

“Teoria e métodos da investigação social”

Johan Galtung

Capítulo I - A Matriz de dados¹

1.1. A forma tripartida dos dados

Obtém-se dados sociológicos quando um sociólogo registra fatos sobre algum setor de realidade social ou recebe fatos registrados para ele. Costuma-se dizer que os dados devem conter algum tipo de referencial empírico do próprio sociólogo e vamos respeitar esta tradição.

Apesar de que os dados da sociologia, devido a enorme extensão do seu objeto, adaptam muito diferentes formas. Há uma estrutura comum que em muitos casos se evidencia rapidamente e em outros impõe com facilidade aos dados. Esta estrutura tem três partes, mas elas aparecem com nomes distintos em diferentes contextos.

Em primeiro lugar, e é o mais importante, temos os elementos ou unidades de análise, tais como seres humanos (questionados) em uma enquete ou em um experimento de laboratório, ou nações em um estudo comparativo de nações ou unidades mais complexas. Essas unidades têm uma característica em comum enquanto objeto da investigação sociológica. Elas consistem em seres humanos ou em produtos dos seres humanos, como é no caso da análise do conteúdo de recortes de jornais.

Em segundo lugar, existem as dimensões ou variáveis que alguém deseja conhecer sobre as unidades, tais como o consumo de energia *per capita* nos estudos sobre as nações subdesenvolvidas ou as respostas de certas perguntas estratégicas

em um estudo de uma eleição. Algumas vezes isto mesmo pode ser formulado com mais propriedade como um conjunto de condições com as quais se deseja estudar as unidades ou como um conjunto de estímulos. Utilizaremos a expressão ‘variáveis’ quando estivermos tratando unidades que sejam imponderáveis, no sentido de que elas não podem (de uma maneira artificial ou natural) ser submetidas a um conjunto de condições, de maneira que o sociólogo possa estudar os resultados. Utilizaremos, também, as expressões ‘condição’ ou ‘estímulos’ no caso que isto seja possível.

Em terceiro lugar, teremos os valores que alcançam as unidades nas variáveis estudadas ou, utilizando as outras expressões, as respostas dos resultados que se obtém quando as unidades expõem aos estímulos ou às condições segundo as quais o sociólogo deseja estudá-las.

Deste modo, teremos as três partes que podem se dizer que constituem a estrutura dos dados sociológicos, expressas na linguagem das variáveis, na linguagem das condições, dos estímulos e respostas (linguagem SR) e no que se pode denominar linguagem das enquetes:

Tabela 1.1.1. As partes dos dados expressos em termos diferentes

Termos gerais	Dimensões	Unidades	Valores
Linguagem de variáveis	Variáveis estudadas	Unidades de análise	Valores das variáveis
Linguagem SR	Condições, estímulos	Unidades de análise	Resultados, respostas
Linguagem das enquetes	Objetos sociológicos	Sujeitos questionados	Respostas

É difícil escolher entre estas terminologias, pela simples razão de que elas não são totalmente sinônimas, de tal maneira que teremos que usar algumas

¹ Tradução publicada, com autorização do autor, do primeiro capítulo do livro: “Teoría y Métodos de la Investigación Social”. EUdeBA, Buenos Aires, 1969. Tradução: Cristina Silveira. Revisão da

tradução: Marisa Helena Degasperí. Colaboração na transcrição: Stephanie da Silva Freitas.

combinações dos termos ou selecionar a linguagem que pareça mais apropriada para a ocasião. Para evitar muitas das dificuldades que surgem disto e evitar diferenças que são, muito frequentemente, desnecessárias, introduziremos alguns símbolos simples:

Tabela 1.1.2. Símbolo para as três partes dos dados sociológicos

Parte dos dados	Dimensões, variáveis	Unidades de análise	Valores, respostas
Símbolo para um elemento	S,V,X	O	R
Símbolo para uma classe de elementos	S,V,X	O	R
Símbolo para o número	n	m, N	r

Para dar um exemplo, a pesquisa corrente do tipo enquete começa propondo a m questionados (O) n perguntas (S) ou ‘objetos psicológicos’ segundo a terminologia de Thurstone, de tal maneira que, para cada pergunta S e para cada sujeito interpelado O há r possibilidades de respostas diferentes, R. Outro exemplo: pode-se fazer um estudo comparativo de m nações, colocando em uma lista seus valores R sobre uma série de n variáveis de desorganização social, S, tais como alcoolismo, doença mental, criminalidade geral, homicídio e suicídio, para um determinado ano. Se, além disso, se estuda a mudança nas escalas ao longo de um período de anos, de modo que nos fatos pode-se obter uma tabela de três dimensões, se conserva a estrutura original tripartida, multiplicando as cinco escalas de desorganização por, digamos, os 15 (quinze) anos, de tal maneira que no total se conseguem 75 variáveis (ou condições). A estrutura dos dados permanece idêntica: se estudam unidades de análise com um número determinado de dimensões.

Se admitimos, no momento, que muitos dados podem se apresentar como se estivessem compostos destas três partes, podemos passar ao próximo conceito: a matriz de dados. A matriz de dados é um modo de organizar os dados de maneira que seja particularmente visível a forma tripartida:

Tabela 1.1.3. A matriz de dados

O diagrama mostra uma matriz de dados organizada em linhas e colunas. As linhas são rotuladas O₁, O₂, O₃, ..., O_i, ..., O_m. As colunas são rotuladas S₁, S₂, S₃, ..., S_j, ..., S_n. Os elementos da matriz são rotulados R₁₁, R₁₂, R₁₃, ..., R_{1i}, ..., R_{1n} na primeira linha; R₂₁, R₂₂, R₂₃, ..., R_{2i}, ..., R_{2n} na segunda linha; R₃₁, R₃₂, R₃₃, ..., R_{3i}, ..., R_{3n} na terceira linha; e R_{m1}, R_{m2}, R_{m3}, ..., R_{mj}, ..., R_{mn} na última linha. O elemento R_{ij} na linha O_i e coluna S_j está circulado.

Neste esquema, R_{ij} é a resposta que da unidade n^oi ao estímulo n^oj ou, em outras palavras, o valor que tem a unidade n^oi na variável n^oj.

Esta é exatamente a forma em que aparecem os dados se há um cartão perfurado para cada unidade, uma coluna para cada variável e uma perfuração para cada valor e se passam os cartões por uma máquina que registra, para cada um, o que está perfurado nela. Além disso, esta é a forma que reveste o tipo de dados que tem um diretor de escola em seu registro, com os nomes dos alunos dispostos horizontalmente e a qualificação obtida pelo estudante n^oj colocado na célula apropriada.

Em termos técnicos, a matriz de dados nos dá, para cada par no produto cartesiano O por S, o valor correspondente em R para cada combinação O, S. A forma da mesma matriz de dados e esta definição nos conduzem, imediatamente, aos primeiros 3 princípios dos dados coletados:

- 1) **Princípio da comparabilidade:** a proposição '(O_i, S_j) dá o valor correspondente em R_k' deve ser verdadeira ou falsa para cada i, j e k.
- 2) **Princípio da classificação:** para cada estímulo S_j, a série das categorias de respostas R_{jk} deve produzir uma classificação de todos os pares (O_i S_j) (i=1, ..., m).
- 3) **Princípio da integridade:** para cada par (O_i S_j) deve achar-se, empiricamente, um valor R_k.

Estes três princípios exigem alguns comentários:

Se observada uma coluna qualquer da matriz M na tabela 1.1.3, se vê que a ideia é, simplesmente, que o mesmo estímulo deve ser apresentado a todas as unidades de análise; por exemplo: deve-se formular a mesma pergunta a todos os pesquisados. O importante é que seja a mesma pergunta; o estímulo não deve ser mudado de um objeto a outro. Isto, entretanto, é só o que já estava na definição da matriz.

O **princípio da comparabilidade** acrescenta, agora, que cada uma das combinações apresentadas de um estímulo determinado com uma unidade de análise deve ter sentido; ou seja, deve ser verdadeiro ou falso que uma combinação (O_i S_j) de um valor determinado R_{jk}. Através desse princípio se podem comparar os estímulos, as unidades e as respostas. Se O_j é uma nação, S_j uma pergunta direta sobre o estado civil e R_{jk} a lista dos elementos "casados, solteiros, viúvos ou divorciados", então qualquer combinação não é certa nem falsa, senão que, simplesmente, não tem sentido, como não teria sentido perguntar sobre o sexo de uma nação.

Portanto, se obtém a comparabilidade quando as três séries se ajustam umas às outras de tal maneira que satisfaça a condição mencionada no princípio nº1. Isto quer dizer, por exemplo, que as perguntas devem ser adequadamente traduzidas numa pesquisa comparativa, mas

não quer dizer que uma pergunta numa pesquisa 'deva ter o mesmo significado' para todos os interpelados. É suficiente que tenha algum significado. A comparabilidade, em nossa acepção, é um requisito menos exigente; apesar disto, significa que uma linha pode ser comparada com outra porque as R nas células não carecem de sentido.

Se for descoberto que duas unidades são diferentes quanto a uma dimensão S, se pode decompor a diferença em dois componentes, um 'real' e outro 'semântico' e atribuir a diferença observada a alguma função destes dois componentes. É possível, inclusive, sustentar que toda a diferença se deve a um componente semântico, como no caso em que exista uma mera discordância verbal e não uma mera discordância dos fatos.

Pode também acontecer o caso de que os componentes semântico e real se anulem reciprocamente de tal maneira, que o resultado final seja uma semelhança observada que pode se qualificar como enganoso. Mas, tudo isto é um problema de interpretação e não pertence ao campo da coleta de dados, salvo se decidido reduzir ao máximo o componente semântico. Deve-se notar, no entanto, que não é evidente por si mesmo que o fator semântico constitua um 'erro' em um estudo do tipo enquete ou que as escalas de criminalidade das nações não possam ser comparadas porque se baseiam em códigos penais diferentes. Uma frase ambígua é, contudo, uma frase, e uma investigação sobre como as pessoas reagem ou uma frase deste tipo pode produzir informação importante sobre como reagiriam a tal frase na vida diária e, ainda, que as escalas de criminalidade variariam se fossem baseadas em um critério universal. Podem-se realizar muitas investigações baseadas na comparação de escalas criadas em definições diferentes, pela simples razão de que as escalas podem nos dizer algo sobre como muda de um país a outro 'isso que se chama crime'.

