

O Uso de Computadores na Análise Qualitativa: Uma Nova Tendência na Pesquisa Sociológica*

Wilma Mangabeira

A década de 80 representa um momento decisivo na expansão do uso de computadores nas ciências sociais. Com a introdução dos micros IBM-PC e seus similares, relativamente baratos e de ambiente "amigável", os cientistas sociais puderam estender o uso de computadores além da análise estatística, adotando-os também no gerenciamento de textos, no registro do trabalho de campo, no acesso a redes de dados, na comunicação através de correio eletrônico e na preparação de aulas.(1)

Essa expansão vem sendo agora acompanhada pelo surgimento de programas e sistemas destinados a processar *dados qualitativos*. O final da década passada marcou o término de uma era em que os programas de computação eram território exclusivo da pesquisa quantitativa e os computadores lidavam apenas com dados numéricos. Não é fácil formular uma definição de "dados qualitativos", na medida em que existem conceitos alternativos que variam segundo a abordagem sociológica adotada (Halpenny 1979).

Neste artigo emprego uma definição principalmente metodológica: dados qualitativos são dados não-numéricos, não-estruturados e de tamanho variável.

Esses dados podem ser obtidos por uma grande variedade de métodos de pesquisa, in-

cluindo-se, por exemplo, entrevistas não-estruturadas ou semi-estruturadas, observação participante e o registro de histórias de vida.

O artigo aborda duas questões relativas à nova tendência delineada acima. Em primeiro lugar, ilustro as novas possibilidades abertas por esse novo desenvolvimento, através de um relato pessoal do emprego que fiz do programa *The Ethnograph*, em minha própria pesquisa de doutoramento. Em segundo lugar, analiso as possíveis implicações dessa nova tendência, tanto para os paradigmas sociológicos, quanto para a prática cotidiana de pesquisa.(2)

Argumentarei que o uso de programas de computação para a análise de dados qualitativos representa uma importante abertura para a pesquisa sociológica, pois não só aumenta a velocidade da análise, como oferece condições para o pesquisador "brincar com seus dados" e explorar diferentes interpretações acerca do seu material. Além disso, esses programas permitem o aperfeiçoamento do exercício comparativo, em virtude de seu sistema de classificação e recuperação de dados. Por outro lado, argumento também que, se a nova tendência não for suficientemente discutida pela comunidade acadêmica, ela poderá dar margem, no futuro, a alguns resultados indesejáveis.

* Esse artigo foi traduzido por Vera Pereira. A revisão técnica foi realizada pela autora.

O artigo divide-se em três seções. Na primeira, apresento um histórico da penetração do uso de computadores nas ciências sociais e faço alguns comentários sobre o que considero ser a fase mais recente desse processo, ou seja, a emergência de programas destinados à análise qualitativa.

Na segunda seção exemplifico as possibilidades de um programa específico denominado *The Ethnograph*, através de minha pesquisa de doutoramento, que se baseou em trabalho de campo entre metalúrgicos de Volta Redonda. Minha intenção é demonstrar as vantagens do uso de programas de computador para análise qualitativa em comparação com o procedimento manual tradicional. Na conclusão, ofereço uma avaliação das implicações positivas e negativas dessa inovação para a pesquisa sociológica.

1. O Uso de Computadores nas Ciências Sociais: Breve Histórico

O início do emprego de computadores nas ciências sociais pode ser localizado no final dos anos 50, quando começaram a ser desenvolvidas metodologias para as ciências sociais, com o apoio da informática. Os computadores eram principalmente usados para fins estatísticos, embora também houvesse outros usos. Predominavam, então, os equipamentos de grande porte. (3) Embora se encontrem algumas referências na literatura sobre a utilização desses equipamentos nas ciências humanas — especialmente na arqueologia e nos estudos lingüísticos —, o seu custo elevado de operação e seu ambiente não-amigável, desestimulavam os pesquisadores, que eram em grande parte “não-alfabetizados em informática”.(4)

Durante os anos 70, a redução de custos tornou os *mainframes* mais acessíveis a centros universitários, o que criou uma demanda de programas mais específicos para as ciências sociais. Um exemplo pioneiro, foi o advento do *Statistical Package for the Social Science* (SPSS), hoje de uso generalizado na análise quantitativa, inclusive no Brasil. A criação de programas de banco de dados também expandiu a utilidade do computador como instrumento para as ciências so-

ciais. Apesar desses avanços, a comunidade de usuários na área de humanas permaneceu bastante reduzida até o final da década de 70.

No início dos anos 80, com a revolução da microeletrônica, a redução do tamanho dos computadores e a significativa queda nos preços tornaram o micro um bem de consumo cada vez mais acessível. A indústria de equipamentos, estimulada pela nova demanda, passou a investir em novos desenvolvimentos para os microcomputadores.(5) O resultante surgimento de programas baratos e de *uso genérico* aproximou os cientistas sociais do computador. Gerenciadores de arquivos, editores de texto, bancos de dados, organizadores de bibliografias e toda uma série de aplicativos de comunicação por computador, permitiram aos cientistas sociais o uso mais intensivo dos computadores, independentemente de sua orientação metodológica.(6)

Até esse momento, os pesquisadores interessados em dados qualitativos que haviam resistido — com toda razão, a meu ver — à “numerificação” de seus dados, eram marginalizados do acesso ao computador. Pode-se dizer que, até pouco tempo, o tipo de pesquisador “incompatível com computadores” ou “resistente ao computador” se localizava principalmente entre aqueles de orientação qualitativa, o que só reforçava os estereótipos de comportamento e uma rígida separação entre “pesquisadores quantitativistas” e “pesquisadores qualitativistas”. O aparecimento de micros acessíveis e a disponibilidade de uma grande variedade de pacotes de processadores de textos começaram a quebrar essa divisão; muitos qualitativistas tornaram-se “alfabetizados em computador”.

