se fosse incorporado ao texto, provavelmente ficaria destoante por ocupar bem mais espaço do que o dedicado à monografia, à dissertação, à tese e ao artigo que são também objetos deste trabalho. Há também um apêndice (II) sobre tipos de revisão a serem evitados. Instrumentos de medida, glossários, documentos de pesquisa, transcrições de entrevistas, são exemplos de materiais incluídos como anexos ou apêndices de trabalhos acadêmico-científicos. A diferença entre anexos e apêndices (UFF, 1998) é que estes são elaborados pelo próprio pesquisador, enquanto os primeiros não. Em ambos os casos complementam, fundamentam, comprovam ou ilustram o trabalho, sem prejudicar sua unidade.

Detalhes técnicos

Aspectos técnicos, como margens, espaços, notas de rodapé, numeração de páginas, legendas de tabelas e figuras, resumo e *abstract*, folha de rosto, dedicatória, agradecimentos, não serão tratados neste texto. O leitor pode facilmente encontrá-los em manuais de redação de trabalhos acadêmico-científicos. Uma outra alternativa é ir à biblioteca e examinar o formato de algumas monografias, dissertações e teses.

O projeto de pesquisa

O projeto de pesquisa é uma proposta específica e detalhada da pesquisa, com o objetivo de definir uma questão de pesquisa e a forma pela qual ela será investigada (Bastos et al., 1994, p. 1). Antes de terem sido realizadas, as pesquisas relatadas em dissertações, teses, relatórios de pesquisas ou artigos científicos passaram por uma fase de planejamento que, uma vez posta no papel, originou um projeto de pesquisa. Essencialmente, o projeto de pesquisa deve dar uma ideia clara do que se pretende

fazer, por que vale a pena fazer e como será feito. As agências financiadoras de pesquisa só concedem subsídios às pesquisas submetidas através de projetos de pesquisa, e em caráter competitivo (os melhores projetos é que levam os recursos financeiros disponíveis). As comissões de pós-graduação exigem projetos de dissertação e projetos de tese que são, em última análise, projetos de pesquisa. Tais projetos, muitas vezes, devem ser defendidos publicamente.

Isso significa que projetos de pesquisa são extremamente importantes no meio acadêmico-científico. Saber redigir projetos de pesquisa é uma questão de sobrevivência. No *Apêndice 1* deste texto de apoio há sugestões específicas sobre a apresentação de um projeto de pesquisa.

Parte II - Recomendações

Como fazer uma boa pesquisa ?

Leia muito! Escolha uma questão relevante! Seja original! Planeje muito! Projete uma pesquisa viável, mas não se contente com pouco! Colete dados de boa qualidade!

Ler é fundamental. Leia tudo que puder sobre o fenômeno de interesse de sua pesquisa. Não se preocupe em logo definir uma questão de pesquisa e sair lançando hipóteses. Leia, estude, faça fichas catalográficas. Peça sugestões a pesquisadores experientes sobre o que ler a respeito de um certo assunto. As leituras ajudar-lhe-ão a definir e delimitar o problema, a questão de pesquisa. Além disso, a revisão da literatura poderá estar bem encaminhada. A propósito, lembre-se que a revisão da literatura não é apenas uma colagem de leituras, é preciso integrá-las, tecer comentários articuladores, de modo a chegar a um certo estado da arte sobre tema em pauta. (Vide *Apêndice II*)

Escolher uma questão relevante não é fácil, mas é primordial. Quem avalia projetos não apóia o que não vale a pena ser pesquisado. Quem avalia artigos não recomenda a publicação de pesquisas irrelevantes. É claro que o conceito de relevante é relativo. Mesmo assim, é preciso buscar uma questão relevante para pesquisar. Tente argumentar com você mesmo por que é importante investigar o que você pretende investigar (aliás, você terá de justificar isso para os avaliadores). Você, certamente, desejará, quererá que sua pesquisa seja muito bem feita. Então, lembre-se que só vale a pena fazer bem feito aquilo que vale a pena ser feito.

Originalidade também é importante. O mesmo problema pode ser abordado de distintas maneiras. Você já leu muito, já sabe como certo problema foi enfrentado, como se tentou responder uma certa questão de pesquisa, então, procure ser original, busque fazer o que ainda não foi feito. Na verdade, não há muito disso na pesquisa em educação científica, ou em educação, de um modo geral. A tendência é a de que, quando alguém tem uma ideia original, muitas e muitas pesquisas sejam feitas em torno dessa ideia. Obviamente, a réplica e a repetição, em outro contexto, são importantes para o progresso e consolidação de uma área, mas o que ocorre muitas vezes é uma circularidade, fica-se andando em círculos durante muito tempo. Por exemplo, quando Ausubel propôs, em 1968, os organizadores prévios com pontes cognitivas entre o que o aluno sabe e o que deveria saber para que o material de aprendizagem fosse potencialmente significativo, foram realizados centenas de estudos experimentais do tipo "com organizador (tratamento experimental) x sem organizador (grupo controle)". Foram estudos essencialmente repetitivos que não levaram a nada e ofuscaram outros conceitos importantes propostos por Ausubel, na mesma época, como o de aprendizagem significativa. Na pesquisa em ensino de Física, na década de 70, ocorreu algo mais

ou menos parecido com as concepções alternativas. Nessa época, praticamente todos os pesquisadores da área investigavam concepções alternativas. Não se pode dizer que tais pesquisas não levaram a nada pois, pelo menos, chegou-se a um grande catálogo de concepções alternativas, mas não deixaram de ser repetitivas. O mesmo aconteceu, na década de 80, com a ideia de mudança conceitual, e o mesmo pode estar acontecendo agora, no ensino de ciências, com o conceito de modelo mental (Johnson-Laird, 1983; Gentner e Stevens, 1983; Moreira, 1996). É preciso, no mínimo, tentar ser original na pesquisa, não só seguir a moda.

Planejamento não é perda de tempo. Em geral, pesquisas planejadas rapidamente resultam em triviais ou enfrentam sérios tropeços no seu desenvolvimento. Um exemplo típico é aquela pesquisa em que o pesquisador decide, sem base teórica e sem saber bem o porquê, aplicar um questionário de atitude ou de opinião a uma certa amostra. Feita a aplicação, é feita a tabulação e o resultado da pesquisa é um histograma que não acrescenta nada ao conhecimento na área e que não contribui para o crescimento da mesma. Certamente, às vezes, algum levantamento deste tipo é necessário em uma fase inicial de uma pesquisa, mas, geralmente, não é o que se vê na pesquisa educacional. Planejar a pesquisa implica definir variáveis, indicadores, construtos, amostras, técnicas e registros observacionais, casos, instrumentos, técnicas de análise de dados, custos, cronograma. Obviamente, tudo o que for planejado pode vir a ser modificado durante o desenvolvimento da pesquisa. Muitas decisões são tomadas ao longo da investigação, mas a existência de uma proposta de pesquisa bem trabalhada facilita muito a tomada de decisões.

