

COMO NÃO MENTIR SEM A ESTATÍSTICA*

HOW NOT TO LIE WITHOUT STATISTICS

Gary King[†]
Eleanor Neff Powell[‡]

Resumo: Neste artigo realçamos e sugerimos formas de evitar uma grande quantidade de equívocos comuns na literatura acerca das melhores práticas na pesquisa qualitativa. Discutimos estas questões em quatro áreas: teoria e dados, estratégias qualitativas e quantitativas, causalidade e explicação, e viés de seleção. Alguns dos equívocos envolvem debates incendiários dentro de nossa disciplina, que são prontamente resolvidos ou diretamente ou com resultados conhecidos em áreas de pesquisa que podem ser desconhecidos dos cientistas políticos. Muitos destes equívocos podem também ser encontrados na pesquisa quantitativa, geralmente com nomes diferentes, e alguns dos quais podem ser fixados com referência a ideias mais bem compreendidas na literatura de métodos qualitativos. O nosso objetivo é melhorar a habilidade de estudiosos com orientação quantitativa e qualitativa, para apreciarem as vantagens de *insights* de ambas as áreas. Assim, no todo, tentamos construir orientações práticas específicas que possam ser usadas para melhorar desenhos reais de pesquisa qualitativa, e não somente as literaturas de métodos qualitativos que tratam acerca deles.

Palavras-chave: Ciência Política. Pesquisa Qualitativa. Pesquisa Teórica e Empírica.

Abstract: We highlight, and suggest ways to avoid, a large number of common misunderstandings in the literature about best practices in qualitative research. We discuss these issues in four areas: theory and data, qualitative and quantitative strategies, causation and explanation, and selection bias. Some of the misunderstandings involve incendiary debates within our discipline that are readily resolved either directly or with results known in research areas that happen to be unknown to political scientists. Many of these misunderstandings can also be found in quantitative research, often with different names, and some of which can be fixed with reference to ideas better understood in the qualitative methods literature. Our goal is to improve the ability of quantitatively and qualitatively oriented scholars to enjoy the advantages of insights from both areas. Thus, throughout, we attempt to construct specific practical guidelines that can be used to improve actual qualitative research designs, not only the qualitative methods literatures that talk about them.

Keywords: Political Science. Qualitative Research. Theoretical and Empirical Research.

* Nossos agradecimentos a Mitchell Duneier, Gary Goertz, Dan Hopkins, Phil Jones pelos comentários muito úteis. Título original: How Not to Lie Without Statistics. Tradução para o português: João Batista de Oliveira Silva Júnior e Ricardo Alaggio Ribeiro.

† Instituto de Ciência Social Quantitativa, 1737, Cambridge Street, Harvard University, Cambridge, MA 02138; <<http://GKing.harvard.edu>>, <king@harvard.edu>, (617)495-2027.

‡ Instituto de Ciência Social Quantitativa, 1737, Cambridge Street, Harvard University, Cambridge, MA 02138; <<http://www.people.fas.harvard.edu>>, <enpowell, enpowell@fas.harvard.edu>.

1 Introdução

Neste artigo realçamos os erros comumente feitos na pesquisa qualitativa e em várias literaturas de métodos qualitativos. Essas literaturas abrangem a ciência política, outras ciências sociais e muitas áreas profissionais ou não disciplinares relacionadas. As sinalizações que levantamos a esses problemas, e as orientações práticas que oferecemos a cada um deles, destinam-se a possibilitar aos estudiosos evitarem as armadilhas comuns e a construírem desenhos de pesquisa mais exatos. Muitos dos problemas que levantamos podem inicialmente parecer específicos à pesquisa qualitativa, mas mesmo assim questões inferenciais subjacentes também afetam a pesquisa quantitativa. Similarmente, embora muitos dos erros que Darrell Huff (1954) expôs em *How to lie with Statistics* [Como mentir com a Estatística] (“o livro de estatística mais lido na história mundial”, STEELE, 2005) sejam únicos à tecnologia de estatística, todas as questões importantes sobre projeto discutidas em seu livro, e muito da literatura quantitativa que veio em seguida, são relevantes também à pesquisa qualitativa. De fato, alguns problemas que identificamos são mais bem compreendidos e resolvidos via referência à literatura de métodos qualitativos e outros à literatura quantitativa; mas as resoluções geralmente se aplicam a ambos.

Os problemas que discutimos incluem erros de interpretação, equívocos e algumas assertivas falsas¹ e abrangem uma ampla gama de questões de projeto de pesquisa. Iniciamos a Seção 2 discutindo o papel da teoria e da evidência em separado e como elas interagem. A Seção 3 trata de problemas relacionados à distinção entre estratégias de coleta de dados quantitativos e qualitativos. E, finalmente, discutimos problemas tanto de causalidade quanto de explicação na Seção 4 e problemas de seleção na Seção 5.

2 Pesquisa teórica vs. empírica

Algumas das divergências mais acirradas entre cientistas sociais podem ser originadas em diferentes gostos para fazer e aprender acerca de pesquisa em determinados locais no continuum da teoria para a evidência. Estes cientistas, com a mente voltada para a teoria, reclamam que “não têm estômago para as intermináveis discussões em seminários para obterem os métodos corretos quando ninguém parece se importar acerca da ideia que está sendo testada e que vale a pena para início de conversa”. Empíricos relatam “achar difícil lidar com uma contemplação teórica a qual não se tem esperança de ser provada como certa ou errada, ou que ignora uma relevante evidência existente”. Como a ciência política é dentre todas as disciplinas acadêmicas a mais diversa, e ela inclui estudiosos de toda gama no *continuum* das ciências humanas, essas divergências se manifestam mais frequentemente aqui. É claro que algumas disputas semelhantes podem ser encontradas dentro de outras ciências sociais, como a educação, saúde pública, direito e outras áreas.

O objetivo da pesquisa de ciência política é descrever, explicar e por vezes melhorar o governo e a política. Para conseguir esta tarefa, devemos reconhecer que nenhuma perspectiva

¹ Evitamos apontar os dedos com citações específicas quando discutimos assertivas falsas e erros metodológicos em pesquisa anterior, visto que nosso objetivo é basear-se em, e não repreender aqueles que vieram antes.

extrema sobre teoria e empirismo é certa ou errada; estas são preferências normativas para qual tipo de conhecimento qualquer um de nós escolha adquirir a qualquer momento. No fim, precisamos tanto da criatividade teórica quanto da validação empírica. Teoria criativa é importante mesmo sem perspectivas imediatas para validação e para alguns propósitos, e mesmo sem a esperança de que a teoria prenunciará ou explicará eventualmente a realidade empírica. Devemos reconhecer o valor de uma ideia em si própria, desassociada dos dados (i.e., informações de qualquer tipo relevante). Mas devemos também reconhecer o teor de validação empírica sólida, pois o progresso requer ambos. E nenhum deles deve ter prioridade ou mesmo precedência sobre outro no processo de pesquisa: precisamos de “implicações empíricas de modelos teóricos” (GRANATO; SCIOLI, 2004), assim como estudo de implicações teóricas de pesquisa empírica. Nas quatro subseções a seguir nós elaboramos e ampliamos este tópico.

2.1 A iteração entre pesquisa teórica e empírica

Um grande componente de pesquisa acadêmica com tópico regular na literatura de métodos qualitativos é a iteração entre o desenvolvimento de teoria e a construção da evidência empírica (GEORGE; BENNETT, 2005). A iteração pode ocorrer em um único trabalho acadêmico ou em publicações em um campo de pesquisa (embora esta prática não seja discutida com frequência em livros e textos de métodos quantitativos, ela é, evidentemente, uma característica bem regular também na pesquisa quantitativa aplicada). A questão que tratamos aqui é de que a perspectiva de se chegar a algum lugar produtivo, interagindo entre desenvolvimento teórico baseado em dados falhos e observação empírica baseada em teoria falha, parece altamente duvidosa. Esclarecemos esta questão fundamental em três passos: primeiro, descrevendo-a, depois realçando e separando um argumento político que ela gerou e, finalmente, oferecendo o primeiro conjunto de condições matemáticas formais sob as quais este procedimento iterativo possa produzir o resultado desejado.

Primeiro, a ideia de iterar entre a coleta de evidência, aceitando alguma teoria e desenvolvimento teórico e também aceitando a veracidade de alguma evidência eventualmente empírica, surge a partir de dois pontos importantes e aparentemente contraditórios:

1. As teorias de ciência social não surgem *ex nihilo*. Todas as teorias úteis são eventualmente baseadas em *algumas* observações empíricas, não importando o quão tênues possam ser;
2. Dados empíricos não podem ser coletados sem pelo menos implicitamente serem feitas algumas suposições teóricas acerca do que deve ser observado.

O problema não é que esses pontos sejam contraditórios, tal como eles não o são: até mesmo as categorias básicas que usamos para mensurar fatos aparentemente brutos e diretos requerem algum tipo de teoria ou tipologia explanatória (ELMAN, 2005), e nenhuma teoria útil da realidade social pode ser construída sem serem aceitos alguns aspectos dessa realidade. Em vez disso, o problema real é que o processo iterativo pode nos colocar no caminho errado quando teorias são baseadas em evidências falhas ou evidências coletadas quando do condicionamento

sobre as suposições teóricas erradas. E, ainda assim, estar certo acerca da veracidade de qualquer observação ou teoria é impossível, e aprender sobre o mundo sem ambas é uma fantasia.

Segundo, o status de incerteza científica de iterar entre teoria e evidência tem levado a um debate improdutivo, recorrente e amplamente político, pois este pode ser resumido por dois argumentos feitos de forma estridente por grupos diferentes, tipicamente formados para fins de confrontar o outro grupo:

1. Se você derivar sua teoria de um conjunto de dados existente, você não pode usar os mesmos dados para validar a teoria empiricamente. Adequar uma teoria a dados o torna invulnerável à prova de que você está errado por causa de tais dados e, assim, incapaz de aprender se a teoria é válida ou não;
2. Podemos melhorar substancialmente uma teoria baseando-a firmemente em características importantes de realidade empírica e ajustando-a continuamente para adequá-la a novas observações.

Ambos os pontos são manifestamente verdadeiros, muito embora eles pareçam contradizer um ao outro; cada um é geralmente ignorado por aqueles que enfatizam o outro. Embora nunca saibamos o final de qualquer dos pontos, devemos sempre reconhecer ambos em toda pesquisa e projetar uma pesquisa com ambos em mente de maneira simultânea. Afinal, construir teorias conhecidas por violar características importantes de realidade empírica é uma perda de tempo (porém, como teorias destinam-se a serem abstrações, elas visam a deixar passar aspectos menos importantes do mundo empírico), e então o segundo ponto certamente é válido. Mas, ao mesmo tempo, é fácil demais pensar que você desenvolveu uma ideia importante quando ao invés disso sua teoria meramente tem mais afirmações do tipo “mas se” como também afirmações como “assim é”, alinhadas a cada dado novo, então o primeiro ponto também é essencial. A qualquer momento que você notar um desses pontos sendo tratado sem a total consideração do outro, pelo menos implicitamente, é provável que você também encontre alguns erros inferenciais graves. Às vezes, a motivação para ignorar um dos pontos é política, mas outras vezes nós estamos naturalmente bastante focados no problema que conseguimos identificar. Esclarecer o panorama geral, que sempre envolverá ambos os pontos, tem o potencial de melhorar muitas pesquisas.

Finalmente, agora reforçamos a base matemática desse procedimento observando que iterar dessa maneira é uma versão qualitativa da amostragem de Gibbs nos algoritmos estatísticos da Cadeia de Markov Monte Carlo (GILL, 2008, Seção 9.3). A ideia é que em estatística e outras áreas geralmente precisamos ser capazes de retirar amostras aleatórias de duas variáveis, digamos x e y (como analogias a dados e teoria, respectivamente), a partir de sua distribuição conjunta bivariada $p(x, y)$ (que indica a lei que rege o quão particular os valores de x e y ocorrem juntos), mas só podemos saber como retirar a partir das duas distribuições condicionais univariadas mais simples, $p(x|y)$ (i.e., como x varia quando y assume um valor específico) e o oposto, $p(y|x)$. A amostragem de Gibbs nos ajuda a resolver o problema começando com uma suposição (mesmo

uma suposição falha) para y , inferindo x da distribuição de x , dado o valor suposto de y , e inferindo um novo x da distribuição de x , dado o valor inferido de y , e continuando a iterar.

Sob as condições corretas, podemos provar matematicamente que esse processo iterativo convergirá a inferências a partir da distribuição conjunta desejada – que sob a analogia ao nosso caso deve fornecer a teoria correta e a evidência correta para a teoria. Então, quais são as condições? Primeiro, a distribuição conjunta $p(x, y)$ deve realmente existir. Em pesquisa qualitativa, isso significa que há um processo comum que rege a conexão, caso exista entre teoria e evidência. Se não há processo comum, então tentar aprender acerca disto com passos irrelevantes em um processo iterativo obviamente irá falhar.

A forma que retiramos de cada distribuição condicional necessita estar fixada ou pelo menos permanecer consistente com o tempo. O ponto aqui é que para um dado projeto cada vez que você for confrontado com a mesma evidência, você precisa ter a mesma visão acerca de qual teoria é provável ser correta; e toda vez que você considerar uma teoria específica, o processo pelo qual você seleciona e avalia dados de uma determinada fonte deve permanecer o mesmo. Se considerarmos a mesma observação como suporte a uma teoria em uma iteração, confrontando-a com outra, violamos esta condição.

E, finalmente, a convergência à distribuição conjunta sob a amostragem de Gibbs requer que iteremos muitas vezes, colemos dados o bastante para que sejam inferidos das teorias e exploremos muito as teorias para que sejam consistentes com os dados. Exatamente, quantas vezes precisamos iterar depende de quanto mais próximo cada iteração nos leve ao nosso objetivo (i.e., quão eficientes são os métodos) e quanto complexa é a teoria que estamos desenvolvendo (mais complexidade requer mais iterações). Na prática, a única verificação real sobre se chegamos à convergência é observar se em uma longa sequência de iterações encontramos a mesma teoria juntamente com observações consistentes. Mas tanto na versão matemática formal da amostragem de Gibbs quanto no projeto análogo de pesquisa qualitativa, nunca podemos ter certeza de que iteramos tempo o bastante. Então, um pouco mais de iteração é sempre útil. Isso meramente reflete o fato de que a inferência é sempre incerta até um ponto, e continuar a iterar – no seu projeto de pesquisa, ou de outros estudiosos como parte de um programa maior de pesquisa ou literatura – pode sempre melhorar nosso conhecimento do mundo que buscamos compreender.