Quanto à análise de dados, contudo, os softwares disponíveis só aceitavam material numericamente codificado e altamente estruturado. Há descrições criativas de adaptações de *gerenciadores genéricos* de bancos de dados para a análise de dados qualitativos, em que se faz a “codificação” e a “busca” de dados textuais e não-estruturados.(7) Entretanto, os gerenciadores de banco de dados

podem ser considerados como um degrau inferior se comparados aos programas *especialmente desenhados* para análise qualitativa, do qual o *Ethnograph* é um exemplo.

A adaptação dos gerenciadores de banco de dados gerais para a análise qualitativa sofre de duas limitações. Primeiro, esses gerenciadores impõem ao pesquisador um gasto inicial de tempo a fim de “formatar” o programa para poder classificar os dados segundo a lógica desejada. Segundo, gerenciadores de banco de dados geralmente limitam o processo de codificação e indexação a palavras, em oposição a segmentos de texto maiores. Eles também não permitem a criação de códigos superpostos ou a marcação de segmentos de texto com fronteiras de códigos em comum.

Em comparação, quando se usa programas *especialmente elaborados* para análise qualitativa, o pesquisador só tem de inserir os dados no programa antes de rodá-lo. Esses programas, além de permitirem a codificação de segmentos maiores de texto, também localizam o contexto dos códigos, respeitando-se, assim, um elemento importante para a interpretação sociológica. Em resumo, uma importante vantagem dos programas *especialmente desenhados* para dados qualitativos, em relação aos gerenciadores de bancos de dados *genéricos*, é que os primeiros permitem dar um tratamento mais adequado à forma e ao conteúdo dos dados textuais.

Os programas *especialmente desenhados* para dados não-numéricos podem ser divididos em duas categorias. A primeira inclui os softwares especificamente criados para codificação e “busca” de códigos de dados qualitativos. Segundo Tesch, em 1990 existiam pelo menos cinco programas desse tipo no mercado.(8) Embora haja variação em alguns aspectos especiais desses programas, todos substituem a atividade manual básica da análise de dados qualitativos: *cortar e colar segmentos de texto*. Outra característica que reúne os cinco programas é a de terem sido formulados por pesquisadores de

orientação qualitativa, com habilidades diferentes no manejo do computador.

A segunda categoria de programa pertence ao ramo chamado de “expert-system” da inteligência artificial (IA). Esses sistemas são capazes de interpretar ou atribuir significados a um texto, ou fazer simulações ou previsões, de acordo com regras previamente especificadas (Hinze, 1987, Sproull, 1982). Em oposição, tanto aos programas de *uso genérico* encontrados no mercado, que têm um multiplicidade de usos dentro e fora do campo das ciências sociais, quanto aos “programas de análise qualitativa” acima citados, os “expert-systems” são elaborados por especialistas em informática em geral, para um projeto institucional com objetivos claramente definidos.(9)

Não irei me deter, no espaço deste artigo, sobre os diferentes tipos de programa, mas pretendo discutir sobre a utilização de um software particular, para a análise qualitativa, como apresento a seguir.

2. O Programa *The Ethnograph* e seu Emprego numa Tese de Doutorado

O *Ethnograph* (10) é um programa de computador que pertence ao primeiro grupo de programas mencionado acima. Foi criado por Seidel, Kjolseth e Seymour, todos três pesquisadores de orientação qualitativa. O programa substitui a tradicional atividade de “tesoura e cola” no momento da pesquisa em que as entrevistas ou observações de campo, por exemplo, já foram transcritas e a análise vai começar. No método tradicional, a análise se inicia pela classificação e codificação dos dados, fazendo-se diversas cópias do material e literalmente “cortando e colando” textos segundo categorias que vão surgindo ao pesquisador.

O procedimento básico para iniciar o trabalho com o *Ethnograph* é importar para o programa — através de um editor de texto — os dados qualitativos já devidamente transcritos. Cada arquivo de texto é importado para o *Ethnograph* através de um procedimento de conversão. O novo arquivo, convertido, chamado de *Ethno File*, é uma cópia

do arquivo de texto original, só que num formato aceito pelo programa.

No processo de conversão, pode-se renomear o arquivo original como recursos de proteção do anonimato dos informantes. Através do comando “numerar arquivos” (*number a file*), o programa enumera todas as linhas do arquivo, uma por uma, da primeira à última linha do texto (é permitido um arquivo com até 9999 linhas, apesar de não ser recomendável trabalhar com um arquivo tão grande).

O *Ethnograph*, então, imprime os arquivos, isto é, faz uma cópia impressa do arquivo com os dados, com as linhas numeradas uma a uma, deixando um espaço de três polegadas na margem direita para a colocação de códigos pelo pesquisador.

O programa não permite a entrada de comentários dentro do corpo original dos arquivos, uma vez que estes tenham sido convertidos em um *Ethno File*. Todas as alterações e notas sobre os dados devem ser feitas ou no arquivo original, antes da conversão para o programa, ou numa folha de papel em separado. Como a elaboração de comentários sobre os dados é, em si, um procedimento importante no processo de análise, a rigidez do programa nesse aspecto é uma limitação do *Ethnograph*.