A viabilidade é essencial para uma pesquisa exitosa. Não basta ter uma questão interessante e uma abordagem original: a pesquisa tem de ser viável. A inexperiência de muitos estudantes

de pós-graduação leva-os, frequentemente, a proporem investigações praticamente inviáveis dentro das condições e recursos disponíveis. Ao projetar a pesquisa, é necessário levar em conta o tempo e os recursos materiais e humanos disponíveis. Por exemplo, se o mestrado deve ser projetado para ser feito em dois anos, a pesquisa que conduzirá à dissertação de mestrado deve ser planejada para realização dentro deste período levando em conta que o mestrando terá outros requisitos, principalmente disciplinas de pós-graduação, a satisfazer ao longo deste mesmo prazo. Para o doutorado, o tempo estimado para estudantes em tempo integral é da ordem de quatro anos. É possível, então, fazer uma pesquisa de mais fôlego para chegar à tese de doutorado. E assim deve ser, pois o doutorado tem por finalidade formar o futuro pesquisador.

Pesquisa viável não implica estudo excessivamente simples, despretencioso, desambicioso. Ao contrário, o pesquisador deve delinear uma investigação factível dentro das condições existentes, mas não menos do que isso. Ele não deve se contentar com um mero levantamento quando, por exemplo, é possível alguma intervenção objetivando uma solução para o que foi levantado. *Principalmente nas pesquisas de doutorado é preciso evitar aquelas que não passam de constatações ou propostas.* O doutoramento, como tem sido insistentemente, visa a preparação do pesquisador independente, e não é fazendo constatações e propostas que se atinge esse objetivo.

Se os dados coletados forem ruins, não há tratamento quantitativo ou qualitativo que salve a pesquisa. Quer dizer, nunca poder-se-á responder honestamente uma questão de pesquisa com dados não confiáveis, independente de sua natureza. Em estudos quantitativos, exige-se que os instrumentos de coleta de dados tenham validade e fidedignidade. Em investigações qualitativas, tais conceitos não têm muito sentido, mas, ainda assim, a qualidade dos dados gerados determina a credibilidade das realidades construídas.

Como identificar a estrutura de uma pesquisa?

Uma vez redigida, digamos, a primeira versão de uma tese, dissertação ou artigo de pesquisa, como saber "se não falta nada"? Como reconhecer a estrutura do trabalho? O mesmo questionamento se aplica à análise crítica de trabalhos de outros autores, por exemplo, quando se faz uma revisão da literatura.

Uma boa maneira de resolver esse problema é responder as seguintes questões propostas por D.B. Gowin (1981, p.88) como método de análise da estrutura do processo de produção do conhecimento:

1. Qual é a questão-básica investigada?
2. Quais são os conceitos-chave usados para formular a questão-básica?
3. Qual é a metodologia usada para responder a questão-básica?
4. Quais são as respostas obtidas?
5. Qual é o valor destas respostas?

A questão-básica (ou as questões-básicas, se for o caso) é aquela que identifica o que, em última análise, foi pesquisado; é a pergunta que, de certa forma, "conta" a pesquisa (Gowin usa o termo "telling question"). A segunda pergunta tem a ver com a estrutura conceitual da pesquisa, com o marco teórico. A terceira se refere ao procedimento utilizado, aos passos seguidos, na busca de respostas à questão básica. Não se trata, portanto, do chamado "método científico". As respostas obtidas são os achados da pesquisa, o conhecimento produzido. Finalmente, a quinta questão é relativa ao valor do conhecimento produzido. Para que serve esse conhecimento?

Para saber se não falta nada em sua tese, dissertação ou artigo, responda a estas cinco questões de Gowin: Está explicitado, claramente, o que foi pesquisado? Está clara a base conceitual, o marco teórico? A metodologia está bem explicada, bem transparente? E os achados, estão nitidamente

identificados, comentados e relacionados à questão-básica? Por último, você fez algum juízo de valor sobre seus achados de pesquisa? Para que servem eles?

Estas mesmas cinco questões podem ser dispostas em um diagrama em forma de V, também proposto por Gowin (1981, p. 107), que talvez permita uma melhor compreensão da estrutura do processo de produção do conhecimento ou, simplesmente, da estrutura de uma pesquisa, independente de sua natureza. A Figura 1 apresenta uma versão simplificada (Moreira, 1990, p. 73) e ligeiramente modificada do "V" epistemológico de Gowin. Nesse diagrama, a produção de conhecimento, ou seja, a pesquisa é concebida como resultante da permanente interação entre pensar e fazer, entre um domínio conceitual e epistemológico e um domínio metodológico. No domínio conceitual estão, antes de mais nada, os *conceitos-chave* da pesquisa; tais conceitos estão relacionados e hierarquizados em uma *estrutura conceitual* que pode ser explicitada em um mapa conceitual (Moreira e Buchweitz, 1993; Moreira, 2010). *Quer dizer, não é suficiente identificar e listar os conceitos-chave, é preciso explicitar também como estão estruturadas, hierarquizados, relacionados e isso pode ser feito traçando-se um mapa conceitual* (op. cit., p. 77). Às vezes, uma investigação envolve mais de uma estrutura conceitual, como é o caso da educação científica que, geralmente, inclui conceitos da matéria de ensino e das áreas de ensino-aprendizagem, de educação, de psicologia e outras. Ainda nesse domínio estão os *princípios* norteadores da pesquisa, ou seja, proposições iniciais da pesquisa, pressupostos da pesquisa. Normalmente, essas proposições fazem parte ou, pelo menos, estão coerentes com determinadas *teorias*. De um modo geral, teorias são tentativas humanas de sistematizar uma área de conhecimento, uma maneira particular de ver as coisas, de explicar e prever observações, de resolver problemas. Teorias direcionam a produção de conhecimento e são fruto do conhecimento produzido. Há uma permanente interação entre

teoria e metodologia. Por isso, toda pesquisa deve ter um marco teórico, conceitual. Não necessariamente uma teoria em particular, mas um conjunto articulado de pressupostos teóricos. Por isso também a revisão da literatura é indispensável, pois, certamente, ao fazê-la, o pesquisador dar-se-á conta de quais os referenciais teóricos que têm fundamentado as pesquisas na área escolhida. Registre-se, no entanto, que assim como a revisão da literatura não deve ser uma mera coletânea de artigos, o marco teórico não pode ser apenas uma colagem de propostas de vários autores. Aliás, é comum encontrar-se, em dissertações e teses, uma sequência de citações a título de fundamentação teórica. Não é disso que se está falando aqui; tal sequência é inútil e revela desconhecimento da área. Finalmente, no topo do lado conceitual do "V", estão as crenças profundas subjacentes à pesquisa. Normalmente, tais crenças não estão explícitas em dissertações, teses e artigos. São os *paradigmas*, as *filosofias*, as *visões de mundo* do pesquisador. Por exemplo, estudos quantitativos em educação subscrevem um paradigma científico. Pesquisas de natureza piagetiana ou ausubeliana estão dentro de uma filosofia construtivista. Como o conhecimento é produzido, como se dá o progresso científico é uma questão epistemológica superordenada em relação às teorias. À medida em que se vai subindo no lado esquerdo do "V", o domínio conceitual da pesquisa vai ficando mais abrangente. Quando se chega ao topo, está-se no nível das filosofias.

