
Algumas Considerações sobre Coleta de Dados para a Pesquisa Qualitativa

Marcos Cáprio Fonseca Soares¹

Resumo: O presente ensaio destina-se a uma breve reflexão acerca de um importante tópico da pesquisa científica: trata-se da questão da amostra. Com efeito, reconhecendo as peculiaridades que circundam as duas orientações que polarizam a aferição da realidade na pesquisa social, quais sejam, a orientação quantitativa e a orientação qualitativa, podemos contemplar como vem se desenvolvendo um enfoque teórico que procura refletir ditas nuances, especificamente no que tange à questão enfocada.

Palavras-chave: amostra; *corpus*; pesquisa qualitativa; amostragem.

Introdução

Desde o seu nascimento, as ciências sociais se encontram em meio à discussão a respeito da precisão de suas inferências, tendo, ao tempo de seu nascedouro, muitos pensadores refutado a verossimilhança de êxito em seu intento². Ocorre que as ciências naturais gozavam já de um *status* calcado na idéia (hoje revista) de certeza, baseada nas experimentações e medições. Sob forte influência desta perspectiva, as ciências sociais marcaram seus primeiros passos justamente pelas elucubrações respaldadas em dados estatísticos (lembramos Durkheim), como forma de se atingir, ou ao menos aproximar-se ao máximo possível, da simétrica precisão que ainda se acreditava apresentar nas ciências da natureza. Nisto se caracterizaram as *pesquisas quantitativas*. Contudo, com as críticas ao paradigma positivista, emerge a

¹ Bacharel em Direito pela UCPel e em Ciências Sociais pela UFPel, Professor da Universidade Católica de Pelotas, Mestre em Sociologia pela UFRGS.

² Nisto, lembramos Tobias Barreto, o qual chegou a assim se referir: “...sociologia é o nome de uma aspiração tão audaciosa quanto vã”. (MARQUES NETO, 2001, p. 158).

abordagem qualitativa nas ciências sociais, apregoando uma forma peculiar de inferência da realidade, preocupada menos com as quantificações estatísticas, mas mais com ilações voltadas a representações, conceitos, definições, etc. Ocorre que entre estas duas formas de abordagem científica da realidade social, ostentadas por diferentes perspectivas teóricas e sob o auspício de distintos argumentos, há um ponto estratégico que tem recebido diminuta atenção por parte dos pensadores especializados: trata-se da seleção de elementos da população que se quer estudar, para se proceder à análise e quantificação ou compreensão, interpretação, definição, em suma, conforme o caso. Tradicionalmente, o resultado de tal seleção é denominado “amostra” e a técnica para se proceder à mesma “amostragem”. As possibilidades, bem como pertinência e a existência ou não de peculiaridades das mesmas nas inferências qualitativas será objeto da exposição seguinte.

Para tal intento, iniciaremos com uma explanação voltada para as teorizações a respeito da amostragem nas pesquisas quantitativas, posto que aí foi onde mais se desenvolveram idéias até aqui. Posteriormente, adentraremos na incipiente discussão que envolve o tema no âmbito das pesquisas qualitativas.

1.Considerações sobre amostragem na pesquisa social

Com efeito, a prática da pesquisa social incide sobre um objeto que, além de mui complexo, se constitui, no mais das vezes, em uma unidade de análise mui extensa. O sociólogo, amiúde, debruça seu olhar sobre coletividades, compostas por subjetividades díspares, heterogêneas entre si. Com isso, o conhecimento que busca a respeito desta coletividade deve ser subsidiado por dados a serem coletados junto à mesma, e para que tais dados apresentem-se úteis ao intento, devem guardar em si ao máximo as características da população, inclusive com aquela mencionada heterogeneidade.