Estando um arquivo com todas as suas linhas “numeradas” e impressas, pode-se partir para a classificação e codificação dos dados. Este é, sem dúvida, um primeiro passo do processo analítico da sociologia qualitativa. Nesta etapa, o pesquisador é convidado a sair da frente do computador, apanhar as cópias impressas e refletir sobre os dados no “modo tradicional”, isto é, com papal e lápis.

Assim, embora o programa tenha sido criado com o objetivo de eliminar tarefas mecânicas da análise qualitativa, a interação entre “pesquisador” e “dados”, nessa etapa classificatória, é feita por um procedimento manual. Essa é uma característica particular do *Ethnograph* e foi intencionalmente formulada pelos seus inventores. Segundo Seidel, “o pesquisador tem de envolver-se na manipulação dos dados a fim de ficar em

contato com eles”. (Seidel et al.: 1-3) Esse aspecto do programa foi bem recebido por alguns usuários e criticado por outros. Devido ao meu estilo pessoal de trabalho, sentime bastante à vontade com esse procedimento, na medida em que me permitiu uma “leitura” completa e comparativa dos arquivos impressos. Contudo, essa é uma opção muito individual; outros pesquisadores preferem que o processo de codificação se dê no monitor do computador, diretamente sobre o texto.

Uma vez terminado o primeiro processo de classificação e codificação manual, o pesquisador volta ao computador e digita os códigos referentes aos respectivos segmentos de texto. Os números que constam da cópia impressa, possibilitam que segmentos de texto sejam localizados e classificados com um ou vários códigos. Isso porque, como se sabe, na interpretação sociológica, um segmento de texto pode ter um ou vários significados.

A entrada dos códigos no programa se faz pela execução da opção “codificar os dados” (*code a file*), e a digitação dos números das linhas de texto e sua respectiva palavra-código ou palavras-códigos.

O programa é bastante aprimorado para permitir que cada segmento de texto possa conter até 12 diferentes códigos. Os segmentos de texto podem também ser sobrepostos ou embutidos um dentro dos outros até sete vezes.⁽¹¹⁾ O processo de recodificação dos arquivos também é muito simples, o programa permite mudanças nos códigos, bem como o acréscimo de novos códigos.

Durante o processo de entrada dos códigos no programa, o pesquisador fica inteiramente afastado do texto. O monitor funciona como uma “tela limpa”, que é ocupada pelos números das linhas e seu respectivo código ou códigos. Algumas críticas têm sido feitas ao procedimento com base na “tela limpa”, sob o argumento de que isso pode implicar a perda do contexto dos dados. (*Symposium on Qualitative Knowledge and Computing*, julho de 1989) Novamente, acredito que a avaliação depende do estilo pessoal de trabalho do pesquisador. De mi-

nha parte, achei que esse método me permitiu criar a distância necessária em relação ao material original. Além disso, o contexto do código é perdido apenas provisoriamente, pois é retomado quando se faz a “busca dos códigos” (*search for codes*) e se obtém um texto impresso contendo os resultados da busca.

Concordo com a idéia, discutida na literatura especializada, de que o processo de codificação é o momento culminante de dois movimentos que se dão ao mesmo tempo no processo de análise. De um lado, o pesquisador deve ter “amadurecido com os dados”, para citar uma expressão de Seidel, que aponta para a necessidade de uma profunda familiarização do pesquisador com seu material de pesquisa. De outro lado, o pesquisador deve ser capaz de manter um certo distanciamento dos dados originais, para que atinja algum nível de abstração e síntese.

Como sugeriu Sproull (1982:284), no processo de “categorizar” seus registros originais, o pesquisador acaba por descrever padrões recorrentes ou estabelecer sínteses. São essas categorias, suas recorrências e diferenças entre si, que se tornam a fonte da análise e não mais o material original. Os dados originais, na íntegra, são postos de lado para serem retomados na fase de redação, como “ilustrações”.

No *Ethnograph*, o processo de “localização de códigos similares entre arquivos diferentes” começa após a codificação e recodificação de dados, quando o pesquisador se considera minimamente satisfeito com a classificação elaborada. O “coração” do programa é o processo de “busca dos códigos” (*Search for codes*), quando o pesquisador pode passar a experimentar as múltiplas possibilidades de interpretação de dados, na busca de significados, padrões e regularidades. O programa permite a busca das ocorrências de um código isoladamente ou de múltiplos códigos simultaneamente. Pode-se também sofisticar a busca através das opções “e” e “não”.

O processo de “busca” pode ser enriquecido por outro recurso do programa, de-

nominado de “folha-de-rostro”, que oferece possibilidades analíticas muito mais complexas. No entanto, antes de passar a uma descrição detalhada desse processo, gostaria de dar uma idéia da minha própria pesquisa, como exemplo da utilidade do programa.

Principais Questões da Pesquisa de Tese

Minha tese de doutorado se intitula *Union Politics and Workplace Militancy: a case study of Brazilian Steelworkers in the 1980s* (Política Sindical e Militância na Fábrica: Estudo de Caso de Metalúrgicos Brasileiros nos anos 80) (Mangabeira, 1991). A tese analisa as relações entre a militância no interior da fábrica e a política sindical, no período posterior ao surgimento do “novo sindicalismo” no Brasil, na década de 80, e discute os problemas e dilemas enfrentados por esse novo tipo de sindicalismo. São dois os objetivos da tese. Primeiro, ela estuda em profundidade as relações entre a política no chão-de-fábrica e a política sindical numa usina siderúrgica. Segundo, faz uma avaliação do desenvolvimento do “novo sindicalismo” no Brasil, dez anos após seu nascimento. A tese discute se esse novo movimento significou, de fato, uma ruptura com os sindicalismos populista e burocrático, e favoreceu o surgimento de formas mais democráticas de política sindical ou se, com o passar do tempo, esse movimento teria sucumbido à “lei de ferro das oligarquias”.