Como o conhecimento é produzido, como se dá o progresso científico é uma questão epistemológica superordenada em relação às teorias. À medida em que se vai subindo no lado esquerdo do "V", o domínio conceitual da pesquisa vai ficando mais abrangente. Quando se chega ao topo, está-se no nível das filosofias.

Na ponta do "V", estão eventos que acontecem naturalmente, ou que se faz acontecer, para responder a questão básica da pesquisa. A seleção de eventos para observação e

registro é influenciada pelo domínio conceitual (conceitos, princípios, teorias, visão de mundo).

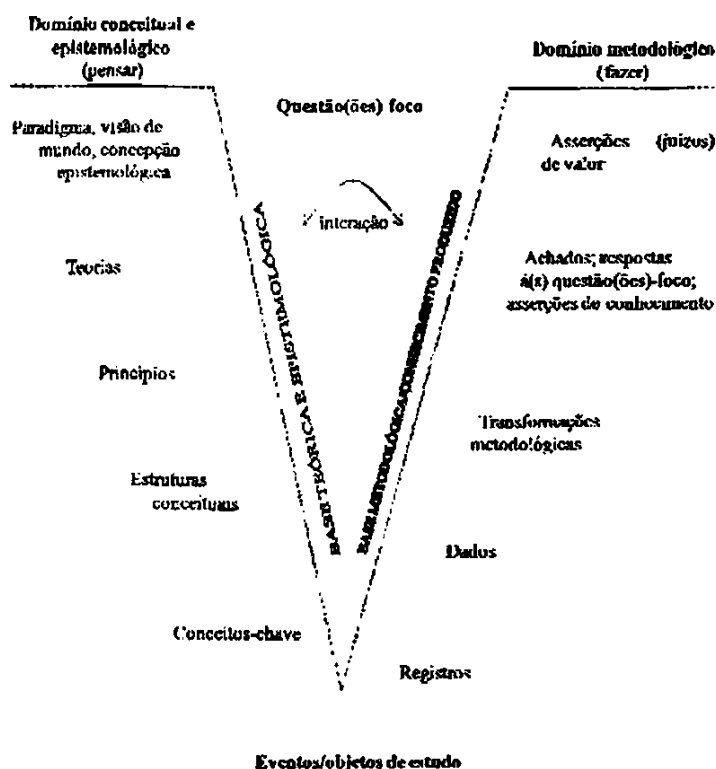


Figura 1 – O “V” epistemológico de Gowin (ou diagrama V).

A fim de estudar esses eventos, registros devem ser feitos. Esses registros geram dados² que são transformados e interpretados à luz do domínio conceitual, conduzindo a possíveis respostas à questão-foco, a *asserções de conhecimento*, aos produtos da pesquisa, por exemplo, correlações, relações causais, explicações, justificativas, compreensões. São os *achados* da pesquisa.

² A distinção entre registros e dados é sutil; poder-se-ia dizer que os registros são dados brutos.

O lado direito do "V", começando pelos registros e chegando até às asserções de conhecimento, é o da metodologia, o do domínio metodológico. Este domínio inclui também as *asserções, ou juízos, de valor*. Tais asserções referem-se à *significância, utilidade, importância, do conhecimento produzido*. É feita alguma *alegação sobre o valor do estudo? Alguma asserção sobre a significância social? Estética? Científica? Educacional? Significante para quem? Para quê?* (Moreira, 1990, p. 75).

Este é o "V"! Se você quiser verificar se o relato de sua pesquisa – seja ele através de dissertação, tese, artigo ou relatório de pesquisa – está completo e bem estruturado, faça o diagrama V de sua investigação. É também uma boa maneira de fazer um sumário, de dar uma visão geral de sua dissertação ou tese, tanto na versão escrita como na apresentação oral na hora da defesa.

Para dar-lhe uma ideia de quais têm sido as questões mais pesquisadas em educação em ciências e com quais bases teóricas, epistemológicas e metodológicas, as Figuras 2 e 3 apresentam (Moreira, 1999), respectivamente, um diagrama "V" e um mapa conceitual para a pesquisa em educação em ciências, segundo minha perspectiva.

Como escolher a metodologia de pesquisa?

Há diferentes tipos de pesquisa e distintas maneiras de classificá-los. Por exemplo, *experimental, teórica, descritiva, histórica, bibliográfica*. Outra distinção usualmente feita é entre *pesquisa básica* (pura, ou fundamental), aquela que busca a ampliação do conhecimento em uma área, sem preocupação de utilizá-lo na prática, e *pesquisa aplicada*, a que procura resultados utilizáveis imediatamente na solução de problemas reais (Marconi e Lakatos, 1982, p. 19). Há algum tempo, uma outra diferença tem permeado a pesquisa em educação: *pesquisa quantitativa* x *pesquisa qualitativa*.