O **princípio de classificação** agrega mais três limitações para a seleção das séries S, O e R. O que diz este princípio é que, para cada estímulo S_j , a série R_{jk} deve formar uma classificação. Por exemplo: para cada par (O_i, S_j) deve haver uma (exaustividade) e só uma (exclusão mútua) R_{jk} em que possa ser localizado. (É ideal que também deva haver um *fundamentum divisionis*). Quando se satisfaz este princípio, na prática, estamos seguros de que há uma e só uma resposta para cada combinação possível de unidades e estímulos. Na construção de questionários, o mesmo se conhece como o princípio segundo o qual cada pesquisado deve marcar uma e só uma resposta ou categoria. Quando se concede ao interpelado a chamada múltipla escolha, se está dando nos fatos uma série de estímulos, um para cada categoria, de tal maneira que há duas respostas para cada estímulo, aceitação ou rejeição. Esta é uma classificação tão boa como qualquer outra; é uma dicotomia.

Dos três princípios mencionados, os dois primeiros se referem à lógica da matriz e as três séries sobre as quais ela se baseia. O terceiro: **princípio de integridade**, se refere ao trabalho empírico de preencher a matriz e diz simplesmente isso: não deixar nenhuma célula vazia. Este *desideratum* se expõe facilmente, mas não se alcança tão facilmente na prática, de tal maneira que tudo o que contém o princípio é uma norma segundo a qual a quantidade de células que contenha ‘nenhuma informação’, ‘nenhuma resposta’, etc., deve ser mantida o mais baixo possível. Somente se poderá cumprir com o terceiro princípio no processo de coleta de dados e com posterioridade, quando se puder cumprir com os dois primeiros princípios antes da coleta de dados, com o cuidado de que os estímulos e as variáveis se apliquem a todas as unidades e que todas as séries de categorias de

respostas sejam classificadas. Aqui entram em jogo coisas como as técnicas para manter baixas as escalas de rejeição através de uma legitimidade adequada da investigação ou as maneiras de rastrear dados quase impossíveis de se obter. Mas existe também a possibilidade *post hoc* de examinar a distribuição das células vazias em M e de eliminar as O e S mais adequadas para liberar-se do máximo de células vazias com o mínimo de supressões. Como regra geral o tipo prático, pode dizer-se que 10% é o máximo absoluto de células vazias admissível em qualquer coluna ou qualquer fila de M e que 5% é um máximo mais aconselhável.

Em alguns casos, contudo, pode haver algum valor em manter uma O ou uma S na matriz, mesmo que infligindo os dois princípios relativos às células vazias; uma razão é que "nenhuma resposta" também pode ser uma resposta e se possa tratá-la como tal.

Generalizando, a técnica consiste, por tanto, em contar o número de ‘DKs’ e de ‘NAs’ para cada O e para cada S e obter sua distribuição². Muito frequentemente, uma pequena porcentagem de O e de S absorve a uma grande porcentagem das ‘DKs’ e das ‘Nas’ e, naturalmente, estas são nossas candidatas óbvias para a rejeição. Somente em raras oportunidades as ‘DKs’ e as ‘Nas’ estarão distribuídas ao acaso na matriz de dados.

Se, como de costume, m é o número de unidades, n é o número das dimensões e r é o número intermediário das categorias de respostas na representação $W = m \times n \times r$, o; melhor escrita deste modo: (m, n, r) , é a característica da matriz. W é um símbolo que dá uma ideia de quanto trabalho há em um processo de coleta dos dados, ainda que deva se dar mais importância a m e a n que a r ; r é uma medida de quanta medição discriminatória se faz e, por tanto, deve ser

*DK’, forma abreviada de “*don’t know*”, “não sei”, ‘NA’, forma abreviada de “*no answer*”, falta de resposta (*nota do tradutor*).

igual, pelo menos, a 2. Se $r = 1$, não pode haver nenhuma discriminação entre diferentes unidades e a dimensão não tem valor descritivo nem analítico, já que o resultado sempre é conhecido e conhecido como sendo sempre igual. Com certeza, r pode variar e em geral varia de variável em variável, de tal maneira que r deve ser concebido como algum tipo de valor intermediário entre as diferenças.

A escolha básica, em qualquer plano de pesquisa terá que se referir aos números de m e n . Um projeto com as dimensões (300, 20) é, obviamente, muito diferente de um projeto de tipo (20, 300). O pesquisador terá que escolher. Chamaremos o par ordenado (m,n) a estratégia de investigação do projeto, o pesquisador tem que escolher sua estratégia para atacar o problema que encontrar e na próxima sessão se dedica a fazer algumas reflexões sobre os diferentes tipos de estratégias.

Por último, diremos algumas palavras sobre como r depende de m e n . Não há nenhuma razão pela qual deva depender de n : adicionar uma variável não muda o número de valores das variáveis anteriores. Contudo, se pode considerar a r como dependente de m . Falando de forma restrita, sempre podemos ter $r \leq m$, haja vista que não necessitamos mais valores que elementos ou unidades para medir. Se tivermos uma unidade, necessitamos um só valor; não há que especificar o resto (portanto, a variável consistirá nesse valor e em todos os outros valores possíveis agrupados). Entretanto, na medida em que m aumenta, r aumentará cada vez menos. Logo estará evidente a amplitude (alcance), e com mais unidades haverá mais dificuldades em discriminar entre elas e maior tendência a agrupá-las. Muitas vezes, a variável utilizada não será estritamente unidimensional, de modo que a localização exata das unidades a respeito às outras unidades será discutível, conflito que somente pode ser resolvido estabelecendo categorias mais grosseiras, isto é, baixos valores de r . O pesquisador procura capturar

a localização genérica e deixa de lado os detalhes e os problemas que lhe vem à mente, precisamente quando se encontra com muitas unidades.

1.2. O número de unidades e de variáveis

Os números m e n são de importância crucial para a avaliação de qualquer processo de coleta de dados, tanto de um ponto de vista teórico como de um ponto de vista muito prático. O valor mais baixo de m é 1, como no caso em que se investiga uma só pessoa ou uma só nação, e o valor mais baixo de n é também 1, como no caso das eleições, em que se formula uma só pergunta (Qual é o partido do seu candidato?). Geralmente, obtemos as seguintes combinações entre os valores possíveis de m e n :

Tabela 1.2.1- As combinações de valores de m e n

N= número de dimensões	muitas	Pesquisa 'intensiva' (Psicologia profunda; estudo de casos)	Psicologia	
	poucas	Jornalismo		Sociologia
	uma		Enquete	Pesquisa 'extensiva' (eleições, Gallup)
	uma	poucas	muitas	

M = número de unidades

Certamente, esta tabela deve ser aceita de forma ponderada. Nada aparece na diagonal principal, ainda que tais combinações possam ocorrer perfeitamente. Os tipos de pesquisa tipicamente intensiva e extensiva aparecem nos extremos, tendo como vizinhas células que caracterizam muito do que se encontra habitualmente nas ciências gêmeas da Psicologia e da

sociologia. Entretanto, existe a seguinte diferença: para o psicólogo a unidade de análise é quase sempre o organismo vivo; para o sociólogo pode ser, além disso, algum tipo de coletivo. Se o coletivo é a nação, talvez se possa colocar o historiador, com sua ênfase nos aspectos mais casuísticos das unidades de análise, no lugar onde se encontra localizado no quadro o psicólogo. Isto não significa que as enfoques localizados no canto superior esquerdo do quadro não possam dar origem a generalizações; contudo, na medida em que se deseja generalizar ao nível da unidade de análise, se faz necessária a repetição da totalidade da matriz de dados. Com uma investigação muito intensiva se podem obter generalizações utilizando as técnicas de amostragem e a inferência estatística ou, simplesmente, dividindo uma amostra grande em sub-amostra mais pequenas e colocando a prova as hipóteses em uma sub-amostra menor e colocando à prova as hipóteses em uma sub-amostra depois da outra. Por outro lado, não se pode dizer muito a partir do resultado de uma análise extensiva, visto que se se conhece tão pouco sobre de cada unidade.

O ideal é a combinação (muitas, muitas) tantas unidades e tantas dimensões que forem possíveis. No entanto, nós supomos que a palavra 'muitas' se utiliza de tal maneira que isto é impossível, por falta de recursos tais como tempo, energia, pessoas e dinheiro. A disponibilidade destes importantes fatores, junto com a habilidade no manejo de diferentes tipos de dados, levará sempre o investigador a direções específicas e é possível que se sintam tão condicionado na direção de um dos tipos de tabela 1.2.1. que virtualmente os outros nem existam para ele. Porém, qualquer combinação de m e n pode ser de algum valor, e uma pesquisa tradicional que inflexivelmente desqualifique uma ou mais combinações somente demonstra suas próprias limitações. Por outro lado, cada célula tem seu estilo próprio e exige

habilidades específicas que só raramente se encontram em uma mesma pessoa, mas que facilmente se se pode encontrar-se em uma equipe de investigação. Na realidade, parece óbvio que um problema dado em muitos casos pode ser enfatizado melhor por uma combinação adequada de métodos que por um método isolado ou combinação (m, n). Isto não é denegrir; o que acontece é que o terreno exige conhecimentos das diferentes enfoques e uma visão tolerante das vantagens e limitações das diversas combinações.

Ao invés de fazer uma lista com argumentos prós e contra de cada uma das novas células, faremos essencialmente o mesmo, mas de tal maneira que, ao mesmo tempo, se destaque o fundamento metodológico da importância da matriz de dados na pesquisa sociológica. Se $m=1$ ou $n=1$, se se diz que a matriz degenerou (em um vetor de linha ou de coluna, respectivamente, ou em uma resposta, se o resultado é que tanto m como n são iguais a 1). Existem boas razões que explicam porque estas matrizes de dados degenerados criam problemas em alguns aspectos.