Uma das metas declaradas do “novo sindicalismo” foi a de estreitar as relações entre os dois níveis de ativismo industrial — o sindicato e o chão-de-fábrica. A tese analisa que medida essa meta foi cumprida e discute a extensão em que a base operária passou a ser representada pelo “novo sindicato”. A discussão central da tese é o debate sobre a democracia sindical e os indicadores adequados para realizar esta avaliação. Meu esquema de análise se apóia em duas perspectivas teóricas: por um lado, o debate acerca da democracia sindical, encontrado na sociologia política e introduzido por Robert Michels (1959) e Roxborough (1984); por outro lado o conceito de Burawoy de *politics of*

production (1987) e as teorias sobre processo de trabalho.

No campo da sociologia política, a tese discute com a literatura sobre o “novo sindicalismo” e procura examinar o desenvolvimento de suas características gerais ao longo dos últimos dez anos. Dado que a literatura definiu esse tipo de sindicalismo como “novo” dentro do movimento trabalhista brasileiro, em virtude de sua militância mais ampla e da maior representatividade de sua liderança, o ponto de partida da tese foi a redefinição da política sindical, de modo a abarcar duas áreas da ação coletiva: o sindicato e o chão-de-fábrica.⁽¹²⁾ A primeira área focaliza as relações formais e institucionalizadas entre os operários manuais, como a base operária, e a liderança sindical. A segunda, analisa a *shop-floor politics*, enfocando, por exemplo, os conflitos cotidianos dos operários diante do processo de produção, das condições de trabalho, das políticas gerenciais e dos modos de mobilização e organização dos operários em torno de objetivos específicos.

Articulando esses dois aspectos da ação coletiva, a tese propõe uma perspectiva complementar, em que a ação política dos operários, no plano micro, fosse analisada junto com os aspectos institucionais do novo sindicalismo.

Dada a ênfase na relação entre liderança e trabalhadores, nos conflitos informais e nas organizações formais, o enfoque qualitativo pareceu mais adequado do que os métodos quantitativos, como, por exemplo, a pesquisa por questionário. Duas razões levaram a essa escolha. A primeira decorre da “natureza sensível” do tópico em questão.⁽¹³⁾ A segunda tem relação com a primeira e refere-se à natureza das respostas que buscava. Acredito que a opção por entrevistas em profundidade permitiu superar dois tipos de problema. Um foi a necessidade de quebrar com o discurso “pronto” de alguns líderes e ativistas politicamente articulados — que dificultava a compreensão de certas questões substantivas, tais como a natureza subjacente das divisões políticas ou os tipos de obje-

vos não-explícitos perseguidos pela liderança. Nesse sentido, as entrevistas em profundidade, orientadas por um roteiro flexível, permitiram-me voltar e, às vezes, insistir no esclarecimento de alguns pontos.

O segundo problema situava-se no outro extremo, nas entrevistas com pessoas pouco articuladas, não-organizadas e às vezes muito reticentes na verbalização de um assunto tão explosivo quanto a sua avaliação do sindicato e dos conflitos vividos ou observados no chão-de-fábrica. No caso desse grupo, entrevistas longas e abertas permitiram criar uma relação de “confiança” entre entrevistados e entrevistadora, bem como a utilização da própria linguagem e expressões dos operários, o que, de certa forma, facilitou a solução do problema de comunicação.

O trabalho de campo foi realizado entre um grupo de metalúrgicos da Companhia Siderúrgica Nacional, em Volta Redonda, Estado do Rio de Janeiro. O grupo foi escolhido para o estudo de caso por constituir um “caso crítico” de “novo sindicalismo”.⁽¹⁴⁾ Embora a fábrica em questão já estivesse em operação desde 1946, foi apenas depois da vitória da oposição sindical, em 1984, que ocorreu a primeira greve dos trabalhadores manuais. Desde então, houve dez greves na fábrica, com um nível de radicalização crescente. Em quatro ocasiões o Exército interveio para reprimir violentamente as greves e ocupações.

Eu vivi em Volta Redonda entre abril e julho de 1988 e voltei ao trabalho de campo entre janeiro e fevereiro de 1990. Durante esse tempo, realizei 72 entrevistas gravadas; 55 foram obtidas junto a metalúrgicos que, ou trabalhavam na fábrica, ou eram ativistas ou sindicalistas do Sindicato dos Metalúrgicos de Volta Redonda. Cinco dessas entrevistas foram realizadas duas vezes com os mesmos informantes; as 12 entrevistas restantes foram feitas com gerentes da companhia, líderes da Igreja Católica, mulheres de metalúrgicos engajadas nos movimentos comunitários, um advogado e assessores do sindicato.

Dentre as entrevistas, 55 foram obtidas pela técnica da “bola de neve aleatória”, iniciada através de duas redes diferentes: contatos no Sindicato dos Metalúrgicos de Volta Redonda e na diocese da Igreja Católica local, que era ponto de encontro da “pastoral operária”.(15) Essas duas redes iniciais depois se expandiram para outras três, baseadas nos bairros de moradia dos trabalhadores. As entrevistas apoiavam-se num roteiro, abarcando um conjunto de temas relativos à história do operário antes de ele entrar na Companhia Siderúrgica Nacional, o histórico de sua família no trabalho manual e como ele foi admitido na companhia.