Por diversas razões se faz profícuo, e por vezes até mesmo cogente, a inferência empírica de parte da coletividade a que temos por meta conhecer (sua totalidade ou algum aspecto que apresente); e é exatamente neste ponto que surge a noção de amostra, posto que, conforme definição de Antônio Carlos Gil, trata-se de *um subconjunto do universo ou da população por meio do qual se estabelecem ou se estimam as características do universo* (população). Universo ou população, por sua vez, é definido como *um conjunto de elementos que possuem determinadas características* (GIL, 1999, pp.99-100). Aliás, nuança de extrema

relevância, não obstante não muito bem delimitada muitas vezes, refere-se a existente entre “amostra” e “amostragem”, pois enquanto a primeira cinge-se aos elementos selecionados da população alvo, “amostragem” corresponde às diretrizes procedimentais voltadas para tal seleção.

Por outro lado, casos há em que não é recomendável a utilização de amostra na pesquisa social, como bem observa Pedro Alberto Barbeta, ao asseverar três situações em que tal ocorreria: quando a população for pequena, pois neste caso a fração tomada por amostra teria que ser muito grande, quando comparada com a população investigada – fala o autor em 80%; em segundo lugar, teríamos a situação onde a variável que se quer observar tem característica de fácil mensuração; ou ainda, quando há mister de alta precisão (saber o número de habitantes, por exemplo), como nos casos em que o IBGE, a cada 10 anos procede ao “censo nacional”(BARBETTA, 2004, p. 44).

Maria Immacolata Vassalo Lopes subdivide a pesquisa científica em diferentes *instâncias*. Neste sentido, fala-se nas instâncias epistemológica, teórica, metódica, e técnica, sendo certo que a questão da amostragem se situa nesta última³.

As amostragens podem ser classificadas em dois grandes grupos: amostragem aleatória (também denominada por alguns de “probabilística”) e amostragem não aleatória (também conhecida como “não probabilística”). O divisor de águas entre ambas é precisamente a forma pela qual serão selecionados os elementos da população para compor a amostra. Nisto, com relação à seleção dos elementos que compõem a amostra, temos a opção do sorteio, o que confere à amostragem a adjetivação de “aleatória”⁴.

A classificação da amostragem aleatória (como será denominada neste ensaio), tende a ser mais ou menos equivalente entre os autores, alguns deles acrescentando uma ou outra espécie, mas podemos dizer que, fundamentalmente, as principais são: amostragem aleatória simples; sistemática; estratificada e de conglomerados (BARBETTA, 2004, pp. 45 e ss.). Conforme já

³ Nesta instância técnica se constrói os dados ou o objeto empírico. Abrange os atos tendentes à coleta das informações e das transformações destas em dados consentâneos à *problemática geral*. O âmbito da técnica, conforme a autora, é o que possui o maior número de operações: três. Com isso falamos em operação de observação, seleção e operacionalização (LOPES, 1990, p. 110-1).

⁴ Pedro Alberto Barbeta, em sua obra já citada anteriormente, enfatiza a amostragem aleatória, por entender que a mesma facilita a análise dos dados e fornece maior segurança ao generalizar os resultados para a população. As técnicas clássicas de inferência estatística são viabilizadas, no dizer do autor, por esta espécie de amostragem (BARBETTA, 2004, p. 45).

dito, estes tipos de amostragem se classificam assim em razão da forma de proceder à seleção dos elementos que a comporão.

A amostragem aleatória simples consiste em associar um algarismo a cada elemento da população, associação esta absolutamente aleatória, e posteriormente, “sortear” um destes algarismos. Desta feita, os elementos vão sendo selecionados ao acaso”(BARBETTA, 2004, p. 45).

Já amostragem sistemática é aquela que dispõe da noção de “intervalo de seleção”. Este intervalo é definido pela fórmula N/n , onde N é igual ao total da população e n é o número de elementos ou unidades da amostra. Após formar uma lista, sorteia-se um nome dentre os que compõem a primeira seqüência equivalente ao intervalo; daí para frente é só acompanhar o intervalo de seleção”(BARBETTA, 2004, p. 48).