Suponhamos que $m=1$, por exemplo, que escolhemos somente uma unidade para o estudo. Se tudo o que desejamos é conhecer esta unidade tanto quanto seja possível, nosso propósito pode estar justificado se tivermos selecionado exatamente a unidade que queremos estudar. Se, ao contrário, nosso propósito é estudar um sistema social e selecionamos um informante com esse fim, a unidade estudada não é a mesma que a unidade selecionada e, então, passa a ser problemático. Porque escolher precisamente este informante, quando sabemos que todos os informantes serão um pouco diferentes? Acontece o mesmo se, ao invés de estudar um sistema social, estamos estudando uma categoria social, uma unidade escolhida aleatoriamente ou sistematicamente será sempre só uma unidade e, ainda que a unidade possa

coincidir com o meio aritmético, ou ser o líder ou qualquer outra coisa, sempre haverá outras unidades diferentes da unidade escolhida.

Em química ou física, frequentemente, parece haver sido resolvido o problema de encontrar o caso puro. Se um químico deseja comprovar uma proposição sobre o ácido sulfúrico, pode tomar uma quantidade qualquer de ácido sulfúrico quimicamente puro (sempre que sua forma cristalina seja irrelevante - para o experimento) e tratá-lo como se fosse um puro e verdadeiro representante do ácido sulfúrico, S³. Se um sociólogo deseja estudar o Eleitor Norueguês, se simplificaria enormemente a investigação se fosse possível encontrar o eleitor puro, uma só pessoa que fosse representativa de todos os eleitores noruegueses, de tal maneira que tudo o que seria necessário fazer seria interrogá-lo ou observar seu comportamento. Na atualidade, parece haver desaparecido completamente a pesquisa social, a crença na possibilidade de encontrar o caso puro tanto no nível de análise individual como no coletivo.

Pode-se fazer a mesma argumentação sobre o caso n=1, por exemplo, quando se trata de uma investigação do tipo enquete, no caso em que se formula uma só pergunta. Vamos supor que se individualizou uma dimensão para seu estudo, tal como 'orientação geral cosmopolita *versus* orientação geral local'. Nos dias de hoje, a maneira normal de fazer isto constituiria em selecionar uma série de perguntas preparadas para 'tocar' esta dimensão, formular todas as perguntas e fundamentar a análise em algum tipo de índice cumulativo que se pode formar a

³ Isto não significa que a solução seja trivial ou que se alcance facilmente. O químico tem em sua mente um caso ideal, 'puro' ácido sulfúrico ou qualquer outra substância que se defina por ter só um tipo de moléculas. Este fato só se pode questionar indiretamente, pegando uma amostra do material e examinando o valor que uma certa quantidade de indicadores (características de fusão, refração, dispersão, etc.) alcança. Entretanto, os indicadores

partir das perguntas formuladas como item; assim não se limita a enquete a uma só pergunta. Na verdade, porque exatamente essa pergunta? Porque não poderia ser outra pergunta? ou por quê, pelo menos, não complementar a informação proporcionada pela primeira pergunta? Contudo, existem exceções a este princípio de muitas perguntas para a mesma dimensão. Quando se busca simples informação sobre o contexto, ou quando se investiga uma dimensão de atitude que ocorreu muito frequentemente no debate diário, quem investiga através de uma enquete parece justificado limitar-se a uma só pergunta. As enquetes de Gallup caem em uma destas categorias ou em ambas, o que as faz metodologicamente justificáveis deste ponto de vista. Por outro lado, no entanto, por esta mesma razão é difícil usar dados tipo Gallup em uma análise em que aparecem conceitos mais abstratos, de um tipo mais teórico e menos de senso comum, em que é necessário coletar dados.

Existem, além disso, outras razões de uma natureza mais positiva para evitar os casos degenerados e estas razões podem ser resumidas da seguinte maneira:

Tabela 1.2.2. Razões para evitar casos e, que M= 1 ou N= 1

Razão para	<u>Mais de uma unidade</u>	<u>Mais de uma dimensão</u>
Razão negativa	Não se crê que exista o caso puros. Por que essa unidade e não outra?	Não se crê que exista a pergunta pura. Por que não esse estímulo e não o outro?
O uso da dispersão	A existência de mais de uma unidade faz possível ver a diferenças, variações e evitar estereótipos	A existência de mais de uma pergunta ou dimensão faz possível ver

podem não discriminar suficientemente ou a amostra pode ser exigente demais. A metodologia não é muito diferente da metodologia das Ciências Sociais, mas o cientista social frequentemente não tem nem sequer um conceito de caso puro em sua mente nem, muito menos, um método para aproximar-se operacionalmente dele.

		padrões de respostas e ver uma resposta o valor ou contexto
O uso da correlação	A correlação de respostas a diferentes perguntas é uma das ferramentas principais de análise e pode ser utilizada somente com um certo número de unidades e pelo menos com duas dimensões.	

A terceira razão é de uma importância angustiante para o analista que está acostumado a trabalhar com dados estatísticos (estas é outra palavra para se referir a dados que podem ser organizados na forma de uma matriz de dados completa), e a série completa de argumento, em certo sentido, é para ele a própria base sobre a qual trabalha. Não há nada que possa substituir o tipo de análise que possa fazer, mas se pode, naturalmente, argumentar que existem outros enfoques valiosos baseados em tipo de dados em que $m=1$ ou $n=1$. Num certo nível, na análise sempre se está estudando só uma unidade, como quando se faz uma enquete sobre o homem norte-americano e sua comportamento sexual. O estudo de comunidade se dirige a uma comunidade, mas dentro desta comunidade é habitualmente aconselhável fazer uma matriz de dados completa baseada em unidades e dimensões individuais adequadamente escolhidas. Porém, o

⁴ Sorokin, nas páginas 57 e seguintes de *Social and Cultural Dynamics*, Boston, Porter, Sargent, 1957, formula a seguinte observação contra a posição de que os eventos humanos são únicos: “Não se pode narrar nenhum processo histórico sem admitir explícita ou implicitamente, que muitas características essenciais se repetem. Tomemos, por exemplo, a religião romana. Se qualquer momento de qualquer evento histórico ou processo sociocultural fosse único, imediatamente surgiria uma dificuldade: Que momento da religião romana tem que se descrever? Deve ser a religião romana das 20h do dia 1º de julho do ano 321 a.C. ou a das 7h do dia 10 de maio do ano 322 a.C.? Há invariantes na “unicidade”, pois, de outra maneira, não seria possível a descrição; mas também há variações, e nosso enfoque proporia cortar o processo em pedaços adequados, unidades e recolher tanta informação quanto fosse possível sobre cada uma,

interesse em uma unidade não implica crença na posição Einmalig Keit⁴.

Um dos exemplos mais famosos da ideia de estudar em fenômenos através do estudo de um caso puro se encontra em *As formas elementares da vida religiosa*, de Durkheim. A introdução contém uma brilhante defesa de seu método, ainda que possa parecer menos aceitável hoje em dia do que quando foi publicado pela primeira vez⁵.

Diz já no começo:

“Propusemo-nos a estudar neste livro a religião mais primitiva e mais simples que se conhece na atualidade, analisá-la e tentar identificar sua explicação. Dizemos de um sistema religioso que é o mais primitivo do que se dispõe quando cumpre as duas condições seguintes: em primeiro lugar, é necessário que se encontre em sociedades cuja organização não é superada em simplicidade por nenhuma outra; além disto, é necessário que seja possível explica-la sem deixar interferir nenhum elemento tomado de alguma religião anterior”⁶.

Mais adiante diz:

“Na base de todos os sistemas de crenças e de todos os cultos, deve necessariamente existir um certo

incluindo o tempo. Cortes diferentes produzirão resultados diferentes e haverá que combinar os resultados. Do mesmo modo, na ciência social sincrônica: se fariam cortes no espaço (veja-se 2.2) e se explorariam, analiticamente, as variações. Portanto, para colocar à prova hipóteses, se obteriam uma amostra das unidades de interesse e se submeteriam, sistematicamente, a prova da hipótese. McClelland, em *The Archeving Society*, contrasta habilmente esta com a metodologia tradicional do encontro-ilustração, dentro do qual se escolhem os casos que estão de acordo com a hipótese e se rejeitam hipóteses se se encontra um só caso desviado. Esta é uma concepção muito ingênua sobre as proposições da Ciência Social. Se só se admitissem correlações perfeitas, a Ciência Social não iria muito longe.

⁵ Paris, 1912.

⁶ *Ibidem*, pág. 1.

número de representações fundamentais e de atitudes rituais que, apesar da diversidade de formas que umas e outras puderam se revestir, tem em todas partes a mesma significação objetiva e cumprem em todas partes as mesmas funções. São estes elementos permanentes os que constituem o que há de eterno e humano na religião; eles são todo o conteúdo objetivo da ideia que se expressa quando se fala da religião em geral”⁷.

Aqui temos duas ideias importantes. Em primeiro lugar, a divisão de uma instituição social em duas partes que podem se chamar o essencial e o acidental e a simples equação:

$$\begin{array}{rcccl} \text{Instituição Social} & = & \text{Parte} & + & \text{Elementos} \\ \text{coisa social} & & \text{essencial} & & \text{acidentais} \end{array}$$

O essencial, evidentemente, representa a condição *sine qua non*, “da ideia que se expressa quando se fala de religião em geral”. Durkheim expressa essa equação com um discurso semelhante ao nosso quando diz:

“A imaginação popular ou sacerdotal não teria sido ainda nem a tempo nem a medida de refinar e transformar a matéria prima das ideias e das práticas religiosas, ela se mostra, por tanto, claro e se oferece por si mesma a observação, e que não tem que fazer maior esforço para descobri-la. O acessório, o secundário, os progressos do luxo não vão sobreviver a um ocultar do principal. Tudo está reduzido ao indispensável, aquele sem o qual não poderia haver religião. Mas o indispensável é também o essencial, é dizer, o que nos importa conhecer diante do todo”⁸.