2.2 Maximizando a Alavancagem

A tensão entre adequar uma teoria aos dados e testar a teoria pode ser resolvida em parte por projetos de pesquisas bem construídos e por um tanto de criatividade. Precisamos condicionar teorias ao máximo de informações das quais podemos ter razoável certeza. Mas, uma vez que o consigamos, a teoria adequa todos os dados conhecidos e não ficamos vulneráveis a provar que estamos errados – o que, evidentemente, é outra maneira de dizer que não podemos descobrir se a teoria é exata ou se, ao invés disso, pusemos uma falácia *posthoc* que parece adequar-se aos fatos, mas não os explica.

Então, a chave é deixar um espaço para a iteração entre a teoria e os dados. Como conseguimos isto? Sugerimos duas respostas que podem ser aplicadas quando plausíveis, sendo

uma teórica e outra empírica: (1) reduzir a complexidade da teoria para que uma teoria mais simples explique os mesmos fatos empíricos ou (2) encontrar novas implicações observáveis da mesma teoria, coletar esses dados e ver se eles são consistentes com a teoria. Ambas aumentam a alavancagem e a quantidade de evidência empírica quanto ao grau de complexidade teórica.

Por que maximizar a alavancagem é tão importante? Considere três motivos. Um é que quando a nossa teorização revela uma nova implicação observável, temos a chance de apoiar a teoria trazendo mais informações com impacto em nosso problema. Este procedimento é vantajoso, quer a teoria existente já esteja condicionada a todos os dados disponíveis ou somente a alguns. É caso útil que a implicação possa ser observada, coletando dados adicionais do mesmo tipo, os mesmos dados em um novo período ou dados inteiramente novos a partir de áreas ou unidades de análises diferentes. Os dados sobre novas implicações observáveis são mais valiosos quando menos relacionados às implicações observadas, pois os novos dados fornecem testes independentes e assim mais informativos. Dessa forma, por exemplo, complementar uma análise quantitativa abstrata de uma grande coleção de países com uma etnografia detalhada em uma cidade poderia ser bem útil se ambas mensurassem implicações observáveis da mesma ideia. Coletar mais alguns países seria também útil, mas provavelmente não tanto, e certamente não tanto se eles forem bem parecidos aos países que já estão nos seus dados. Você deve sempre levar os dados para onde conseguir obtê-los, mas se o mesmo esforço puder produzir dados que venham de uma fonte muito diferente ou que sejam – por outro motivo – não relacionados aos dados que você tem, e que ainda são uma implicação da mesma teoria, isso geralmente seria preferível.

Uma segunda vantagem de maximizar a alavancagem é que dados nas ciências sociais, e de fato em muitas ciências, são geralmente escassos quanto à enorme criatividade teórica de determinados estudiosos e a comunidade acadêmica como um todo. Criar teorias para adequar qualquer observação empírica pode ser feito de modo tão rápido que é muito fácil se enganar ao pensar que descobriu algo, mesmo quando não o fez. Quando geralmente os seres humanos são impressionantemente exatos em reconhecer padrões, somos em reconhecer a ausência de padrões. Num piscar de olhos, podemos detectar padrões em manchas de tinta ou formações de nuvens, mas somos piores em detectar teorias sem uma base empírica. Se você não tiver certeza disto, de alguma forma tente o seguinte experimento com o seu colega do escritório ou com o seu cônjuge: invente um “fato” sobre qualquer assunto (por ex.: A Rússia acaba de invadir a Islândia! A Presidente reservou um espaço na TV para um pronunciamento importante! etc.) e veja quanto tempo leva para você ouvir uma explicação. Esteja preparado para contar em milissegundos, já que não existe praticamente demora detectável. Devemos, portanto, sempre permanecermos vigilantes em colocar nossas teorias em risco e continuar a confrontá-las com novas fontes de evidência. Nós aprendemos quando tentamos provar que estamos errados. E ao julgar o trabalho acadêmico pela extensão a qual ele coloca suas assertivas em risco de provarem que está errado, podemos às vezes evitar este obstáculo na pesquisa acadêmica.

Um motivo final pelo qual maximizar a alavancagem é importante, é fundamentalmente biológico. Muitos assuntos são altamente complicados; e nós seres humanos somos criados para que possamos somente manter uma pequena fração dessas complexidades em nossas mentes a

cada vez. Como tais, e de várias maneiras diferentes, muitos definem o conceito de “explicação” como requerendo uma simplificação – como resumir, compreender ou explicar muitos fatos com poucos. Sem este processo de simplificação, sem teorias que possam maximizar a alavancagem, nós não conseguimos entender ou passar aos outros a natureza dos dados e seus padrões subjacentes.

2.3 Parcimônia relativa vs. absoluta

A seção anterior explica que teorias com alta alavancagem são valiosas, porque elas explicam muito sobre fatos que de outra forma seriam desconexos, porque elas nos ajudam a testar nossas assertivas e por causa de fatos biológicos acerca de como os seres humanos pensam. Embora preferíssemos teorias que fossem relativamente mais parcimoniosas aos fatos empíricos que elas explicam e a outras teorias que explicam os mesmos fatos, não existe motivo para valorizar uma teoria, porque ela é simples em um sentido absoluto. Como resultado, muitas assertivas acerca da parcimônia na literatura são relatadas de maneira falsa.

Por exemplo, deveríamos ficar felizes em algumas circunstâncias ao *adicionar* complexidade a uma teoria se, assim fazendo, obtivéssemos uma gama desproporcionalmente maior de observações empíricas. A parcimônia é, portanto, importante somente quanto aos fatos que ela busca explicar. Diferente de assertivas implícitas na literatura que parecem tratar da parcimônia absoluta como uma lei misteriosa da natureza, e se uma teoria parcimoniosa é mais provável de ser correta (ou útil) do que uma teoria mais complexa, esta é uma proposição totalmente empírica. Testar esta teoria requer novos dados, que servem como novas implicações observáveis.

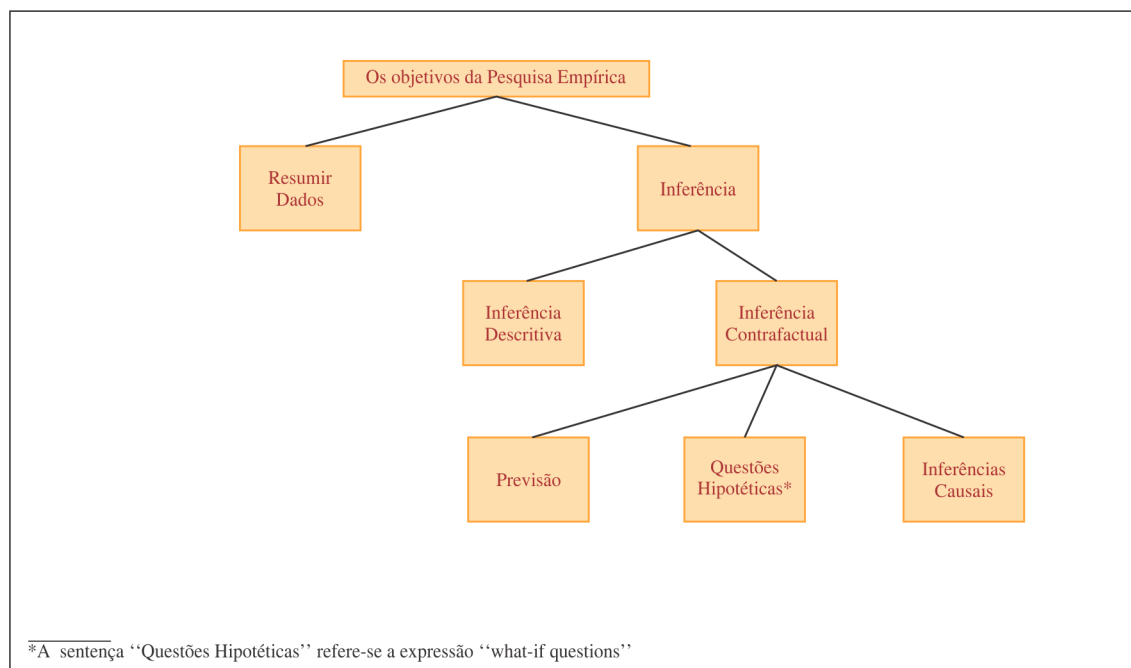
2.4 Os objetivos da pesquisa empírica

Os objetivos da pesquisa empírica envolvem pelo menos três distinções fundamentais. Todos os projetos de pesquisa se deparam com essas distinções como escolhas básicas no processo de pesquisa. Elas não são sempre consideradas tão explicitamente como estamos para fazê-lo, mas elas estão sempre presentes. Nós as colocamos aqui para dar aos leitores um senso dos objetivos do empreendimento e para orientar os seus trabalhos no contexto mais amplo da pesquisa e para criar as condições para o restante deste artigo. Mostramos essas três distinções nos pontos de ramificação da Figura 1.

Um ponto chave da Figura 1 é que *nenhuma* das caixas ou distinções envolve uma oposição entre pesquisa quantitativa e qualitativa. De fato, cada caixa da página mostra um objetivo da pesquisa empírica que pode ser perseguido via pesquisa quantitativa ou qualitativa.

A primeira distinção da figura, no ponto de ramificação superior, é aquela entre *resumir dados* [o que King, Keohane e Verba (1994) chamam de “resumir detalhe histórico”] e *inferência*. Inferência é simplesmente usar fatos que temos que aprender acerca de fatos que não temos (ver Seção 3.1). Em contraste, resumir dados envolve somente examinar e resumir as observações perante nós, ao invés de tentar aprender sobre fatos não observados. Para qualquer projeto com um objetivo de algum tipo de inferência, aconselha-se primeiramente examinar os dados que temos.

Figura 1: Os Objetivos da Pesquisa Empírica. Observe que a distinção entre estilos de coleta de dados quantitativos e qualitativos não aparece, pois a figura inteira é a mesma para ambos



Este passo pode revelar problemas de mensuração, sugerir novos alvos inferenciais ou às vezes ser útil em si e por si só. Para ser justo, para manter a distinção e distinguir entre fatos que conhecemos e fatos que gostaríamos de conhecer seja crucial para reduzir vieses em pesquisa, toda observação requer alguma teoria, e então qualquer quantidade de resumo e de observação dos dados brutos sempre envolverá alguma inferência; de fato, a discussão acerca de iterações entre a teoria e a evidência na Seção 2.1 também se aplica à iteração entre a observação dada a alguma teoria sobre o que nós estamos observando e a inferência dada a alguma observação que pensamos que tenhamos feito.

A segunda distinção da Figura 1 (lendo de cima para baixo) é a escolha entre *inferência descritiva* e *inferência contrafactual*. Inferência descritiva é o processo de aprender sobre fatos que existem, mas são atualmente desconhecidos para o pesquisador, ao passo que inferência contrafactual é tentar aprender acerca de fatos que existiriam ou existiram em alguma outra época ou em uma situação hipotética. O termo mensuração é às vezes usado para significar e resumir dados; ele é geralmente reservado como um sinônimo de inferência descritiva. O que os norte-americanos pensam do presidente é um fato que não é totalmente conhecido, e assim pode ser assunto de uma inferência descritiva. Às vezes, mensuração se refere a aprender o que cada norte-americano pensa do presidente, o que por si só não é totalmente conhecido ou nem mesmo conhecível com certeza por um pesquisador ou um entrevistador ativo, com a inferência descritiva se referindo a aprender acerca de todos os norte-americanos [ver Adcocke Collier (2001), para definições mais sofisticadas].

A distinção final, na parte de baixo da figura, são os três tipos de inferência contrafactual (ver KING; ZENG, 2007). *Previsão* envolve fatos que existirão no futuro (quando a época estiver em um valor contrafactual, i.e., futuro). *Questões do tipo “E se?”* perguntam acerca de fatos que existiriam se o mundo fosse diferente de alguma forma; já que o mundo não é diferente

desta forma, esses fatos são desconhecidos. E, finalmente, uma *inferência causal* é a diferença entre algum detalhe factual e a resposta a uma questão “e se?”. Por exemplo, o voto que um(a) candidato(a) recebe, menos o voto que esse(a) candidato(a) teria recebido se ele(a) tivesse uma postura política diferente. A inferência contrafactual (e cada de suas três partes) também inclui o amplo objetivo da *explicação*, um conceito que geralmente é confundido com a inferência causal (uma distinção que levantamos na Seção 4.2).

Embora nenhum dos objetivos listados na Figura 1 seja inerentemente mais importante do que os outros, alguns são mais valorizados em algumas áreas. Por exemplo, na saúde pública a inferência descritiva, sob a bandeira de mensuração, é valorizada enormemente. É crucialmente importante descobrir exatamente onde os problemas estão para ajudar a priorizar o financiamento, a pesquisa e esforços de melhoria. Em contraste, a ciência política e a maioria das ciências sociais são primordialmente preocupadas com a inferência causal e menos interessadas em mensuração. Essas não são verdades inabaláveis e elas não são necessariamente corretas. Elas são simplesmente práticas convencionais e preferências normativas de grupos de pessoas. Mas inferências causais podem também ser altamente valiosas na saúde pública, e a mensuração pode agregar grande valor e conhecimento às pesquisas científicas sociais. De fato, quando se olha através da ampla gama de campos científicos, parece claro que muitas áreas das ciências sociais têm subenfocado a mensuração básica, que pareceria implicar extraordinárias oportunidades para os pesquisadores empreendedores.

3 Estratégias de coleta de dados quantitativos vs. qualitativos

Nesta seção discutiremos aspectos da divisão entre estilos de pesquisa quantitativa e qualitativa, bem como estratégias de coleta de dados. Pesquisa qualitativa é qualquer investigação que não aquela que requer medidas numéricas. Exemplos são as partes de métodos etnográficos (tais como observação participativa e entrevistas), trabalho de arquivo, análise histórica e outros que não usam medidas quantitativas. Em contraste, a pesquisa de estilo quantitativo inclui medidas explícitas de algum tipo. As medidas podem ser ordinais, de intervalo ou de proporção, ou podem incluir rótulos, tais como dados categóricos não ordenados. Geralmente, há uma unidade fixa de análise, tal como a pessoa, o país, a díade, a relação social, etc., sobre a qual cada medida é obtida, mas mesmo isso pode mudar ao longo de um conjunto de dados.

Em ambos os tipos de pesquisa as unidades de análise podem ser uniformes ou variadas ao longo da evidência disponível. Visto que conduzir uma pesquisa sem quaisquer discernimentos qualitativos é impossível. Mesmo com medidas quantitativas, concluímos que toda pesquisa é qualitativa, e um subconjunto também é quantitativo.

3.1 Teorias de inferência

O progresso científico confiável na ciência política requer pelo menos uma compreensão de uma teoria de inferência coerente com a qual métodos específicos em determinadas aplicações possam ser construídos, derivados e adaptados. Tal teoria existe para a análise quantitativa.

