



UNIVERSIDADE
FEDERAL DO RIO
GRANDE DO SUL

Reitora

Wrana Maria Panizzi

Vice-Reitor

e Pró-Reitor de Ensino

José Carlos Ferraz Hennemann

Pró-Reitor de Extensão

Fernando Setembrino

Cruz Meirelles

Vice-Pró-Reitora de Extensão

Renita Klüsener

EDITORA DA UFRGS

Diretor

Geraldo F. Huff

CONSELHO EDITORIAL

Antônio Carlos Guimarães

Aron Taitelbaun

Célia Ferraz de Souza

Clovis M. D. Wannmacher

Geraldo Valente Canali

José Augusto Avancini

José Luiz Rodrigues

Lovois de Andrade Miguel

Luiza Helena Malta Moll

Maria Cristina Leandro Ferreira

Geraldo F. Huff, presidente

PARTE III
5.1
leitoma
OBRIGATORIA

Métodos de Pesquisa Social Empírica e Indicadores Sociais

ACHIM SCHRADER

organizadores:

Clarissa Eckert Baeta Neves

Emil Albert Sobottka

Editora da UFRGS • Av. João Pessoa, 415 - 90040-000 - Porto Alegre, RS - Fone/fax (51) 3316-4090, 3226-4953 e 3224-0327 - editora@ufrgs.br - www.ufrgs.br/editora • Direção: Geraldo Francisco Huff • Editoração: Paulo Antonio da Silveira (coordenador), Carla M. Luzzatto, Maria da Glória Almeida dos Santos, Rosângela de Mello; suporte editorial: Fernando Piccinini Schmitt e Sílvia Aline Otharan Nunes (bolsista) • Administração: Najára Machado (coordenadora), José Pereira Brito Filho, Laerte Balbinot Dias e Mary Cirne Lima; suporte administrativo: Ana Maria D'Andrea dos Santos, Felipe Brito Vasconcellos (bolsista), Jean Paulo da Silva Carvalho, João Batista de Souza Dias e Marcelo Wagner Scheleck • Apoio: Idalina Louzada e Laércio Fontoura.


UFRGS
EDITORA

© de Achim Schrader
1ª edição: 2002

Direitos reservados desta edição:
Universidade Federal do Rio Grande do Sul

Capa: Vera Lúcia Gliese
Figura da capa: Detalhe da obra *Der Naschmarkt in Wien* de Carl Moll
Revisão: Luciane Leipnitz
Maria da Glória Almeida dos Santos
Editoração eletrônica: Sílvia Aline Otharan Nunes

Achim Schrader. Professor titular de Sociologia na Universidade de Münster, Alemanha, Instituto de Sociologia (Achim.Schrader@uni-muenster.de) e professor Associado do Programa de Doutorado em Sociologia da Universidade Federal do Rio Grande do Sul.

Apresentação

É com grande satisfação que apresento esta coletânea de textos do professor Achim Schrader sobre metodologias de pesquisa social empírica.

Docente da Westfälische Wilhelms Universität-Münster/Alemanha, o professor Schader, é associado ao Programa de Pós-Graduação em Sociologia da UFRGS. Sua contribuição tem sido especialmente importante no campo da metodologia. Através de seminários e *workshops*, abordou, de modo aprofundado, a metodologia da construção de índices, a função dos indicadores no planejamento social e de desenvolvimento, e os problemas da medição e da utilização em estudos comparativos internacionais.

A coletânea, que ora publicamos, é o resultado desse trabalho e, certamente, será de grande utilidade para pesquisadores, analistas e planejadores sociais que necessitam ter um conhecimento rigoroso sobre a construção e o uso de indicadores sociais.

O primeiro texto discute os *indicadores sociais na Sociologia, suas bases teóricas e considerações práticas*. Os indicadores sociais, há algumas décadas e, recentemente com mais intensidade, vêm ganhando importância como instrumento da pesquisa social empírica, e são largamente utilizados por governos e agências internacionais para a medição de tendências sociais, econômicas e culturais. Documentos do Banco Mundial, bem como das Nações Unidas, têm grande repercussão principalmente quando classificam os países, como é o caso do Human Development Index.

Os indicadores constituem-se numa medição agregada e complexa que permite descrever ou avaliar um fenômeno, sua natureza, estado ou evolução, articulando e correlacionando variáveis. Indicadores sociais, são expressões numéricas de fenômenos quantificáveis, representando fenômenos sociais politicamente relevantes, que não podem ser medidos diretamente. Quando regularmente pesquisados e publicados, permitem tirar conclusões sobre o desenvolvimento dos fenômenos sociais em questão.

Através da construção de indicadores sociais busca-se contribuir para a reconstrução da realidade em duas perspectivas complementares: a científica e a política. Mas a utilização de indicadores envolve também problemas metodológicos, técnicos e de conteúdo relacionados com sua própria construção. Eles dizem respeito tanto à falta de exatidão e representatividade do levantamento dos dados, como a problemas de processamento e de combinação de indicadores parciais.