Em segundo lugar, temos a ideia da parte essencial se encontra em sua forma mais pura nas sociedades primitivas e, em geral, enquanto mais primitiva a sociedade, melhor. Em outras palavras, a ideia é que a sociedade, de alguma maneira, parte com a “matéria prima”, e a esta se agrega, por processos de evolução e difusão “o acessório, o secundário, os progressos de luxo”.

Já que nada parece ter refutado uma das proposições básicas dos evolucionistas, que a mudança social tende a produzir-se em direção a formas mais complexas e não a formas menos complexas (com divisão do trabalho e sua concomitante, a complexidade crescente da rede de *status*, como uma das dimensões), não se negará que algumas instituições sociais podem ser mais simples em níveis mais baixos e a organização social em geral. Para isto, inclusive sendo mais de mais possível defesa do que o normal, estamos dispostos a admitir que não é o mesmo que dizer que estas instituições são mais puras. A simplicidade não é o mesmo que a pureza.

Das ideias muito fundamentais que podem facilitar a aceitação da ideia de Durkheim de trabalhar com o caso puro, é que conduziu o estudo intensivo da tribo de Los Arunta com dados de segunda mão. Em primeiro lugar, a clássica diferença entre as universais e os particulares, que corresponde a nossa distinção entre o essência e o acidental, uma diferença entre a ideia geral e as características adicionais não essenciais (tal como a famosa ‘cor’ do cavalo, exatamente ‘que’ objetos se definem como sagrados, etc.). Se agora introduzimos a igualmente clássica discussão entre realistas e nominalistas, podemos pensar em um contínuo de atitudes neste conflito, ao invés dos dois extremos tradicionais; a saber: o extremo realista *ante rem* (os universais têm uma existência independente anterior as coisas, que são suas manifestações) e o extremo

⁷ Ibidem, pág. 6.

⁸ Ibidem, pág. 8.

nominalistas *in rebus* (os universais não tem tal existência, são somente nomes de abstrações obtidas a partir das coisas). As seguintes seriam algumas posições intermediárias entre ambos os extremos.

Pode-se adaptar a posição de que “as coisas” (por exemplo, as indicadas da ideia, dos universais) podem ser organizadas de acordo com as suas particularidades, partindo desde as aproximações mais puras à própria ideia e adicionando a elas particulares, “o secundário”, para utilizar a terminologia ampliada, até chegar às formas mais complexas que exibem mais elementos acidentais. Se admitida a possibilidade chegar a uma organização linear ou parcial das ‘coisas’, isto só seria interessante se esta variável (grau de ‘pureza’) estivesse altamente correlacionada com alguma outra variável. Neste ponto se introduz a segunda ideia cultural básica: a ideia de que a variável está correlacionada com as etapas da evolução, qualquer coisa que ela seja. Ainda que Durkheim não tivesse esta ideia, teve a ideia de “sociedades em que a organização não é superada em simplicidade por nenhuma outra” e a ideia de que é aqui onde haveria de buscar para encontrar ‘as formas elementares da vida religiosa’. A dificuldade está em que a ‘simplicidade’, ainda que possa funcionar como uma característica global de uma sociedade, não parece funcionar para todas as partes da sociedade, como demonstram os sistemas de parentesco das sociedades primitivas. Apesar disto, a posição é interessante como uma espécie de posição *ante rem* modificada; no *ante* todas as instituições sociais, senão *ante* a existência das instituições religiosas em nossas sociedades.

Outra posição intermediária pode tomar como ponto de partida a equação

apresentada mais acima, mas não concebida como no caso anterior, como uma relação ‘temporal’, em que as características acidentais se agregam ao decurso do tempo e sim como uma reação ‘especial’ em que as características acidentais se sobrepõe às essenciais e são ambas suscetíveis de distinção, não apenas analiticamente, mas também empiricamente. Qualquer proprietário de uma motocicleta chegará a saber, o que é essencial e o que não é essencial em seu veículo a medida que este vá se deteriorando, para também chegar à conclusão de que sempre é possível atribuir às partes de sua motocicleta algo a mais que a propriedade de ser essencialmente funcional. Por exemplo, pode chegar à conclusão que é possível fazer funcionar o veículo sem o tanque de gasolina como tendo um galão conectado ao carburador com um pedaço de mangueira, mas ainda nesse caso, o predicado ‘galão’ não é uma característica essencial. Por último, pode haver certos graus dentro dos quais o acidental pode ser empiricamente retirado, de maneira a revelar o essencial em sua forma descoberta; “ela se mostra então, revelada”.

A utilidade científica destas posições é subestimada, em geral, hoje em dia. A posição atualmente dominante é a nominalista e se aplica o termo escarnecedor ‘reificação’ a qualquer tendência a atribuir uma existência independente aos universais. Trata-se de uma independência *ante rem*, ou concebida segundo o evolucionismo primitivo ou se utilize aos universais como categorias empíricas⁹. Todas estas posições incluirão a análise lógico e/ou empírico como uma ferramenta na busca do essencial. Em contraste com isto está a prática atual de distinguir entre a denotação e a conotação de um conceito e de incluir na conotação

⁹ Uma exposição deste ponto, no nível introdutório, se encontra em *Methods in Social Research*, New York, McGraw-Hill, Good y Hatt, 1952; Cap. 5 e, especialmente, pág. 42.

somente as características que são necessárias para a inclusão de um elemento de denotação. Unidas todas as propriedades que formam a conotação, constituirão a condição necessária e suficiente para a inclusão na denotação.

Não é necessário atribuir a Durkheim as etiquetas de ‘quase realista’ e ‘evolucionista’. Pode-se ser antirrealista ou antirrevolucionista, qualquer coisa que isto signifique e, no entanto, proceder quase exatamente como o fez Durkheim, embora possivelmente com uma introdução diferente, pondo ênfase na simplicidade mais que na pureza (“matéria prima”). Porém, outro procedimento surge mais facilmente dos enfoques, geralmente aceitos, sobre a relação entre as ideias e as coisas indicadas anteriormente: estudar uma amostra de unidades que se incluem todas sob a ideia de ‘instituições’ religiosas e utilizam sua variação como uma ferramenta analítica. Isto pode ir acompanhado de análises de casos, que incluíram não só os Arunta, mas também os sistemas religiosos muito modernos.

Consequentemente, a crítica contra o tipo de metodologia de Durkheim contém pelo menos estes elementos: 1) (‘antirrealista’) que a instituição religiosa ou qualquer outra instituição deve ser vista em seu contexto social, que invariavelmente envolve ‘acidentes’ e é estudada em seu contexto, não é isolada, e 2) (‘anti-evolucionista’) que características globais como ‘primitivo’ podem caracterizar sociedades em seu conjunto, enfatizando a simplicidade de certas instituições. Mas, estas características não se transferem necessariamente para outras partes da sociedade. Se analisarem muitas unidades, não é necessário nenhum julgamento sobre a pureza, nem tão pouco é necessário separar, de modo mais ou menos arbitrário, o essencial e o accidental.

Dando uma olhada para trás, pode dizer-se que a metodologia moderna da investigação extensiva ou pesquisa com um alto número de unidades, em geral

pressupõe recursos facilmente disponíveis, tanto em termos de dinheiro, como de energia, de tempo ou de pessoal. A metodologia intensiva, compatível com o tipo de raciocínio de Durkheim, também é compatível com o modelo de estudioso de biblioteca do século passado. Portanto, há aqui interrelações entre a metodologia, a ideologia fundamental e a estrutura social da investigação social, que podem chegar a sérios conflitos.

Ao longo da história da metodologia sociológica, a metodologia da ciência social em geral investiu em uma considerável quantidade de ideologia em diferentes células da tabela 1.2.1., ou em conjunto de célula, sustentando que em tais células representam as combinações ótimas na busca de conhecimentos científicos sobre a realidade social.

Parece que quanto mais se separa a posição defendida da diagonal do quadro, maior é o investimento. Preferimos uma distribuição mais equivalente do investimento ideológico, posto que ele parece estar mais de acordo com o que realmente acontece na investigação social contemporânea.

Parece que o próprio processo científico leva invariavelmente ou maior número de células, quando se seguem os cânones usuais da pesquisa científica. Assim pode perfeitamente acontecer que, na primeira suspeita germinal que conduza a um projeto de pesquisa na mente do sociólogo esteja localizada no canto inferior esquerdo do quadro, sob a forma de uma resposta dada por uma pessoa respondendo uma pergunta. A investigação pode, na sequência, se estender legitimamente pela horizontal, fazendo um certo número de pessoas a mesma pergunta e, pela vertical, formulando à mesma pessoa um número de perguntas diferentes. Mais cedo ou mais tarde, as duas dimensões das unidades e das variáveis, provavelmente terão ser unidas, pelo menos explicitamente e isto conduz o processo em direção à diagonal, a um nível

alto ou baixo, dependendo mais dos recursos do que da ideologia metodológica.

Pode ser que a última comprovação deva ser obtida por meio de projetos localizados na diagonal ou a próxima dela, mas também pode ser que as primeiras ideias (e talvez as perduráveis) se obtenham próxima dos eixos. Portanto, não confrontamos somente as estratégias de investigação extensivas com as intensivas, como também confrontamos as estratégias assimétricas com as simétricas. Uma investigação muito intensiva terá, provavelmente, um baixo rendimento em valor comprobatório como proposições gerais, e se supõe que o propósito perseguido é, de certa forma, genérico, devido ao baixo número de unidades. Da mesma maneira uma investigação muito extensiva não conterà um número suficiente de variáveis para controlar condições relevantes e colocar à prova pelo menos algumas de muitas hipóteses alternativas que terão que aparecer na mente do investigador. Sendo assim, a filosofia mais frutífera em relação à tabela parece ser uma combinação de dois ‘tanto/quanto’: tanto uma investigação de tipo extensivo quanto uma investigação do tipo intensivo e com tal fim, tanto uma forma m grande como uma n grande.

