DOI: 10.1177/0094306116664521
http://cs.sagepub.com

FEATURED ESSAY

Toward a Nightmare-Resistant Sociology

JOHN LEVI MARTIN
University of Chicago
jlmartin@uchicago.edu

What Do Sociologists Do All Day?

American sociologists have a number of things in common. They find, or create, data. They focus on some parts of the data in some structured projections, often by comparing cases to one another. They interpret the patterns that they produce in the data. And then they develop stories about why certain things have happened. This is a reasonable way to spend one's days. The problem is that at night, they toss and turn, trying to banish from their minds a vision of the way the world may be, one that, or so they assume, would undermine all that they do in their daylight hours.

The Nightmare

Sandra Sociologist is sleeping fitfully. She is almost done with her comparative historical dissertation on the causes of political revolution. But in this incredibly vivid dream, she has been transported to a doppelganger planet, Htrae. All human history on Htrae has unfolded the exact same way it did on Earth. . . until April 16, 1917. But on Htrae, when Lenin steps out of the train and starts trying to gather people around him, few are interested. Russia stabilizes with a bourgeois liberal democracy. Sandra awakes bolt upright. Convinced of the significance of this vision, she resolves to abandon comparative historical approaches, which deny contingency. Instead, she embraces the importance of "stochastic" or "unpredictable" or "agentic" factors and vows to reproduce the causal sequence of her case in close detail.

Her brother, Steve, has a different nightmare. He has done a regression using an incredible new data set and has a better estimate of a parameter than anyone before. It fits a general theory about human action. He's almost done with *his* dissertation. But in this dream, he has read an interesting book about the American South, and so he decides to split his sample by region. He finds that the coefficients in the South and non-South are quite different, and only in the latter does the theory seem to hold. To figure out what is going on, he splits the South into East and West, and again, finds the resulting coefficients to differ even *more* than those from the North and the South did! He keeps splitting his sample, looking for the true relation that will allow him to have a scientific finding, but every time he splits, he finds increasingly different results. Just as he is about to put each individual in a unique category, he shoots up in bed, staring madly ahead, in a cold sweat. He resolves never to split samples, but instead to do a test for dispersion.

Their sister Sarah also is dreaming. She has been interviewing medical students, women and men, about their feelings about family and career—especially how this affects their understanding of what a "future" is. In her dream, she sees Steve interviewing the *same* people she did. And though he phrases the questions a bit differently, he's clearly asking about the same issues she is. And he's doing a pretty good job. There are a few things she could teach him, perhaps, but she can learn from him too. But then she realizes—he's going to come to the *exact opposite conclusions* as she did! She slams awake, hyperventilating. She tries to banish the memory by working on a "methodological appendix" in which she cites other researchers and a few dime-store social theorists as to the "situatednessitality" and the "positionalityness" of all knowledge.

What Is the World We Study

The nightmares that our imaginary compatriots experienced can be seen as due to the possibility, one which we cannot dispel, that the world may have certain characteristics. The first is that events may be *unrepeatable*, or at least, not *necessarily* repeated. If we somehow restarted the world, it could progress the same way up until a point and then diverge from the course it has taken. This is compatible with a fundamentally stochastic nature of the world, but that particular assumption isn't necessary (we might, instead, ascribe the difference to *freedom*). Although we have no good reason to be confident that necessary repetition is a characteristic of the world, for many analytic purposes, we have banked all our money on assuming that it is so. This is because we take the particularities of *outcomes* as extremely weighty in guiding our theoretical constructions.

The second part of our ontology is that there is an end to *heterogeneity*: sooner or later, as we slice up groups, we get to something that allows us to use inter-individual comparisons to make individual-level (if not intra-individual) statements. If we split the world over and over and it always shears off into pieces with different internal relationships, all of our so-called findings are meaningless aggregates, a bit like reporting how many pounds of flour, nails, and wide-lined looseleaf paper you have in your house. So either we need to claim that there are some people who are fundamentally interchangeable—if you know one college-educated, white evangelical with above-average income, you know them all—or that our linear models are correctly specified to be able to adjust for any particular person who might stand a bit between others. But if you think about the people you actually know, you might be struck by how *similar* very ''different'' people are (this Tibetan lama reminds me of Richie in third grade!) and how very *different* two seemingly similar people can be.