S377m Schrader, Achim

Métodos de pesquisa social empírica e indicadores sociais / Achim Schrader; Clarissa Eckert Baeta Neves e Emil Albert Sobottka (orgs). – Porto Alegre : Ed. Universidade/UFRGS, 2002.

1. Sociologia – Pesquisa empírica. I. Neves, Clarissa Eckert Baeta. II. Sobottka, Emil Albert. III. Título.

CDU 301.15

Catálogo na publicação: Mônica Ballejo Canto – CRB 10/1023
ISBN 85-7025-629-9

No Brasil, os organismos que produzem estatísticas oficiais oferecem, cada vez mais, novos e variados dados e indicadores sociais, que não têm sido plenamente utilizados em estudos e análises e no planejamento social. Ressalte-se que o ensino dos métodos da pesquisa social empírica não inclui, sistematicamente, a temática dos Indicadores Sociais. Disso resulta a situação paradoxal dos pesquisadores sociais, tanto nas universidades como em centros de pesquisa, que têm acesso a significativo volume de dados e de relevantes indicadores sociais, mas não dispõem de bases metodológicas para seu processamento, sua reelaboração e reconstrução.

No segundo texto, *Pesquisa social empírica: sobre a vantagem de estudar os seus métodos*, o professor Schrader discute quatro teses que estimulam a pesquisa empírica na Sociologia, a saber: *o método empírico não é o único, mas o melhor método de pesquisa, na Sociologia; a pesquisa social empírica não é um fim em si mesma; sua razão de ser é contribuir para melhorar a teoria sociológica; e, por fim, o aprendizado dos métodos da pesquisa social empírica é mais fácil para aqueles que cedo superam certos problemas cognitivos.*

E conclui: "a pesquisa social empírica segue sendo o que sempre foi: um trabalho árduo".

A Aula Magna de despedida do professor Schrader, quando de sua aposentadoria, proferida no Instituto de Sociologia da Universidade de Münster em junho de 1999, compõe o terceiro texto desta coletânea

O professor Schrader, que na maior parte de sua atividade universitária, se dedicou ao ensino e ao desenvolvimento de metodologias na pesquisa social empírica especialmente as que se utilizam do tratamento estatístico dos dados, na Aula magna desenvolveu um tema inusitado para a Sociologia, *Loas ao passeio*, um método menosprezado na pesquisa social empírica.

Será possível adquirir conhecimentos sociológicos também a pé? O andar a pé ou quiçá o ir passear pode ser considerado uma metodologia de pesquisa social empírica?

Nesta aula o professor Schrader convidava seus ouvintes e, através desta coletânea, os leitores "a verificar ... através da história da Sociologia ... que a pesquisa social a pé, constitui-se num método digno de ser levado a sério".

A presente publicação, certamente, vem preencher uma lacuna na literatura de ciências sociais e contribuirá para o conhecimento das metodologias e dos problemas e perspectivas de sua utilização na pesquisa social empírica.

CLARISSA ECKERT BAETA NEVES

Professora no Programa de Pós-Graduação em Sociologia/UFRGS

Sumário

Indicadores sociais na sociologia: sobre suas bases teóricas e considerações práticas	9
Fundamentos teóricos e o relacionamento ciência versus política	11
Conceitos básicos e a sua operacionalização	21
"Minicensos" e pesquisas sociais nacionais gerais	33
Comparação internacional	47
Obtenção e tratamento dos dados	55
Discussão de alguns indicadores brasileiros	59
Pesquisa social empírica: sobre a vantagem de estudar os seus métodos	73
Loas ao passeio: um método menosprezado na pesquisa social empírica	95

“Minicensos” e pesquisas sociais nacionais gerais

Censos e censos por amostragem

QUADRO INTERNACIONAL

Muitos países, além dos censos oficiais gerais, vêm realizando com certa regularidade levantamentos, oficiais ou oficiosos, na base de uma amostragem. Embora a de (*Mikrozensus*), quase todos os estados-membros da OCED, isto é, os países chamados “industrializados”, têm alguma forma de censo por amostragem. É notável que países não-membros da OCED, como o Brasil, também há muito tempo já apliquem as pesquisas censitárias por amostragem. As eventuais desvantagens de uma amostragem compensam bem as desvantagens sofridas pelos censos completos. Estes são muito demorados na apuração e não há como estimar o volume dos erros na coleta dos dados, como é possível em amostragens. As modalidades amostrais, aliás, são as mais diversas:

- a participação da população pode ser obrigatória ou voluntária;
- os levantamentos podem ser anuais ou em intervalos mais curtos e mais longos; e,
- as amostras podem variar em volume e desenho.