Esta teoria tem sido desenvolvida, elaborada e ampliada dentro das disciplinas de estatística, filosofia e de muitas outras ciências sociais. Ela tem levado ao desenvolvimento de uma ampla gama de métodos para a análise de determinados tipos de dados e informação, e ainda tem possibilitado avaliar e melhorar inúmeras abordagens.²

Que teoria de inferência reforça a pesquisa qualitativa na ciência política? Atualmente, duas opções estão em oferta. A primeira é apelar para a bem desenvolvida teoria de inferência usada nas literaturas estatística e filosófica (como sugerem KING; KEOHANE; VERBA, 1994). A segunda é esperar por algum tipo de teoria de inferência singular e coerente surgir a partir da própria literatura de métodos qualitativos. Embora pesquisadores qualitativos às vezes façam objeção em ter seu trabalho avaliado por uma teoria de inferência tão intimamente associada à pesquisa quantitativa, nenhuma chamada “teoria de inferência qualitativa” surgiu, nenhum argumento coerente se desenvolveu para apoiá-la e nenhum esforço está em andamento para torná-la viável. Obviamente, há pouco motivo para desenvolver uma nova teoria de inferência, já que a existente, embora desenvolvida em grande parte dentro de e para a pesquisa quantitativa, não requer que a informação seja quantificada e possa ser usada na maioria das situações diretamente na pesquisa qualitativa, sem quantificação. De fato, o grau pelo qual a teoria de inferência tem de ser modificada e adaptada para se aplicar à maioria de ambientes de pesquisa qualitativa não é maior do que o que já ocorre rotineiramente quando da sua aplicação a novas formas de dados quantitativos.

Felizmente, a aplicabilidade e a validade da teoria reinante de inferência conforme dados qualitativos são ampla e crescentemente reconhecidas e respeitadas por muitos pesquisadores qualitativos. Mesmo aqueles inflamados por analogias à inferência estatística regularmente se dedicam e contribuem à aplicação de noções desta teoria de inferência como viés de variável omitida, viés de seleção e muitas outras questões. No entanto, quando pesquisadores não conseguem reconhecer esta teoria de inferência, eles às vezes agem como se fossem livres para inventar métodos sem restrições ou avaliação, exceto a introspecção ou intuição. Isso, é claro, pode ser um verdadeiro engano, pois a intuição falha com bastante frequência em fazer inferências (ver Seção 3.4).

De fato, reconhecer a base inferencial de métodos qualitativos existentes e futuros é essencial, já que sem ela o progresso cumulativo é extremamente improvável. Certamente, quando não há uma teoria de inferência comumente reconhecida, até mesmo compreender propostas diferentes para métodos qualitativos pode ser difícil. Esta questão é exacerbada pela propensão de cada pesquisador qualitativo a inventar uma nova linguagem para descrever o que geralmente são as mesmas questões. Inúmeros exemplos seriam fáceis de citar, como a equivalência dos projetos dos “sistemas mais semelhantes” e o “método da diferença” de Mills. Múltiplas linguagens são também usadas através dos diferentes subcampos metodológicos de

² O que nós chamamos de a teoria de inferência estatística é na verdade um conjunto de teorias fundamentalmente distintas que diferem matematicamente de formas importantes. Estas incluem teorias bayesianas, de probabilidade, robustas e não paramétricas, entre outras. No entanto, a partir da perspectiva mais abrangente desta discussão, e da literatura qualitativa em geral, elas estão todas suficientemente e intimamente relacionadas que é razoável tratá-las como uma única teoria (IMAI, KINGE LAU, 2007). De fato, contribuidores a cada uma, geralmente reconhecem e frequentemente se baseiam nas contribuições das outras teorias de inferência.

várias disciplinas substantivas, embora elas tenham as mesmas representações matemáticas subjacentes para ajudar com a tradução.

3.2 Subcampos metodológicos quantitativos vs. qualitativos

A natureza dos campos profissionais de metodologia da pesquisa quantitativa e qualitativa não poderia ser mais diferente. O núcleo da metodologia de pesquisa quantitativa é a disciplina de estatística. Este campo básico é apoiado pelos subcampos de métodos quantitativos intimamente relacionados que existem dentro da maioria das disciplinas de ciências sociais aplicadas e áreas profissionais não disciplinares. Estes subcampos incluem metodologia política dentro da ciência política, psicométrica e outras análises estatísticas dentro da psicologia, a bioestatística, bem como a epidemiologia dentro da saúde pública, a econometria dentro da economia, metodologia sociológica dentro da sociologia, e muitos outros. Estas áreas estão unidas por representações matemáticas comuns ou intimamente relacionadas de modelos e abordagens estatísticas. Os subcampos de métodos inovam por desenvolver métodos que tratam de problemas de dados novos e quantidades de interesse, e a disciplina de estatística inova com provas rigorosas e até certo ponto, mas as teorias subjacentes de inferência estatística que originam estes modelos são compartilhadas através de todos estes campos. O progresso intelectual a partir do esforço coletivo tem sido notadamente rápido e inequívoco.

Um resumo dessas características de métodos quantitativos está disponível ao se olhar como essa informação é ensinada. Através dos campos e das universidades, o treinamento geralmente inclui sequências de cursos, logicamente feitos em uma ordem, abordando matemática, estatística matemática, modelagem estatística, análise de dados e gráficos, mensuração e métodos ajustados a diversos problemas de dados e direcionados a muitos alvos inferenciais diferentes. A sequência específica de cursos difere nas universidades e nos campos, dependendo do conhecimento matemático esperado dos alunos ingressantes, dos tipos de aplicações substantivas e da profundidade do que será ensinado, mas o arcabouço inferencial, estatístico e matemático subjacente é notadamente sistemático e uniformemente aceito.

Em contraste, a pesquisa em métodos qualitativos parece mais com um saco de surpresas de ideias do que com uma área disciplinar coerente. Como uma medida desta assertiva, em nenhum departamento de ciência política do qual tenhamos conhecimento os cursos de métodos qualitativos são ensinados em uma sequência, com um se baseando em, e necessário para, o próximo. Em nosso próprio departamento, mais de um terço do corpo docente mais antigo ensinou – em um momento ou outro – a uma turma acerca de algum aspecto de métodos qualitativos, nenhum com um curso qualitativo como um pré-requisito obrigatório.

Talvez isso mude através dos esforços do Consórcio sobre Métodos de Pesquisa Qualitativa e de seu programa popular de treinamento de verão, para promover o ensino de cursos qualitativos nas ciências sociais. Mas é provável que tenha mais sucesso, como esse grupo também enfatiza, somente se esses cursos forem integrados com cursos regulares de estatística, por causa da importância conceitual da teoria de inferência estatística e outros tópicos frequentemente abordados em aulas de estatística.

3.3 Múltiplas fontes de dados, não multimétodos

Uma bandeira popular entre pesquisadores qualitativos na ciência política em anos recentes é a “pesquisa multimétodo”. A expressão está até incluída na recém-renomeada “Seção de Pesquisa Qualitativa e Multimétodo”, da Associação Norte-americana de Ciência Política. A expressão multimétodos é um chamado para o pluralismo na escolha de estilos de pesquisa, fontes de dados e métodos analíticos. É também um chamado que tenta criar um pouco de espaço para abordagens que não envolvem análises estatísticas. Mas, para ser claro, da perspectiva de aprender acerca do mundo, o significado literal da expressão “pesquisa multimétodo” não faz muito sentido e é facilmente confundido com o objetivo desejável de ter múltiplas fontes de dados (exceto, é claro, quando se refere a diferentes métodos de coleta de dados ao invés de analisá-los; LIEBERMAN, 2005).

Ou seja, para um determinado tipo de informação coletada a partir de uma única fonte de dados, um método quase-ótimo (ou uma gama de métodos distintos somente por suposições inverificáveis) está disponível ou pode ser construído. É geralmente fácil chegar a um acordo quanto a este método ou gama de métodos, mesmo entre estudiosos com formações divergentes. Ademais, acrescentar análises baseadas em métodos fora deste escopo, por questões de analisar os mesmos dados, pode somente então significar usar abordagens deficientes.

Em contraste, coletar fontes de informação adicionais e diversas, que são implicações da mesma teoria, é uniformemente benéfico. Isso favorece diretamente o empreendimento da pesquisa ao nos permitir a maximização da alavancagem. Mais dados, nas mais diversas formas possíveis (i.e., tal que cada nova fonte de informação esteja minimamente ligada às fontes existentes), são sempre o melhor neste sentido. Assim, você pode às vezes necessitar de múltiplos métodos para lidar com múltiplas fontes de dados, *mas são fontes de multidados e não de multimétodos o valor que deve ser maximizado.*

Para ser mais específico: a diversidade vem em vários sabores. A diversidade de fontes de dados ajuda a certificar contra um viés da fonte, contanto que as fontes não sejam relacionadas. A diversidade de tipos de dados (tais como um estudo etnográfico detalhado de uma cidade vs. resumos abstratos quantitativos de vários países) pode garantir contra algum viés, mas eles são úteis principalmente por aumentarem a eficiência, visto que geralmente trarão implicações observáveis de uma teoria menos provável de estar relacionada, e assim mais provável de ser adicionalmente informativa, quando comparada aos dados existentes. Em contraste, a diversidade de método, dado um determinado conjunto de dados quantitativos ou qualitativos, somente nos leva para longe de quaisquer método ou métodos ótimos nesta situação. Aplicar diferentes métodos aos mesmos dados só é útil quando do estudo de dependência de modelos (ver mais na Seção 5.4).

Um ponto relacionado se refere à noção popular de enfatizar a coleta de evidência tanto quantitativa quanto qualitativa no mesmo trabalho acadêmico, especialmente em dissertações. Como deve estar claro agora, combinar ambas as fontes é bom somente se aumentar a diversidade e a quantidade de implicações observáveis. Mas não há valor *agregado* quando se enfatiza ambas: não ocorre mágica se elas estiverem misturadas nas proporções corretas para além do ponto

em que a coleta de dados possa aumentar a alavancagem de alguém. Pode ser verdade que as perspectivas de seu trabalho possam ser aprimoradas se você fizer ambas em alguns subcampos, isso pode estar ligado ao avanço de uma carreira individual, mas só funciona às vezes; não é necessariamente a melhor forma de fazer avanços científicos. Da perspectiva de aprender acerca do mundo, a alocação ótima de suas fontes escassas, incluindo seu próprio tempo, sugere coletar a maioria dos dados e as mais diversas formas de dados. Isso pode sugerir acrescentar dados quantitativos a qualitativos, casos qualitativos a dados quantitativos, ou isso pode significar coletar formas adicionais dos mesmos tipos de dados. A questão é o que maximiza a alavancagem, e não se você consegue coletar uma folha de toda árvore.

3.4 Dados quantitativos ou qualitativos são melhores para a sua pesquisa?

Notadamente, esta é uma pergunta com uma resposta extremamente bem justificada, baseada em uma pesquisa de apoio considerável – pesquisa que parece ter sido completamente ignorada pela literatura de métodos qualitativos. O fato de que a resposta possa ser incendiária para alguns não a torna menos verdadeira. Lidar com a resposta deve melhorar bastante todo tipo de pesquisa em nosso campo. Explicamos a resposta através de dois fatos em separado:

O primeiro fato é:

Quando informações insuficientes acerca de um problema tiverem sido quantificadas para tornar as análises estatísticas úteis, e existir informação qualitativa adicional, o julgamento e análise qualitativos serão geralmente superiores à aplicação de métodos estatísticos. A evidência que apoia este ponto não é discutida seriamente. Nenhuma aplicação de métodos estatísticos, não importa quão sofisticada seja, pode superar uma fonte de informação inadequada. Sim, a teoria pode melhorar bastante uma análise estatística, mas a teoria deve em algum ponto ser baseada em fato empírico, do contrário a mesma história permanece: se a informação quantificada disponível for inadequada, análises estatísticas funcionarão inadequadamente e geralmente de forma pior do que análises qualitativas.

Os pesquisadores devem entender que é – e provavelmente sempre será – impossível quantificar a grande maioria da informação no mundo. Da última vez em que você entrou em uma sala de aula, você decidiu instantaneamente que seus alunos não o comeriam para o jantar. Quando você acordou esta manhã, você decidiu rapidamente que não havia emergência e provavelmente você descobriu em que cidade você estava sem muita demora. Quando lhe servem uma refeição, você consegue detectar com uma cheirada se ela está estragada com uma confiabilidade imperfeita, mas alta. Se não fôssemos capazes de tomar decisões instantâneas como estas com um determinado grau de precisão, a principal realização da humanidade teria sido servir de alimento para tigres dente de sabre. Em outras palavras, não importa quantos recursos estatísticos você tenha escolhido, decidir coletar dados e realizar uma regressão em momentos como esses, e em muitas áreas de pesquisa científica, não seria útil.

O segundo fato, muito menos amplamente conhecido em nosso campo do que o primeiro, é:

Quando a informação suficiente acerca de um problema puder ser quantificada (uma qualificação crucial!), uma análise estatística de alta qualidade é bem superior a julgamento qualitativo. A matemática e a estatística permitem que os seres humanos raciocinem de maneira adequada mesmo quando o raciocínio humano informal falha. O raciocínio humano, por sua vez, falha de maneiras altamente previsíveis, que especialistas qualitativos não têm conseguido superar mesmo quando o campo da estatística o tem. Julgamentos qualitativos por especialistas em disciplinas são rotineiramente melhorados, realizados de maneira melhor, raciocinados de maneira melhor e previstos de maneira melhor por abordagens estatísticas de força bruta. Isto é verdade mesmo quando os analistas de dados conhecem pouco acerca do problema substantivo de que dispõem e a informação quantificada parece impressionantemente incompleta para especialistas em disciplinas.