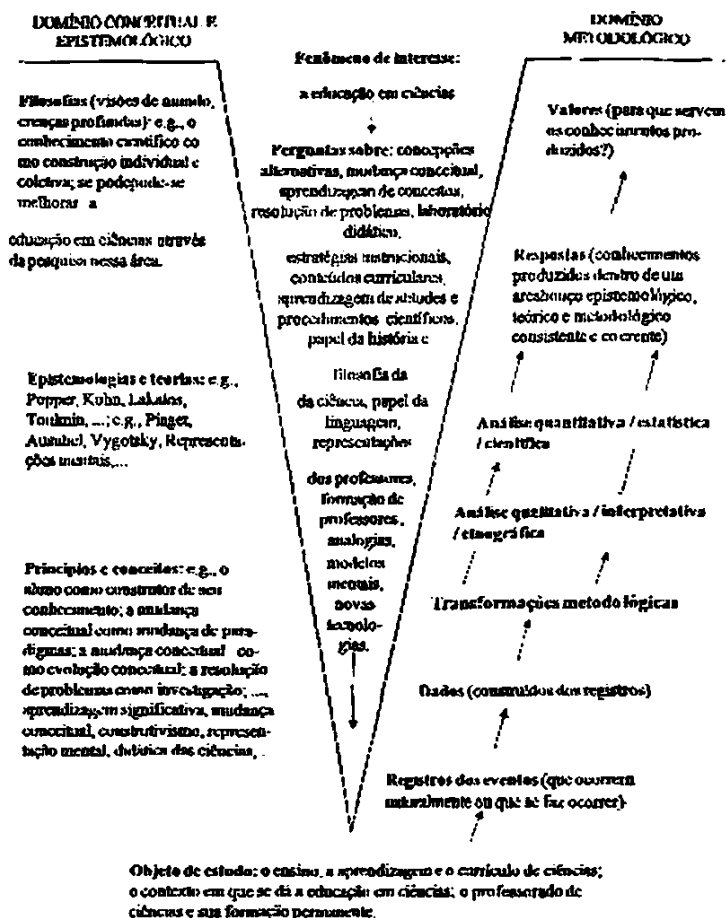


Figura 2 – Um diagrama “V” para a pesquisa em educação em ciências: uma visão detalhada (M.A. Moreira, 1999).

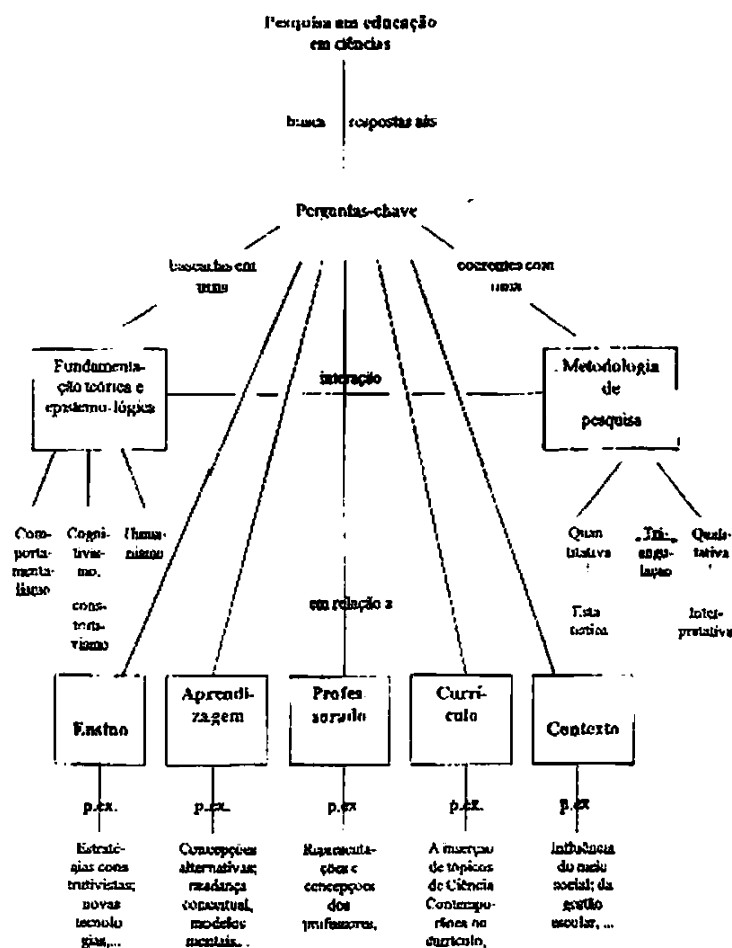


Figura 3 – Um mapa conceitual para a pesquisa em educação em ciências (M.A. Moreira, 1999).

A investigação quantitativa procura estudar o fenômeno de interesse da pesquisa em educação geralmente através de

estudos experimentais ou correlacionais caracterizados, principalmente, por medições objetivas e análises quantitativas. De um modo geral, esse enfoque está baseado no modelo de pesquisa, dito científico, das ciências exatas, do qual a Física é, provavelmente, o melhor exemplo (Moreira, 1997, p. 7). Por outro lado, investigação qualitativa é um termo que tem sido usado alternativamente para designar vários enfoques à pesquisa educacional, tais como pesquisa etnográfica, participativa observacional, estudo de caso, fenomenológica, interpretativa. Cada um desses enfoques forma um todo coerente que engloba suposições internamente consistentes sobre natureza humana, sociedade, objeto de estudo e metodologia; no entanto, compartilham muitas semelhanças e por questão de simplicidade, são chamadas comumente de pesquisa qualitativa. (op. cit., p. 24). O interesse central da investigação qualitativa está nos significados que as pessoas atribuem a eventos e objetos, em suas ações e interações dentro de um contexto social, e na elucidação e exposição desses significados. Nos estudos quantitativos, o interesse está no estabelecimento de generalizações, relações causais, correlações, entre variáveis.

Há duas ou três décadas, a maioria das investigações em educação, e em ensino de ciências em particular, era de natureza quantitativa. Hoje, predominam os estudos qualitativos. Contudo, uma pesquisa feita nos dias de hoje não tem que, necessariamente, seguir uma metodologia qualitativa. A metodologia depende da questão de pesquisa, da clara formulação do problema. A pesquisa começa com a existência de uma pergunta que se deseja responder através de um procedimento – um conjunto de passos que permitam obter a informação necessária para respondê-la. Além destes dois requisitos (pergunta e procedimento), Luna (1997, p. 27) argumenta que uma pesquisa deve preencher um terceiro: a indicação do grau de confiabilidade na resposta obtida, segundo

os parâmetros do referencial adotado. "Em outras palavras, é necessário haver um problema de pesquisa (o que não significa uma hipótese formal), um procedimento que gere informação relevante para a resposta e, finalmente, é preciso demonstrar que esta informação decorre do procedimento empregado e que a resposta produzida por ele não é apenas uma resposta possível, como também é a melhor nas circunstâncias (o que inclui, certamente, o referencial teórico)." (ibid.). O preenchimento destes três requisitos não compromete a pesquisa com nenhuma tendência metodológica particular. "As decisões metodológicas são pura decorrência do problema formulado e este só se explica devidamente em relação ao referencial teórico que deu origem a ele" (op. cit., p. 32).

Voltamos assim ao "V" de Gowin: *pesquisa é produção de conhecimento a partir de uma questão básica (problema) formulada no marco da interação entre teoria (domínio conceitual) e metodologia (domínio metodológico)*. A escolha da metodologia de pesquisa depende do referencial teórico. O domínio metodológico não pode ser definido *a priori*.