Chama a atenção a alta participação voluntária, entre 80% a 95%, na maioria dos países; exceção são os Países Baixos, com apenas 62%, e a Suíça com 71%. A participação tendeu a diminuir nas últimas décadas, enquanto cresceu a tendência de fazer levantamentos em intervalos mais curtos (mensais ou trimestrais) (Esser e colaboradores 1989, p.IX-X). Além da área da OCED, as Nações Unidas ajudaram a iniciar esse tipo de levantamento em 80 países “em desenvolvimento” (1989, p. 76). Também os países no Leste Europeu, antes dominados pela União Soviética, começaram a fazer levantamentos desse tipo. Hungria e República Checa parecem estar na frente neste desenvolvimento. Às vezes as amostras são muito grandes e abrangem até mais de 2% da população (em 8 países), como também podem abranger menos de 0,1% (Estados Unidos) da população (Esser e colaboradores 1989, p.94).

A União Européia está realizando um minicenso da mão-de-obra em todos os seus estados associados em alguns meses ou durante o total do pri-

meiro semestre de todos os anos, com participação obrigatória ou facultativa dos entrevistados, com quotas de rejeição entre 0 e 25% (Áustria) e com os mais variados métodos de levantamento, inclusive os métodos Capi e Cati.⁷

Censo Demográfico e Pesquisa Nacional por Amostragem de Domicílios (PNAD) no Brasil

No Brasil, o primeiro censo demográfico foi realizado em 1872. Seguiram depois os censos de 1890, 1900 e 1920. A partir de 1940 e até 1980 os Censos foram realizados a cada 10 anos. O último é de 1991.

Desde 1960 o censo não é mais censo em sentido estrito. Aplica-se desde então o método da amostragem. Em 1991 foram registrados todos os domicílios (excluída a população indígena no Norte do país) através de visitas e a partir disto se escolheu aleatoriamente uma amostra ponderada de domicílios a serem entrevistados. A ponderação teve o objetivo de garantir a todas as pessoas igual probabilidade de serem escolhidas, considerando a densidade populacional. A participação das pessoas nas entrevistas era obrigatória por lei. O tamanho da amostra foi de aproximadamente 3% da população alvo, isto é, mais ou menos 4,5 milhões de pessoas. O conteúdo das questões feitas seguiu o padrão internacional, dedicando, aliás, mais detalhamento à situação domiciliar. De especial interesse sociológico são os dados sobre a estrutura familiar, a distribuição da renda, religião, migração, educação e desemprego/atividade econômica (Fibge, 1996).

A Pesquisa Nacional por Amostragem de Domicílios – PNAD – só superficialmente é parecida com os censos demográficos. As diferenças se devem a:

- frequência dos levantamentos;
- método da amostragem;
- cobertura regional; e,
- conteúdo.

Os levantamentos da PNAD são bem mais frequentes que os do censo demográfico. Entre 1967 e 1970 foram trimestrais. Desde 1971 são anuais, excluindo os anos nos quais foi realizado o censo demográfico.

A amostragem se faz por estratos (amostragem estratificada), escolhendo municípios, depois áreas censitárias e, finalmente, os domicílios. A cobertura

⁷Emmerling e Riede 1997; CAPI: Computer Assisted Personal Interviewing, entrevistas pessoais apoiadas pelo computador; CATI: Computer Assisted Telephone Interviewing, entrevistas telefônicas apoiadas pelo computador.

do País se desenvolveu em etapas, até que em 1996 abrangeu quase todo o território nacional; faltam ainda as áreas rurais do Norte brasileiro (IBGE, s. d.). As temáticas das PNAD na maioria dos anos eram específicas, mantendo sempre o conjunto das variáveis mais básicas⁸ (ver tabela 6).

TABELA 6
Temáticas das PNADS

Anos	Tópicos
1981	saúde
1982	educação
1983	mão-de-obra e previdência
1984	fecundidade feminina
1985	situação do menor
1986	anticoncepção, acesso a serviços de saúde, suplementação alimentar e associativismo
1988	participação político-social e estoque de aparelhos utilizadores de energia
1989 e 1990	trabalho
1992	migração, fecundidade e nupcialidade
1996	mobilidade social

Para o uso da pesquisa sociológica, vale mencionar a existência dos estudos de emprego e desemprego, realizados mensalmente em várias regiões do Brasil (São Paulo desde 1985, Belém 1988, Brasília 1991, Porto Alegre 1992, Curitiba 1994, Belo Horizonte 1995, Salvador 1997). Além dos indicadores de emprego e desemprego há neles outras características das pessoas no mercado de trabalho que merecem uma análise sociológica secundária.