Este fato será surpreendente somente para aqueles não familiarizados com a literatura quantitativa, mas muitos exemplos desse ponto existem agora em tantos campos que ele não é mais discutido seriamente. Por exemplo, em um concurso frente a frente dois cientistas políticos com um modelo estatístico rudimentar de seis variáveis previram o resultado de casos da Suprema Corte dos Estados Unidos (sem chegarem a lê-los) mais acuradamente do que um grupo de 83 professores de direito e outros especialistas da área raciocinando qualitativamente e com acesso a muito mais informações e décadas de experiência jurisprudencial (MARTIN et al., 2004). Como outro exemplo, cientistas políticos há muito têm mais sucesso em prever as eleições presidenciais do que os gurus experts, agentes de pesquisa de opinião e outros (CAMPBELL, 2005; GELMAN; KING, 1993). Tetlock (2005, p. 64) mostrou que a maioria de seus 284 especialistas experientes, articulados e altamente eruditos, preveem muitos aspectos do futuro político com “menos habilidade do que simples algoritmos de extrapolação”. De forma semelhante, dois cientistas políticos sem treinamento médico construíram um modelo estatístico que se sai melhor que os médicos (avaliando causas de morte individuais) em determinar taxas de mortalidade específicas quanto à causa (KING; LU, 2008). Estes são somente quatro das centenas de exemplos em muitas áreas acadêmicas. De fato, pelo menos desde Meehl (1954), inúmeros concursos e comparações semelhantes têm ocorrido nos vários campos de estudo e da prática. O resultado não é sempre o mesmo, mas a mesma tendência muito forte que favorece as estimativas quantitativas é onisciente (GROVE, 2005). Existe até mesmo um livro bem abrangente e popular dedicado ao assunto (AYRES, 2007). A literatura de métodos qualitativos na ciência política precisa lidar com esses fatos. Uma previsão ou estimativa acurada é terrivelmente difícil na maioria dos campos sem uma precisão de mensuração.

A assertiva não significa que análises estatísticas ineficazes ou inadequadas, das quais existem muitas, sejam de alguma forma superiores ou mesmo necessariamente úteis. Conduzir análises quantitativas é difícil e desgastante, requer preparação e treinamento e podem facilmente ser feitas de maneira errada, e geralmente assim o são. A assertiva não é que análises estatísticas

sejam feitas mais frequentemente de forma correta do que análises qualitativas, somente que análises estatísticas de alta qualidade, baseadas em dados adequados, são geralmente superiores a uma análise qualitativa profissional, quando informações suficientes foram quantificadas. A mera presença de números em um artigo ou livro não traz necessariamente uma garantia de algo mais.

O fato de que muitas análises estatísticas são feitas de maneira ruim deveria nos levar a concluir que abordagens quantitativas não têm valor prático? Isto seria o caso sem uma abordagem razoavelmente unificada à teoria de inferência. Com ela, diferentes analistas, chegando a diferentes conclusões a partir de diferentes abordagens de análise de dados ou fontes de informação quantitativa, podem convergir a respostas semelhantes ou idênticas. A teoria de inferência fornece um arcabouço, um padrão comum que pode ser aplicado a abordagens aparentemente contraditórias. Isto é o que boas abordagens quantitativas têm a seu favor agora; isto é o que os estudiosos qualitativos às vezes se beneficiam agora. Mas, mesmo quando mais estudiosos qualitativos desenvolvem suas abordagens com referência a teorias comuns de inferência, métodos quantitativos ainda serão superiores a métodos qualitativos quando (e somente quando) informação suficiente tiver sido quantificada.

Se dados quantitativos ou qualitativos são melhores para a sua pesquisa, dependem, então, de quanta informação está disponível, o grau ao qual ela pode ser sistematizada e quantificada, e quanto tempo você pode dedicar ao problema. Quantificação por sua própria causa é uma perda de tempo (como bem sabem os estudiosos forçados pelos revisores a incluírem análises estatísticas sobre bebês em seus trabalhos qualitativos e também os orientadores de teses!). Mas, no pequeno número de situações em que você é capaz de quantificar a informação essencial, preparar-se para fazer uma análise estatística adequada e dedicar tempo e recursos necessários, vale a pena ir adiante porque a elevada precisão provavelmente produzirá resultados bem mais apurados. Se todas estas condições não permanecerem válidas, então é melhor proceder qualitativamente.

3.5 Qual pesquisa a inferência estatística não pode representar?

Agregar conhecimento acerca de um problema de pesquisa conduzindo análises estatísticas não é sempre fácil, eficiente, aconselhável ou útil, mas geralmente é possível, pelo menos em princípio. As críticas regulares na literatura de métodos qualitativos acerca de problemas particulares inevitáveis à pesquisa quantitativa parecem ser baseadas em determinadas aplicações de pesquisa quantitativa de aproximadamente 1970 ou de quando os críticos estavam na pós-graduação. Estudiosos qualitativos têm geralmente argumentado que a pesquisa quantitativa é incapaz de lidar com variáveis dependentes dicotômicas, variáveis colineares explanatórias, erro de mensuração, interações, caminhos multicausais, dependência de caminho, formas não lineares funcionais, seleção sobre a variável dependente, modelos sem formas específicas funcionais, problemas “superdeterminados”, análises sem modelos e inúmeros outros padrões e questões.

Essas assertivas, e muitas outras como elas, são falsas. De fato, iríamos adiante e faríamos a seguinte afirmação alternativa:

Toda afirmação inferencial, todo padrão empírico e toda noção de incerteza podem ser representados suficientemente bem, para questões de análise da ciência social, pela teoria estatística da inferência.

Alcançar uma abordagem estatística formal para qualquer ideia arbitrária não será sempre fácil, e de fato milhares de metodologistas em vários campos acadêmicos estão fazendo o mesmo com seus próprios problemas, mas isto deve ser sempre possível. No mínimo, nenhum teorema de impossibilidade foi afirmado ou provado. Para alguns exemplos recentes de novas abordagens em nosso campo relacionadas a discussões deste trabalho, ver Braumoeller (2003) e Glynne Quinn (2008).

Obviamente, somente porque uma afirmação inferencial possa ser dada em termos formais estatísticos não torna outras versões qualitativas inúteis. Elas podem ser tecnicamente desnecessárias, já que existem outros termos para os mesmos conceitos, mas ênfases diferentes podem ser muito úteis em orientar o conhecimento a novas abordagens e questões subapreciadas. Considere o seguinte exemplo.

Exemplo: O que é *dependência da trajetória*? A visão mais ampla de dependência de trajetória na pesquisa qualitativa é a ideia simples e importante de que a história tem uma importância. Nesta versão ampla, tudo discutido na literatura qualitativa acerca de dependência da trajetória tem sido ou pode facilmente ser formalizado dentro da extensa literatura quantitativa existente sobre análise de séries temporais (HAMILTON, 1994), ou, em outras palavras, literalmente nada é novo. Em outras formulações, dependência de trajetória se refere à ideia mais específica de processos históricos que têm “retornos crescentes”, que é quando eventos aparentemente pequenos acabam por ter consequências maiores (ou pelo menos permanentes), conforme o passar do tempo (PIERSON, 2000; HALL, 2009, trabalho vindouro). Esta noção mais específica de dependência da trajetória é também extremamente bem estudada na literatura de séries temporais sob os nomes processos “não estacionários” ou “raízes unitárias”, então, tecnicamente, nada é novo aqui também. No entanto, as ênfases das duas literaturas são quase os opostos exatos, com quase nenhuma referência cruzada e assim pouca possibilidade de se basear no trabalho de cada uma. A falta de contato entre estes campos complementares apresenta uma oportunidade substancial; as ênfases diametralmente opostas servem como um propósito importante em encorajar estudiosos para focarem em ideias que possam ser perdidas em outro campo.

Para observar este ponto, observe que qualquer processo histórico ou de série temporal pode ser decomposto em componentes estacionários e não estacionários. Os componentes estacionários, não importam quão complicados, são aqueles que seguem os mesmos padrões probabilísticos sempre que eles ocorrem, para que, por exemplo, os efeitos de eventos ou choques a séries temporais não cresçam sem limite com o tempo. Os componentes não estacionários são aquelas partes com retornos crescentes ou decrescentes, ou qualquer característica dependente de um determinado período histórico. O ponto chave aqui é que *enquanto a literatura qualitativa sobre dependência da trajetória coloca ênfase primária no componente não estacionário, a literatura estatística geralmente vê o componente não estacionário como um problema a ser corrigido (ou “diferenciado para longe” na sua linguagem) e em vez disso focaliza no componente*

estacionário. Sua ênfase às vezes faz com que pareça como se eles tratassem aspectos importantes dos dados como um “estorvo” ao invés de uma disciplina válida de estudos (BECK; KATZ, 1996).

A ideia de dependência da trajetória não é errada; ela pode ser supérflua, mas dizer a mesma coisa em uma linguagem diferente é obviamente importante neste contexto. Encorajamos estudiosos qualitativos a integrarem seu trabalho de maneira mais eficaz com a literatura existente de séries temporais estatísticas, pois ela pode tornar assertivas qualitativas mais eficazes, poderosas, abrangentes e precisas. Ao mesmo tempo, estudiosos quantitativos poderiam tornar seu trabalho mais relevante, abrangente e influente ao desenvolver suas ferramentas para analisar os componentes não estocásticos de séries temporais, e não os tratar como meros estorvos a serem corrigidos. E, é lógico, não existe qualquer motivo que seja para ignorar o componente não estocástico em séries temporais quantitativas ou o componente estocástico em pesquisa qualitativa.

Uma das melhores coisas acerca da estatística é que o campo tem feito um progresso impressionante, com desenvolvimentos aparecendo mais rápidos a cada ano. Os vários campos da estatística estão se multiplicando, com novos métodos sendo desenvolvidos e aplicados a gamas cada vez maiores de problemas. A marcha da quantificação através dos campos da ciência tem continuado rápida e constante. A julgar pelos artigos publicados em nosso principal periódico, cerca de metade da pesquisa em ciência política desde fins da década de 1960 tem sido quantitativa. Na Economia a porcentagem é maior; na Sociologia é um tanto menos. A Medicina é agora “baseada em evidência”. A tendência mais quente na pesquisa da escola do direito é quantitativa (que eles chamam de “pesquisa empírica”). A Biologia era outrora altamente qualitativa, mas a bioestatística e a bioinformática e vários outros subcampos têm se desenvolvido para possibilitar muitas análises outrora nunca possíveis. Os governos e as empresas agora conduzem inúmeros experimentos randômicos em larga escala.

Há campos de estudo que ainda não se revolucionaram com o aumento da quantificação e da estatística moderna, mas é uma predição fácil de que isso ocorra eventualmente, onde os estudiosos sejam empreendedores o bastante, onde quer que isto seja útil (e infelizmente, em outros períodos também!). Certamente, as oportunidades para a arbitragem intelectual são enormes. Para tomar como exemplo, por questões de clareza fora de nosso campo, considere a arquitetura. De longe, as decisões mais dispendiosas que as universidades tomam são acerca de prédios e sua planta física. Porém, a arquitetura como um campo é composta principalmente de engenheiros que conservam prédios e sujeitos criativos qualitativos que inventam novos *designs*: cientistas sociais quantitativos geralmente não arranjam trabalho em escolas de *design*. Imagine em vez disso quanto progresso poderia ser feito por até mesmo uma simples coleta de dados e uma análise estatística direta.

Algumas questões relevantes, com variáveis explanatórias associadas poderiam ser: Os corredores ou as suítes fazem os alunos e o corpo docente produzirem e aprenderem mais? A circulação vertical funciona tão bem quanto a horizontal? Devemos colocar o corpo docente bem próximo de outros que trabalham com os mesmos projetos ou deveríamos maximizar adjacências interdisciplinares? (Os alunos de pós-graduação aprendem mais quando eles são

privados por uma falta de janelas que mostram o mundo lá fora durante o dia?) E se o propósito de uma universidade é mais ou menos maximizar o número de unidades de conhecimento criadas, disseminadas e preservadas, então coletar medidas não seria difícil, tais como contagens de citações, o número de novos docentes contratados ou de diplomas expedidos, a qualidade da colocação dos alunos no mercado ao ser formarem etc. Um pouco de análise social quantitativa em arquitetura poderia até ajudar a colocar estas decisões mais dispendiosas em uma base científica sólida.

Seria fácil continuar essa história, pois a estatística afeta muitas outras áreas de relações humanas. Na ciência política, a análise quantitativa está se juntando à análise qualitativa em campo após campo. Outrora, somente a política norte-americana incluía qualquer análise quantitativa. Agora, política comparativa e relações internacionais são quase tão quantitativas quanto a política norte-americana.

O ponto não é que tudo eventualmente será quantificado. Não será, visto que essa extremidade descartaria a grande maioria de informações disponíveis. O ponto é que inferência estatística não é um conjunto fixo de ferramentas metodológicas. É um conjunto de procedimentos para desenvolver ferramentas adaptadas a problemas conforme eles surgirem. Afirmar que alguma necessidade inferencial não é abrangida por inferência estatística pode não ser apropriado em um determinado período no tempo, mas somente até que alguém veja a necessidade e desenvolva uma nova ferramenta para tal propósito. Mas, simultaneamente, não importa quão longe a quantificação e a mensuração sistemática proceda, a análise qualitativa sempre será parte de toda análise. Para realizar o progresso mais rápido, precisamos de melhores conexões entre estes novos campos bem desenvolvidos.

4 Causalidade e explicação

Objetivos fundamentais de grande parte da ciência social quantitativa e qualitativa incluem causalidade e explicação, como distintos de descrição e inferência descritiva. Geralmente nós perguntamos *por quê?* A maioria de nossas teorias e a grande maioria de nossas análises empíricas busca ir além de mensuração para compreender a estrutura causal de fenômenos sociais. Cientistas políticos com frequência nem mesmo aceitam evidência preditiva sem alguma história causal plausível acerca de como as variáveis explanatórias escolhidas poderiam levar às variáveis dependentes. Causalidade parece definir a natureza essencial de muitas disciplinas de ciências sociais e distingui-las de campos profissionais relacionados.

Infelizmente, apesar do papel central de causalidade e explicação na maioria das áreas substantivas de pesquisa, existe um grande equívoco dentro da literatura de métodos qualitativos sobre estas questões e um equívoco em separado e tão grande quanto esse dentro da literatura quantitativa. O erro em uma área não é feito na outra, então às vezes há esperança de que uma comunicação adicional entre pesquisadores quantitativos e qualitativos possa resolver o problema. Apresentaremos a confusão qualitativa sobre a definição de efeitos causais na Seção 4.1 e a confusão da literatura quantitativa sobre o significado de explicação na Seção 4.2. A Seção 4.3 discute então as confusões em ambas as áreas sobre como estimar efeitos causais.