Indagava sobre sua educação e treinamento, e solicitava uma descrição das suas tarefas atuais e passadas, a rotina do trabalho, sua percepção do conjunto do processo de produção — “você sabe como é que se faz o aço?” — e os níveis de autonomia e decisão experimentado no seu posto de trabalho. Solicitava também uma avaliação sobre seu trabalho, sobre o significado de trabalhar numa fábrica estatal, sobre as relações com os companheiros e a gerência. Um conjunto importante de perguntas destinava-se a obter informações sobre os conflitos no interior das seções no chão-de-fábrica de que o operário tinha conhecimento, ou dos quais tivesse participado — especialmente a natureza e a dimensão desses conflitos, seus protagonistas e como se encaminharam as soluções.(16)

Outro grupo de questões voltava-se para a participação do informante no sindicato e suas opiniões sobre este. Perguntava sobre a história da “sindicalização” e procurava reconstituir esse relato até os dias de hoje. Quando o entrevistado não era sindicalizado, eu tentava descobrir suas razões para isso. A presença cotidiana do sindicato na fábrica e o conteúdo da ação dos ativistas também eram explorados. Algumas perguntas buscavam saber as opiniões dos entrevistados a respeito das dez greves na companhia, após 1984, e sobre seu comportamento nessas ocasiões: de adesão ou rejeição.

Eram igualmente exploradas sua participação em associações não-sindicais, como a Igreja e organizações comunitárias, assim como as preferências políticas partidárias.

Finalmente, havia um conjunto de questões específicas, dirigidas aos operários que eram líderes. Procurei traçar a trajetória de sua “politização”, a sua afiliação a partidos políticos e organizações, quando era o caso, bem como preoquepei-me em obter um julgamento geral a respeito da administração sindical após o surgimento do “Grupo de Oposição Sindical”.

Tentei ainda explorar a posição dos entrevistados em relação às diferenças centrais de trabalhadores da época — a CUT e as duas CGT's — além dos objetivos de curto e longo prazo dos líderes. Algumas perguntas foram formuladas envolvendo as diferentes facções dentro e fora do sindicato, tentando compreender a posição pessoal dos líderes acerca das divisões que se davam. Um outro grupo de questões referia-se à administração e ao cotidiano do sindicato, o processo decisório, as formas de prestação de contas à base operária e os modos de definição das prioridades.

A Codificação dos Dados Qualitativos

De volta a Londres, iniciei a difícil e trabalhosa tarefa de transcrição de todo esse material no computador. Usei o programa *The Ethnograph* para o grupo de 60 entrevistas com os operários metalúrgicos e sindicalistas. As 12 restantes foram deixadas de lado para serem tratadas como material complementar e para triangulação das informações. No processo de codificação das 60 entrevistas, criei 64 códigos diferentes. Após a codificação e a mudança das classificações tornou-se claro que esse processo guardava uma certa lógica subjacente, ainda que esta tivesse permanecido inconsciente.

O processo de classificação foi desenvolvido a partir de três níveis de abstração, produzindo três “tipos distintos de códigos”. O primeiro tipo pode ser chamado de “código de informação”, de nível muito concreto; tratava-se de segmentos de texto relativos a da-

tas, nomes de pessoas ou lugares e quantidades. O segundo tipo baseava-se nas categorias, percepções e crenças dos informantes. São os códigos baseados nas *categorias nativas* dos informantes. A categoria “peão” é um exemplo desse tipo de código: todos os trabalhadores se identificam com essa categoria. Isto é, todo trabalhador manual era um “peão”. Por oposição, apenas alguns operários eram “profissionais” e era esse o critério de diferenciação entre os “peões”.

O terceiro grupo de códigos pode ser chamado de *analítico* — e resultou das percepções da pesquisadora sobre os elos não-evidentes do material empírico, constituindo sínteses e abstrações. Prosseguindo no exemplo acima, “identidade” e “diferenciação” são códigos analíticos de “peão” e “profissional”. Outro exemplo de “código analítico” são os “tipos políticos” criados pela pesquisadora, que se sugere existir em outros contextos de “novo sindicalismo” no Brasil. Esses tipos são “o sindicalista paternalista”, o “radical reformista” e o “revolucionário”.

Vale notar que, na minha experiência, os tipos de código ligados a certos segmentos de texto mudavam conforme se desenvolvia a análise. Alguns “códigos de informações”, por exemplo, eram mais tarde transformados em “categorias” ou “códigos analíticos”. Um relato completo do processo de produção do aço elaborado por um informante, por exemplo, podia ser lido inicialmente como um “código de informação” mas também pode ser entendido como um saber estruturado a respeito do processo de produção e um indicador expressivo do nível de qualificação do trabalhador. Nesse caso, o mesmo segmento de texto, que fora definido como um “código de informação”, se torna também um “código de categoria”, refletindo uma nova compreensão da pesquisadora acerca dos mesmos dados.

Como foi dito anteriormente, o processo de codificação é uma primeira etapa da análise, que apesar de muito importante, deve ser seguida por uma segunda fase: a de “busca de ocorrência dos códigos” (*search for codes*). No processo de “busca”, o pes-

quisador pode testar a riqueza de um *insight*, os níveis de recorrência de padrões, o contexto em que esses padrões variam e as demais sínteses analíticas próprias do ofício do cientista social.