Disponibilidade dos dados

Em muitos países ainda é difícil, ou pelo menos muito caro, o acesso dos cientistas sociais aos dados. O fato de serem órgãos públicos que conduzem os

⁸Até 1980 e em 1987 foram levantadas somente as variáveis básicas. Em 1970, 1980, 1991 e 1994 não foi realizada a PNAD. Essas informações se encontram em um comentário no CD pelo qual o IBGE está divulgando os dados da PNAD de 1995 (IBGE s.d.).

levantamentos resulta em uma série de obstáculos. Alguns países, tais como a Alemanha, além disso, introduziram, nos anos passados, disposições legais ou até constitucionais para proteger os dados dos cidadãos ("autodeterminação informacional"). Nesse caso, através dos esforços da comunidade científica, recentemente foi introduzida a possibilidade de cientistas firmarem convênios com o serviço estatístico estatal, pelo qual podem receber os dados "de fato anonimizados". Essa expressão significa que a identificação de um entrevistado é possível somente através de um enorme esforço de sondagens de tipo "detetive". Para fins de pesquisa, os cientistas podem usar o total dos dados, que representam 70% da amostra total, mas devem formalmente obrigar os seus assistentes a guardar o sigilo dos dados. No ensino podem usar apenas uma amostra de 50%. O signatário é pessoalmente responsável pelo cumprimento do convênio e pode ser processado em caso de transgressão dos termos do acordo. Além disso, os problemas técnicos não são fáceis de resolver. O CD tem quase 200 MBytes e para a expansão dos arquivos são necessários temporariamente 500 MBytes de memória no disco rígido.

Pesquisas sociais nacionais gerais

O "minicenso" e, no caso do Brasil, a PNAD, não podem ser confundidos com as "pesquisas sociais nacionais gerais" ("Nationwide General Social Survey"). Elas não são levantamentos oficiais no sentido de que um órgão ou uma repartição pública as conduza, embora elas sempre dependam de um financiamento público. Nos Estados Unidos da América, por exemplo, é o General Social Survey - GSS -, na Inglaterra o British Social Attitudes Survey - BSA -, na Alemanha a Allgemeine Bevölkerungsumfrage der Sozialwissenschaften - Pesquisa geral das ciências sociais na população, Allbus -⁹. Os países cooperam neste campo através do International Social Survey Programme - ISSP.

Assim como nas ciências exatas, como por exemplo na meteorologia e na física nuclear, os cientistas sociais aproveitam bases de dados coletivas, elaboradas segundo vários propósitos e para vários usuários (*data sharing*). O termo "dados", neste caso, se refere à legibilidade dos mesmos em discos, disquetes, fitas magnéticas e outros meios, prontos para o uso por programas tais com Spss, NSDStat+, SAS etc.

⁹ A sigla é formada em parte por uma palavra do jargão dos pesquisadores sociais empíricos alemães, que chamam de bus (ônibus) o questionário multitemático, porque nele são admitidas todas as propostas de perguntas que chegarem a tempo. All significa "geral".

Nisto eles se distinguem da forma como os serviços estatais de estatística apresentam os seus resultados: tabelas, programas de visualização (por exemplo Disi) e arquivos em CD ou disquetes legíveis em computadores pessoais ou bancos interativos na Internet, como o do IBGE.

Superestruturas nacionais e internacionais da pesquisa social empírica

INTERNATIONAL SOCIAL SURVEY PROGRAMM (ISSP)

Para melhor integrar os programas nacionais, pelo menos no que se refere aos conteúdos dos questionários, criou-se em 1985 o International Social Survey Programm (ISSP) para

- cooperar no desenvolvimento dos módulos temáticos que se referem a áreas importantes das ciências sociais;
- incluir estes módulos nos levantamentos nacionais regulares (ou, eventualmente, como levantamento próprio) como um suplemento aplicável em 15 minutos;
- incluir um núcleo extensivo de variáveis de fundo;
- colocar os dados quanto antes à disposição da comunidade científica das ciências sociais.¹⁰

Entretanto mais de vinte países cooperam no ISSP. Uma consulta recente (outubro de 1997) à *homepage* daquela organização¹¹ mostrou que os seguintes países realizam este tipo de levantamento com regularidade: Alemanha, Austrália, Áustria, Bangladesh (em construção), Bulgária, Canadá, Chile (em construção), Chipre, República Checa, Espanha, Estados Unidos da América, Filipinas, França, Grã-Bretanha, Hungria, Irlanda, Israel, Itália, Japão, Letônia (em construção), Países Baixos, Polônia, Portugal, Nova Zelândia, Noruega, Rússia, Eslováquia, Eslovênia e Suécia. Note-se que a América Latina está quase completamente ausente desta lista, enquanto muitos dos países ex-socialistas perceberam a utilidade de tais dados para acompanhar e monitorar a mudança social acelerada decorrente da queda do comunismo.