Conclusão

Este trabalho assemelha-se a uma monografia. Entretanto, é fruto muito mais de minha longa experiência como professor, orientador, avaliador e pesquisador do que de uma revisão da literatura sobre o assunto. Decidi escrevê-lo depois de ler duas dezenas de monografias sobre teorias de aprendizagem que solicitei como trabalho de conclusão de uma disciplina de pós-graduação nessa área. Meus alunos provinham todos da área científica e muitos eram iniciantes na redação de trabalhos dessa natureza. As dificuldades que percebi, mesmo entre os mais experientes, levaram-me a escrever algo para oferecer-lhes uma melhor ideia do que seria uma monografia. Mas logo lembrei que, talvez, viessem também a ter dificuldades ao elaborar o projeto

de dissertação/tese, ao redigir a dissertação/tese e ao preparar artigos sobre a pesquisa relatada na tese. Assim, o texto, que inicialmente foi pensado para esclarecer o que se espera de uma monografia, acabou incluindo dissertação, tese, projeto de pesquisa e artigo científico, e quase se transformou em uma monografia sobre trabalhos acadêmico-científicos.

Como insisti ao longo de todo o texto, trabalhos dessa natureza não devem ser apenas uma coleção de citações, definições, recomendações ou coisa que o valha. Eles devem constituir um todo articulado, com uma *introdução*, na qual se defina o tema, os objetivos e o âmbito, um *desenvolvimento* – que é o núcleo do trabalho – e uma *conclusão*, em que se faz um resumo, uma síntese, ou análise integradora final.

Chegado, então, o momento da conclusão, o que dizer a título de síntese, de recomendação final? Creio que nada melhor do que o que está na epígrafe: o novo conhecimento é produzido a partir de novas questões; uma vez que se aprende a formular questões apropriadas e substanciais, se aprende a aprender, i.e., se aprende a pesquisar e se adquire autonomia. Quer dizer, o mais importante é a questão de pesquisa, a pergunta-foco. De acordo, mas é preciso planejar como respondê-la – aí entra o projeto de pesquisa – e, depois, relatar a resposta – o que se faz através de trabalhos acadêmico-científicos. Este texto pretende servir de apoio nessa parte, não tão difícil nem tão nobre como a da formulação da questão de pesquisa, mas indispensável no meio acadêmico.

Referências bibliográficas

Além das referências efetivamente usadas e citadas no texto, foram agregadas outras para exemplificar como podem ser referidas as diversas fontes bibliográficas de um trabalho acadêmico. Foram seguidas, aproximadamente³, as normas do

³ Pelas normas da APA, não é necessário dar o número total de páginas dos livros, nem o nome da cidade onde é publicada a revista no caso de artigos

manual da APA (American Psychological Association), as quais são bastante aceitas internacionalmente.

Alves-Mazzotti, A.J. & Gewandszajder, F. (1998). *O método nas ciências naturais e sociais: pesquisa quantitativa e qualitativa*. São Paulo: Pioneira. 203p.

American Psychological Association. (1983). *Publication manual of the American Psychological Association* (3rd. Ed.). Hyattsville, MD: American Psychological Association. 208p.

Ausubel, D.P. (1968). *Educational psychology: A cognitive view*. New York: Holt, Rinehart and Winston. 685 p.

Bastos, L.R., Paixão, L., Fernandes, L.M. & Deluiz, N. (1994). *Manual para a elaboração de projetos e relatórios de pesquisa, teses, dissertações e monografias* (4ª Ed.). Rio de Janeiro: Guanabara. 96 p.

Gentner, D. & Stevens, A.L. (Eds.). (1983). *Mental models*. Hillsdale, NJ: Lawrence Erlbaum Associates. 348 p.

Gowin, D.B. (1981). *Educating*. Ithaca, N.Y.: Cornell University Press. 210p.

Johnson-Laird, P.N. (1983). *Mental models*. Cambridge, MA: Harvard University Press. 513 p.

Luna, S.V. (1997). O falso conflito entre tendências metodológicas. In I. Fazenda (Org.), *Metodologia da pesquisa educacional* (pp. 23-33). São Paulo: Cortez Editora. 174 p.

Marconi, M.A. & Lakatos, E.M. (1982). *Técnicas de pesquisa*. São Paulo: Editora Atlas. 205 p.

Moreira, M.A. & Buchweitz, B. (1993). *Novas estratégias de ensino e aprendizagem: os mapas conceituais e o Vê epistemológico*. Lisboa: Plátano Edições Técnicas. 114 p.

Moreira, M.A. (1996). Modelos mentais. *Investigações em Ensino de Ciências*, Porto Alegre, 1: 193-232.

Moreira, M.A. (1999). A pesquisa em educação em ciências e a formação permanente do professor de ciências. In Sánchez, J.M., Oñorbe, A. & Bustamante, I. (Eds.) *Educación Científica*. Alcalá de Henares: Servicio de Publicaciones Universidad de Alcalá. 250p.

Moreira, M.A. (1999). *Investigación en enseñanza: aspectos metodológicos*. Universidad de Burgos, Programa Internacional de Doctorado en Enseñanza de las Ciencias, Texto de Apoyo N° 1.

Moreira, M.A. (2010). *Mapas conceituais e aprendizagem significativa*. São Paulo: Centauro Editora. 80p.

Moura, M.L.S., Ferreira, M.C. & Paine, P.A. (1998). *Manual de elaboração de projetos de pesquisa*. Rio de Janeiro: Ed. UERJ, 134 p.

Postman, N. & Weingartner, C. (1969). *Teaching as a subversive activity*. New York: Dell Publishing Co. 219 p.

Universidade Federal Fluminense. (1998). *Apresentação de trabalhos monográficos de conclusão de curso* (3ª Ed.). Niterói, RJ: Editora da UFF. 69 p.

Apêndice I

Apresentação de um Projeto de Pesquisa em Educação⁷

Introdução

A preparação de um projeto ou de um termo de referência⁸ de pesquisa é uma tarefa que todo pesquisador ou estudante de pós-graduação necessariamente enfrenta – às vezes com relativa frequência – ao longo de sua trajetória. Trata-se de um trabalho imprescindível, porém, muitas vezes, indevidamente supervalorizado. Na verdade, é uma decorrência natural do planejamento de pesquisa:

Se a ideia da pesquisa foi bem trabalhada, o projeto ou termos de referência surgem claramente. Ao contrário, se o pesquisador não sabe o que quer nem onde pretende ir, não

⁷ Apêndice do Texto de Apoio nº9 – Sobre monografias, dissertações, teses, artigos e projetos de pesquisa: significados e recomendações para iniciantes da área de educação científica, Moreira, M.A. (1999) – do Programa Internacional de Doutorado em Ensino de Ciências. Universidade de Burgos/Espanha, UFRGS, Porto Alegre, Brasil.