No processo de “busca”, o recurso da “folha-de-rosto” (*face sheet feature*), que o *Ethnograph* oferece, permite explorar variações nos dados através de variáveis nitidamente delineadas. Conforme assinala Seidel (1988: 14-s1), o recurso da “folha-de-rosto” permite que o pesquisador acrescente uma nova dimensão à análise de dados, complementando os códigos usados na definição dos segmentos de texto.

Para usar “folhas-de-rosto” no processo de busca, é necessário criar primeiro uma espécie de gabarito, ou seja, uma lista com os nomes das variáveis com que se deseja trabalhar, e detalhar se estas são numéricas ou textuais. Criado o gabarito, pode-se elaborar uma “folha de rosto” para cada um dos arquivos. Isso significa que o programa pedirá ao pesquisador para “completar” quais são os valores de cada variável, para cada um dos arquivos respectivos. Feito isso, a “folha de rosto” pode ser usada durante o procedimento de “busca de códigos” com até seis diferentes variáveis, aumentando assim, as possibilidades de análise.

Em minha pesquisa utilizei folhas-de-rosto com “variáveis”, tais como idade; anos de escolaridade; se o pai também foi trabalhador da CSN; nome do posto de trabalho atual do informante, seção em que trabalha na fábrica, etapa da linha de produção a que se liga o posto; se sindicalizado ou não, em quem votou na última eleição sindical; se recorreu à justiça do Trabalho contra a CSN, quantas vezes e quando; em quem votou nas últimas eleições nacionais e filiação partidária, se fosse o caso.

Um exemplo simples de busca seletiva com o emprego de variáveis de “folha-de-rosto”, seria “buscar” os segmentos de texto codificados como “participação sindical”, para cada informante, com a variável de “folha-de-rosto” para idade. Se poderia observar então, se, e em que medida, o aspecto gera-

cional ajuda a explicar o engajamento de alguns em atividades de militância sindical.

Os resultados do estudo de caso revelam que a contribuição do movimento do novo sindicalismo foi especialmente significativa no uso político da CIPA (Comissão Interna de Prevenção de Acidentes do Trabalho) e na forma inovadora com que se utilizou o antigo recurso à justiça do Trabalho. Essas duas práticas foram interpretadas como ilustrativas das tentativas de ampliação dos direitos dos trabalhadores, bem como da criação de novas bases de garantia de sua aplicação. O estudo de caso concluiu também que a dinâmica interna do movimento do “novo sindicalismo” ainda contém elementos que podem ser considerados antidemocráticos, e que isso se desdobra numa série de novos problemas e dilemas para o sindicalismo no Brasil nos anos 90.

Espero que essa breve descrição tenha demonstrado as enormes possibilidades abertas pelo uso de um programa de computador que respeita e potencializa as especificidades dos dados qualitativos. Penso que o programa *The Ethnograph* me proporcionou possibilidades de análise que teriam sido extremamente difíceis, ou mesmo impossíveis, caso tivesse de manipular artesanalmente as 60 entrevistas transcritas.

3. Conclusão

Este artigo ilustrou o uso de um software particular — *The Ethnograph* — para a análise de dados qualitativos numa pesquisa sociológica. Programas como este podem ser de grande utilidade para a análise de dados qualitativos, pois permitem o tratamento adequado de dados que são não-numéricos, não-estruturados e de tamanho variado.

O uso desse tipo de programa constitui, na minha opinião, não somente uma mudança de grau — no sentido de que o pesquisador pode experimentar mais livremente com seu material de pesquisa, livrando-se dos aspectos mecânicos e enfadonhos da análise —, mas representa também uma mudança de qualidade. Acredito que esse tipo de programa não só aumenta a velocidade e eficiência da análise, como oferece novos modos de o-

lhar os dados, já que permite ao pesquisador trabalhar simultaneamente com uma grande variedade de dados.

Finalmente, pode-se sugerir que a expansão do uso dos novos programas qualitativos possibilitará ampliar os exercícios comparativos entre monografias e estudos de caso, posto que a classificação e recuperação dos dados podem ser efetuadas com maior rigor e facilidade.

Já que indicamos vários aspectos positivos do advento desses programas para dados qualitativos, convém indicar alguns dos problemas que podem surgir com essa nova tendência.

O primeiro problema diz respeito às concepções do criador do software quanto ao ofício sociológico e seu modo de trabalhar, que não são evidentes para o usuário mas estão subjacentes na estrutura de uso do programa. No programa de Seidel, por exemplo, existe um “convite compulsório” para que o pesquisador saia do computador e classifique e codifique seus dados a partir de cópias numeradas dos mesmos. Similarmente, o processo de digitação desses códigos ao programa é feito numa “tela limpa”, longe das transcrições das entrevistas. O pesquisador pode também quantificar determinados aspectos de seus dados através do uso da “folha-de-rosto” e da atribuição de frequências.

Todas essas características retratam opções tomadas quando o software foi criado e refletem tendências e maneiras de trabalhar dos autores. Acredito que isso aconteça com todos os programas, mais ainda com os “expert-systems”, nos quais o próprio programa busca ter capacidade de interpretar e produzir inferências ou previsões.