¹⁰ (a) jointly develop topical modules dealing with important areas of social science, (b) field the modules as a 15-minute supplement to the regular national surveys (or as a special survey if necessary), (c) include an extensive common core of background variables, and (d) make the data available to the social science community as soon as possible (Davis, e colaboradores, 1994, p.22-23).

¹¹ <http://www.issp.org/surveys.htm>.

As finalidades deste tipo de pesquisa originalmente eram as seguintes:

- obter dados cientificamente confiáveis, através de levantamentos contínuos, para monitorar os processos sociais;
- colocar os dados à disposição de todos os institutos interessados e, eventualmente, chegar a uma ação conjunta dos institutos com efeitos sinérgicos (Assim surgiu o GSS nos Estados Unidos; Davis e colaboradores, 1994, p.18).

O ISSP, além de servir como ponto de encontro daqueles institutos que produzem os relatórios sociais nacionais gerais, também está gerenciando um programa de levantamento de dados que se insere nos levantamentos nacionais. Conforme pode ser visto na tabela 8, as temáticas são específicas e variam de ano para ano, sendo que determinadas temáticas são repetidas com certa regularidade, para permitir estudos diacrônicos ou históricos.

TABELA 8.....
Temáticas dos levantamentos do ISSP

1985	<i>Role of Government</i>	Papel do governo
1986	<i>Social Networks</i>	Redes sociais
1987	<i>Social Inequality</i>	Desigualdade social
1988	<i>Family & Changing Gender Roles</i>	Família e mudança nos papéis de gênero
1989	<i>Work Orientations</i>	Orientação para o trabalho
1990	<i>Role of Government II</i>	Papel do governo II
1991	<i>Religion</i>	Religião
1992	<i>Social Inequality</i>	Desigualdade social II
1993	<i>Environment</i>	Meio ambiente
1994	<i>Family & Changing Gender Roles</i>	Família e mudança nos papéis de gênero
1995	<i>National Identity</i>	Identidade nacional
1996	<i>Role of Government III</i>	Papel do governo III
1997	<i>Work Orientations II</i>	Orientação para o trabalho II
1998	<i>Religion II</i>	Religião II
1999	<i>Social Inequality III</i>	Desigualdade social III

O ISSP entretanto permite estudos sociológicos inéditos, tal como aquele de Höllinger (1989), que mostrou diferenças entre as relações pessoais que se mobilizam em caso de emergência. Nos países de língua inglesa (Estados Unidos, Grã-Bretanha) e na Alemanha, além dos familiares mais imediatos, especialmente os pais, conta-se com os amigos ou até pessoas estranhas; na Áustria e na Hungria espera-se pouco de pessoas mais distantes, e na Itália espera-se menos do esposo ou da esposa do que dos demais familiares em caso de emergência.

Neste contexto é importante mencionar os bancos de dados encontrados em vários países. São os principais distribuidores dos dados levantados para os Relatórios Sociais Nacionais Gerais e os do ISSP. Eles foram criados para permitir as análises secundárias. Os arquivos recebem os dados e a documentação pertinente dos “produtores” e analisadores primários, os quais são colocados à disposição dos usuários secundários em forma legível para programas estatísticos e com uma documentação padronizada. Às vezes o acesso é limitado porque o produtor dos dados pode exigir determinadas condições para o uso secundário de seus dados. Esses arquivos têm grande importância didática tanto no ensino da metodologia como da sociologia e na divulgação de dados entre pesquisadores fora dos grandes centros de investigação social empírica. Por outro lado, vale até certo ponto a crítica de Davis e colaboradores (1994, p.18) que esses dados às vezes são “ultrapassados, reciclados e de amostragens locais questionáveis”.

Com o objetivo de ajudar a comunidade científica internacional das ciências sociais, foi criada a International Federation of Data Organizations for the Social Sciences – IFDO –, a federação internacional de organizações (de distribuição) de dados para as ciências sociais. “Organizações de dados” podem ser vistas como resultado de desenvolvimentos na ciências sociais que tiveram seu objetivo primário na coleta, no armazenamento e na distribuição de dados numéricos para análise secundária. Oferecem recursos de laboratório para ciências sociais da mesma forma como os laboratórios de física tradicionalmente fazem para as ciências físicas.

Com o rápido desenvolvimento na tecnologia dos computadores e os crescentes conhecimentos nos métodos analíticos, os fundadores da IFDO acharam que seria proveitoso coordenar os esforços para fomentar a pesquisa nas ciências sociais. Hoje, os arquivos da IFDO formam uma rede que permite o intercâmbio de dados e de informações em escala mundial, juntando organizações regionais, tais como o CeSSDA na Europa e Candoc no Canadá.¹²

Os membros mais importantes da IFDO são: África do Sul (SADA), Alemanha (ZA), Austrália (SSDA), Áustria (Wisdom), Bélgica (BASS), Canadá (DL, LSDB, NAC, CUDC), Dinamarca (DDA), França (BDSP), Grã-Bretanha (ESRC), Hungria (Tarki), Índia (Issrc), Israel (IL-SSDA), Itália (Adpss), Países Baixos (STAR, NHDA), Nova Zelândia (Nzsrda), Noruega (NSD), Suécia (SSD), Suíça (Sidos), Estados Unidos da América (DPLS, Icpsr, Irss, Issrucla, NORC, Roper). Observe-se a ausência absoluta da América Latina.