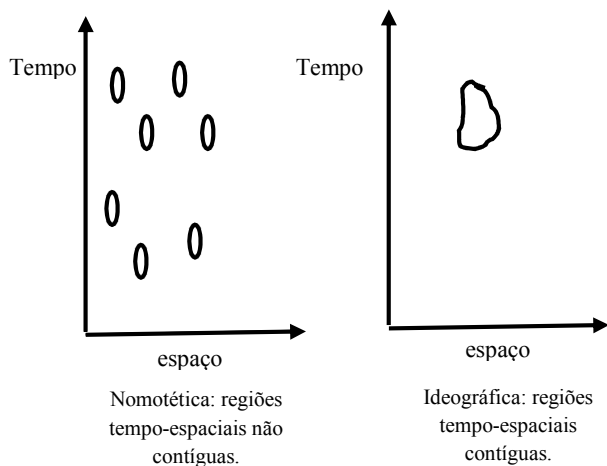
Se nosso argumento de que as diferentes ideologias referentes à escolha da estratégia de investigação representam diferentes fases no processo científico é correto, deve-se esperar que os sustentadores destas diferentes ideologias tenham algum tipo de conexão com estas diferentes fases. Não dispomos de dados empíricos a respeito, senão somente a impressão mais bem intuitiva de que o investigador “para a ideia”, em busca mais de perspectivas do que de comprovações, tende a argumentar em favor das estratégias que estão distantes da diagonal, e mais frequentemente em favor das investigações intensivas, estudos de casos, etc. Da mesma maneira, o investigador ‘orientado para a

comprovação, menos interessado nas ‘perspectivas’ e mais interessados nos ‘difíceis fatos’, se encontra constantemente argumentando em favor de sua especialidade, na esquina superior direita do quadro e, com frequência, também a favor das estratégias extensivas. Neste último caso, o investigador pode, muitas vezes, ser derrotado por seus próprios argumentos, posto que em um número escasso de variáveis dificilmente proporcionará a firme base necessária para explorar alternativas e, eventualmente, descartar $n-1$ hipótese a favor da hipótese n . Seja o que for, é de pouca importância, posto que o que nos interessa é a metodologia normativa e não a metodologia descritiva, nos interessa o que os investigadores devem fazer e não o que os investigadores fazem realmente. Destacamos somente a ocorrência de que as ideologias simétricas neste campo aparecem ser difíceis de acolher, por razões bem conhecidas para os estudiosos da tolerância da ambivalência e, particularmente, porque o folclore da ciência contemporânea parece haver imposto certo tipo de oposição entre ambas: com frequência parece estar a favor da investigação intensiva ou a favor da investigação extensiva, estar orientado para o ‘caso’ ou para a ‘enquete’.

Isto mesmo se aplica diferentemente entre as chamadas ciências nomotéticas (‘generalizadoras’) e as ciências ideográficas (‘singularizadoras’): tal diferença constantemente é exagerada. Ambos os tipos de ciências desenvolvem proposições e ambos os tipos de ciência conectam as proposições em teorias por meio de relações de implicação. Em outros termos os dois tipos têm como meta tanto a descrição como a explicação. Não obstante, enquanto as proposições das ciências ideográficas se referem a regiões limitadas e contíguas tanto no tempo quanto no espaço (França submetida a Luís XIV, a cidade-estado italiana durante o Renascimento, etc.), as proposições das

ciências nomotéticas se referem a fenômenos que (pelo menos a princípio) podem ser encontradas em uma vasta área de regiões não contíguas no espaço e no tempo ('A relação entre a industrialização rápida e a estabilidade política'). Portanto, se pode representar a relação da seguinte maneira:

Diagrama 1.2.3. – A relação entre as ciências nomotéticas e ideográficas



Frases feitas tais como a diferença entre 'dizer muito pouco sobre muito e dizer muito sobre muito pouco' não captam o ponto essencial: as proposições das ciências ideográficas estão pragmaticamente limitadas a uma região contígua (habitualmente um intervalo na história de uma nação ou de alguns aspectos selecionados de uma nação), enquanto as proposições das ciências nomotéticas não conhecem tais limites. As regiões que são relevantes para ser colocada à prova, habitualmente se encontram dispersas por um amplo campo e não existe necessidade de preencher com conhecimento os espaços entre elas.

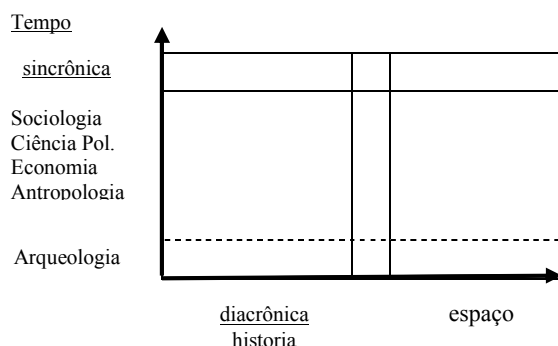
Obviamente, o campo total de uma monografia ideográfica, pode ser mais amplo que o campo total de uma monografia nomotética como quando uma pessoa escreve a história da Revolução Industrial na Grã-Bretanha e outra pessoa põe à prova hipóteses sobre as relações

entre o trabalho e o empresário em uma certa quantidade de estabelecimentos, em um dado período a uma dada nação. Obviamente também, os limites entre ambas de nenhuma maneira são precisos. Ao contrário, representam dois enfoques diferentes: a descrição concreta, detalhada, realista e a explicação de algo que é significativamente inter-relacionado, a colocação à prova de proporções gerais, menos realistas, mais abstratas sobre algo que se encontra disperso e que não está inter-relacionado no mesmo sentido.

Como acontece muitas vezes, as dicotomias tem uma tendência de polarizar não só o pensamento mas também as pessoas. É tão obvio que ambas se completam, que não são necessárias justificativas: a ciência generalizadora pode dar ideias à ciência particularizante sobre o que tem que buscar que tenha interesse teórico; a ciência particularizante pode dar a ciência generalizadora, dados para colocar à prova suas hipóteses. Entretanto, a diferença não se encontra simplesmente ao longo de um eixo explicação-descrição. A ciência ideográfica tem a vantagem de dar explicações que são realistas e como no caso particular, no sentido de considerar um número aproximadamente mais suficiente de fatores; a ciência generalizadora tem a vantagem descritiva de examinar um campo mais amplo, razão pela qual pode desenvolver categorias que tem sentido em um contexto mais amplo.

Isto é particularmente importante, em relação a muitas vezes discutida relação entre a ciência social sincrônica (que trata de fenômenos que acontecem em qualquer local ou espaço, mas dentro de um intervalo de tempo relativamente estreito) e a ciência social diacrônica (que trata fenômenos que acontecem em qualquer ponto do tempo, mas dentro de uma região limitada do espaço). Talvez se possam representar as relações entre elas da seguinte maneira.

DIAGRAMA 123. A relação entre as ciências nomotéticas e ideográficas



No entanto, a diferença entre, por exemplo, a Sociologia e a história não se baseiam meramente em termos de sincronia *versus* diacronia. A Sociologia tende a ser mais nomotética, a história a ser mais ideográfica. O conjunto total do trabalho sociológico pode cobrir relativamente bem o espaço com um conjunto de monografias interconectadas, uma referente às relações familiares em qualquer lugar; outra, a estrutura do poder em qualquer lugar, etc. O que falta é ciência ideográfica, sincrônica: conjunto realmente bom de tratados sobre a Sociologia de uma região dada no espaço, por exemplo uma nação. Isto se encontra em antropologia e nos estudos de comunidade, mas talvez à margem das ideias gerais. Da mesma maneira, o conjunto total do trabalho histórico e de grande parte do trabalho antropológico pode preencher tanto o tempo como o espaço, reunindo trabalhos sobre regiões adjacentes e contíguas no espaço e no tempo. O que falta aqui é a ciência nomotética, diacrônica, que conecta todo este conhecimento dentro dos marcos teóricas que não conhecem limitações no espaço e no tempo.

É um deplorável estado de coisas o que mantém separadas a Sociologia e a História institucionalmente nas universidades. Separa, assim, não só os enfoques sincrônicos e diacrônicos mas também, em uma medida considerável, os enfoques nomotéticos e ideográficos. Seria

possível imaginar que estas paredes fossem derrubadas para abrir campo inteiro das relações humanas no tempo e no espaço, sem nenhuma limitação, de acordo com o atual estado da humanidade, com a emergência rápida do homem internacional e de uma audácia tempo-espacial até agora desconhecida que leva a abrir inclusive o cosmos. É óbvio que isto exige o trabalho de equipe e a cooperação institucional, e é provável que se faça mais facilmente em universidades novas, mas talvez inclusive em países em desenvolvimento, em que as infelizes práticas acadêmicas do passado não sedimentaram em uma crosta muito grossa, de modo que só um terremoto intelectual possa derrubar as muralhas. Teoricamente, seria possível adiantar uma ciência do homem mais geral, que explora relações tanto no espaço como no tempo, em regiões contíguas e não contíguas, nos níveis explicativos e descritivos, tudo isto dentro do mesmo projeto de pesquisa. Com certeza, isto requereria uma fundamentação muito melhor em termos de ciência social sincrônica, descritiva (Sociografia, Sociologia descritiva) e de ciência social diacrônica, descritiva (historiografia, história descritiva) da que temos hoje em dia. Entretanto, não deveria ser muito difícil fazê-lo com os modernos métodos de armazenamento de dados.