Por seu turno, a amostragem estratificada pode ser de duas espécies: estratificada proporcional e estratificada uniforme. Característica das amostragens estratificadas é que elas pressupõem a fragmentação da população em subgrupos; enquanto na proporcional a amostra deve preservar o percentual que o subgrupo representa no total, na uniforme cada subgrupo terá o mesmo número de elementos tomados para amostra, independentemente de haver ou não nuances na sua representatividade quando relacionado ao todo”(BARBETTA, 2004, pp. 51-2). Como exemplo, podemos pensar em uma população de 50 indivíduos integrantes de um Tribunal: 30 desembargadores; 10 procuradores e 10 secretários. Supondo que eu quisesse saber qual a concepção de Direito nutrida pela instituição. Sabendo que 60% da população é composta por desembargadores, a amostra deverá ser constituída, em 60% por desembargadores, e assim sucessivamente. Por outro lado, na estratificada uniforme, teríamos o mesmo percentual para cada subgrupo na amostra.

A amostra por conglomerados, por seu turno, é aquela em que se agrupam elementos da população em unidades definidas como “conglomerados”, selecionando-se, num primeiro momento, alguns conglomerados para, posteriormente, ou proceder a seleção de elementos destes conglomerados selecionados, ou ainda, inferir todos os elementos integrantes dos conglomerados selecionados. No primeiro caso teríamos uma amostragem de conglomerados em um estágio, e no segundo caso, amostragem de conglomerado em dois estágios”(BARBETTA, 2004, p. 52). Como vantagens desta espécie de amostragem, Barbetta afirma seu baixo custo em relação à demais espécies, sobretudo quando a população está dispersa em grandes territórios”(BARBETTA, 2004, p. 54), porém, como

desvantagem, apresenta um menor grau de precisão no resultado. Como exemplo de aplicação de tal modelo, o autor menciona a pesquisa de domicílios de um município, onde cada quarteirão forma um conglomerado.

Para alguns adeptos da clássica tradição da pesquisa estatística, as amostragens não aleatórias, ou não probabilísticas, são vistas com algumas ressalvas, como no caso de Antônio Carlos Gil, que assevera não apresentarem as mesmas, fundamentação matemática ou estatística, dependendo unicamente de critérios do pesquisador, em contraposição às aleatórias, que seriam “rigorosamente científicas”, no dizer do autor (GIL, 1999, p. 101). Porém, há também um reconhecimento teórico destinado a tais tipos de amostragens, como no caso, por exemplo, de André Pierre Contandriopoulos *et alli* em *Saber Preparar uma Pesquisa*. Para quem, estas amostragens têm como apanágio tornar mais difícil a generalização das observações em relação à população alvo, mas ressaltando que não se trata de uma deficiência (CONTANDRIOPOULOS, BOYLE, CHAMPAGNE, et al, 1994, p. 66).

Podemos delimitar a nuance entre as amostragens aleatórias em relação às não aleatórias com a assertiva de que enquanto as primeiras caracterizam-se pela utilização do acaso para lograr êxito no intento de constituir um subconjunto da população alvo, guardados os mesmos apanágios desta última, as amostragens não aleatórias valem-se do *raciocínio* para tanto (CONTANDRIOPOULOS, BOYLE, CHAMPAGNE, et al, 1994, p. 64).

Em termos gerais, os teóricos apresentam como sendo espécies de amostragem não aleatória, precipuamente, as seguintes: amostragem por cotas; por julgamento (que alguns denominam por tipicidade); e por acessibilidade.

Tal como na amostragem estratificada proporcional (que é aleatória) na amostragem por cotas a população é vista de modo segregado, dividida em diversos subgrupos. Daí, seleciona-se para compor a amostra uma “cota” de cada subgrupo. Tal cota deve ser proporcional ao tamanho do subgrupo de onde é extraída. A nuance é que aqui a seleção não é aleatória como alhures (BARBETTA, 2004, p. 56).

Por outro lado, temos a amostragem por julgamento, que é aquela em que os elementos “escolhidos” para comporem a amostra são os *considerados* típicos da população a ser estudada. Haveria caracteres que tornariam os elementos que os possuem representativos por excelência da população que se quer estudar.