¹² Maiores informações em <http://www.za.uni-koeln.de/cooperation/en/IFDO/index.htm>.

NECESSIDADE DE UMA SUPERESTRUTURA DA PESQUISA SOCIAL EMPÍRICA

O fato de os dados produzidos precisarem servir a muitos usuários coloca muita responsabilidade naqueles que desenvolvem os questionários e, ao mesmo tempo, excluem os usuários posteriores na elaboração dos mesmos. Para poder chegar a um consenso em nível nacional ou internacional é preciso criar certas instituições aptas para coordenar os interesses divergentes. Disso resulta a necessidade de os cientistas sociais saírem de suas “torres de marfim” e se integrarem em estruturas institucionalizadas da comunidade científica. Os problemas a serem resolvidos são inúmeros, entre eles pode-se enumerar os seguintes:

- ritmo dos levantamentos (anuais, cada dois ou cinco anos?);
- inclusão de grupos etários (a partir de 16 ou 18 anos, codificação por data de nascimento ou em grupos?);
- equivalência do texto das perguntas, do desenho do questionário em várias (sub)culturas, etnias ou nações;
- tipo padronizado do comportamento do entrevistador;
- tipo de amostragem (aleatória ou por quotas?).

Para que se chegue a séries históricas e análises dinâmicas é necessário que as mensurações ocorram com regularidade. Nem sempre é válido repetir as perguntas da mesma forma como foram aplicadas nos levantamentos anteriores, já que os conteúdos ou a fraseologia podem mudar. Na Alemanha, por exemplo, onde o relacionamento com minorias recebe muita atenção, ocorreu, ao longo dos anos, uma mudança na forma oficial ou politicamente correta de chamar os imigrantes. A denominação passou pelas seguintes mudanças:

- “trabalhador estrangeiro” (*Fremdarbeiter*);
- “trabalhador hóspede” (*Gastarbeiter*);
- “concidadão estrangeiro” (*ausländischer Mitbürger*);
- “migrante por motivo de trabalho” (*Arbeitsmigrant*) ou simplesmente;
- “estrangeiro” (*Ausländer*).

A solução do Allbus foi ficar com o segundo conceito até que, a partir de 1994, essa denominação foi usada numa metade dos questionários e a última na outra metade. Outro exemplo de denominações sujeitas à forte alteração é o conceito de “socialismo” ou “socialista” e outras alusões a conflito leste-oeste.

Mudanças na estrutura administrativa dos países podem resultar em graves problemas. No desdobramento de Municípios e Estados se perde a oportunidade de fazer comparações se os endereços dos entrevistados não tenham sido registrados detalhadamente – uma informação eventualmente sujeita às leis de proteção de dados do cidadão.

O consenso internacional apresenta ainda problemas de outro vulto. Tanto a abrangência das instituições dedicadas ao manejo das pesquisas sociais nacionais gerais como a tecnologia da amostragem foram criadas em nível dos estados nacionais. Disto resulta que uma amostragem da população da União Européia, por exemplo, necessariamente será uma amostragem estratificada, que produz uma margem de erro maior do que uma amostragem não-estratificada, em que todos cidadãos europeus teriam exatamente a mesma chance de serem escolhidos.

Análise secundária de microdados oficiais

O levantamento de dados por agências oficiais é uma atividade que tem uma tradição tão longa como a nossa civilização. Basta lembrar da estória natalina no Novo Testamento, na qual o motivo da migração de José e Maria para Belém foi o censo. A coleta de dados pelos governos sempre teve uma função exclusivamente política, que quase sempre se opõe à participação dos cientistas. Além disso, as próprias ciências sociais não chegaram a um consenso interno: existe, por um lado, uma doutrina de “sociologia pura”, hostil ao empiricismo, e, por outro, um estilo de pesquisa que se exaure na obsessão com o detalhe empírico “socioestatístico” e que carece de uma base teórica (Esser, 1989).

MICRODADOS

“Microdado” é um conceito oriundo na “gíria” dos cientistas sociais empíricos. O termo “micro” quer indicar que os dados representam características das unidades mínimas do levantamento, por exemplo: pessoas, empresas, ônibus, e não os dados agregados em níveis superiores, por exemplo: municípios, setores econômicos, frotas de transporte urbano. Dados agregados não podem ser cruzados ou reprocessados. E o processo de agregação é irreversível para quem não dispõe de dados originais. Por isso é fundamental para o cientista social conseguir, dos serviços estatísticos, os dados em sua forma original, para poder cruzá-los, construir índices e fazer as suas próprias agregações.