The third part of our ontology is that when we interact with the world to gather our data, we can ignore the part that we play in it. If we have done our job right, there is no *interactivity* between ourselves and the information that we get. The good researcher gets the right story, 'nuff said. It seems to me that we can best summarize this ontology according to three slogans, none of which any of us will defend:

1) Whatever is, had to be.
2) You all look alike to *me*.
3) It was like that when I got here.

I go on to argue that these planks are less plausible than their contraries, and accepting them leads to bad science.

What Need We Face?

Whatever Is, Had to Be

As a casual observer, I have been struck by the fact that sociologists in general seem to be no better at predicting future social, political, or economic occurrences than anyone else. I do not think that is because they are stupid or haven't learned what they should. I think it is because of the misleadingness of the assumption *whatever happened, had to be*. It may well be true—perhaps the course of things is fixed, and we have to fulfill the book—but still, the logic encapsulated in this slogan misleads us into thinking that we should be able to use the past to predict future outcomes. Yet we all are familiar with the fact that even when we understand a set of *dynamics*, we may be unable to successfully predict *outcomes*. On the other hand, it must be acknowledged that there are times when the reverse is true—when a system has what Fritz Heider called ''equifinality,'' in that what is fixed is where the system will end up, and not how it will get there. He bids us to consider the difference between a rock rolling *down* a mountain (as it bumps into various other rocks, it will go left or right, like a quincunx, and follow at best a probability distribution) versus a human hiking *up* a mountain (as he comes to obstacles, he will deviate in whatever way he thinks necessary to allow him to reach the summit).

This strongly implies that the first thing we must do is to determine what sort of system we have—the kind where dynamics must be privileged over outcomes (call these ''downhill''), or the kind where outcomes must be privileged over dynamics ("uphill"). It seems that we have some reasonable guidance here. First, despite an endless series of claims to the contrary, there is precious little evidence that history is a recognizably uphill-type process. Second, organic life, in contrast, has many uphill aspects, and its close analogues—organizations, groups, perhaps even stratification systems—also often seem to have uphill aspects. However, it is far easier to make mistakes regarding (putatively) uphill processes than about downhill ones, because, as the outcome is independent of the precise dynamics, we free ourselves from the necessity of close observation. We are unlikely to build a serious science (science understood here in its broadest senses) by focusing our attention on uphill processes. Instead, we would do better to concentrate on downhill dynamics.

You All Look the Same to Me

Rather than develop such defensible models of dynamics, we have often relied on statistics to search for non-zero relationships of covariance, without any strong faith that such an examination would reach invariants. (I turn to non-statistical approaches momentarily.) Indeed, we have learned to scoff at the idea of invariants. But if we aren't pursuing invariants, it is very difficult to know what we *are* doing. All of our statistical tests are about inference—that we can extrapolate a pattern observed within a data set to a wider population. But what if 2017 is different from 2016? If November is different from October? Wednesday different from Tuesday?

Imagine that we are simply trying to estimate the association between employees' having (as opposed to not having) a college degree and being remunerated above the average (as opposed to below the average) and our world is blissfully free from any confounders, so that our measure is, or so we think, a measure of dynamics. Any such measure we make is based on the 2×2 cross-classification of these dichotomies, leading to four sub-totals, and then some function (such as an odds-ratio, or tetrachoric correlation) turns these into a single number. Are we really attempting to estimate "the" true parameter? One person in one cell retires—and the number accordingly changes. Another drops out of high school and gets a job—and the number changes again. A dip in the stock market—and the numbers change again.