⁸ Alguns autores usam as expressões “projeto de pesquisa” e “termo de referência” indistintamente para designar um documento preliminar da pesquisa, outros reservam a expressão “termo de referência” para o que se poderia chamar de “anteprojeto da pesquisa”. O termo de referência seria uma espécie de carta que descreve o que se pretende fazer, por que fazer e como fazer, enquanto o projeto conteria os mesmos elementos de maneira mais elaborada e formal.

adianta nada rebuscar ou remendar o projeto; é necessário repensar o que pretende fazer. Em outras palavras, o planejamento da pesquisa é uma etapa crucial; a redação dos termos de referência, um corolário subalterno. (Moura Castro, 1976, p. 14).

O projeto de pesquisa deve, essencialmente, dar uma ideia clara do que se pretende fazer, por que fazer e como fazer. Três tipos básicos de informações devem estar contidas no projeto ou serem transmitidas através dos termos de referência:

1. Temos que caracterizar a importância do problema que estamos examinando, seja em termos da sua atualidade, implicações práticas, relevância da população estudada ou contribuição à teoria. 2. É também preciso explicitar a originalidade da contribuição pretendida. 3. Finalmente, é preciso indicar da melhor maneira possível a viabilidade da pesquisa; não basta ter um problema interessante e original, é essencial que o pesquisador seja capaz de enfrentá-lo com sucesso. (op. cit., p. 20)

Entretanto, embora sejam requisitos básicos de um projeto de pesquisa, tais informações geralmente não determinam o formato do projeto:

Importância do tema, originalidade e viabilidade do projeto são os três elementos fundamentais a serem considerados. Todavia, dividir a discussão rigorosamente segundo estes três critérios não se mostra um bom formato de redação. Diríamos que, na feitura do projeto, estes três critérios estarão guiando o desenvolvimento do tema e que, por sua vez, os juizes lerão o documento com as mesmas preocupações. Mas a apresentação deve seguir uma lógica ligeiramente diferente. (ibid. p. 20)

Isso posto, sugere-se, a seguir, um roteiro para elaboração do projeto. Não é um formulário a ser preenchido, é apenas um roteiro sugerido.

Roteiro

I- Título do projeto – Deve ser razoavelmente conciso, mas descritivo; deve incluir elementos essenciais do estudo, palavras-chave da pesquisa. É, provavelmente, a frase isolada mais importante de todo o projeto.

II - O problema

- II.1 Natureza. Nesta parte do projeto, deve-se não só enunciar claramente o problema, como também os objetivos do estudo a ser feito, o produto que se espera obter. Uma maneira de definir os objetivos é formular questões que a pesquisa, em princípio, responderá. (Não devem ser incluídos objetivos que não poderão ser atingidos!) Além disso, deve-se, também aqui, delimitar o escopo do problema, deixando claro quais os aspectos do mesmo que serão investigados e quais não serão pesquisados. O projeto deve, antes de mais nada, explicitar precisamente o que vai ser investigado.
- II.2 Justificativa. Qual a significância do problema e do projeto? Por que o problema merece ser investigado? Como a pesquisa proposta contribuirá para sua solução?
- II.3 Contexto. Qual é a situação na qual se origina ou se insere o problema? Quais são as audiências, os tipos de escolas, os programas, os professores, os sistemas sociais, culturais, etc. nele envolvidos? Qual é o contexto da pesquisa?
- II.4 Revisão da literatura. Quão extensa é a literatura nessa área? Em que local esta se encontra? Quais são os trabalhos mais significativos na área? O que

dizem eles? Quais são as dificuldades apontadas, na literatura existente, em relação ao problema em pauta? Trata-se, aqui, de resumir as contribuições mais importantes sobre o assunto escolhido, e não de apresentar uma extensa lista de referências, muitas vezes remotamente relacionadas ao tema da pesquisa com a finalidade de impressionar juízes (Este é um erro muito comum!). A melhor maneira de não ter "problemas com o problema", isto é, com o enunciado, a justificativa, o contexto e a revisão bibliográfica, é ler o máximo possível sobre o assunto. É impossível definir um problema sem conhecer o assunto. Para isso, deve-se começar lendo aspectos gerais do tema e, progressivamente, chegar-se a artigos, em revistas especializadas, diretamente relacionados com o que se pretende investigar.

III - Referencial teórico - A falta de um referencial é provavelmente a maior falha de muitos projetos de pesquisa em educação. Uma pesquisa feita sem fundamentação teórica, sem um sistema de referência teórico, é uma pesquisa feita ao acaso e, muitas vezes, não passa de um simples levantamento. Qual é a teoria, ou teorias, que fundamenta(m) as hipóteses formuladas em relação ao problema estudado? Como os resultados obtidos poderão realimentar a teoria, no sentido de contribuir para sua comprovação, enriquecimento, modificação? Quais são os conceitos-chave envolvidos na pesquisa? Qual é o sistema conceitual que fundamenta a pesquisa proposta? É possível traçar um mapa conceitual para a investigação?

- III.1 Hipóteses. Devem ser formuladas de um modo natural, da maneira como o pesquisador pensa que ocorrerão. Não devem ser formuladas como

hipóteses estatísticas. Quando testadas, como contribuirão para a solução do problema? São coerentes com o referencial teórico? Tal como estão formuladas, contêm tanto as variáveis independentes como as dependentes? Estas variáveis podem ser medidas com os recursos disponíveis para a pesquisa?

IV – Procedimentos

- IV.1 **Suposições e limitações.** Quais são as suposições subjacentes à pesquisa que está sendo proposta? Qual é a evidência de que são justificadas? Quais são as limitações do estudo?
- IV.2 **Delineamento.** Qual é o tipo de delineamento? Experimental? Qual? Quase-experimental? Qual? Outro? Qual?
- IV.3 **Definições.** Qual é o significado atribuído aos diferentes termos que serão usados na pesquisa? Qual é o significado das variáveis escolhidas? Dos construtos?
- IV.4 **Coleta de dados.** Qual é a fonte de dados? Se for uma amostra, qual é a sua composição? Qual é o tipo de amostragem? Quais são os instrumentos necessários? Como serão validados? Serão feitas entrevistas? Gravações? Haverá um trabalho de campo? Como os dados responderão às questões das hipóteses?

V – Análise dos dados – Que métodos serão usados? Como serão apresentados os resultados da análise? Quais são as possíveis interpretações? Quais são as técnicas estatísticas que serão utilizadas? A análise será qualitativa? Como?

VI – Exequibilidade – Que experiência tem o proponente na área de pesquisa proposta? Com a população-alvo proposta? Que

dados já estão disponíveis? Que tipo de cooperação tem assegurada por parte de agências e pessoas? Em que quantidade? Quais são os recursos financeiros disponíveis para coleta e análise de dados, se for o caso? Para viagens, material de consumo, digitação, etc.?