Como exemplo, temos o caso de se pretender estudar a produção científica dos departamentos de ensino de uma Universidade; o pesquisador poderia escolher os departamentos que ele entende melhor representar tal Universidade (BARBETTA, 2004, p. 56). A não utilização aqui, de uma amostra aleatória, se deveria ao fato de a população ser pequena; e um estudo de todos os departamentos despenderia muito tempo.

A teoria da amostragem foi, durante muito tempo, desenvolvida por autores que se respaldavam em uma concepção de grande apreço às inferências estatísticas, convictos no ideal da certeza e da previsibilidade, conforme já dito. Nisto, a exposição das espécies de amostragem se faz, no mais das vezes, sob argumentos imersos nestes valores. Todavia, traço marcante das amostragens não aleatórias é o reconhecimento de que as mesmas não se equívalem, nestes aspectos, às amostragens aleatórias. Na senda desta lógica, as amostragens por acessibilidade, por exemplo, são apresentadas por Antônio Carlos Gil como sendo o menos rigoroso dentre todos os tipos de amostragem, sendo, por esta razão, indicado para estudos exploratórios ou *qualitativos*, onde não se exigiria elevado grau de precisão. Com efeito, nesta espécie, o pesquisador seleciona os elementos a que tem acesso, admitindo que estes possam representar o universo (GIL, 1999, p. 104).

2. O caso das pesquisas qualitativas

De acordo com as observações precedentes, as construções teóricas a respeito da pesquisa qualitativa ainda se fazem incipientes no que tange à questão alusiva a seleção de elementos de uma população, para pesquisa. Aliás, a própria expressão “amostra”, não é de emprego pacífico nesta modalidade de pesquisa, dada a proposição de Martin Bauer acerca do “corpus”, cuja melhor apreciação passaremos em breve.

Conforme Teresa Haguette afirma, as pesquisas qualitativas fornecem uma compreensão profunda de certos fenômenos sociais, ancorados no pressuposto *da maior relevância do aspecto subjetivo da ação social* (HAGUETTE, 1999, p. 63), diante da configuração das estruturas sociais, com o que, sua especificidade resta absolutamente demarcada, em relação às pesquisas quantitativas. Bauer e Gaskell destacam que na pesquisa social a seleção de elementos em termos de dados para análise corresponde à seleção de evidências para argumentar. A tradição das ciências sociais seria calcada na “amostragem estatística aleatória” para efeitos desta seleção. Porém, destacam os autores que nem sempre tal modelo é

aplicável, ou seja, fundamentalmente em muitos casos das pesquisas qualitativas não haveria lugar para o mesmo. Daí é que emerge a proposta de se proceder à “construção do *corpus*”, em contraposição à “amostragem”, implicando em conceitos diferentes, correspondendo este último à “amostragem estatística aleatória” e o primeiro à escolha sistemática de algum racional alternativo (BAUER e GASKELL, 2002, p. 39). Tal nuance, contudo, cinge-se ao plano estrutural, posto que funcionalmente seriam equivalentes, segundo advertência dos autores precitados (BAUER e GASKELL, 2002, p. 40).

Ensinam tais autores que o interesse por excelência do pesquisador qualitativo é na “tipificação da variedade de representações das pessoas no seu mundo vivencial” (BAUER e GASKELL, 2002, p. 57), o que pressupõe que se constitui em tarefa impossível a operacionalização de listagens às mesmas alusivas, como ocorre no caso da maioria das pesquisas aleatórias supra citadas, especialmente no caso das aleatórias simples e sistemáticas⁵. Face a este desconhecimento *a priori*, um dos princípios norteadores da inferência qualitativa seria exatamente o de “maximizar a variedade de representações”, ou seja, colocar-se como princípio a busca de novas representações no momento da construção do corpus.

Dada a peculiaridade da abordagem qualitativa, os autores entendem que as variedades que compõe o universo que se deseja captar com a pesquisa são desconhecidas, o que impossibilita o emprego de amostragem de acordo com um racional de representatividade, razão pela qual, adotam a sugestão dos lingüistas, no sentido de que devemos proceder, primeiramente, a uma seleção “preliminar”, depois, analisando a variedade contida em tal seleção e, finalmente, ampliar o corpus de dados até que não se descubra mais variedade (aqui, particularmente, evidencia-se o critério da saturação) (BAUER e GASKELL, 2002, p. 55).