A análise secundária de dados oficiais não é novidade na pesquisa social empírica. Muito pelo contrário, ela é a mais importante fonte de dados dos pioneiros, tal como Émile Durkheim. Seu trabalho sobre o suicídio se baseou na análise de registros de óbitos em comunidades francesas e alemãs.

Da história da pesquisa empírica conhecemos também os graves erros que ocorreram ao cruzarem-se dados já agregados. Trata-se da falácia ecológica.¹³ Por exemplo, se alguém quisesse provar que a criminalidade depende da desigualdade na sociedade, poderia avaliar os bairros de uma cidade com base numa escala de qualidade dos domicílios e estudar a frequência dos atos criminais de seus moradores. É fácil imaginar qual seria a conclusão a que chegaria. Através de um levantamento junto a uma amostra de todos os habitantes daquela cidade, cruzando a renda das pessoas com a atividade criminal por bairro, no entanto com muita probabilidade se chegaria à seguinte conclusão: nos bairros mais abastados são os mais pobres que mais frequentemente cometem crimes; nos bairros mais populares, são os mais ricos. A hipótese sobre uma correlação direta entre *status* social e frequência criminal não pode ser verificada (Schrader, 1978 p.137).

O uso de dados gerados pelos serviços estatísticos oficiais nos últimos anos tem sido facilitado tanto de parte dos provedores, que permitem maior participação da ciência, como de parte dos consumidores, cuja capacidade tanto de seus computadores como de seus programas de análise melhorou consideravelmente. Nos Estados Unidos estão à disposição amostras de uso público (*public use samples*) dos censos decenais, com um volume de até 5% da população (até 13 milhões de pessoas!), os levantamentos correntes da população (*current population surveys*) com os dados de 60 mil domicílios, o levantamento saúde (*national health interview surveys*, National Center of Health Statistics), educação (*high school and beyond*) e segurança (*national crime surveys*) (Alba e colaboradores, 1994, p.58). Novas formas de análise social se apresentam ao cientista pela grande representatividade e a continuidade dos levantamentos. No que se refere à continuidade, a mudança social agora pode ser medida por estudos longitudinais. A grande representatividade, por outro lado, permite evitar a falácia ecológica, porque o grande número de pessoas incluídas permite fazer correlações entre as características individuais e aquelas das coletividades (Alba e colaboradores, 1994, p.59).

¹³ Quando Robinson criou o termo *ecological fallacy* nos anos 50, ele se referiu ao espaço em geral, em sentido amplo, e não especificamente ao meio ambiente. Um estudo recente mostrou que Durkheim provavelmente também cometeu a falácia ecológica. Com dados holandeses comparáveis àqueles usados por Durkheim, no Leste da França e no Sudoeste da Alemanha, Poppel e Day chegaram à seguinte conclusão: "Achamos completamente explicável a diferença católico-protestante nas taxas de suicídio. Deve-se ao costume de classificar como 'morte súbita' ou 'morte por uma causa mal identificada ou não específica' uma grande parte dos óbitos ocorridos entre católicos, os quais teriam sido classificados como suicídios caso tivessem ocorrido entre protestantes". "We find the Catholic-Protestant differential in suicide rates to be explicable entirely in terms of practice of categorizing as 'sudden deaths' or 'deaths from ill-defined or unspecified cause' a large proportion of deaths among Catholics which would have been categorized as suicides had they occurred among Protestants" (1996, p.500).

VANTAGENS E DESVANTAGENS DA ANÁLISE SECUNDÁRIA DE MICRODADOS OFICIAIS

Para qualquer tipo de dados, a análise secundária tem as mesmas vantagens e desvantagens, a saber (Alba e colaboradores, 1994, p.63-64):

– a produção de conhecimentos deve ser aberta à crítica e às abordagens competidoras e deve garantir que se possa replicar e reanalisar a mesma. Nisso a análise secundária tem um papel extremamente crítico, porque a decisão entre interpretações alternativas se efetua pela reanálise dos dados;

– a análise secundária representa um uso eficiente dos recursos;

– a fraqueza mais importante é que os dados não correspondem à natureza do problema da pesquisa secundária.

Essas condições têm importância ainda maior quando se tratar de dados de fontes oficiais.

– dados oficiais frequentemente obedecem às prioridades do governo, e exatamente por isso o seu uso não deve ser monopolizado; eles devem ser analisados sob várias perspectivas;

– o levantamento de dados oficiais envolve altos custos e os dados frequentemente não recebem uma análise satisfatória nos órgãos oficiais. A análise não-oficial é justificada já por considerar os custos;

– uma vez que os dados oficiais foram gerados dentro do sistema referencial político e administrativo, às vezes não correspondem às exigências científicas.