Um dos possíveis problemas dos “expert-systems” é o de implicitamente impor à sociologia modelos de racionalidade próprios dos especialistas de informática, antes que estes sejam plenamente compreensíveis para a disciplina. Esse perigo se deve ao fato de os “expert-systems” trazerem “lógicas” subjacentes estranhas aos paradigmas sociológicos. Conceitos como os de “consis-

tência”, “inconsistência” e “contradições” possuem significados muito distintos para um sociólogo e um especialista em informática. A saída para esse problema parece encontrar-se no acesso a informações mais amplas acerca dos diferentes programas disponíveis e na discussão apropriada, no campo das ciências sociais, a respeito de suas vantagens comparativas.

Outro problema diz respeito ao risco de se cair num processo de “dominação do computador”. Refiro-me à situação quase caricatural em que a generalização do uso de programas qualitativos pelos cientistas sociais se torne uma “moda”, e as pesquisas que tiverem utilizado esses programas ganhem legitimidade imediata, em oposição ao que passaria a ser visto como uma análise manual “antiquada e tradicional”.

Sou favorável a uma abordagem pluralista, que sustenta a decisão do pesquisador de usar ou não o computador. É necessário também garantir a circulação de informação a respeito dessa nova tendência, seus sucessos e seus problemas, de modo a permitir que o pesquisador decida ou não pelo uso do computador, ou escolha o programa mais adequado para sua pesquisa. Na posição pluralista, as decisões são tomadas de acordo com critérios de rigor e qualidade no ofício de pesquisa, pois, como afirmaram Frantzich e Purtkitt (1987:501), a marca de um bom pesquisador é ele saber procurar o tipo mais adequado de instrumento de trabalho.

*Recebido para publicação
em maio de 1992.*

Notas

1. Este artigo é uma versão revista do trabalho que apresentei à Associação Internacional de Sociologia e que recebeu um prêmio da ISA na competição “jovens sociólogos”. Madrid, 1990.
2. Foi extremamente proveitosa minha participação no estimulante seminário sobre “Pesquisa Qualitativa e Computação”, realizado na Universidade de Surrey, em 11-12 de julho de 1989. O seminário foi organizado pelos doutores R. Lee e N. Fielding. Durante as sessões, vários pesquisadores da Inglaterra, Estados Unidos e Austrália tiveram oportunidade de debater e avaliar essa nova tendência. Veja livro, Fielding e Lee (1991).
3. Trata-se do *main-frame*. Refere-se aos grandes computadores desenvolvidos durante as décadas de 50 e 60. Esse também é o termo geral para computadores de uso múltiplo, destinados a atender às necessidades de uma grande empresa, em oposição aos computadores pessoais, de emprego mais limitado. Veja *QUE's Computer Dictionary* (1990, p. 285).
4. Para mais detalhes veja Evans e Bernard (1983); Patton e Holoien (1981) e Rahtz (1987).
5. Embora o microcomputador estivesse disponível comercialmente desde o início dos anos 70, a difusão de seu uso entre os cientistas sociais é um fenômeno mais recente. Nos Estados Unidos, por exemplo, uma pesquisa realizada pelo American Council of Learned Societies, descobriu que em 1980 apenas 2% dos acadêmicos tinham acesso a um micro. Já em 1985, 95% passaram a utilizar um computador. Em Hinze (1987, pp. 440).
6. É importante ressaltar que existe uma resistência por parte de alguns pesquisadores, especialmente nos Estados Unidos, à expansão do emprego de computadores nas ciências sociais, já que se responsabiliza o computador pela excessiva ênfase na pesquisa quantitativa. (Sproull e Sproull, 1982, p. 288) É difícil comentar com mais detalhes esta reação. Purkitt (1987, p. 486), no entanto, sugere que a variável geracional pode explicar um aspecto dessa reação. Segundo o autor, os professores mais velhos resistem mais à difusão do computador do que seus alunos.

7. Alguns exemplos interessantes são relatados em Evans e Bernard (1983) e Garson (1986). Sobre experiências mais recentes, veja Guillet (1985), Hudson e Hudson (1986) e O'Neil (1987).
8. Os cinco programas, mencionados por Tesch, são: *The Ethnograph*, *Qualpro*, *TAP*, *Textbase* e *Hyperqual*. Todos eles importam textos de editores de texto e executam "buscas" de segmentos codificados. Mais detalhes desses programas são encontrados em Tesch (1991).
9. Dois exemplos dessa tendência incluem um programa denominado "N-ACT", que utiliza o modelo dramatúrgico de Goffman para a análise da integração social, criado por Brent, E. (1986) e um programa incompleto criado por A. C. Finkelstein e M. D. Fischer (1989), que se propõe "a descrever comportamentos sociais e teorias sobre o comportamento social e a fazer previsões verificáveis a partir das descrições"; o programa foi aplicado a uma etnografia de uma comunidade punjabe no Paquistão. O primeiro projeto está ligado à Universidade de Missouri, Estados Unidos e o segundo vincula-se a uma pesquisa conjunta da Universidade de Kent e do Imperial College, de Londres.
10. O programa *The Ethnograph* roda em equipamentos IBM-PC, XT ou AT e compatíveis. Ele requer uma memória de 256K.
11. No *The Ethnograph* "segmentos sobrepostos" são trechos codificados que dividem uma ou mais linhas do arquivo de dados. "Segmentos embutidos" são trechos codificados que ficam completamente contidos dentro de um segmento codificado maior. Veja Seidel (et al.), (1988, pp. 1-2).
12. Para fins deste artigo, apresento o debate de modo muito simplificado. O "novo sindicalismo" foi definido como um movimento militante: por sua maior propensão à atividade grevista, pelas formas de luta que escolhe, pelo uso politizado de canais legais e pela introdução de questões que atingem diretamente a organização da produção. Os sindicatos que integram esse novo movimento são definidos como mais representativos de sua base, como tendo propiciado a emergência de novos líderes e encorajado a representação nos locais de trabalho. Eles tentam alargar a base de participação e favorecem eleições mais competitivas. Foi assim que a literatura sobre o tema definiu e caracterizou o movimento. Veja em Alvaro Moises (1982), Humphrey (1980, 1982), Tavares de Almeida (1983) e Keck (1989), dentre outros.
13. Para uma definição detalhada de "tópicos sensíveis", veja Lee e Renzetti (1990).
14. "Caso crítico" (*critical case*) foi definido por Yin como uma situação empírica que atende a todas as condições de teste de uma teoria, quer ela seja confirmada, contestada ou estendida. Veja Yin (1989, pp. 47).
15. Para uma discussão detalhada sobre a técnica da "bola de neve aleatória", veja Biernacki e Waldorf (1981, pp. 155).
16. As entrevistas foram realizadas na casa dos trabalhadores, no sindicato e no escritório da Diocese. Foram entrevistas semi-estruturadas. O roteiro de entrevista foi mais usado como uma "lista de controle" das questões que eu gostaria de ver comentadas pelos entrevistados do que como um questionário de perguntas fechadas.