Visto que o enfoque deste livro é principalmente nomotético, pode valer a pena dar mais substância a estas ideias, destacando o que significaram em termos de cartões IBM. A unidade básica da ciência social sincrônica tem sido o indivíduo. Dá-se um grande passo adiante agregando unidades de nível mais alto, províncias, nações, etc. Se dá outro grande passo adiante através da análise de pares, conjunto de três, etc. unidades em qualquer nível (individual ou coletivo), já que isto permite a inclusão de variáveis relacionais (ver 2.1). Deste modo, o passo que sugerimos agora é simplesmente adicionara dimensão tempo. Assim, se pode considerar uma nação como

uma unidade espacial, assim como ‘um fragmento tempo-espacial’. É possível, assim, tomar os últimos 150 anos de história da nação (sempre que tenha existido durante esse período) e dividi-lo em intervalos de, por exemplo, 5 anos, chegando a um total de 30 unidades espaço-temporais. Se há possibilidade de fazer o mesmo para, por exemplo, todas as nações latino-americanas, se obteriam cerca de 600 unidades de análises. Se fizer o mesmo para todos os pares de nacionalidade latino-americanos, inclusive dentro de um período somente de 50 anos, se obteria $10 \times (20_2) = 1.900$ unidades de análise e poderiam explorar a interação de tempo e espaço sobre as relações internacionais. Mantendo constante a nação, se obteria análise diacrônica, mantendo constante o tempo, se obteria análise sincrônica, permitindo a variação de ambas as dimensões, se obteria o tipo de análise (variedade nomotética) para o que a ciência social, sem dúvida, se está dirigindo.

Em consequência, os argumentos a favor de ter mais de uma unidade de análise sem nenhum sentido são somente argumentos aplicáveis às enquetes a aos estudos de opinião pública - mesmo que se possa dizer-se que estes estudos mostraram o caminho, e que eram mais fáceis de realizar. Melhores são argumentos a favor de uma maior imaginação na escolha das unidades dentro da região tempo-espacial possíveis de conter os dados e os argumentos a favor da destruição das barreiras institucionais na escolha de tais unidades.

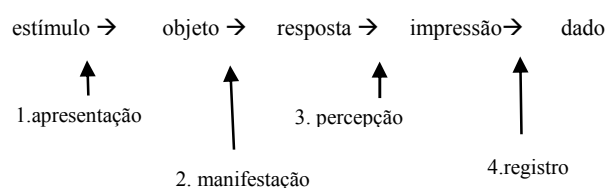
1.3. Observação e inferência

Os dados nas Ciências Sociais, como em outras ciências, se fundem em observações dos sentidos e em um critério para organizar os dados segundo sua qualidade se refere à proximidade do dado. Uma marca em um quadro vazio colocado ao lado de uma resposta é um dado e não é a inferência de que o sujeito pesquisado está a favor ou contra algo. Um sorriso é um

dado e não a inferência de que a pessoa sorridente está de bom humor, etc. Um dado é o que é observado, manifesto ou fenotípico, não é o que é inferido, latente ou genotípico. O que nos interessa nesta explicação de coleta de dados são sempre os dados manifestados e nunca o que está latente, aquilo que pode ser inferido.

Embora a palavra observação se refira explicitamente à percepção visual, é utilizada aqui, e em outras partes, para referir-se a todas as formas de percepção por meio dos sentidos, utilizados para registrar as respostas da maneira como se apresentam aos nossos sentidos. Contudo, fazemos uma diferença entre uma resposta e um dado: uma resposta é certo tipo manifestação de ação; um dado é o produto do registro da resposta. O caminho a partir da resposta ao dado é distorcido e apresenta, entre outras coisas, dificuldades tais como variações sensoriais intra e interpessoais, variações no uso de símbolos para registrar as impressões das respostas (o aspecto semântico), etc. Portanto, é um objetivo evidente fazer tão curta possível a última parte da seguinte sequência:

Tabela 1.3.1. A sequência estímulo-dado



Uma resposta é observável, um dado é observado

As observações científicas, ao contrário de outras observações, devem cumprir duas importantes condições, e estas condições delimitam até certo ponto o que pode ser considerado propriamente como dados manifestados. Suponha-se que o fenômeno observado é repetitivo e constante (por exemplo, um homem que dá a mesma resposta à mesma pergunta, uma pessoa que reage da mesma maneira à

mesma situação experimental, etc.) Parece muito razoável exigir que as observações repetidas de tal objeto produzam dados constantes, independentemente do fato de que as observações sejam realizadas por um observador ou por observadores diferentes. Se repetidas observações de um fenômeno constante pelo mesmo observador produzem dados constantes, se diz que a observação é intersubjetiva ou confiável. Se observações repetidas de um fenômeno constante por observadores diferentes produzem dados constantes, se diz que a observação é intersubjetiva (Deve-se notar que em ambos os casos estamos tratando com definições por redução).

Estes dois requisitos: a confiabilidade e a intersubjetividade, podem ser empregadas juntas como *definiens* do termo ‘objetividade’, tão vagamente empregado. Tais requisitos permitem excluir da esfera da observação científica todas as observações que são instáveis, precipitadas e vagas, insuficientemente padronizadas e todas as observações que pertençam à percepção subjetiva de uma pessoa determinada somente e que não podem ser compartilhadas por outras pessoas. Obviamente, é necessário falar de graus de confiabilidade e de intersubjetividade e talvez seja melhor decidir em cada caso particular por onde passa exatamente a linha divisória.

A confiabilidade e a intersubjetividade se referem aos dois componentes do processo da observação que denominamos percepção e registro e não se referem nem à representação de estímulo nem à manifestação das respostas que estão no sujeito interpelado. Se um estímulo constante produz uma resposta constante ou não é um problema inteiramente diferente: se refere (por definição) à constância do objeto (à unidade de análise) e não tem nada a ver com o

processo da observação. Frequentemente se considera desejável encontrar estímulos que produzam facilmente respostas constantes, mas esse *desideratum* tem certa raiz de irrealidade. Em primeiro lugar, os estímulos podem parecer constantes e não ser, porque o contexto em que eles aparecem relevantes para os objetos não é constante; por exemplo, quando se formula a mesma pergunta pela segunda vez em um estudo do tipo ‘painel’. Em segundo lugar, se os estímulos oferecem a respostas constantes, isto pode indicar que os estímulos são muito grosseiros já que as respostas não refletem as pequenas oscilações que, podemos presumir sem risco, ocorrem no espírito humano. Este último, além disto, é um argumento para utilizar mais de um estímulo para abranger uma mesma ‘dimensão’, visto que, provavelmente, não existe um *desideratum* que exija constância também para os objetos.

Em resumo: além dos princípios para a construção de uma matriz de dados, agregamos dois novos princípios¹⁰:

4) Princípio de intersubjetividade ou confiabilidade: observações repetidas das mesmas respostas pelo mesmo observador devem produzir os mesmos dados.

5) Princípio da intersubjetividade ou objetividade: observações repetidas das mesmas respostas por observadores diferentes devem produzir os mesmos dados.

Neste ponto, muitas vezes, se menciona um princípio de validade, com a ideia de que uma observação é válida se foi observado o que se queria observar. Em consequência, a validade, obviamente, se refere à relação entre o manifestado e o latente. Visto que, por definição, as observações se fazem ao nível manifestado,

¹⁰ Uma exposição deste ponto pode se encontrar em Johan Galtung, “An Inquiry into the Concepts of

‘Reability’, ‘Intersubjectivity’ and ‘Constancy’, Inquiry, 1959, pp. 107-125.

podemos formular o princípio da seguinte maneira:

6) Princípio de validade: devem ser obtidos dados de tal natureza e tal maneira que possam ser feitas inferências legítimas sobre o nível latente a partir do nível manifestado.

Já que, na realidade, os conceitos de manifesto e latente não foram definidos, deixaremos para mais adiante outros argumentos e não faremos uso do princípio de validade. Certamente, os dados sempre são válidos em algum contexto, com referência a alguma dimensão latente; isto debilita um pouco o conceito de validade. A validade se transforma em uma pergunta sobre se os dados podem ser utilizados exatamente ou não para conseguir o propósito para os que foram concebidos. Mais uma vez, é um ponto ao qual deve-se ter consciência antes da coleta de dados e que podem ser de grande importância, mas também podem desempenhar um papel menor, como é o caso de estudos mais descritivos e exploratórios.

1.4. Classes de variáveis e de modelos de investigação

Nas análises da Ciência Social, as unidades de análise mais frequentemente utilizadas são os indivíduos. O indivíduo pode ser caracterizado - e de fato tem sido caracterizado - por meio de um número imenso de variáveis. Tal número é alto por duas razões: o indivíduo é o nível habitual da interação humana e, conseqüentemente, é o nível em que a necessidade de discriminar na percepção e na comunicação se sente mais agudamente e está mais desenvolvida; em segundo lugar, a possibilidade da interação verbal com indivíduos faz possíveis as respostas verbais, com sua imensa variedade.

Em meio a essa variedade se requer algum tipo de organização, alguma tipologia, Sugerimos dois critérios para

classificar as variáveis que caracterizam os indivíduos.

Em primeiro lugar, o valor que um indivíduo tem é uma variável determinada pode ser pública ou privada. Entendemos por uma 'variável pública' aquela em que os valores individuais são conhecidos por outras pessoas e se sabe que são conhecidos. Alguns exemplos: a idade, o sexo, a raça, a religião (frequentemente), a ocupação, a renda (pelo menos em termos gerais: veja-se a prática estabelecida pelo acesso público e registros de impostos), o endereço, a família de origem, os dados de fertilidade sobre a família, etc. Por 'variável privada' entendemos uma variável em que os valores individuais, provavelmente, podem ser conhecidos por outros, mas os outros não podem ter pretensões legítimas de conhecer tais valores. Alguns exemplos são o coeficiente de inteligência, os projetos para o futuro, se uma pessoa é autoritária ou não, as atitudes e os padrões de comportamento de diferentes tipos. Uma pessoa pode tornar pública sua opinião a respeito da Comunidade Econômica Europeia ou da OTAN, mas isto não implica que sua atitude seja uma variável pública, porque também pode mantê-la fora da interação humana.

Em segundo lugar, o valor que uma pessoa tem em uma variável pode ser permanente ou temporal. Com este último queremos dizer que o conhecimento que se tenha de valor que uma pessoa alcança em tal variável em um momento determinado proporciona uma base relativamente fraca para formular um prognóstico sobre tal valor que tal pessoa terá na variável em um momento posterior; ao por outro lado, se tal valor é permanente, a base para a prognóstica é muito forte. Obviamente, a linha divisória entre o que pode chamar-se mais propriamente um valor 'temporal' e o que pode ser chamado um valor 'permanente' está longe de ser clara, e o mesmo se aplica à linha divisória entre o valor 'público' e o 'privado': Quantas pessoas devem ter uma pretensão legítima de conhecer um valor para que esse valor

seja público? O coeficiente de inteligência, conhecido pelo psicólogo da escola, é uma variável pública? Relativamente, a quais sistemas deve atribuir a ‘publicidade’? Etc. No entanto, tais diferenças têm um grande valor interpretativo.