Is it *bad* that the numbers change? Does it undermine our quest to use data to understand the world? Hardly. A world in which regression coefficients, and similar statistics, corresponded to true invariants is a world so monstrously insane and unpredictable that I feel sure we would prefer to quit it at the cost of our lives. We are unlikely to build a serious social science pretending that the opposite is true.

Ethnographers face a homologous problem, and they often recognize that there is no homogenous "group" and that talking about "the" group requires an imputation of stability. But they are less likely to problematize the idea that there is "an" individual. What *this* fellow does or says *today* must tell me something about him in general, and, indeed, about people *like* him more generally: members of this particular informal group, or perhaps his class, his nationality, his race, or whatever. But it is also true that personalities generally involve variation—the same person acts one way in *one* situation, and then, later, acts a different way—in the same sort of situation! And personalities themselves change over time. One reason that field workers often fear showing their work to their subjects is that they suspect, rightly, that the subjects will feel trivialized because in the admirable quest of scientific simplification we have snipped away the bits of their selves that did not fit our claims. And just as we don't want to live in a world where regression coefficients accurately explicate dynamics, so we wouldn't want to be the sorts of one-dimensional and predictable people that we often describe.

It Was Like That When I Got Here

Humans may, as thinkers, indeed be simpler than they flatter themselves to be. But what they perhaps lack in complexity they make up for in ambivalence, vagueness, and confusion. That means there is every reason to think their responses are extremely sensitive to the precise constellation of methods used to elicit them. When they speak, they generally speak for a reason, and the reason is an irreducibly social one. People do not simply *answer*—a predetermined response to an applied stimulus—they *explain*. This means that depending on their theory of their interlocutor, they use different words and different arguments, even if they are trying to communicate the same general principle.

Of course, the degree of sensitivity to the particularities of the social situation varies dramatically. There is relatively low (but, for good reason, non-zero) variability in their reports as to the number of children that they have had (even the number of live births!) and very high variability regarding opinions where they have multiple, potentially contradictory sentiments, or limited information, or that involve confusing generalities. Further, some facts that we might want to treat as most ''objective''—Is this person married? What is her race? Did she commit suicide?—are actually the outcomes of *negotiations* between researchers and ''subjects.''

Finally, all but the most minimally socially adept are sensitive to context in their behavior (though not, interestingly, the mere presence of a single observer). Young Samoan women would likely be one way together when alone with a young and delightful Margaret Mead, and quite another way in the presence of their fathers—or an old and decidedly un-delightful Derek Freeman. We have learned that attempts to deny these differences and insist on the one and true story increase the distortions of our science; they do not make it *more* objective.

In sum, there is no good reason to believe these three planks, and good reason to expect the opposite. They seem like a very shaky place to begin a science. And we all know this. How do we respond?

Our Responses

Regarding the first assumption (repeatability), it seems that those who are most likely to accept the flaws in the notion of ''whatever is, had to be'' quite reasonably reject the attempt to deduce the presence of causal factors through rough comparison of what we can call ''elephant causes'' (e.g., a state has run a negative budget for ten years). And they may even be more likely to accept the importance of ''flea causes'' (e.g., a revolution had much to do with a delayed grain shipment). But in attempting to focus on these small causes, and the microhistory, they become even *more* committed to the notion that they should be able to determine why April 19th unfolded the way it did. As John Goldthorpe in particular has emphasized, there is no a priori reason to think that sociology gives us tools to answer these sorts of questions, and good reason to think that it cannot.

Regarding the second assumption (homogeneity), we seem to have a similarly inconsistent approach. Let's first take the case of numerical analyses. While there are simple tests for overdispersion that, in a very few cases, can be used to force us to reject assumptions of homogeneity, these are almost never carried out by sociologists. And that's for good reason—we generally *assume* that our models are incompletely specified. Instead, we allow the present constellation of hostile critics to lead us (a little bit) down one path in the infinite space of predictors, gathering up a few into our basket and stopping when our critics (or the editors) are exhausted. What we don't know can't (in the daylight hours) hurt us.