4.1 Definindo efeitos causais: confusão na literatura qualitativa

Estudiosos da literatura de métodos qualitativos às vezes escrevem como se existisse uma controvérsia em algum lugar sobre a definição fundamental do que é um efeito causal, e eles rotineiramente apresentam novas formulações — às vezes razoáveis, às vezes logicamente inconsistentes, mas geralmente desnecessárias — para tentar caracterizá-lo. De fato, a definição de efeito causal dada em King, Keohane e Verba (1994, cap.3) — agora amplamente conhecida como estrutura de resultados potenciais (ou o “modelo causal de Rubin”) — tem desde então se tornado a posição próxima de um consenso em quase todos os campos acadêmicos onde tais assuntos têm sido discutidos. Estatísticos atribuem esta definição a Neyman, Rubin e Holland; cientistas computacionais a Pearl; economistas a Granger e outros; epidemiologistas a Robins; e filósofos a Aristóteles, Locke, Hume, Mill ou Suppes. Alguns estudiosos preferem solicitar que a estrutura de resultados potenciais seja complementada com outras características, outros ajustam levemente a definição principal de algumas formas, e outros ainda estão mais interessados em conceitos ao invés de efeitos causais. Mas os cientistas políticos devem reconhecer que qualquer que seja a versão sobre a qual nos concentramos e para quem quer que seja que nós a atribuímos, a definição básica de resultados potenciais está agora bem estabelecida e geralmente acordada. De fato, é difícil pensar em muitos conceitos sofisticados que eram outrora assunto de tal discordância bem difundida, que estão agora tão acordados como a definição de resultados potenciais de causalidade.

4.1.1 A definição básica

A melhor maneira de compreender o âmago da definição de resultados potenciais de causalidade é no contexto de uma única unidade (pessoa, país ou outra observação). Suponha por simplicidade que o fator causal de interesse, o qual rotulamos de T_i para a variável de “tratamento” aplicada à unidade i , seja aplicado ($T_i = 1$) ou não aplicado ($T_i = 0$) à pessoa i a qualquer momento específico. Em algumas versões o tratamento deve ser manipulável, tal como implementar uma determinada política pública em alguns estados e não em outros ou um partido político escolher se endossa um titular à reeleição ao invés de um não titular; contudo, a maioria dos cientistas sociais permitem agora também o tratamento para incluir um atributo das unidades, por exemplo, gênero ou região, que é por questões práticas não manipulável (GOLDTHORPE, 2001).

Então suponha que observamos o valor da variável de resultado para a pessoa i , Y_i , quando exposta ao tratamento estatístico, a qual rotulamos $Y_i(1)$, e então nós obviamente não observamos e não podemos observar o valor da variável de resultado para a pessoa i ao mesmo tempo quando não exposta ao tratamento, $Y_i(0)$. O efeito causal de T sobre Y para a pessoa i é então a diferença entre os dois resultados potenciais, como $Y_i(1) - Y_i(0)$. Sendo $Y_i(0)$ não observável, o efeito causal nunca é conhecido com certeza e deve sempre ser estimado. Nós podemos tentar estimar $Y_i(0)$ em um momento diferente para a pessoa i ou para uma pessoa diferente semelhante a i que não foi exposta ao tratamento ao mesmo tempo, mas estas são estratégias de estimativa que requerem suposições que somente às vezes são válidas; no entanto,

de qualquer maneira, elas não têm nada a ver com a definição do efeito causal que é para uma pessoa i em um ponto único no tempo. Às vezes estudiosos estão interessados neste efeito causal com uma média sobre todas as unidades em alguma amostra ou população escolhida, mas a definição central é mais bem compreendida para uma unidade por vez. (Esta definição central é também complementada por alguns, para solicitar informações adicionais ou modelos formais acerca de aspectos de como a causa tem seu efeito; HECKMAN, 2008.)

Como um exemplo da definição básica, considere o efeito causal (digamos) da eleição de George W. Bush ao invés de Al Gore como presidente dos EUA em 2000 sobre o produto interno bruto do país (PIB) no ano de 2004. O efeito causal é a diferença entre o PIB real em 2004 (com Bush como presidente) menos o PIB que os EUA teriam tido em 2004 se Gore tivesse ganhado em 2000. Este segundo valor não observável do PIB é conhecido por um resultado potencial, já que é uma situação contrafactual que poderia ter sido observada em circunstâncias diferentes (i.e., se Gore tivesse sido eleito), mas não o foi.

Essa é uma definição razoavelmente específica de um efeito causal, mas ela não é ainda totalmente definida até que também detalhemos precisamente como o mundo contrafactual de Gore sendo eleito poderia ter acontecido. Uma forma seria imaginar que a Suprema Corte decidiu *Bush vs. Gore* de forma contrária. O efeito causal baseado nesta definição do efeito de tratamento é bem definido, visto que é fácil imaginar esta situação contrária tendo realmente acontecido. Podemos conjecturar que o efeito sobre o PIB com a Suprema Corte decidindo de maneira diferente seria semelhante ao efeito sobre o PIB com Gore sendo eleito por causa da ausência dos problemas com a cédula borboleta [usada nas eleições presidenciais de 2000, no Condado de Palm Beach, Flórida, e que causou grande confusão entre os eleitores], ou Ralph Nader decidisse não concorrer, ou por uma variedade de outras “pequenas” mudanças possíveis, mas estes são efeitos causais diferentes. Por exemplo, o efeito da decisão da Corte seria plausivelmente muito diferente de alguma outra situação contrária, como Gore sendo eleito porque o exército levou tanques de guerra através de Washington e forçou sua posse como presidente.

Porém, caso a situação contrafactual, e assim o efeito causal, seja definido, ela deve ser delineada antes que possamos começar a discutir estimativas. De fato, um ponto chave é que o efeito causal deve ser definido sem referência a qualquer método de estimativa – tais como as abordagens insatisfatórias de dizer que o efeito é o que acontece após controlar-se variáveis de controle desconcertantes em potencial, ou apelar para assertivas do tipo “todas as coisas são iguais”, ou confundir o efeito causal com um coeficiente de regressão, ou definir o efeito causal como a diferença entre Y antes e após um evento etc. Isto é essencial para coerência lógica e somente para isto, visto que uma quantidade alvo real de interesse deve existir antes de e em separado de nossos esforços para estimá-la.

Esse exemplo do efeito do PIB dá a versão mais básica de um efeito causal dentro da estrutura de resultados potenciais. Mas, embora a estrutura seja muito simples, ela tem provado ser uma abordagem notadamente durável e fértil com aplicabilidade extremamente ampla. Muitas versões mais sofisticadas de efeitos causais têm sido construídas sobre esta estrutura simples em muitas literaturas, incluindo estudos em muitos tipos de dados quantitativos e

qualitativos; seja para observações ao longo do tempo, através do espaço e variando entre ambos; seja para estudar efeitos intermediários e efeitos causais que ocorrem em diferentes níveis de análise. Um grande número de aplicações também apareceu, todas começando com esta distinção básica (MORGAN; WINSHIP, 2007).

Os estudiosos, aplicando métodos qualitativos, precisam chegar a um acordo quanto a essa definição bem mais do que já o têm feito. A precisão de esboçar a situação contrafactual envolvida em uma assertiva causal pode ser tremendamente esclarecedora de uma perspectiva teórica, pode ajudar a identificar dados relevantes para estimativa e outras implicações observáveis, e pode eliminar dados que não têm relevância à assertiva disponível. Como vemos nas aulas, aparentemente compreender a estrutura de resultados potenciais não é o mesmo que ser capaz de aplicá-la a um determinado problema. Como qualquer outra coisa, deve-se praticar para conseguir fazê-lo certo. Consegui-lo tem benefícios enormes, já que uma grande gama de implicações para a compreensão, a definição e a estimativa de efeitos causais tem sido trabalhada e poderia ser coletada por estudiosos qualitativos para o seu trabalho.

4.1.2 Para além da definição: as condições necessárias e suficientes

Muitos estudiosos, especialmente na literatura de métodos qualitativos, estão interessados em ideias tais como condições necessárias e suficientes. Estes são conceitos importantes, mas muita confusão seria eliminada se estudiosos reconhecessem que elas não são efeitos causais. Este fato não as torna sem importância ou indignas de estudo.

Começamos com uma definição de condições necessárias. Para simplicidade expositiva, suponha (como T) a variável de resultado Y que somente aceita o valor de 1 ou 0 para a presença ou ausência, respectivamente, de alguma característica. Então, uma condição necessária é que Y não pode ocorrer quando T estiver ausente: $p(Y = 1|T = 0) = 0$. Assim, uma condição necessária é um alvo inferencial bem definido, pois é sempre ou verdadeiro ou falso e é claramente à parte de qualquer método que possa ser usado para inferir se ele é válido. No entanto, a definição de uma condição necessária não envolve qualquer situação contrária e então não é uma afirmação causal (ou pelo menos é diferente da, e não relacionada à, definição de efeitos causais dada na Seção 4.1.1). A probabilidade $p(Y = 1|T = 0)$ é de fato uma quantidade *descritiva*, que requer inferência descritiva e não contrafactual (ver Figura 1).

Para ser mais específico, no contexto do resultado dicotômico nesse exemplo, o efeito causal pode ser escrito como $p(Y = 1|T = 1) - p(Y = 1|T = 0)$ (para uma única unidade). Porém, se T for necessário para Y , para que o segundo termo seja zero, o efeito causal se torna o primeiro termo, $p(Y = 1|T = 1)$, o que, óbvio, é uma quantidade inteiramente diferente de, e de forma alguma restrita por, $p(Y = 1|T = 0)$. [De fato, se $p(Y = 1|T = 0)$ for uma quantidade descritiva, então quando definida para a mesma unidade para a qual $T = 0$, $p(Y = 1|T = 1)$ trata-se de uma quantidade contrafactual]. Isto prova que efeitos causais e condições necessárias são alvos inferenciais em separado, ambos consistentes logicamente, ambos de algum interesse substantivo e nenhum dos quais deve ser confundido com o outro.

A mesma lógica pode ser aplicada facilmente às condições suficientes, visto que uma condição necessária sempre implica em pelo menos uma condição suficiente. Por exemplo, se

a condição necessária $p(Y = 1|T = 0) = 0$ for válida, então a condição suficiente equivalente $p(Y = 0|T = 0) = 1$ também é válida (já que a soma das duas é um). Muitas combinações criativas e sofisticadas e ampliações de condições necessárias e suficientes têm sido também desenvolvidas (MAHONEY, 2008), e é óbvio que o mesmo ponto se aplica a todas elas: elas são alvos inferenciais dignos de nota, mas não efeitos causais. Um pesquisador pode desejar escolher entre as duas para aplicar a um determinado problema aplicado, mas aqueles interessados em métodos qualitativos nunca precisam escolher uma ou outra, visto que elas são objetos de pesquisa relevantes e diferentes.

Pesquisadores quantitativos às vezes menosprezam condições necessárias e suficientes, pois não vivemos em um mundo determinístico. Como tal, com pouco erro de mensuração ou variabilidade estocástica as condições nunca parecem ser válidas na prática. Porém, pelo menos no contexto de erro de mensuração os conceitos são usados frequentemente (GOERTZ, 2003) e bem definidos, e eles podem ser facilmente usados como assuntos para inferência descritiva (BRAUMOELLER; GOERTZ, 2000).

4.1.3 Ainda para além da definição

Estudiosos qualitativos e quantitativos parecem regularmente confusos com termos relacionados a efeitos causais, tais como *mecanismos causais*, *efeitos superdeterminados*, *equifinalidade* e *multicausalidade*. A definição de um efeito causal é dada na Seção 4.1. Estes termos não são obrigatórios nem necessariamente úteis para compreender a definição, aplicá-la ou estimar efeitos causais. É claro que eles são às vezes úteis em encorajar estudiosos a enfatizar diferentes abordagens de coleta de dados, procurar diferentes tipos de padrões e em alguns casos identificar relações específicas em dados, mas dar uma definição coerente de cada um é mais bem conseguido por se basear na, ao invés de ignorar, definição de resultados potenciais de efeitos causais.

Por exemplo, a recomendação de buscar por mecanismos causais após estabelecer um determinado efeito causal pode ser muito útil em motivar um programa de pesquisa produtivo, produzindo uma melhor compreensão do mundo, ou gerar ou confirmar uma teoria mais poderosa do fenômeno sendo estudado (GERRING, 2007, p. 185). Todavia, *definir* causalidade somente em termos de mecanismos causais não faz sentido lógico. No exemplo corrente da seção anterior, uma hipótese acerca de um mecanismo causal é que Gore não teria ido à guerra no Iraque, então o governo dos EUA teria gastado menos em suprimentos de guerra. Há alguma evidência de que gastar em guerra tem menos efeitos positivos sobre a economia do que outros tipos de gastos, e assim a eleição de Bush, ao invés de Gore, pode ter decrescido o PIB dos EUA. Este é um assunto de estudo interessante e talvez importante, em que a estrutura de resultados potenciais engloba inteiramente (GLYNN; QUINN, 2008), mas não é uma definição autossuficiente do efeito causal do resultado da eleição sobre o PIB. O motivo é regressão infinita: se você tentar definir o efeito causal em termos do mecanismo, com o mecanismo meramente sendo outra hipótese causal, então a pergunta imediatamente se torna qual é o mecanismo do mecanismo. Em nosso exemplo, precisaríamos então do mecanismo pelo qual gastos em guerra não têm efeito tão positivo sobre a economia quanto gastos em bem-estar social – pois muitos dos produtos

são enviados para outros continentes ou destruídos por explosão. Mas, então, precisamos do mecanismo pelo qual isto acontece etc. Esta regressão infinita pode definir bem uma agenda produtiva de pesquisa, mas não oferece uma definição coerente de um efeito causal.

Para outra confusão semelhante, fenômenos são geralmente descritos na literatura como sendo superdeterminados, significando que mais de uma causa poderia ter produzido o mesmo resultado. Como verificado, esta frase é basicamente vaga porque por definição quase todo evento é superdeterminado. Ele morreu porque seu coração parou, ou ele tinha muito pouco sangue, ou sua aorta foi perfurada, ou a bala entrou no seu corpo, ou o gatilho foi puxado, ou ele foi baleado, ou o assassino realizou seu intento, ou o assassino foi induzido por seus amigos, sua formação, ou a sociedade, ou por causa da forma como se deu o processo biológico da evolução, ou por causa da física do *Big Bang*. O conceito de ser superdeterminado é incoerente.

Pelos mesmos motivos, cada evento e variável podem ser descritos como um exemplo de causalidade múltipla ou equifinalidade. Estes termos – como superdeterminação e mecanismos causais – não distinguem uma causa de qualquer outra. Eles podem ajudar em estabelecer certas ênfases, mas em conduzir pesquisa empírica real para o propósito de definir e estimar efeitos causais (sejam quantitativos ou qualitativos), o foco deve ser sobre a definição de efeitos causais dada na Seção 4.1, precisamente como sua hipótese casual e a situação contrária subentendida podem ser definidas, e que suposições podem ser necessárias para estimá-la.

4.2 Definindo explicação: confusão na literatura quantitativa

Efeitos causais estão bem definidos, mas eles não são o mesmo que explicações. Este fato é mal compreendido em toda a literatura quantitativa, mas é razoavelmente bem compreendido entre estudiosos qualitativos. Esclareceremos aqui estes pontos.