Os dados oficiais frequentemente têm as seguintes características:

– grande número de casos porque os dados oficiais frequentemente foram derivados de censos completos ou amostras grandes. Assim é possível chegar a resultados para grupos sociais pequenos que numa amostra de, por exemplo, duas mil unidades dificilmente podem ser identificados;

– levantamentos oficiais têm uma quota menor de respostas recusadas porque os dados foram gerados em procedimentos administrativos ou os entrevistados foram obrigados por lei a participar. Mesmo assim, cada levantamento oficial pode ter a sua própria seletividade;

– a quota de recusas em levantamentos comerciais e científicos é relativamente grande e, com toda razão, se fala de uma tendência a favor da classe média (*middle-class bias*) cujos membros mais prontamente obedecem às expectativas dos entrevistadores. Os dados oficiais mais completos podem servir para estimar a defasagem dos levantamentos não-oficiais;

– os dados oficiais frequentemente descrevem tanto o indivíduo como o contexto social no qual está inserido. Os levantamentos por domicílio, por exemplo, obtêm tanto os dados de cada membro da família como também as características da moradia, da unidade doméstica etc. Nos dados prove-

nientes dos serviços de emprego/desemprego constam informações tanto sobre a pessoa como sobre suas condições de trabalho;

– os levantamentos dos órgãos oficiais freqüentemente têm uma longa continuidade e, às vezes, seguem as recomendações das agências supranacionais. Assim, a comunidade científica pode analisar as tendências de longa amplitude tal como, nos países industriais, as mudanças sociais desde a industrialização, ou, no Brasil, as mudanças sociais desde a época de Vargas. As comparações internacionais tornam-se mais importantes na medida em que o ingresso dos Estados-nação em integrações supranacionais e na famosa “globalização” têm efeitos mais imediatos sobre as suas populações;

– os dados oficiais tendem a corresponder às definições legais de fenômenos sociais. Assim, uma relação estável entre homem e mulher que não consta nos registros de casamentos pode ser ignorada;

– os dados oficiais distribuídos aos cientistas, por motivo de confidencialidade, são anonimizados quanto aos endereços e outros dados sobre a localização da moradia.

Análise de dados transversais e longitudinais

Dados de igual conteúdo, levantados por amostras consecutivas em épocas diferentes (meses, anos) são denominados “transversais”, em oposição aos dados de igual conteúdo levantados pela mesma amostra em épocas diferentes ou em “ondas”, denominados “longitudinais”. O primeiro tipo também é conhecido como análise de *trend* ou tendência, enquanto o segundo se chama análise de *panel* ou de painel. Na análise de tendência, o que importa é que os dois ou mais levantamentos de dados ocorram com todas as mesmas definições operacionais, desde a formulação das perguntas até o comportamento do entrevistador. Ideal seria que as diferenças se limitassem ao fato de que se trata de duas amostras tiradas do mesmo universo, de modo que se pode calcular o erro decorrente para corrigir as freqüências e os demais cálculos. Embora jamais se obtenha a situação ideal, essa é a forma mais comum nos relatórios sociais gerais.

Os levantamentos longitudinais, por outro lado, prometem evitar alguns problemas, porque as pessoas entrevistadas são sempre as mesmas. Eles permitem o estudo de cortes. No caso do painel surgem, no entanto, novos problemas (Engel e Reinecke, 1996, p.4-5), entre eles:

– dados sobre a mesma pessoa, levantados em épocas consecutivas, podem mostrar uma diferença, embora a pessoa, de fato, não tenha mudado e *vice versa*;

– mesmo que o registro de mudanças tenha validade, não se sabe como a pessoa mudou no intervalo de tempo entre os dois levantamentos;

– o fato de um elemento da amostra não responder a uma pergunta ou a toda a entrevista sempre é deplorável, mas no caso dos painéis é “desastroso” porque o volume da amostra se reduz a partir dali para todas as ondas posteriores.

Entretanto existem várias proposições para solucionar estes e outros problemas (Engel e Reinecke, 1996), mas ainda não há soluções consensuais. Também a manutenção de um painel é muito cara. Em primeiro lugar, o painel é muito maior em número de elementos do que uma amostra. Enquanto bastam em média 2 mil pessoas para representar o todo de uma sociedade, os painéis facilmente chegam a 60 mil unidades (como no caso do painel europeu de domicílios). O tamanho maior é necessário para melhor estimar os erros decorrentes da “mortalidade” e das ausências de respostas individuais. A “mortalidade” em sentido estrito e em sentido mais amplo também tem que ser tratada. Pessoas que saem do painel por morte ou emigração devem ser substituídas por novos elementos que se equivalem em todos os aspectos aos mortos ou emigrados.