When it comes to data collected by the sociologist, such as via observation or discussion, the same assumptions are made: heterogeneity must be neutralized or ignored. If a subject is changing, so that he seems to be resisting my attempt to characterize him (by attaching predicates—he *is* this or that), our response is to make this itself a characteristic of *him* (he is mercurial or inconsistent) . . . and not of our methods.

Regarding the third assumption (non-interactivity), with a few notable exceptions, those who are most likely to recognize that we may in fact have an interactive role in the production of data seem least interested in doing anything about it. Anything, that is, other than *talking* about it, the way we do with those things, like death and taxes, or the weather, that have to be simply endured. We seem a bit like surgeons who shake their heads sadly about the danger of sepsis and don't bother washing their hands.

On the other side, users of already-collected data are generally allowed to have absolutely no idea how the data were created. Data that are patent fictions are used without comment to make strong theoretical arguments, and claims that make implausibly strong assumptions about the *error-free* nature of the data are allowed without a second thought. Again, with some notable exceptions, we treat ready-made data like sausage: we appreciate being able to consume it, and the less we know about how it was made, the better.

The Nightmare

Thus our responses to these nightmares are the sort that work well in the day, when the demons are far away. They don't make any particular sense, but so long as we all do them together, we feel safe. Still, this is not enough to banish the specters. The principles behind these nightmare scenarios—unrepeatability, heterogeneity, and interactivity—remain more reasonable than their opposites, and our "responses" are irrelevant to the actual challenges.

And so, to banish the nightmare, we comfort ourselves with the rituals resulting from our fetishizing of nineteenth-century methods, by treating career success as a huggable stuffed animal, and by developing an epistemology that is a philosophic version of Ambien. Most important, we simply don't take our mission seriously enough to proceed in the absence of the convenient assumption that the regularities we seek will be found gift-wrapped by history, by simple correlational statistics, or in the mouths of our respondents. The challenge for the future is to have enough faith that the world possesses regularity that we can let go of the teddy bear and reach out to reality.

I go on to sketch a few general classes of responses to these nightmares and then some tactics that can, it seems to me, be used to great advantage in producing a nightmare-resistant sociology.

Possible Responses

What would it mean to have a social science that was not vulnerable to these nightmares? It would mean, first, resisting the impulse to explain what cannot be explained. It would mean rewarding researchers for *not* having a full story, or a complete model, when the data don't support one. It would mean, second, determining for what (possibly few) purposes humans can be treated as interchangeable. It would mean, third, closely studying the social processes that comprise our investigation and developing some theories of the way in which these interactions produce data. Survey researchers have done a fair amount of this, and conversation analysts continue to do this. But otherwise, most of us simply invent stories about how our presence affects our data, stories that are usually self-congratulatory, narcissistic, and made in the absence of any knowledge of human cognition. Let me amplify these points briefly.

I noted that the lemma of repeatability has encouraged us to focus on outcomes as opposed to dynamics. A focus on the latter is more likely to produce valid results, because the former tend to fold in the complex interaction of many different dynamics, as well as contamination and dumb luck. To examine dynamics, we need to do what other sciences do—to simplify. While some aspects of these dynamics are transportable to laboratory situations, not all are. To handle out-in-the-world dynamics, we must do two things. First, and most important, we must choose *plausible* dynamics to investigate. We are unlikely to be able to get a precise understanding of an imprecise notion, yet many of our ideas (for example, those having to do with conceptual change) are so indefensibly vague that almost anything can be said about them. Second, we must develop *ontologically defensible* models. There is no virtue in forming parsimonious models that contradict what we would claim to be true about human beings. If it isn't true for *you*, it probably isn't worth making it into the heart of your theory about the world.

Where have we seen successes in this approach? For one, in studies of the physical movement of human beings (and also fishes and birds) using models of the animals or humans as self-propelled particles (or, nearly equivalently, akin to a fluid as in the work of Dirk Helbing). Why is this such an effective strategy? First, because it seizes upon limited questions that allow for simple modeling, and second because fish, birds and humans *are* self-propelled. The model is adequate for the question—just as in other sciences—because it looks like for *these* purposes, people, like molecules, can be treated as interchangeable.