O problema parece se originar em uma declaração no influente artigo de Holland (1986, p. 945) sobre a estrutura de resultados potenciais:

A ênfase aqui será em *mensurar os efeitos de causas* porque isto parece ser um local onde a estatística, que é preocupada com mensuração, tem uma contribuição a fazer. É minha opinião que uma ênfase sobre os efeitos de causas ao invés de as causas de efeitos é, em si própria, uma consequência importante de trazer o raciocínio estatístico a ter efeito sobre a análise de causalidade e contrapõe-se diretamente às análises mais tradicionais de causalidade.

Mensurar os efeitos das causas é de fato a noção por trás da definição de resultados potenciais de causalidade. A variável de tratamento é definida e então o efeito causal é a diferença entre dois resultados (ou “efeitos”). O que Holland quer dizer com “as causas de efeitos” é raciocinar de trás para frente a partir do efeito ou evento (ou processo, pessoas, perspectiva, papel, comunidade etc.) a uma ou mais causas daquele evento. Definir este processo é logicamente impossível de uma maneira bem definida, e tem os mesmos problemas de regressão infinita como tem o conceito de “superdeterminado”, discutido acima: mais causas sempre existem

para qualquer evento. Tudo o que você precisa fazer é olhar para níveis próximos ou distantes diferentes.

A decisão de Holland em se concentrar sobre os efeitos de causas é um foco perfeitamente razoável para um artigo, e de fato para literaturas por inteiro, mas não significa que procurar pelas causas de efeitos não faz sentido. De fato, o que é amplamente chamado de causas de efeito na literatura quantitativa é chamado mais geralmente na literatura qualitativa de *explicação*. Estatísticos sabem que têm menos probabilidade de ter sucesso em criar uma definição precisa e bem apresentada de explicação, pois isto envolve descobrir as causas de efeitos, o que não se encaixa bem na estrutura de resultados potenciais porque a própria causa não está bem definida, e para a qual nenhuma outra definição precisa ainda foi desenvolvida. Porém, somente porque estatísticos não criaram uma maneira precisa e coerente de fazê-lo, não significa que devemos esperar que os cientistas sociais parem de tentar explicar fenômenos sociais! [Trabalho relacionado que toca em questões de explicação de formas um tanto diferentes existem na filosofia, Little (1991), e na economia, HECKMAN, 2008].

Se isso ainda não soar convincente, como se explica o conselho para considerar teorias rivais quando ao projetar uma pesquisa? Manter uma mente aberta para a possibilidade de estar enormemente errado enquanto persegue as implicações observáveis de sua teoria favorita atual é difícil de argumentar, e então o conselho parece altamente útil. Mas uma versão específica de teorias rivais são causas rivais, e causas rivais significam exatamente procurar pelas causas de efeitos. Novamente, o fato de que isto não é prontamente acessível à análise sob uma estrutura de resultados potenciais não significa que devemos ignorar explicações rivais.

Então, o que é exatamente explicação? Bem, não é exatamente algo, visto que sua definição nunca foi precisa. Mas, basicamente, é um conjunto de causas tais que os efeitos destas causas sobre o fenômeno de interesse são “grandes” e ao nível “correto” de proximidade ao evento. Tais causas devem também ter algum outro atrativo, mas existem propriedades difíceis de serem afirmadas, tais como as concernentes às disciplinas ou áreas aceitáveis sendo estudadas. Então, comece com um resultado – o PIB dos EUA foi 11,4 trilhões de dólares em 2004 – e em seguida retroceda para encontrar uma variável de tratamento que tenha um grande efeito sobre esse resultado. Uma resposta possível seria a decisão da Suprema Corte dos EUA em *Bush vs. Gore*. Deve estar claro que isto é de fato um tipo de resposta que buscamos quando procuramos pela explicação do PIB dos EUA em 2004 e porque ele é diferente do que poderíamos ter esperado. No entanto, seu *status* como “a explicação” é altamente impreciso e ambíguo.

Para uma resposta, outra explicação perfeitamente razoável seria o preço do petróleo, e outra seria o estado da economia mundial, e outra poderia ser os incidentes terroristas de 11 de setembro de 2001. Outra, em um nível diferente de proximidade, seriam o tamanho e a abertura da economia norte-americana. Há milhares de candidatos explanatórios também, e nenhum sentido em determinar que o conjunto destes seja mais apropriado como “a explicação” do que outros. De fato, mesmo todas as causas juntas em que possamos pensar não constituem necessariamente a explicação “completa”, já que tal coisa ou não existe ou é extremamente mal definida. Afinal, uma “boa” explicação por estas regras pelo que o PIB dos EUA foi de 11,4 trilhões em 31 de dezembro de 2004 é que ele estava somente um pouco abaixo em 30 de

dezembro de 2004. Talvez isso “explique” o resultado, mas por outros motivos não estudados pode não ser muito satisfatório (ver também HALL, 2006).

Você consegue ver para onde esta conversa está indo: Estudiosos qualitativos compreendem que a verdadeira questão que os cientistas sociais têm em muitos casos é o que explica um evento: Por que aconteceu a crise cubana dos mísseis? Por que, em eleições livres e justas na República de Weimar (Alemanha), os eleitores escolheram eleger Hitler? Qual foi a causa da Revolução Francesa? Por que os terroristas explodiram aqueles prédios no dia 11 de setembro? Estes são todos temas pedindo explicações, pelas causas de efeitos. Quando estudiosos quantitativos ouvem questões como estas eles às vezes desdenham de sua imprecisão, que é uma reclamação justa, mas também sem objetivo. Ou seja, nenhuma destas questões está definida tão claramente como a estrutura de resultados potenciais quando da procura por efeitos de causas.

Estudiosos quantitativos podem começar a definir e avaliar efeitos causais via resultados potenciais para cada fator de candidatura explanatória – que é sua sugestão mais útil em casos como estes –, mas eles não definiram ou formalizaram adequadamente a definição de uma explicação ou como estimá-la. Eles precisam compreender que explicações são um alvo legítimo de pesquisa científica social, apesar de serem vagas. A comunidade estatística pode não gostar muito deste ponto ou, de qualquer forma, ainda não fez algo a respeito. Uma agenda importante de pesquisa dentro desse campo deve ser formalizar a “explicação” e mostrar tanto aos estudiosos quantitativos quanto aos qualitativos como produzir explicações de maneiras mais exatas e eficientes. Até lá, a literatura qualitativa liderará, mesmo que de maneira imprecisa.

4.3 Inferência causal sem premissas de regressão

Existe uma noção estranha dentro da literatura de métodos qualitativos, e em algumas áreas de pesquisa quantitativa, de que estimativas quantitativas de inferências causais requerem alguma forma de análise de regressão. Esta noção é falsa. A literatura quantitativa não só inclui inúmeros tipos de métodos inferenciais para estimar efeitos causais, mas alguns métodos são muito próximos àqueles usados na literatura de métodos qualitativos. Nesta seção, discutiremos um método chave – emparelhamento – que pode ser facilmente descrito como tanto um método quantitativo quanto um método qualitativo. De fato, métodos modernos de emparelhamento são altamente intuitivos e não paramétricos, e não requerem qualquer linearidade ou premissas de normalidade (HO et al., 2007). Ambos os campos podem se beneficiar de um maior aprendizado acerca destes métodos, pois eles podem ser vistos como formalizações de algumas ideias de inferência causal na literatura de métodos qualitativos, e as formalizações e os métodos específicos podem ser usados para construir melhores desenhos de pesquisa quantitativa e qualitativa. De fato, todas as formalizações que usam regressão para intuição em King, Keohane e Verba (1994) podem ser descritas no contexto de emparelhamento sem premissas de regressão.

Na sua formulação mais simples, a ideia de emparelhamento tem paralelo bem próximo aos resultados potenciais que necessitam ser estimados. Suponha que estejamos interessados no efeito da forma de governo de uma cidade, tal como um prefeito eleito vs. um gestor municipal profissional, sobre alguma medida de capital social. Todos nós sabemos que meramente comparar

a quantidade de capital social em cidades com prefeitos eleitos à quantidade de capital social em cidades com um gestor profissional pode levar a um viés de “variável omitida”. Por exemplo, suponha que gestores municipais sejam mais prevaletentes em uma região do país que por acaso tem mais capital social bem antes de a forma de governo ter sido decidida. Em sendo a região uma causa anterior à forma de governo, e podendo também a região afetar o capital social após levar em consideração a forma de governo, é provável um viés de variável omitida.

Evitar esse problema via emparelhamento é fácil. Por simplicidade, considere esse efeito causal somente para cidades com um prefeito eleito. O efeito causal é então definido como a quantidade de capital social observada (i.e., com um prefeito eleito) menos a quantidade de capital social que teria sido observada naquelas mesmas cidades se elas, ao invés disso, tivessem um gestor profissional. Este segundo resultado potencial não é observável e então o estimamos encontrando, para cada cidade com um prefeito eleito, uma cidade real com um gestor profissional que seja emparelhado na região (i.e., que seja da mesma região). Em seguida tiramos a diferença no capital social para estas duas cidades; repetimos o mesmo procedimento de emparelhamento para cada cidade com prefeitos eleitos; e, em seguida, tiramos a média aritmética dos efeitos causais estimados.

É óbvio, emparelhamento, como regressão e todos os outros métodos quantitativos e qualitativos usados para estimar efeitos causais a partir de dados não experimentais, requer que o investigador identifique todas as variáveis que possam causar confusão para se controlarem, não somente uma única variável conveniente como em nosso exemplo. Como sempre, os potenciais fatores de confusão incluem todas as variáveis que satisfazem três condições: são existentes causas anteriores ao tratamento, relacionadas ao tratamento e afetam a variável de resultado após o tratamento. Evitar o viés de variável omitida é um problema difícil em dados observacionais, mas é lógico que ele é bem conhecido em toda a disciplina e então nós pelo menos temos maneiras de pensar nele e atacá-lo. Emparelhamento e outras abordagens não habilitam os pesquisadores a se evadirem do viés de variável omitida, somente a evitarem fazer tudo menos as mínimas premissas adicionais após as variáveis omitidas serem identificadas e mensuradas.

Assim, essa simples versão de emparelhamento e o método de Mills da diferença são quase o mesmo procedimento. A diferença vem quando nós percebemos que os dois grupos de cidades diferem mais em princípios do que somente em região. De fato, é fácil pensar em dezenas de variáveis potenciais omitidas nesse exemplo, que são casualidades anteriores à forma de governo e que podem afetar o grau de capital social. Como resultado, encontrar até mesmo duas cidades que diferem em sua forma de governo, exatamente emparelhadas sobre todas essas outras variáveis, se torna rapidamente impossível. Mas é exatamente onde os métodos modernos de emparelhamento começam a se diferenciarem dos métodos existentes e começam a fazer a verdadeira diferença, visto que eles permitem que se emparelhem aproximadamente das melhores maneiras possíveis sobre inúmeras variáveis. Métodos de emparelhamento também não são somente um análogo quantitativo da pesquisa qualitativa, já que eles podem ser usados bem diretamente na pesquisa qualitativa para selecionar casos, desenhar estratégias de pesquisa e realçar heterogeneidade inaceitável.

De forma semelhante, estudiosos quantitativos também têm muito a aprender a partir dos

usos qualitativos de emparelhamento: em muitos conjuntos de dados quantitativos a identidade de cada assunto é conhecida (tal como em política comparada, onde a unidade é o país ou país-ano, ou em relações internacionais, onde a unidade é o conflito ou díade-ano). Isto significa que uma quantidade considerável de informação não numérica existe em conjuntos de dados quantitativos que pode ser usada para melhorar a qualidade de algoritmos de emparelhamento. Até mesmo um simples passo de listar as correspondências (parelhas) para os leitores verem geralmente, realça algumas correspondências que podem ser melhoradas. Para isso, a melhor maneira adiante é que pesquisadores quantitativos adicionem algo deste tipo de pesquisa qualitativa ao seu repertório.

Então, regressão não é simplesmente sinônimo de inferência causal. De fato, mesmo naqueles casos quando regressão é usada com emparelhamento, este torna as inferências causais muito menos sensíveis às premissas que subjazem a regressão e outras formas de modelagem estatística. Parece que tanto os estudiosos qualitativos quanto os quantitativos podem se beneficiar muito destes métodos.

5 Seleção

5.1 O absurdo da seleção aleatória para desenhos de estudos de caso

A aleatoriedade tem um *status* singular na literatura de métodos qualitativos. Afinal, é um conceito extremamente poderoso, possivelmente o desenvolvimento mais importante nos métodos de pesquisa dos últimos cem anos. Estudiosos na literatura de métodos qualitativos invocam a ideia em discutir a seleção de casos, mas quase todas as invocações são irrelevantes ou inúteis. Felizmente, aqueles que aplicam métodos qualitativos sem pesquisa empírica parecem conhecer intuitivamente que seleção aleatória é inadequada para desenhos de pesquisa com um pequeno número de observações, e de fato assim o é. Eis o porquê.

O que a aleatoriedade faz na pesquisa experimental é possibilitar a seleção de casos e atribuir valores à variável de tratamento, de uma forma que nos proteja, não só de variáveis que possam causar confusão e sejam conhecidas, mas também de todas as variáveis que possam causar confusão e que não conhecemos. O que é possível fazer algo para nos proteger de males desconhecidos é um desenvolvimento surpreendente e impressionante.

Para ser mais específico, podem-se fazer inferências relativamente automáticas com (1) seleção aleatória de observações a partir de uma população conhecida, (2) atribuição aleatória de valores da variável causal principal e (3) um grande n . Todos devem estar presentes ou o estudo se torna um estudo essencialmente observacional com todas as premissas e problemas potenciais de qualquer pesquisa.

Mas, dados os benefícios de inferências relativamente automáticas livres de premissas, quem não quereria entrar nesta onda de aleatoriedade? A resposta é: qualquer pessoa que possa selecionar somente um conjunto bem pequeno de casos. King, Keohane e Verba (1994, p. 124-128) dão um exemplo com três observações, onde a seleção aleatória dá a você a resposta incorreta em dois terços das vezes. De fato, isto é bem geral: se você puder somente se dar o luxo de selecionar um pequeno número de casos, a seleção aleatória será geralmente pior do

que a seleção por outros meios. Como explica a Seção 5.2, se você puder somente coletar um pequeno número de casos, então qualquer método de seleção, incluindo aleatoriedade, não o deixará generalizar com algum nível útil de certeza. Nesse ponto, a seleção aleatória foge à questão. Você precisa ou de mais informação ou deve focar nos casos disponíveis.