Conclusivamente, as nuances entre corpus e amostragem são sintetizadas na idéia de que enquanto a construção do primeiro ajuda a tipificar representações desconhecidas, a segunda descreve a distribuição de representações já conhecidas na sociedade (BAUER e GASKELL, 2002, p. 58).

⁵ Aliás, esta impossibilidade ontológica de abarcar *a priori* todas as representações que se quer inferir, é o que levou Bauer à proposição da teoria do corpus para as pesquisas qualitativas, posto que nestas, a exemplo do que se dá na lingüística, seu objeto está em meio ao que ele denomina “sistema aberto”, posto que, no caso das primeiras, passam pela questão alusiva à situações e às ações (BAUER e GASKELL, 2002, p. 43).

Dado o princípio da maximização da variedade, o mister de se adotar um critério de finalização na coleta de dados se impõe: trata-se da noção de saturação – *investigam-se diferentes representações apenas até que a inclusão de novos estratos não acrescente nada de novo*.

Convém destacar a importância central que possui os *corpora* na teorização de Bauer *et al*, uma vez que, segundo estes, na investigação científica haveria duas questões nodais: formulação de proposições e prestação de contas pública (BAUER e GASKELL, 2002, p. 481). Dentro da prestação de contas pública, Bauer *et al*, identificam duas categorias que forneceriam a base para a garantia de qualidade: confiabilidade e relevância. Ambas captariam a essência da avaliação de qualidade de uma pesquisa, e se aplicariam tanto às quantitativas quanto às qualitativas. É neste ponto que ganha relevância a questão do “*corpus*”: sua construção seria indicadora tanto da confiabilidade como da relevância, para as pesquisas qualitativas (BAUER e GASKELL, 2002, p. 482). Pela confiabilidade, os leitores e receptores da pesquisa teriam a “confiança” de que os resultados da investigação representam a “realidade”, sendo, assim, algo mais do que a singela imaginação fértil do pesquisador. Por seu turno, a “relevância” da pesquisa expressa tanto a utilidade como a importância. Aqui há o inesperado e a surpresa. Novas compreensões e representações no caso das pesquisas qualitativas, e o teste de hipótese no caso das quantitativas (BAUER e GASKELL, 2002, p. 482). Por estas colocações, Bauer *et al* se situam dentre aqueles que pleiteiam critérios de qualidade para a pesquisa social, valendo-se de critérios que apresentem equivalência funcional em relação aos critérios da tradição da pesquisa quantitativa (BAUER e GASKELL, 2002, p. 479-0).

Considerações Finais

Percebe-se, pelo exposto, que o cerne da questão passa pelo caráter da pesquisa qualitativa, o qual implica na busca da tipificação da variedade de representações dos indivíduos, atrelada ao fato de tais representações não serem conhecidas pelo pesquisador. Naturalmente, tal desconhecimento gera a impossibilidade de se proceder a uma listagem ou catalogação do todo para, subsequentemente, de forma aleatória, selecionarmos parte sua para o estudo. Mas, convém ressaltar que mesmo Bauer *et al* admitem a utilização de *amostragem* na pesquisa qualitativa, apenas ressaltando que isto nem sempre é verossímil, posto que o

desconhecimento acima referido pode não se dar em alguns casos, o que nos leva à ilação de que um dos elementos concorrentes nesta questão é o próprio objeto de pesquisa. É bem verdade, no entanto, que Bauer *et al* propõem a utilização de um termo (“corpus”), a partir da observação dos lingüistas, pois poderíamos nos valer da expressão “amostragem” para sintetizar as idéias ali constantes, desde que atentos para as diretrizes indicadas pelo autor, precedentemente elencadas. Apenas como ilustração alusiva a esta questão da terminologia, trago o exemplo de André Pierre Contandriopoulos *et alli*, os quais trazem como uma das modalidades de “amostra não probabilística” a “amostra por escolha racional”, e como subespécie desta, a amostra por julgamento do especialista. A respeito, dizem os autores:

A seleção de elementos típicos se baseia nos conhecimentos dos especialistas. Visa reproduzir a seção mais homogênea da população alvo. Trata-se, então, de determinar a características mais comuns na população e de selecionar elementos que lhe correspondam. A seleção dos casos típicos maximiza a aplicabilidade dos resultados⁶.