VII – Cronograma – Quando cada etapa da pesquisa será completada? O cronograma é realista? Está atrelado ao recebimento de recursos? Se os recursos atrasarem ou forem cortados, quais serão as consequências no cronograma? Cabe aqui lembrar que a razão planejamento/execução é muito importante. Em geral, as pesquisas mais triviais são aquelas planejadas em uma semana e executadas durante meses ou anos a fio.

VIII - Orçamento – Quais são os recursos financeiros necessários? Estão discriminados de acordo com as rubricas usuais (remuneração de serviços pessoais, material de consumo, material permanente, equipamento, serviços de terceiros, etc.)? Há previsão de desvalorização da moeda? O orçamento é realista? Não está subestimado ou superestimado (erros muito comuns)? É compatível com o cronograma?

IX – Relatório – Como será divulgado o trabalho feito? Através de artigos? De uma dissertação ou tese? De um relatório para a agência financiadora? Quando? (Está previsto no cronograma?)

Conclusão

O roteiro aqui apresentado não é uma receita, mas apenas uma sugestão que pode ser adaptada a cada caso. O importante é que o projeto seja coerente, apresentando-se como um conjunto equilibrado que diga claramente o que vai ser feito, como e quando será feito, com que recursos e por que é importante fazê-lo. Afora orçamentos incompatíveis com os recursos da agência financiadora, erros técnicos ou falta de evidências de viabilidade,

a ausência de relevância é, provavelmente, a maior causa de não aprovação de projetos.

Bibliografia

Becker, Fernando et al. (1984). *Apresentação de trabalhos escolares*. 7^a ed. Porto Alegre: PRODIL. 67 p.

Luft, Celso P. (1979). *O escrito científico: sua estrutura e apresentação*. 4^a ed. Porto Alegre: Lima. 54 p.

Moura Castro, Cláudio. (1976). *Estrutura e apresentação de publicações científicas*. São Paulo: McGraw-Hill do Brasil. 70 p.

2. Colcha de retalhos. Parece ser o tipo preferido dos principiantes. Caracteriza-se por um conjunto não articulado de ideias, pesquisas, resultados de diversos autores, sem um fio condutor capaz de guiar o leitor. Os estudos e pesquisas são meramente arrolados sem qualquer elaboração comparativa ou crítica, o que, geralmente, indica que o próprio autor se encontra tão perdido quanto seu leitor. Às vezes, percebe-se a existência de um roteiro, mas, como nos filmes de suspense, nunca se sabe bem aonde se quer chegar e o final ou é inesperado ou não consegue convencer. Quem faz esse tipo de revisão faz também um exercício de ventriloquismo, pois só fala pela boca dos outros: os parágrafos se sucedem, monotonamente, sempre começando por "para fulano", "segundo beltrano", "conforme sicrano", "de acordo com mengano", ou expressões equivalentes, sem comparações entre os citados, nem análises críticas, tomadas de posição ou resumos conclusivos.

3. "Light". É a revisão em que assuntos complexos são tratados de modo superficial, geralmente a partir de fontes secundárias, de preferência tipo enciclopédias de mesa, introduções ao assunto, edições de bolso ou obras tipo "Piaget para leigos". Esse tipo de revisão se presta também para dar um "jeitinho" de sempre citar autores que estão na moda, gerando o que se poderia chamar de "autores curinga", pois encaixam-se em qualquer revisão de literatura, seja qual for o tema estudado. Quer dizer, é tudo tão superficial, tão "leve", que sempre dá para conciliar o inconciliável, podendo também resultar no conhecido *coquetel teórico*, tão do agrado dos iniciantes: mistura-se todos os autores disponíveis e serve-se rapidamente antes que se perceba que muitos não são miscíveis ou que fique intragável. Esta união de esforços é muito sedutora para tentar dar conta de uma possível indisciplina de dados: cada teórico encarregar-se-ia de explicar uma parte dos dados.

4. Fontes anônimas. Em revisões dessa natureza, os autores garantem anonimato às fontes, valendo-se de expressões do tipo "sabe-se que", "tem sido observado", "vários estudos têm mostrado", "muitos autores têm afirmado" e outras semelhantes. Pode refletir inexperiência de quem faz, mas se promove também a apropriação indébita, i.e., plágio de textos de outros autores. O leitor fica sem saber o que é contribuição do autor da revisão e o que é das fontes não identificadas. Lembra algumas más reportagens jornalísticas.

5. Inútil, para constar. É o tipo de revisão que o pesquisador faz porque julga (corretamente) que uma pesquisa deve ter revisão da literatura e/ou teórica, mas depois nunca mais volta a ela. Quer dizer, não a utiliza na interpretação dos dados ou em qualquer outra parte da investigação.

6. Fiel às origens. A característica básica desse tipo de revisão é a fidelidade, ou lealdade, a determinadas origens. Por exemplo, quem fez doutorado na França cita somente pesquisas e autores franceses, ignorando, por exemplo, estudos e autores norte-americanos, e vice-versa. Também encaixa-se nesse tipo a *revisão xenófoba*, i.e., aquela que cita somente autores nacionais, mesmo que estes apenas repitam o que foi dito por estrangeiros, ainda que a produção nacional seja insuficiente. Oposta a esta é a revisão que menciona somente pesquisas e autores estrangeiros, omitindo a produção nacional sobre o tema.

- No caso de siglas, na primeira vez em que forem usadas, o nome deve precedê-las e elas colocadas entre parênteses; depois, basta a sigla.

ex.: Física Moderna e Contemporânea (FMC); Exame Nacional do Ensino Médio (ENEM)

Palavras do inglês parecidas com palavras do português, mas com outro sentido:

absolutely	(certamente)
actually	(realmente)
appointment	(encontro)
cigar	(charuto)
cigarette	(cigarro)
comics	(quadrinhos)
disparate	(diverso)
eventually	(por fim, finalmente)
exquisite	(requintado)
gas	(gasolina)
just	(mal, apenas)
lecture	(conferência)
library	(biblioteca; livraria é <i>bookstore</i>)
parents	(pais)
physician	(médico; físico é <i>physicist</i>)
policy	(política, prudência; polícia é <i>police</i>)
prejudice	(preconceito)
pretend	(fingir)
realize	(perceber)
record	(registrar, gravar)
relatives	(parentes)
retire	(aposentar-se)
scholar	(pessoa ilustrada, erudita)
spirits	(bebidas alcoólicas)
faculty	(corpo docente)

college (faculdade)
 college students (estudantes universitários)
 journal (revista científica, periódico; jornal é *newspaper*)
 professor (professor universitário, titular; professor de ensino básico é *teacher*)
 significant (significativo, mas no caso de aprendizagem significativa usa-se *meaningful learning*)

Algumas expressões úteis em inglês e espanhol

Aprendizagem significativa: *meaningful learning*; *aprendizaje significativo*

Formação de professores: *teacher education*; *teacher preparation*; *formación de profesores*.