Quando coletar mais informação é inviável, como geralmente o é para pesquisadores de estudos de caso qualitativos, vá para onde está a informação: dentro de seus casos. Quando é possível coletar mais casos ou informações ao mesmo nível de análise, então um projeto estratificado evitará um pouco da variabilidade induzida pela aleatoriedade. A ideia de estratificação é selecionar casos dentro de categorias de uma ou mais variáveis explanatórias de pré-tratamento (“pré-tratamento” se refere a variáveis explanatórias que são causativamente anteriores à sua variável causal principal). Por exemplo, se você selecionar países aleatoriamente, você poderia por acaso conseguir em sua determinada amostra somente países democráticos, mas se você estratificar sobre democracia, a sua amostra incluirá exatamente quantos países democráticos ou não que você escolher *ex ante*. Estratificação é uma estratégia bem mais eficiente quando viável (ver IMAI; KING; STUART, 2008). De fato, pode-se pensar em estratificação como um emparelhamento (conforme a Seção 4.3) anterior à atribuição de tratamento.

5.2 Preocupe-se menos em selecionar casos e mais em analisá-los

O problema da seleção é crucial em toda pesquisa, mas estudiosos na literatura de métodos qualitativos tendem a focar muito pouco em seleção *dentro de* estudos de caso e relativamente um tanto demais sobre que casos estudarem. Ou seja, eles escrevem acerca de como selecionar (por exemplo) seus cinco estudos de caso, mas às vezes parecem agir como se não existisse regra (ou talvez “o pesquisador saiba melhor”) quando se trata de coletar informação dentro de seus casos (DUNEIER, 2008). É lógico que ambos são crucialmente importantes e podem determinar as respostas às perguntas apresentadas. Mas pesquisadores qualitativos não deveriam se sentir culpados quando da seleção de uma pequena amostra de forma não aleatória ou não representativa. Inferências nesta situação serão necessariamente limitadas à amostra ao invés de alguma população mais ampla, mas aprender algo acerca de alguma parte do mundo em algum período prescrito da história é geralmente valioso por si próprio. Vale a pena reconhecer isto especialmente na pesquisa qualitativa, na qual compreender eventos ou processos singulares é geralmente um objetivo mais desejável do que generalização a casos não correlatos (MAHONEY; GOERTZ, 2006, Seção 1). Certamente assim o é em inúmeras outras áreas, especialmente quando – como com a seleção de casos, mas não necessariamente a rica informação dentro de casos – as amostras são limitadas a pequenos números.

De fato, a grande maioria da pesquisa de todo tipo – quantitativa e qualitativa, observacional e experimental, e em cada campo de estudo – seleciona suas principais unidades de análise baseadas em conveniência, interesse pessoal e recursos disponíveis. É lógico que há sempre alguma atenção a questões de generalização, mas o principal objetivo de muita pesquisa é tipicamente obter informação útil acerca de alguma amostra. Mesmo os ensaios clínicos mais caros de alta qualidade e aleatórios na medicina geralmente selecionam sujeitos experimentais baseados em

quem por acaso esteja em determinado hospital, atendendo a um anúncio de jornal, que esteja querendo se inscrever em um estudo, e impressione o pesquisador a pensar que esta pessoa responderá ao tratamento. A maioria dos pesquisadores observacionais coleta qualquer dado que seja conveniente. Em estudos de política comparativa, se mais países estivessem disponíveis ou dados adicionais fossem fáceis de coletar por mais anos, eles seriam quase certamente incluídos. Até mesmo levantamentos de amostras não envolvem seleção verdadeiramente aleatória a partir de populações conhecidas e enumeradas, embora eles pelo menos geralmente o tentem. Na maioria de outras pesquisas, estudiosos nem mesmo o tentam. Ao invés disso, eles focam em aprender o que conseguem acerca das observações disponíveis, e não deve haver vergonha em fazê-lo.

Pesquisadores estatísticos em geral tentam explicitamente estimar somente o *efeito de tratamento de média da amostra* – o efeito de alguma causa na amostra disponível – ao invés do *efeito de tratamento de média da população* – o efeito causal em alguma população específica mais ampla (IMAI; KING; STUART, 2008). Fazer inferências acerca do primeiro é em geral consideravelmente mais fácil, menos incerto e mais bem definido do que o último. Se outros pesquisadores fazem inferências baseadas em amostra acerca de diferentes áreas ou períodos temporais, então o produto coletivo de todos esses estudos pode ser consideravelmente maior do que a pesquisa de uma pessoa, o que lógico é a contribuição assinalável, senão a definição, de uma comunidade científica. E caso você esteja ou não interessado na população de todas as crises internacionais importantes ou todos os incidentes terroristas, um estudo que somente nos conte acerca de por que a Crise Cubana dos Mísseis ou os incidentes terroristas de 11 de setembro aconteceram pode ser excepcionalmente importante. Não há motivo para se sentir inferior sobre a falta de representatividade de seu estudo: se o pequeno número de casos que você coletou for verdadeiramente importante, então estes casos não são representativos quase por definição.

Além disso, mesmo se fosse possível coletar uma amostra verdadeiramente aleatória de um número muito pequeno de observações, a seleção aleatória poderia não ser bem aconselhável. Afinal, um pequeno n significa que a incerteza devida somente à variabilidade de amostragem seria enorme. Mesmo sem erro de mensuração ou qualquer outro problema inferencial, não aprenderíamos muito. Isto, lógico, não significa que estudos qualitativos com poucos casos são insignificantes. Pelo contrário, só significa que a grande maioria da informação aprendida neles está dentro dos casos acerca do efeito de tratamento da amostra em vez de ao nível do caso acerca de efeitos de tratamento da população.

Assim, uma vez que se estreite o alvo inferencial a aspectos de seus casos particulares, deve-se ter um extremo cuidado para selecionar informação a ser mensurada e registrada de maneira sistemática e imparcial. Aqui – dentro de casos – é onde a grande maioria do pensamento acerca de questões de seleção deve ser colocada por pesquisadores qualitativos. Se isto não acontecer, então não só você pode perder a habilidade de generalizar a alguma outra população, como você pode perder a habilidade de aprender até mesmo acerca de seus casos. Nesse ponto, há pouco sentido em ter conduzido o estudo de qualquer forma. E *dentro de casos, todas as mesmas questões de viés de seleção se aplicam que estão agora focadas em seleção de caso para inferir a*

uma população maior. Então deve haver um esforço considerável dedicado a convencer-se e os leitores de que o método de selecionar informação dentro de casos é representativo desses casos. Não é suficiente dizer que você foi entrevistar várias pessoas sem esclarecer como elas foram selecionadas, ou que você perambulou ou “impregnou e investigou” ou escolheu o seu caminho para perambular, onde você impregnou e qual foi o processo pelo qual você investigou. O mesmo vale para as regras pelas quais você poderia conduzir observação participante: a pesquisa não é sobre você; é acerca do mundo que você está estudando que requer compreensão do processo pelo qual você selecionou evidência dentro de seus casos. A seleção casual dentro de casos não é necessariamente representativa daquele caso. Leitores precisam de evidência documentada que apoiem exatamente como a informação dentro do caso foi coletada, como ela pode ser vista como representativa desses casos e o que poderia ter sido passado despercebido.

Por esses motivos, se você puder somente coletar poucos casos para estudar, então você deve considerar razoável seguir em frente e selecionar os casos que lhe interessem, ou aqueles que lhe ajudarão a responder ou gerar suas perguntas teóricas, ou aqueles que são mais importantes, ou que informariam melhor a literatura. Não espere que um pequeno número de casos seja representativo de uma população muito maior, mas não se preocupe demais com isto também. Mas tenha certeza absoluta de conseguir seus casos corretamente, e de planejar, compreender, executar e comunicar o processo pelo qual seus dados foram coletados.

5.3 Como e porque optar por Y

“Não opte pela variável dependente” é uma regra bem conhecida na literatura de métodos quantitativos e qualitativos. É também adequada, visto que optar por Y pode influenciar inferências descritivas e causais (KING; KEOHANE; VERBA, 1994, cap. 4). No entanto, a mesma regra causa grande consternação entre pesquisadores qualitativos de estudo de caso, que compreensivelmente não querem perder a Crise Cubana dos Mísseis quando da seleção de suas cinco crises para estudo ou a Segunda Guerra Mundial quando da seleção de conflitos internacionais. Então, o que é que há? Eles realmente precisam selecionar casos sem consideração às questões mais importantes com que eles se importam?

Seria prudente que pesquisadores qualitativos aprendessem mais sobre *estudos de controle de caso*. Este projeto de coleta de dados possibilita a você optar por sua variável de resultado, como algum número de crises e o mesmo número de não-crisis. Então, a regra geral que proíbe alguém de optar pela variável dependente não se aplica se você não optar pela variável explanatória e você corrigir de determinada forma. (Se você selecionar casos baseados nos valores de tanto as variáveis dependentes quanto as explanatórias, nada sobra para aprender a partir dos dados, então você deve ter certeza de não optar por ambas.) Se você selecionar seus dados optando por Y , ao invés de aleatoriamente por X , ou de outra forma, você deve aplicar uma correção de controle de caso, tal como descrito em King e Zeng (2001). Na pesquisa quantitativa, estes métodos podem poupar 99% ou mais dos custos da coleta de dados – para que, por exemplo, não se precise mensurar uma nova variável para todos os pares de países para todos os anos do último século (com um n de aproximadamente 1 milhão); ao invés disso, é

possível conseguir aproximadamente o mesmo resultado empírico coletando somente os 1.000 ou pouco mais disto de conflitos internacionais e um conjunto de 1.000 ou 2.000 díades não conflitantes. A mesma questão se aplica na pesquisa qualitativa, mesmo se não for possível enumerar todos os exemplos possíveis dos casos de interesse.

Pesquisadores qualitativos não serão capazes, é lógico, de fazerem essa correção quantitativamente, mas compreender a simples correção estatística facilitará seguir a direção e a magnitude geral do viés e o que puder ser aprendido a partir deste projeto de coleta de dados. Não é difícil e pode poupar enormes recursos de coleta de dados quando os eventos de interesse forem raros.

5.4 Métodos formais para evitar alargamento conceitual

Com todo o esforço gasto em aumentar o número de implicações observáveis de uma teoria, vale a pena parar para garantir que cada observação esteja intimamente relacionada à teoria em questão. Incluir observações só vagamente relacionadas à teoria ou à inferência causal é uma perda de tempo e de recursos na melhor das hipóteses e pode induzir a viés, na pior das hipóteses. Na literatura qualitativa, o termo *alargamento conceitual* refere-se à “distorção que ocorre quando um conceito não se adequa aos novos casos” (COLLIER; MAHON, 1993, p. 845). Para evitar alargamento conceitual, estudiosos qualitativos tentam selecionar casos que se adequam a suas categorias e cuidadosamente ajustam suas categorias ou conceitos para adequar seus casos. Estiramento conceitual não é somente importante em desenhar pesquisa, mas também está no palco central de uma grande disputa intradisciplinar gerada por severa crítica que estudiosos qualitativos mobilizaram a estudos estatísticos transnacionais ao longo dos últimos cinquenta anos (“nenhum ramo da ciência política tem estado em agitação mais extrema do que a política comparada durante os últimos quinze anos”; ver LA PALOMBARA, 1968, p. 52 e GIROSI; KING, 2008, Seção 1.5). Desde esses dias iniciais, estudiosos qualitativos têm ampliado a partir de estudos de área para serem mais comparativos de maneiras mais cuidadosas, e estudiosos quantitativos têm desenvolvido métodos para evitar observações que não pertencem aos mesmos dados. As conexões entre estas duas abordagens não são bem conhecidas nas duas literaturas, mas merecem sê-lo.

Por exemplo, King e Zeng (2006, 2007) provaram matematicamente que quando uma quantidade estatística de interesse estiver longe dos dados, as inferências são mais *dependentes de modelo*, o que significa que mudanças pequenas e indefensáveis em premissas estatísticas podem levar a diferenças inaceitavelmente grandes em conclusões empíricas. Esta prova foi feita para trabalho quantitativo, mas ela também se aplica diretamente ao problema de alargamento conceitual no trabalho qualitativo. Ela dá alguma precisão à noção qualitativa de que quanto mais você estirar um conceito aplicando-o a novos casos a partir de terras distantes (conceituais), mais indefensáveis são as premissas que precisariam ser defendidas e justificadas a fim de ajudar a apoiar uma alegação de que as inferências de alguém ainda são válidas.

Baseando-se nesses resultados, métodos estatísticos agora foram desenvolvidos para avaliar quão longe uma situação contrária ou outro alvo de inferência está a partir de algum

dado ou de todos os dados de alguém, e em alguns casos eles indicam exatamente quão longe é longe demais (HO et al., 2007; KING; ZENG, 2007). Estudiosos quantitativos agora ou mudam rotineiramente a quantidade de interesse a uma que é possível aprender sem muita dependência do modelo ou eles podam seus conjuntos de dados de observação que causam a dependência do modelo. Em algumas situações, este processo de poda muda a quantidade de interesse de maneira semelhante ao que fazem estudiosos qualitativos quando ajustam os dados e os conceitos para se adequarem um ao outro. Parece que muito benefício resultaria tanto para a literatura qualitativa sobre alargamento conceitual quanto para a literatura quantitativa sobre dependência de modelo se estudiosos com orientação metodológica, e pesquisadores aplicados, em ambas as áreas, se tornassem mais conscientes dos desenvolvimentos paralelos e complementares nas duas áreas.

5.5 Documentando regras de seleção

Até mais importante do que métodos de seleção imparcial é a documentação, e disponibilizar a leitores exatamente como você coletou os dados que está oferecendo como evidência. Dessa forma, se você coletar dados de maneira parcial, tudo não está perdido: você ainda pode corrigir e produzir inferências imparciais.

Por exemplo, se você estiver entrevistando autoridades públicas atuais e antigas, nós sabemos que as respostas destas autoridades tenderão a ser parciais de forma a favorecerem eles próprios e as suas administrações. Se ao analisar ou discutir os resultados das entrevistas você não revelar para que administração cada entrevistado trabalhava, então há uma óbvia fonte de viés inferencial. Se, ao invés disso, nós meramente disponibilizássemos esta informação, então pelo menos um pouco do viés poderia ser corrigido. É claro, atentando para a relação entre onde as pessoas sentam e onde elas ficam de pé, nós o fazemos quase automaticamente toda vez. Mas só é possível corrigir esse viés se soubermos do processo pelo qual os dados foram coletados. E se a análise foi qualitativa ou quantitativa é irrelevante; as mesmas questões se aplicam.

Esse ponto é muito geral: Uma inferência válida não pode ser feita a partir de um determinado conjunto de dados sem conhecer toda a cadeia de evidência que leva do mundo às conclusões qualitativas no trabalho final escrito. Qualquer quebra na cadeia de evidência, qualquer informação que esteja faltando acerca de como os assuntos foram escolhidos ou como a informação foi descoberta, pode levar a grandes vieses para um determinado alvo inferencial.