Naturalmente, percebe-se que no caso em epígrafe o pesquisador possui um conhecimento prévio da população, mas a partir daí, temos que a possibilidade de se proceder a uma inferência qualitativa é manifesta. Neste sentido, as amostras aleatórias também não são de um todo excludentes de apreciações qualitativas, mas apenas pode-se dizer que são mais consentâneas às pesquisas quantitativas.

Ainda na senda das últimas colocações efetivadas, ou seja, no caso de termos um conhecimento prévio acerca da população a ser estudada, podemos enunciar uma regra, válida para as pesquisas qualitativas em geral, no sentido de que quanto mais heterogêneos se apresentarem os subgrupos, maior terá de ser a amostra que se pretenda em relação aos mesmos. Além disso, quanto mais variáveis eu estiver trabalhando, maior será a exigência de uma amostra mais extensa. Esta lógica nos leva à necessária distinção no momento da atribuição de pesos aos diferentes estratos da população. Pensemos no exemplo de possuímos uma população composta por três estratos (a, b, c), sendo que um deles, o estrato “a”, apresenta uma heterogeneidade interna muito maior em relação aos outros dois, os quais apresentam certa equivalência

⁶ CONTANDRIOPOULOS, André Pierre et alli. *Saber Preparar uma Pesquisa*. São Paulo-Rio de Janeiro: Editora Hucitec Abrasco, 1994, p. 65.

neste sentido. Nesta conjectura, a amostra extraída do subgrupo “a” terá de ser maior do que as amostras extraídas dos outros dois subgrupos (“b” e “c”). Porém, no momento de proceder às inferências, ou seja, análise dos dados coletados e conclusão, a respeito da população alvo, deveremos atribuir peso diferenciado, levando em conta exatamente a proporção de cada estrato em relação ao todo.

Abstract: The present essay aims a brief contemplation about an important topic of scientific research: its on the question of sample. Recognizing the peculiarities around the two orientations that polarize the verification of reality on social research, the quantitative orientation and the qualitative one, we may see the development a theoretical view that tries to ponder on such nuances, specifically on what concerns to the focalized thesis.

Key words: Sample; *corpus*; qualitative research; sampling.

Referências Bibliográficas

BARBETTA, Pedro Alberto. *Estatística aplicada às Ciências Sociais*. 5ª ed. Florianópolis: Editora da UFSC, 2004.

BAUER, Martin e GASKELL, George. *Pesquisa qualitativa com texto, imagem e som: um manual prático*. Petrópolis: Vozes, 2002.

CONTANDRIOPOULOS. André Pierre; BOYLE, Pierre; CHAMPAGNE, François; DENIS, Jean-Louis; POTVIN, Louise. *Saber Preparar uma Pesquisa*. São Paulo-Rio de Janeiro: Editora Hucitec Abrasco, 1994.

GIL, Antônio Carlos. *Métodos e Técnicas de Pesquisa Social*. 5 ed. São Paulo: Editora Atlas S.A, 1999.

HAGUETTE, Teresa Maria Frota. *Metodologias qualitativas na sociologia*. 6 ed. Petrópolis: Vozes, 1999.

LOPES, Maria Immacolata Vassalo. *Pesquisa em comunicação – formulação de um modelo metodológico*. São Paulo: Loyola, 1990.

NETO, Agostinho Ramalho Marques. *A Ciência do Direito. Conceito, Objeto, Método*. 2 ed. Rio de Janeiro: Renovar, 2001.