Licenciandos (futuros professores): *pre-service teachers*; *futuros profesores*

Professores de Ensino Médio: *high school teachers*; *secondary school teachers*; *profesores de enseñanza media*

Professor de Ensino Fundamental: *elementary school teachers*; *maestros*

Palavras do espanhol parecidas com palavras do português, mas com outro sentido:

Vaso (*copo*)

Exquisito (*de fino sabor*; *de qualidade extraordinária*)

Pesquisa (*averiguação*; *investigação policial*)

Bachillerato (*escola secundária*, *pré-universitária*)

Críticas a projetos

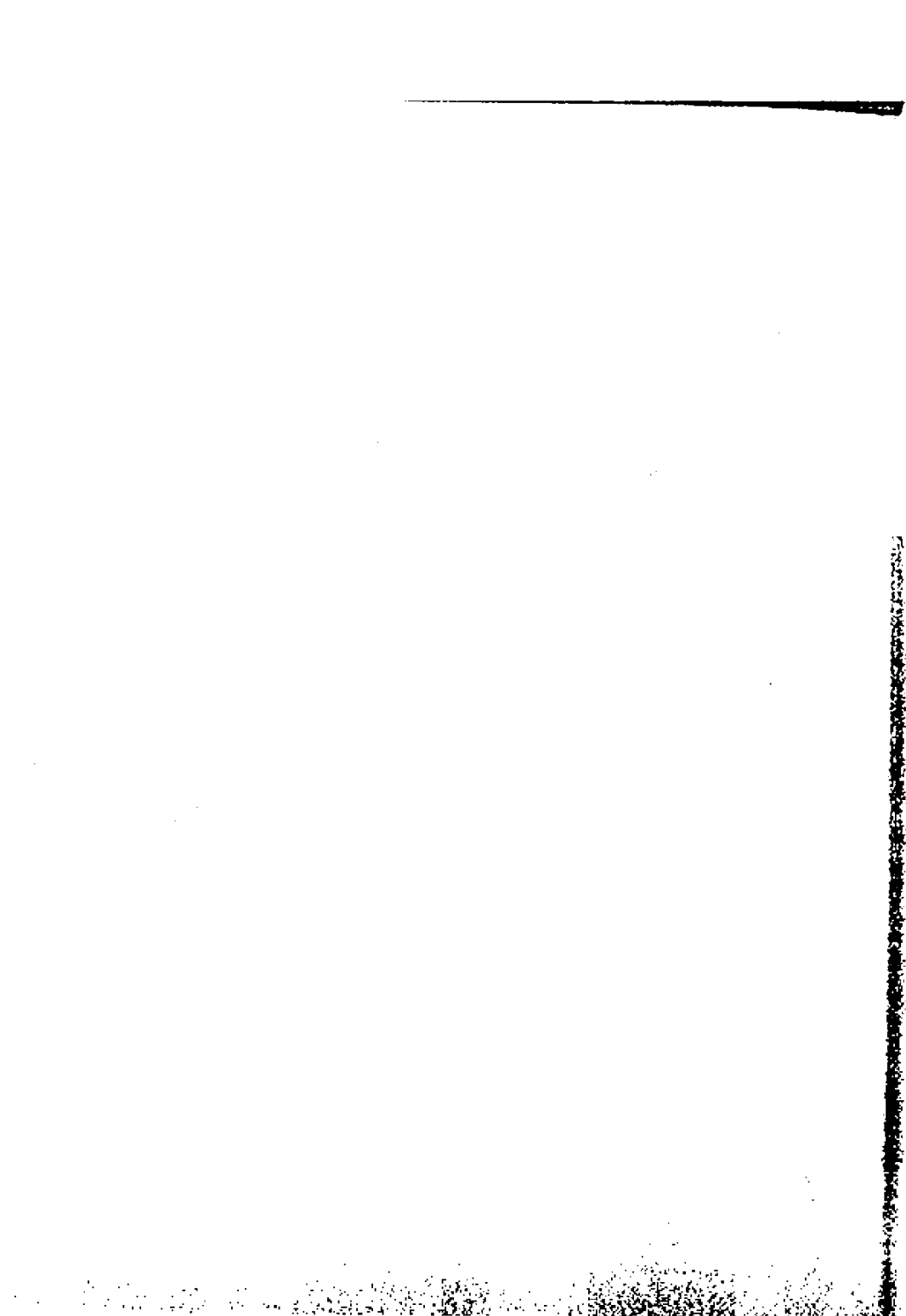
- Alguns projetos não têm, de fato, referencial teórico só discurso.
- Alguns projetos não se preocupam em avaliar se, de fato, haverá alguma aprendizagem ou mudança atitudinal.
- A metodologia de pesquisa em alguns casos se restringe a mencioná-la, classificá-la.

Dicas para projetos

- Estrutura:
Introdução/Problematização/Contextualização/Questão de Pesquisa
 - Estudos relacionados/Revisão da Literatura
 - Referencial teórico/Fundamentação teórica
 - Metodologia (tipo de pesquisa, coleta de dados, variáveis, tratamento/análise dos dados)
 - Recursos necessários/Viabilidade/Cronograma
 - Divulgação dos resultados/Produto
 - Referências
- O projeto não pode ter erros.
- Não escrever mais (nem menos) do que é preciso, não ser prolixo.

Dicas para artigos

- Figuras, tabelas e gráficos sempre com legendas e referidos no texto.
- Não usar "abaixo", "a seguir"; usar sempre o número da figura, tabela ou gráfico.
- Numerar Anexos e Apêndices
- Preferivelmente, usar normas internacionais nas referências.
- Evitar problemas éticos (e.g., identificação de alunos)



Este livro é uma recopilção de textos de apoio sobre metodologias de pesquisa em ensino, particularmente de ciências. Foram escritos para serem usados independentemente uns dos outros. Por consequência, apresentam certas superposições.

O primeiro capítulo é mais abrangente, enfocando tanto aspectos quantitativos como qualitativos da metodologia de pesquisa em ensino. O segundo trata somente de métodos qualitativos enquanto o terceiro se ocupa apenas de métodos quantitativos. O quarto, por sua vez, aborda a redação de trabalhos científicos, já que a pesquisa deve sempre gerar publicações científicas. Ainda que alguns tópicos possam estar repetidos nos diferentes capítulos, os textos são complementares e podem auxiliar na diferenciação progressiva de alguns conceitos.

LF



EDITORA

www.livrariadafisica.com.br

ISBN: 978-6578655303



9 786578 611101