And this brings us to the next problem, that of heterogeneity, and the real possibility that there aren't any real parameters to be estimated in the first place, only meaningless weighted averages—nothing that qualifies for the sorts of investigations of dynamics that we might want to privilege. I am not suggesting that we *do* treat what we laughingly call "models" as if they were invariants, nor that we should immediately abandon making such parameters. However, we might do well to refrain from calling any regression-type analysis a "model," retaining this term for actual claims about process that we are willing to defend. The way in which such analyses should inform our theorizing is strictly via falsification. All they can do is shift the burden of proof by knocking out some stories, as opposed to estimating the parameters from one particular story.

Of course, this doesn't help with nonnumeric research. Here we have allowed a proliferation of case studies that we interpret with an equally schizoid consciousness. Fieldworkers will acknowledge that every site is different but at the same time say that they are convinced their results are generalizable (though it is rare that anyone says *to what* they plan to generalize). Instead, we need to respect the virtue of a case study in itself, and seek not only to replicate studies with different researchers, but to explore minor variations. Rather than look for a "very new" site no one has explored, far better to investigate one just a wee bit different from the last person's dissertation. With this, we can develop an understanding of the analytic dimensions on which our cases fall, as opposed to seeing them as either replications of identical worlds or bafflingly incomparable universes of meaning.

Since we cannot avoid using social interaction to gather the lion's share of the data we use, it behooves us to understand, as best we can, the relation between our findings and this social interaction—and to understand it in concrete, prosaic, and scientific terms, not in fantasized, hoity-toity, or narcissistic terms. This cannot be done in ignorance of findings from other fields, but neither can it wholly rely on them. Rather, we need continued methodological investigation in two forms. First, we can conduct laboratory experiments, or have some observers study *other* researchers. But practicing researchers themselves can often conduct *in situ* experiments to see, for example, how the answers they receive in response to questions change over time or across settings, how different members of a group see the same issue or the same event, and in some cases, like William Whyte, to attempt to intervene in ethically unobjectionable ways to test hypotheses.

When it comes to numerical analysis of found data, we are no less needful of scientific understanding of the production of our data. That means taking seriously processes of recall and response, as well as the formation, coding, and, in many cases, negotiation of items, even of our supposedly "hardest" data. And we must also scientifically understand the production of our *results*—which is not the same thing as using them to make an imputation to the world. Especially as researchers reach for more complex (or seemingly complex) methods of analysis, it is necessary that they determine what sorts of patterns in the data might be responsible for the series of numbers that they receive from the computer. In all too many cases, the answer is, "these numbers could result from the processes you think hold in the data, or any one of a number of extremely plausible error processes." Indeed, unpoliced, our techniques will always tend to take us to the realms of social life in which, by the nature of our methods, we can turn ignorance (for example, I can't really tell the difference between one educated college boy and another) into a "finding" (therefore, if two friends share similar tastes, there must be an influence process!). The only way to determine the robustness of our conclusions is to simulate data that would arise under plausible competing processes.

Overall, I believe that this approach can be summarized in terms of a consistent pragmatism. Rather than attempt to compare our knowledge to some "truly real, really true'' God's knowledge and rejecting claims that fall short of that ideal, we begin with the knowledge that we, as everyday actors, *do* have. We want to do two things. First, we want to try to *improve* existing knowledge, to the extent that we can, using the systematic techniques of social science. This seems to me to be so obviously a practicable endeavor that it in itself justifies continued efforts in the line of sociology. Second, we need to develop specialist knowledge to aid in this first task.