Bibliografia

- Alvaro Moises, José.
1982 "What is the Strategy of the New Syndicalism?". *Latin American Perspectives*. N.º 35, Vol.9 Fall.
- Biernacki, Paul. e Waldorf, Dan.
1981 "Snowball Sampling. Problems and Techniques of Chain Referral Sampling."

Sociological Methods and Research. Vol. 10, N.º 2, novembro.

- Burawoy, Michael.
1987 *Politics of Production*. Londres, Verso Press.
- Evans, M. e Bernard, R.
1983 "New Microcomputer Techniques for Anthropologists." *Human Organization*. Vol. 42, N.º 2, Summer.
- Fielding, Nigel. e Lee, Ray.
1991 *Using Computers in Qualitative Research*. Londres, Sage Publications.
- Garson, Elihu.
1983 "Computing in Qualitative Sociology: An Approach to Structures Text." *Qualitative Sociology*. N.º 9, Vol. 2, Summer.
- Guillet, David.
1985 "Microcomputers in Fieldwork and the Role of the Anthropologist." *Human Organization*. Vol. 44, N.º 4, Winter.
- Halfpenny, Peter.
1979 "Analysis of Qualitative Data." *Sociological Review*. Vol. 27, N.º 4.1.
- Hinze, Kenneth.
1987 "Computing in Sociology: Bringing Back the Balance." *Social Science Microcomputer Review*. N.º 5, Vol. 4, Winter.
- Hudson, R. e Hudson, N.
1986 "Computers for Anthropological Fieldwork." *Current Anthropology*. Vol. 27, N.º 5.
- Humphrey, John.
1980 "As Raízes e os Desafios do Novo Sindicalismo da Indústria Automobilística." *Estudos CEBRAP*, N.º 26.
1982 *Capitalist Control and Workers' Struggle in the Brazilian Auto Industry*. Princeton, Princeton University Press.
- Keck, Margareth.
1989 "New Unionism in the Brazilian Transition" in *Democratizing Brazil. Problems of Transition and Consolidation*. Alfred Stepan (ed.). Oxford, Oxford University Press.
- Lee, Ray. e Renzetti, Claire.
1990 "Problems of Researching Sensitive Topics." *American Behavioral Scientist*. Vol. 33, N.º 5, Maio-Junho.
- Mangabeira, Wilma
1991 *Union Politics and Workplace Militance: A Case Study of Brazilian Steelworkers in the 1980s*. Tese de Doutorado, London School of Economics and Political Science – LSE.
- Michels, Robert.
1959 *Political Parties*. Dover Publications, New York, (primeira edição 1915).
- O'Neil, D.
1987 "Anthropological Computing in the Mid- 1980s." *Social Science Microcomputer Review* 5:4, Winter.
- Patton, P. e Holoien, R.
1981 *Computing in the Humanities*. Hampshire, Gower Publishing.
- Purkitt, Helen e Frantzich, Stephen.
1987 "Computers in Political Science." *Social Science Microcomputer Review*, N.º 5, Vol. 4, Winter.

- 1990 *QUE's Computer User's Dictionary*. Indiana, QUE Corporation.
- Rahtz, Stephen. (ed.).
 1987 *Information Technology in the Humanities. Tools, Techniques and Applications*. New York, Halsted Press.
- Roxborough, Ian.
 1984 *Union and Politics in Mexico: The Case of the Automobile Industry*. Cambridge, Cambridge University Press.
- Seidel, John Kjolseth, Rolf. and Seymour, Elaine.
 1988 *The Ethnograph. (Manual do Programa)*. Colorado, Qualis Research Associates (primeira versão de 1985).
- Sproull, Lee. Sproull, Robert.
 1982 "Managing and Analyzing Behavioral Records: Explorations in Non-numeric Data Analysis." *Human Organization*. Vol. 41, N.º 4, Winter.
- Tavares de Almeida, Maria Hermínia.
 1983 "Novas Demandas, Novos Direitos: Experiências do Sindicalismo Paulista na Última Década." *DADOS*. Vol. 26.
- Tesch, Renata.
 1991 *Qualitative Research. Analysis and Types of Software Tools*. London, Falmer Press.
- Yin, Robert.
 1989 *Case Study Research. Design and Methods*. London, Sage Publications.