Se as combinarmos, obtemos a seguinte tabela 2x2:

Tabela: 1.4.1. - Uma tipologia de variáveis para caracterizar indivíduos

Valores Públicos	Valores permanentes (Variáveis de base)	Valores temporais Impossível
Valores Privados	(Variáveis de personalidade)	(Variáveis de atitude e de comportamento)

Esta tipologia apresenta quatro células, mas uma delas está vazia devido a uma proposição sociológica que pode ser formulada da seguinte maneira: a interação institucionalizada está fundada em variáveis relevantes ‘públicas’ (por exemplo, variáveis em que os valores de um indivíduo específico são conhecidos publicamente, e por tanto, podem ser legitimamente utilizados na interação), mas se as variáveis tem que proporcionar uma base para uma interação duradoura, devem ser previsíveis, o que significa que os valores individuais tem que mostrar certo grau de ‘permanência’. Os valores devem ser conhecidos e reconhecidos, o que significa que as variáveis públicas não podem ser temporais. Deve-se acrescentar que um grande número de interações humanas não estão institucionalizadas e estão sobre a base de indícios de um caráter mais evasivo e efêmero. Uma característica da interação é que ela pode partir fundamentada em valores bastante privados e inclusive temporais, mas quando a interação se inicia e se torna institucionalizada, há uma mudança no caráter dos valores utilizados com base, em direção aos valores públicos e permanentes. Um valor público jamais

pode ser temporal, porque um valor temporal não proporciona uma base suficiente para a interação pública.

Foram colocados alguns termos nas três células restantes. Eles não devem ser considerados como completamente coextensivos como o significado que dado a estas células, por definição. Mais que isto, o significado de uma ‘variável de base’ é o de uma variável que é pública e permanente dentro de um dado sistema de interação, por exemplo, a sociedade norueguesa. Ser o presidente de uma organização local não é um valor público, quando se toma como sistema de referência a Noruega como um todo, mas ‘a ocupação’, é, devido à sua relevância em contexto nacional. No contexto da organização local, contudo, pode ser irrelevante a ocupação, enquanto ser o presidente pode ser muito relevante. Variáveis de base típicas são: a idade, o sexo, a localização geográfica, a ocupação, etc. Quando dizemos que são permanentes, não negamos a possibilidade de que mudem; trata-se somente do que são ‘permanentes dentro de um período’.

As variáveis que são ‘públicas e permanentes’ colocam o indivíduo na matriz social ou, se deseja usar outro termo, na estrutura social. O conjunto de valores nas variáveis de base dá, para cada indivíduo, sua personalidade social. Por esta razão, estas variáveis, frequentemente se chamam ‘variáveis estruturais’, termo que preferimos não utilizar porque pode levar a confundir-se os níveis de análise. As variáveis de base caracterizam indivíduos, mas quando recebem o nome de ‘estruturais’ se pensa na estrutura das sociedades.

Por ‘variáveis de personalidade’ entendemos aquelas variáveis que podem ser desconhecidas inclusive, para o próprio indivíduo e que exigem a habilidade de um psicólogo para ser descobertas. Sua característica essencial é a permanência. Visto que, habitualmente, se sustentem que as variáveis permanentes, são de um caráter

mais latentes, enquanto que as variáveis manifestadas dão lugar a valores menos permanentes para os indivíduos. Normalmente, as variáveis de personalidade coincidirão com o que denominamos variáveis latentes. Deve-se destacar, novamente, que ‘permanência’ não significa que toda mudança seja impossível, e sim que há intervalos suficientemente longos para ser importantes, quer dizer, lapsos dentro dos quais as flutuações são insignificantes.

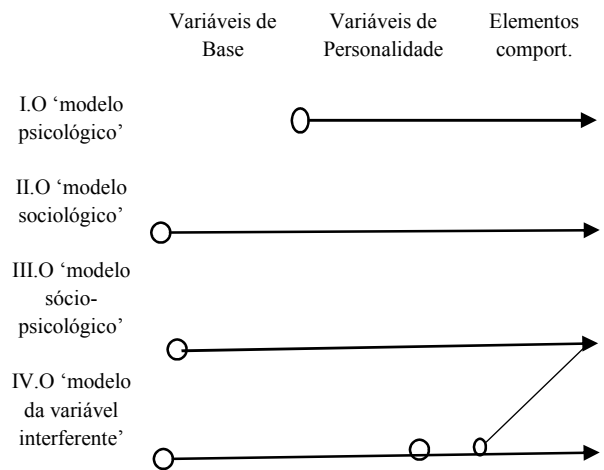
A terceira categoria contém o resto, ou seja, todos os tipos de variáveis atitudinais e comportamentais e outras variáveis que podem ser utilizadas para caracterizar um indivíduo e que nem são públicas, nem são permanentes. Muitos dos dados que se obtém por meio de questionário ou de uma entrevista entram nesta categoria. Haja vistas que, conforme o esquema proposto, o único fator que estas variáveis têm em comum é seu caráter de variáveis temporais, não dispomos de uma qualificação melhor que ode ‘elementos comportamentais’. Portanto, as respostas a entrevistas e a questionários se classificam como comportamento (reações de palavras, comportamento verbal).

Consequentemente, o valor desta tipologia reside em permitir uma classificação das variáveis que caracterizam a indivíduos em variáveis de base, variáveis de personalidade e elementos comportamentais. Na realidade, pode-se utilizar um esquema muito semelhante, por exemplo, a respeito das nações; contudo, a aplicação mais frutífera deste esquema está no campo de análise de unidades que sejam seres humanos.

Habitualmente, tais análises se realizam com um certo modelo mental, que pode haver sido explicitado pelo pesquisador ou não.

O raciocínio típico se desenvolve considerando os elementos comportamentais que qualificamos por ‘temporais’, como produzidos ou causados ou sendo, de alguma maneira, o resultado das variáveis de personalidade ou de ambas. Essencialmente, os seguintes parecem ser os quatro modelos mais frequentes da literatura.

Tabela 1.4.2. Modelos para as relações entre os tipos de variáveis



O pesquisador deve ter uma imagem relativamente clara do que tem na mente. Os termos escolhidos para qualificar os diferentes modelos não são, mais uma vez, coextensivos com o conteúdo dado aos modelos pelo diagrama, mas são com mais frequência verdadeiros do que falsos¹¹.

A diferença entre os modelos III e IV é a seguinte: no modelo IV a ideia é que

oposta a proposições ou teorias em termos de valores, com diversos graus de internalização. Hoje em dia, esperamos, se está menos interessado na definição do reino de uma Ciência Social particular e mais interessado no desenvolvimento de modelos de pesquisa que combinem as virtudes dos enfoques cultivados dentro de uma disciplina.

¹¹ Detrás do modelo sociológico se encontra uma importante tradição, formulada em termos normativos por Durkheim: “A causa determinante de um fato social deve ser buscada entre os fatos sociais que lhe precederam e não entre os estados da consciência individual”, *The Rules of Sociological Method*. Glencoe, Free Press, 1950. Hoje em dia isto se chamaria de uma proposição ou teoria ‘estrutural’.

uma certa condição social, definida por um conjunto de variáveis de base, produz (na maioria dos casos) certa estrutura da personalidade e esta estrutura da personalidade por sua vez, produz (na maioria dos casos) certos elementos comportamentais. Por outro lado, no terceiro modelo das características de personalidade, não se consideram como degraus em escada, por assim dizer, senão como causas independentes, que dão origem aos elementos comportamentais, juntamente com a característica do contexto social.

Suponha-se que se desenvolveram índices muito poderosos de base social e de personalidade, de tal maneira que ambas podem ser analisadas com sucesso por meio de uma só variável (por exemplo, por meio de alguma medida da dimensão) ‘centro-periferia’ para contexto social e da dimensão ‘mente abertamente fechada’ para a característica da personalidade. A seleção de tais dimensões, partindo de outros pontos de vista, incorpora grande parte do pensamento contemporâneo. O seguinte seria um gráfico poderoso para o estudo dos elementos comportamentais tanto manifestados (ações) como latentes (atitudes).

Tabela 1.4.3. Gráfico para o estudo dos elementos comportamentais

		Variável de base social						
		1	2	3	...	o	Total	
Variáveis Personalidade	1							
	2							
	3							
	.							
	.							
	b							
Total								

O gráfico teria que estudar como variam os elementos comportamentais

conforme varie, tanto o contexto social como a personalidade e, para fazê-lo, é necessário um projeto $a \times b$, supondo que a variável social tenha a níveis e a variável da personalidade tenha b níveis. O gráfico conduz por si só, tanto a interpretação do tipo III, como do tipo IV.

Para apreciar em que consiste a ideia do gráfico proposto, considerem-se alguns típicos gráficos incompletos que se usam tanto. Modelos do tipo I e II utilizariam somente as margens horizontais e verticais, respectivamente, do gráfico. Não se pode ir contra isto, ainda que as correlações serão provavelmente mais baixas do que poderiam ser, devido a que em cada nível social haverá uma apreciável quantidade de variação na personalidade que não se considerou, e em cada nível da personalidade haverá uma apreciável quantidade de variação social que não se considerou. Contudo, se podem obter tendências gerais em termos de estrutura social ou de estrutura da personalidade. O que é realmente passível de objeção é fundar a análise em modelos do tipo I ou II e utilizar amostras somente de algumas das colunas ou filas no gráfico. Na investigação sociológica isto significa pegar amostras somente entre pessoas com certas características da personalidade; na pesquisa psicológica, significaria pegar amostras somente entre pessoas com certas características de contexto social. O sociólogo salva mais frequentemente essa falácia devido à dificuldade de critérios de utilizar dados psicológicos como critérios de amostragem, mas o psicólogo sentirá frequentemente a tentação de pegar as amostras mais facilmente acessíveis ou mais extraordinárias, selecionando em excesso as categorias sociais centrais e periféricas, respectivamente. Este é um dos pontos principais da importante crítica metodológica “A personalidade autoritária” feita por Hyman e Sheatsley¹². Tal como

¹² Hyman, Herbert e Sheatsley, Paul B. “The Authoritarian Personality: a Methodological

Critique”, em Christie, Richard y Jahoda, Marie, ed., Studies in the Scope and Method of “The

eles dizem sobre “pegar uma amostra da população nacional heterogênea”.