For this first effort of pushing knowledge away from inadequacy, the single best thing that we can do is to return to the principles of falsificationism as a theory of science. Sociologists scornfully rejected falsificationism as a philosophy of science because it was descriptively invalid for the physical sciences (as well as seeming on the fuddy-duddy side). But for all that, it may have been *prescriptively valid* for the *social* sciences. Falsificationism is an impressively persuasive system for a science that wants to use rigorous methods to build theory but is unable to construct defensible models. A regression equation of income on education does not give us a number that corresponds to a numerical process in the world. But it *can* be used to shift the burden of doubt from one argument to another. Anyone who claims that income has nothing to do with credentials will need to gather further evidence if this coefficient is significantly positive, while one who claims that income is simply remuneration for skills will need to scramble if the coefficient is found to be near zero. Falsification is a reasonable and defensible set of procedures to organize our techniques where our knowledge and the world finally meet. The sociological rejection of falsification comes, I believe, not because of its obvious inadequacy, but ours—we reject falsificationism because otherwise we fear the danger of having our thoughts falsified. But we should want to *maximize* this danger: if the world lacks the leverage over our thoughts to force us to abandon them even when we don't want to, we cannot expect our thoughts to be relevant to the world.

We do not, however, want to confine ourselves to pushing everyday knowledge away from inadequacy. We also want to hone our experience of the world to make it possible to ask new questions (the second task), questions that only arise because we find that our scientific approach to everyday problems is blocked because of a lack of knowledge regarding these new, second-order, scientific questions. To do this requires that we mathematize those dynamics that can be mathematized and organize the descriptions of those that cannot. But this is still in service of the improvement of general knowledge, and such improvement need not wait for answers to these second-order questions.

Whereto

I close with reference to some families of tactics that I think can help build a sociology that has all the rigor that we now ape in ritualized fashion, while being invulnerable to the nightmarish scenario in which our key ontological assumptions prove to be false. First (and starting with the third nightmare), such a sociology would be one that *begins* from social psychology and refuses to rest comfortably with models of the actor that we are not willing to defend literally. We do not need to accept the arguments made by our sister disciplines, but we should not reject them lightly; rather, we should only reject them (or their implications) if we, on the basis of our *own* research, find the weight of the evidence pointing in a different direction.

Second, there are exciting opportunities for collaborating with psychologists, neurologists, and cognitive scientists, not for the phony-baloney sciencism of watching whether this or that brain area lights up when a subject is asked a question, but regarding the accumulation of defensible and practical models of how people respond to questions, and, even more importantly, *how observers see and notate* what they experience. We have taken the word ''schema'' from Frederic Bartlett but not the key insight about the nature of remembering and its strong implications for social research.

Further, this study is one that must push against any facile assumption of universality (thus bringing us to the third point). Unfortunately, rather than fall back on our convenient assumption of ''culture'' as that marvelous answer-all to problems of variation (where there is great variation between, and next to none within, cultures), a critique of universality requires an inductive examination of differences between—and differences *within*—individuals. This is something that we have only scratched the surface of, but something which we are well qualified to examine. It seems to me that it would be fascinating to have dissertations on the different ways in which the same individuals say what their job is, or how they got married. Further, we need to return to a proposal made only half in jest by Herbert Hyman sixty years ago—rather than immediately have a single interviewer interview multiple subjects, we should have multiple interviewers interview the *same* subject. We certainly should have multiple ethnographers—especially from generally hostile or intellectually competing teams—study the same groups or communities.

Finally, we can move toward techniques that are fundamentally *cartographic* and that attempt to reduce complexity, but that do not distill everything down to a single model. A parameter from a wrong model is meaningless. The reciprocal positions in a map, on the other hand, even if a simplification, are not only meaningful to analysts who may still lack an *explanation*, but can be used to answer a number of different questions with equal success by persons who may hold to very different models of reality. I should note that this in no way assumes the use of numerical methods. Ethnography—the art of drawing a people—began exactly as such a cartographic exercise.

These are certainly not the only ways of building a nightmare-resistant sociology. But they would, if widely adopted, I believe, do much to make sociology a more *serious* endeavor. Further, it is of course obviously true that they are being done all the time by many scholars. Thus we know they are *possible*. My argument is not that they are unheard of now, but that they are not universally prescribed, not generally taught, and not sufficiently rewarded.