Pesquisadores quantitativos têm agora uma boa ciência dessa questão e têm se movimentado para consertá-la de maneiras institucionais. O movimento de replicabilidade ou de compartilhamento de dados tem se espalhado ao longo das ciências físicas, naturais e sociais. Este movimento apoia o que é conhecido como “O *padrão de replicabilidade* [que] afirma que existe informação suficiente com a qual se pode compreender, avaliar e basear em trabalho prévio se um terceiro puder replicar os resultados sem informações adicionais do autor” (KING, 1995). Alguns periódicos pedem aos autores para incluírem detalhes o bastante em seu trabalho publicado acerca da fonte de dados e dos critérios de seleção. Em outros tipos de saberes, limitações de página tornam isto inviável; assim, muitos periódicos solicitam, e outros incentivam fortemente, estudiosos a criarem e distribuírem “conjuntos de dados de replicabilidade” juntamente com

cada artigo publicado (FIENBERG; MARTIN; STRAF, 1985; KING, 2007). Muitos destes conjuntos de dados de replicabilidade são frequentemente analisados por outros estudiosos; e alunos, até mesmo em trabalhos escolares, têm a possibilidade de irem direto à vanguarda da pesquisa acadêmica via atividades de replicabilidade e geralmente eles próprios publicam novos artigos (KING, 2006).

É claro, juntar conjuntos de dados de replicabilidade para projetos de pesquisa qualitativa pode ser muito mais difícil do que transferir um conjunto de dados numéricos, alguma documentação ou um código de computador. Muitas das vantagens de explorar um arquivo, conduzir uma etnografia ou fazer uma observação participante vêm de aprender as questões mais interessantes para se perguntar ou de descobrir fontes de informação não previstas *ex ante*. Estas dificuldades podem explicar porque pesquisadores qualitativos parecem dedicar menos (e geralmente insuficiente) tempo e esforço para tornar pública a cadeia de evidência para seus estudos. E poucos disponibilizam conjuntos de dados de replicabilidade.

Porém, pesquisadores qualitativos podem disponibilizar facilmente alguns tipos de anotações de campo, entrevistas orais e *videotapes*, e agora existem procedimentos bem estabelecidos para compartilhar seletivamente este tipo de informação com pesquisadores, não importa quão confidencial, patenteada, sensível ou secreta essa informação possa ser (ver em <http://thedata.org>). Arquivos profissionais foram estabelecidos para aceitar e preservar permanentemente essa informação (tais como *ESDS Qualidata* e *Arquivo de Pesquisa Murray*).

Assim, disponibilizar dados qualitativos é certamente possível, embora a natureza profunda, diversa e extensa da evidência em muitos estudos qualitativos dificulte preparar um conjunto completo de dados de replicabilidade. Mesmo assim, fazer o esforço em cada caso, desenvolver facilidades para ajudar pesquisadores a disponibilizarem seus dados e ensinar aos alunos a fazê-lo é crucial por dois motivos fundamentais. Primeiro, e mais obviamente, uma pesquisa pela qual a cadeia de evidência esteja irreparavelmente quebrada é em si mesma falha: se os leitores não conhecem o processo pelo qual o pesquisador chegou a observar os dados apresentados, eles não podem avaliar por si próprios se acreditam ou não nas conclusões inferidas com base nesses dados. E, acreditar: a impossibilidade de verificação não é uma base séria para um empreendimento científico.

Segundo, desenvolver procedimentos mais fáceis para pesquisadores disponibilizarem dados qualitativos é crucial, não somente para pesquisadores qualitativos, mas também para pesquisadores quantitativos. Já que todos os estudos quantitativos são também qualitativos até um determinado grau, todos os cientistas sociais poderiam utilizar grande ajuda em aprender como gravar, documentar, sistematizar, indexar e preservar a informação qualitativa associada a seus estudos, e os pesquisadores qualitativos estão bem posicionados para liderar esse esforço.

De fato, a tecnologia está transformando muitos tipos de esforços quantitativos sem estudos em mais qualitativos e mais significativos. Por exemplo, imagine informações contínuas de localização temporal a partir de aparelhos celulares conectados a respostas de investigações, informações acerca de proximidade com outros usuários de celulares, dados transformados de voz para texto e assim analisados, fotografias tiradas a partir das câmeras de celulares, *e-mails* e outras mensagens de texto enviadas e recebidas, e chamadas telefônicas feitas. A

profunda informação contextual disponível a partir deste tipo de informação, e muitas outras que surgem como ela, precisa enormemente de estudo a partir de uma perspectiva da replicabilidade, compartilhamento de dados e preservação – assim como fazer as perguntas corretas e coletar a evidência mais útil. Como outro exemplo, investigações amostrais podem gravar entrevistas, e nós podemos preservar a fita de áudio. Mesmo os resumos transnacionais mais básicos de dados quantitativos comparados têm por trás inúmeras histórias qualitativas não totalmente documentadas acerca de como cada dado foi criado e mensurado. Pesquisadores qualitativos e quantitativos precisam entender como melhor documentar, preservar e distribuir todos esses e outros tipos de informação qualitativa, pois todos formam partes cruciais da cadeia de evidência para análises de ciência social.

6 Conclusões

As questões e erros que realçamos na literatura de métodos qualitativos e nos usos de métodos qualitativos nos campos substantivos aplicados de ciência política, têm análogos na ciência política quantitativa. Todas essas questões que tratamos são prontamente fixadas, entendendo a teoria de inferência subjacente a todo o nosso trabalho, e na maioria dos casos conectando trabalho em métodos qualitativos com métodos quantitativos mais bem desenvolvidos ou por trabalho em métodos quantitativos com abordagens qualitativas mais bem desenvolvidas. Falamos um pouco acerca de questões de dependência da trajetória, seleção de caso, viés de seleção, implicações observáveis, obstáculos à inferência, causalidade, seleção, formação de teoria, implicações de viés, a relação entre métodos qualitativos e quantitativos, e várias outras. Mas estes são somente exemplos. Muitas outras tais questões ainda permanecem, e nós encorajamos estudiosos a descobrirem-nas, realçar as semelhanças e ligações através de nossos campos díspares, e a continuarem a melhorar a pesquisa empírica em toda a nossa disciplina.

Referências

ADCOCK, Robertand; COLLIER, David. Measurement validity: a shared standard for qualitative and quantitative research. *American Political Science Review*, v. 95, n. 3, Sept. 2001, p. 529-546.

AYRES, Ian. *Supercrunchers*. New York: Random House, 2007.

BECK, Nathaniel; KATZ, Jonathan. Nuisance vs. Substance: specifying and estimating time-series-cross-section model. *Political Analysis*, v. VI, p. 1-36, 1996.

BRAUMOELLER, Bear F. Causal complexity and the study of politics. *Political Analysis*, v. 11, 2003, p. 209-233.

BRAUMOELLER, Bear; GOERTZ, Gary. The methodology of necessary conditions. *American Journal of Political Science*, v. 44, n. 4, Oct. 2000, p. 844-858.

Conexão Política, Teresina v. 4, n. 2, 9 – 45, jul./dez. 2015

CAMPBELL, James E. Introduction: assessment softer 2004 presidential vote forecasts. *PS: Political Science & Politics*, v. 38, 2005, p. 23-24.

COLLIER, David; MAHON JR., James E. Conceptual “stretching” revisited. *American Political Science Review*, v. 87, n. 4, Dec. 1993, p. 845-855.

DUNEIER, Mitchell. *How not to lie with ethnography*. Technical report Princeton University, 2008.

ELMAN, Colin. Explanatory typologies in qualitative studies of international politics. *International Organization*, v. 59, n. 2, Spring 2005, p. 293-326.

FIENBERG, Stephen E.; MARTIN, Margaret E.; STRAF, Miron L. *Sharing research data*. Washington, D.C.: National Academy Press. 1985.

GELMAN, Andrew; KING, Gary. Why are american presidential election campaign polls so variable when votes are so predictable? *British Journal of Political Science*, v. 23, n. 1, Oct. 1993, p. 409-451. Disponível em: <<http://gking.harvard.edu/files/abs/variable-abs.shtml>>.

GEORGE, A. L.; BENNETT, A. *Case studies and theory development in the social sciences*. Cambridge, MA: Mit Press, 2005.

GERRING, John. *Case study research: principles and practices*. New York: Cambridge University Press, 2007.

GILL, Jeff. *Bayesian Methods: a social and behavioral sciences approach*. 2. ed. Chapman & Hall/CRC, 2008.

GIROSI, Federico; KING, Gary. *Demographic fore casting*. Princeton: Princeton University Press, 2008. Disponível em: <<http://gking.harvard.edu/files/smooth/>>.

GLYNN, Adamand; QUINN, Kevin. *Non-parametric mechanisms and causal modeling*. Technical report Harvard, 2008.

GOERTZ, Gary. The substantive importance of necessary condition hypotheses. In: *Necessary conditions: theory, methodology, and applications*. GOERTZ, Gary; STARR, Harvey (Ed.). Lanham, MD: Rowman & Littlefield, 2003.

GOLDTHORPE, J. H. Causation, statistics, and sociology. *European Sociological Review*, v. 17, n. 1, 2001, p. 1-20.

GRANATO, Jim; SCIOLI, Frank. Puzzles, proverbs and omega matrices: the scientific and social significance of Empirical Implications of Theoretical Models (EITM). *Perspectives on Politics*, v. 2, n. 2, 2004, p. 313-323.

GROVE, William M. Clinical versus statistical prediction: the contribution of Paul E. Meehl. *Journal of Clinical Psychology*, v. 61, n. 10, 2005, p. 1233-1243.

HALL, Peter A. Forth coming: path dependence. In: KING, Gary; SCHOLZMAN, Kay; NIE, Norman (Ed.). *The future of political science: 100 perspectives*. New York: Routledge. 2009.

_____. Systematic process analysis: when and how to use it. *European Management Review*, v. 3, n. 1, Spring 2006, p. 24-31.

HAMILTON, James Douglas. *Time series analysis*. Princeton: Princeton University Press, 1994.

HECKMAN, James J. Econometric causality. Technical Report 13934. *The National Bureau of Economic Research*, Cambridge, MA, 2008. Disponível em: <<http://www.nber.org/papers/w13934>>.

HO, Daniel et al. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, v. 15, p. 199-236, 2007. Disponível em: <<http://gking.harvard.edu/files/abs/matchp-abs.shtml>>.

HOLLAND, Paul W. Statistics and causal inference. *Journal of the American Statistical Association*, v. 81, 1986, p. 945-960.

HUFF, Darrell. *How to lie with statistics*. New York: WW Norton & Company, 1954.

IMAI, Kosuke; KING, Gary; LAU, Olivia. Toward a common framework for statistical analysis and development. Disponível em: <<http://gking.harvard.edu/files/abs/z-abs.shtml>. 2007>.

IMAI, Kosuke; KING, Gary; STUART, Elizabeth. Misunderstandings among experimentalists and observation a lists about causal inference. *Journal of the Royal Statistical Society*, Series A171, part 2, p. 481-502. Disponível em: <<http://gking.harvard.edu/files/abs/matchse-abs.shtml>. 2008>.

KING, Gary. Replication, replication. *PS: Political Science and Politics*, v. 28, n. 3, Sept. 1995, p. 443-499. Disponível em: <<http://gking.harvard.edu/files/abs/replication-abs.shtml>>.

_____. An introduction to the data verse network as an infrastructure for data sharing. *Sociological Methods and Research*, v. 36, n. 2, 2007, p. 173-199. Disponível em: <<http://gking.harvard.edu/files/abs/dvn-abs.shtml>>.

_____. Publication, publication. *PS: Political Science and Politics*, v. 39, n. 1, Jan. 2006, p. 119-125. Disponível em: <<http://gking.harvard.edu/files/abs/paperspub-abs.shtml>>.

_____; ZENG, Langche. Explaining rare events in international relations. *International Organization*, v. 55, n. 3, Summer 2001, p. 693-715. Disponível em: <<http://gking.harvard.edu/files/abs/baby0s-abs.shtml>>.

_____;_____. The dangers of extreme counter factuals. *Political Analysis*, v. 14, n. 2, 2006, p. 131-159. Disponível em: <<http://gking.harvard.edu/files/abs/counterft-abs.shtml>>.

Conexão Política, Teresina v. 4, n. 2, 9 – 45, jul./dez. 2015

____; _____. When can history be our guide? The pit-falls of counterfactual inference. *International Studies Quarterly*, Mar. 2007, p. 183-210. Disponível em: <<http://gking.harvard.edu/files/abs/counterf-abs.shtml>>.

KING, Gary, Robert O. Keohane and Sidney Verba. *Designing social inquiry: scientific inference in qualitative research*. Princeton: Princeton University Press. 1994. Disponível em: <<http://www.pupress.princeton.edu/titles/5458.html>>.

KING, Gary; LU, Ying. Verbal autopsy methods with multiple causes of death. *Statistical Science*, v. 23, n. 1, 2008, p. 78-91. Disponível em: <<http://gking.harvard.edu/files/abs/vamc-abs.shtml>>.

LA PALOMBARA, Joseph. Macrotheories and microapplications in comparative politics: a widening chasm. *Comparative Politics*, Oct. 1968, p. 52-78.

LIEBERMAN, Evan S. Nested analysis as a mixed method strategy for comparative research. *American Political Science Review*, v. 99, n. 3, Aug. 2005, p. 435-452.

LITTLE, Daniel. *Varieties of social explanation: an introduction to the philosophy of social science*. Oxford: West view Press, 1991.

MAHONEY, James. Toward a Unified Theory of Causality. *Comparative Political studies*, v. 41, n. 4/5, Apr./May 2008, p. 412-436.

____; GOERTZ, Gary. A tale of two cultures: contrasting quantitative and qualitative research. *Political Analysis*, v. 14, n. 3, 2006, p. 227-249.

MARTIN, A. D. et al. Competing approaches to predicting Supreme Court decision making. *Perspectives on Politics*, v. 2, n. 4, Dec. 2004, p. 761-767.

MEEHL, Paul E. (1954). *Clinical versus statistical prediction: a theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press, 2004.

MORGAN, Stephen L.; WINSHIP, Christopher. *Counterfactuals and causal inference: methods and principles for social research*. Cambridge: Cambridge University Press, 2007.

PIERSON, Paul. Increasing returns, path dependence and the study of politics. *American Political Science Review*, June 2000, p. 251-268.

STEELE, J. Michael. Darrell huff and fifty years of how to lie with statistics. *Statistical Science*, v. 20, n. 3, 2005, p. 205-209.

TETLOCK, Philip E. *Expert political judgment: how good is it? how can we know?* Princeton: Princeton University Press. 2005.