“... tal população tem uma virtude, com frequência passada por alto, que haveria melhorado a qualidade de análise e que suscita algumas dúvidas sobre certas conclusões de “A personalidade autoritária”. A razão é que uma amostra semelhante revela a extensão e a localização social das atitudes. Sem este conhecimento, a interpretação é difícil e frequentemente perigosa. Assim, pode-se perfeitamente perceber em uma amostra nacional que uma atitude específica é quase universal, e apesar disto, está mesma atitude estudada e um pequeno grupo homogêneo pode aparecer como desviada. Ao não se dispor das normas nacionais, seria possível tentar explicar o último achado em termos de algum processo da personalidade idiossincrático quando, na realidade, está muito correlacionado com um fato social principal. [...] mas o pesquisador, devido ao gráfico de sua amostra, não pode dizer se uma determinada atitude é normativa para a população em geral, pode ir buscar a explicação em um lugar indevido”¹³.

Pode-se contrariar o raciocínio apresentado aqui se a palavra ‘grupo’ na frase: ‘pequeno grupo homogêneo’ se refere a um grupo social e não a uma categoria social, então, é estranho também o desviado, inclusive se é comum em contextos sociais. No entanto, o argumento genérico é importante e se estende, não só às escalas e às porcentagens, mas também às correlações, por exemplo, a organização das atitudes:

“[...] as pesquisas sobre a organização dos sentimentos não podem permitir ignorar as considerações relativas às

amostragens durante a preparação do gráfico de estudo e à análise de resultados...”. Parece-nos que o esquecimento deste assunto por parte dos autores significa uma metodologia superficial e, partindo do ponto de vista prático, tende a perpetuar a implicação de que o nível da organização dos sentimentos é uma espécie de universal, um processo intrapsíquico que tem escassa relação com as condições do contorno”¹⁴.

Demonstraremos, mais adiante, o que o analista pode, em muitos casos, obter quase qualquer correlação que deseje entre as atitudes, distorcendo adequadamente sua amostra.

Na mesma ideia de pegar uma amostra de uma categoria social específica, por exemplo, os famosos estudantes que seguem um curso introdutório de psicologia, encontra-se um caso, muito facilmente reconhecível, de pensamento do tipo ‘o caso puro’ ou do tipo ‘uma espécie de universal’ como o expressam Hyman e Sheatsley. Em seguida, os autores citam um grande número de pesquisas que demonstram que as correlações de atitude variam com fatores tais como: a idade (as crianças chegam a ser mais consistentes em suas características morais conforme vão crescendo) e a educação (uma investigação que demonstra uma correlação crescente entre o preconceito e o autoritarismo com a educação). Obviamente, não são admissíveis as generalizações de uma só categoria social a toda a sociedade; deste ponto surge a importância do gráfico completo.

Em tal gráfico, não há nenhum implícito suposto, no sentido de que o contexto social e a personalidade sejam independentes; ao contrário, uma investigação tradicional muito considerável na Psicologia social demonstra que elas são

Autoritarian Personality”, Glencoe, III, The Free Press, 1954, pp. 50-22.

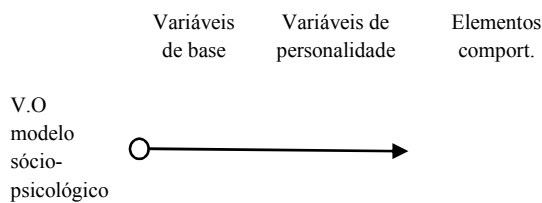
¹³ Ibidem, pág. 67-68.

¹⁴ Ibidem, pág. 60.

dependentes. O que importa é tratar de obter dados que preencham todas as células, inclusive quando o projeto deve ser manejado de alguma das maneiras extremas (modelos I e II). Se os dados são coletados sobre ambas as variáveis básicas para todas as unidades, existe sempre a possibilidade de controlar o efeito da segunda variável, de tal maneira que se possa ir em direção aos modelos III e IV.

A relação entre as variáveis podem ser explorada de melhor maneira por meio de um modelo de investigação de um V tipo:

Tabela 1.4.4 Um modelo para as relações entre os tipos de variáveis



Não incluímos este modelo entre os outros, visto que não toca diretamente nos elementos comportamentais, e, obviamente, terá que estar baseado em investigações realizadas de acordo com os modelos III ou IV, já que devemos presumir que as variáveis de personalidade são latentes e que os elementos comportamentais são manifestações e que só podemos obter dados manifestados. Este modelo deve resultar numa distribuição das características da personalidade ao longo das categorias sociais e, portanto, deve produzir um tipo de conhecimento que é indispensável nas investigações que são concluídas conforme o modelo IV.

Deve-se destacar que a direção das flechas nos cinco modelos não deve ser considerada mais do que uma simples regra de caráter prático. Assim, pode-se considerar a personalidade como uma variável independente e variáveis de base, tais como a ocupação e a educação como variáveis dependentes mesmo que se

possam conservar os modelos mais facilmente ou, inclusive, de forma mais frutífera, definindo a busca da ocupação e da educação com um elemento comportamental. Em geral, os modelos refletem o estado atual de pensamento nas ciências sociais, mas eles terão que ser refinados nos capítulos seguintes, de maneira que sirvam como guias para a investigação.

Os modelos III e IV são interdisciplinares e exigem a colocação, em comum, das habilidades do sociólogo e do psicólogo, quando as unidades são indivíduos. Portanto, são modelos para equipes de investigação, a menos que se possa presumir que um mesmo investigador possua conhecimentos adequados, ou que o sociólogo possa facilitar para o psicólogo, e vice-versa, instrumentos suficientemente bons, e que possam ser usados por indivíduos não especializados.

Geralmente, se faz uma diferença entre variáveis independentes e dependentes. Em matemática, estes termos se usam, com frequência, para indicar certa assimetria na maneira como se apresenta uma relação funcional, mas significa muito pouco em termos de significação matemática.

Nas ciências empíricas os termos somente podem ser utilizados se existe algum modelo, incluso do tipo mais indefinido no que diz respeito às relações entre as variáveis estudadas. Os modelos podem tomar muitas formas; por exemplo, os quatro modelos básicos na investigação sociológica mencionadas mais acima. Contudo, eles habitualmente têm em comum a ideia de ‘alguma coisa que, de alguma maneira, produz alguma outra coisa’. Nesta ideia se veem, claramente, os dois elementos que conduzem às concepções das variáveis dependentes e independentes. É possível dar três precisões sucessivas:

1) Se o modelo é de algum dos quatro tipos discutidos mais acima (ou de qualquer outra espécie que envolve tipos diferentes dos três tipos de variáveis) então a variável independente é a variável mais à esquerda, com a dependente à direita.

2) Se o modelo é casual, então a variável casual é a variável independente e o efeito é a variável dependente.

3) Se existe uma ordem no tempo, entre as variáveis, então a variável precedente é a variável independente.

Por último, temos que mencionar um ponto relativo à eleição das variáveis na Sociologia Aplicada. A Ciência Social Aplicada deve dar uma base para tomar resoluções a favor da mudança social quando algo ‘anda mal’ e a favor da estabilidade social quando algo ‘anda bem’, mas as resoluções em si mesmas podem ser formuladas por outras pessoas. A Ciência Social Aplicada se diferencia do discurso sociológico geral somente em uma coisa: deve formular suas conclusões em termos de variáveis que sejam manejáveis. Uma relação de causa em que não existe possibilidade de manejar a variável independente por razões técnicas, econômicas, éticas ou de outro tipo, é de escasso valor para o prático. Nos fatos, este é um argumento contra a pesquisa social aplicada que se realiza de acordo com o modelo I indicado anteriormente, haja

vistas que os fatores da personalidade são tão difíceis de modificar¹⁵.

Entretanto, não é um argumento muito eficiente contra os modelos II e IV, particularmente em uma cultura em que a ‘engenharia social’, de alguma maneira, é considerada mais praticável que a ‘engenharia humana’¹⁶.

Qualquer que seja a filosofia que se tenha sobre este ponto, a escolha do modelo básico deve ser considerada cuidadosamente antes de começar um projeto, e deve-se examinar o propósito geral do estudo, antes de formular decisões irreversíveis sobre a escolha de variáveis independentes.

¹⁵ Este fato está estreitamente vinculado com a ideologia do nosso tempo, nos países industriais. Não titubeamos quando se chega a mudar a pauta da conduta de uma pessoa através da instalação de um contexto social diferente, como quando essa pessoa muda seu estilo de vida depois (ou antes) de uma promoção. Entretanto, só parece aceitável a mudança no comportamento produzido por uma mudança de personalidade se esta mudança é lenta e não eficiente demais, como o que se produz durante um processo de educação e de socialização ou quando se utiliza a psicoterapia. Os enfoques bioquímicos sobre a mudança de personalidade encontram resistências, a

menos que se possam definir como ‘curas’. As descobertas que utilizam variáveis de personalidade como variáveis independentes ou são inúteis ou são úteis demais: o espaço entre ambos os extremos é muito estreito.

¹⁶ Pode-se facilmente imaginar uma cultura dominada pelos ypgis de Koestler (como extremos opostos aos seus delegados) ou pelos ideacionais de Sorokin (como extremos opostos a seus sensitivos), em que as variáveis de personalidade sempre se utilizariam como variáveis independentes. Um problema diferente é se essa cultura produziria Ciência Social.